Model-based analysis of time-dependent consolidation

A Thesis
Presented to
The Division of School of Psychology
Cardiff University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy (Psychology)

Tamas Andrei Foldes

2023

Approved for the Division
(Psychology)

Professor Penny Lewis

Dr Jiaxiang Zhang

Acknowledgements

This thesis represents not just a solitary voyage but a collective expedition across uncharted academic waters. It is with heartfelt appreciation that I recognize the steadfast crew who have steered this journey toward a safe harbor. First and foremost, I extend my thanks to my supervisor, Dr. Penny Lewis. Her willingness to take a chance on me provided the opportunity to further my education. Her enthusiasm for research, combined with her unwavering patience and feedback, have been cornerstones of my doctoral experience. I am indebted to Dr. Neal Morton, whose expertise in fMRI analysis has greatly enriched my understanding of the field. The comments provided through his review of parts of this thesis have improved the quality of my work. Our conversations have left a lasting mark on my approach to research and analysis. My gratitude equally extends to Dr. Jiaxiang Zhang, who, as my second supervisor gave comprehensive feedback during my mock viva examinations. I owe a debt of gratitude to Dr. Aline Bompas, whose incisive and constructive criticism during the mock viva reflected her thorough engagement with the chapters I presented. Her insightful comments have resonated with me far beyond what I could articulate at the time.

To my colleague, Yihe: Embarking on a PhD journey is no small feat, and I cannot imagine a better buddy to have had by my side, particularly when delving into the complexities of modeling—a field I knew (and sadly still know) so little about. Your friendship and expertise have been invaluable. I haven't forgotten that I owe you a hungarian souvenier! Lorena, your willingness to engage with my lengthy emails and WhatsApp messages about transitive inference have meant a great deal. Your thoughtful responses and encouragement have been a source of clarity. I was also fortunate to have two trusty interns, Alex and Stella, assisting me with the online experiments. Your hard work were instrumental to getting that data, and I sincerely hope that the experience has inspired rather than dissuaded you from the pursuit of online research.

Having people to converse with during the workday can be incredibly energizing. Eva, Isabella, Lucy, Simon, Viviana — without your presence, CUBRIC transforms

into an eerie, self-automated entity with windows that open and close as if by its own volition! Special thanks to Jay and Marta for reading parts of the thesis!

To my beloved sisters, Andrea and Bianca, Words cannot express the depth of my gratitude for your love throughout my life. Andrea, your selfless dedication and assistance have been a constant presence in my journey, from my earliest memories to this very day. Without your guidance and support, I may not have had the opportunity to pursue higher education, let alone achieve the milestone of a PhD. The list of things I am thankful to you for continues to grow with each passing year. Bianca, your resilience and determination in the face of adversity have been an inspiration to me. Witnessing you overcome numerous obstacles and achieve every goal you set for yourself has been truly remarkable. Despite your own challenges, you always find the time to reach out to me, offering encouragement and support. Your presence in my life has been an invaluable source of strength. I am eternally grateful and honored to have you both by my side.

Finally, to Sofia — unquestionably, my most reliable and significant discovery throughout my PhD. Your kind support was indispensable; I sincerely doubt I could have, or indeed would have, completed this without you. Thank you for enduring my highs and lows. I vow to take a respite from pursuing any further PhDs in the foreseeable future.

Preface

The work presented in Chapter 3 related to Experiment 1 and 3 has been previously published as follows:

Foldes, T., Santamaria, L., & Lewis, P. (2023). Sleep-related benefits to transitive inference are modulated by encoding strength and joint rank. Learning & Memory, 30(9), 201.

Chapter 3 is based on my original work published in Learning & Memory, co-authored with Lorena Santamaria and Penny Lewis, with contributions as follows: I was responsible for the data analysis of Experiment 1 and the data collection and analysis of Experiment 3 and wrote the initial draft of the manuscript.

List of Abbreviations

BB2022	Behrens & Bird (2022)
ES	Encoding strength
FD	Frame displacement
GM	Gray matter
HPC	Hippocampus
JR	Joint rank
JRE	Joint rank effect
MP2021	Matorina & Poppen (2021)
MVPA	Multivariate pattern analysis
OR	Odds ratio
RE	Random Effect
ROI	Region of interest
RSA	Representational similarity analysis
SD	Symbolic distance
SDE	Symbolic distance effect
SDMC	Sleep-dependent memory consolidation
TI	Transitive inference
fMRI	Functional magnetic resonance imaging

Table of Contents

List of Tables

List of Figures

Abstract

This dissertation investigates offline reactivation-dependent consolidation and generalization following episodic learning, with a focus on the transitive inference paradigm. Computational modeling, behavioral experiments, mini meta-analysis, and neuroimaging elucidate the mechanisms underlying time and sleep-dependent changes in relational reasoning. A comparative analysis of REMERGE and MINERVA2 models demonstrates the significance of model-based approaches in deciphering similarity-based generalization as a direct consequence of reactivation dynamics. Behavioral experiments examine encoding strength effects on delayed transitive inference, finding sleep-related improvements dependent on pre-sleep encoding. A mini meta-analysis integrates findings across studies, suggesting a positive relationship between sleep and inference though not reaching significance. An fMRI study reveals stronger distance-based representations for remote versus recent latent hierarchies in PPC, mOFC, IFG and PHC during a stimulus localizer. The hippocampus exhibits an inverse pattern, interpreted either as ongoing consolidation of the recent hierarchy or distance-based pattern separation of the remote hierarchy. The shift towards integrated cortical representations is argued to support emergent model-based inference over time, aligning with systems consolidation theory. Overall, results demonstrate reorganization of relational structure over time towards more integrated mnemonic representations, highlighting model-based fMRI analyses for tracking memory transformation. This multifaceted methodology facilitates a rigorous exploration of consolidation and inference, furthering our comprehension of adaptive generalization.

Dedication

I dedicate this thesis to my family, who have always encouraged me to pursue my dreams. I am grateful for their sacrifices and love. The dare say the journey to this point has been long and demanding.

Introduction

Reactivation and replay in both biological and artificial agents offer significant computational advantages. Research has demonstrated that these processes can lead to accelerated learning, reduced forgetting, and the reorganization or augmentation of experiences, further supporting planning and generalization. These reactivations can transpire during both online and offline periods. Specifically, online reactivation pertains to the immediate reactivation of neural activity patterns during wakefulness. In contrast, offline reactivation takes place during rest or sleep intervals, when the brain is disengaged from external tasks. In these moments, the brain revisits and reactivates neural activity patterns initially established during wakeful states.

A significant gap exists in our understanding however of how the brain determines which information to reactivate during its limited offline periods. Specifically, the conditions under which offline reactivation contributes to adaptive generalization remain ambiguous, especially as recent reviews have highlighted the limitations of the extent of offline benefits (Cordi & Rasch, 2021; Lerner & Gluck, 2019a). This uncertainty is not trivial; given the pivotal role offline periods play in memory consolidation, comprehending the mechanisms underlying adaptive consolidation and generalization is paramount. Such understanding can offer insights into optimizing learning strategies and determining the best times to take breaks. Furthermore, it can guide the development of artificial agents designed for continuous learning and might on day serve as reliable assistants in our daily activities. Additionally, a deeper grasp of healthy memory consolidation processes can pave the way for identifying early indicators of consolidation breakdowns, whether due to aging, mental health conditions, or other factors. This knowledge could also usher in innovative applications, such as facilitating learning or unlearning during rest periods, where memory might be more malleable than during wake.

The central aim of this dissertation is to investigate the merit of a model-based analysis of offline reactivation-dependent consolidation following episodic learning. This exploration seeks to understand the potential mechanistic contributions of hippocampal offline reactivations to the generalization observed in animals when engaged in episodic and serial learning tasks. Another pivotal objective is to identify moderating variables that can account for why extended post-learning retention intervals, which encompass “offline reactivations,” sometimes result in noticeable generalization benefits, while in other instances they either have no effect or even lead to decreased performance upon delayed retrieval. Augmenting this empirical research with a quantitative meta-analysis can further validate claims regarding these moderator variables. With a deeper understanding of these conditions, it is anticipated that advancements in pattern analysis of generalization, combined with neuroimaging techniques like fMRI, can be employed to pinpoint not only the “loci” of generalization but also the specific nature and evolution of generalization over time.

This dissertation, at its core, encapsulates my journey in learning diverse methodologies and approaches with the overarching aim of producing research that is both reproducible and replicable. The significance of this work extends beyond its immediate findings, offering broader implications for the field at large. The cognitive modeling of ostensibly simple tasks, such as transitive inference, holds promise. It not only facilitates precise communication through mathematical paradigms among researchers from disparate disciplines but also engenders interdisciplinary adversarial collaborations. Such collaborations can catalyze the design of experiments situated at the intersection of contending formal theories, thereby fostering incremental advancements in the field. The adoption of online experimentation, particularly within the domain of time and sleep-dependent consolidation, represents a relatively nascent approach. Our endeavors in this realm are anticipated to pave the way for future studies, characterized by enhanced statistical robustness. In tandem with this, our concise meta-analysis of the extant datasets serves as a precursor to more expansive meta-analytical endeavors, poised to elucidate the moderators of generalization and provide direction for subsequent research. Furthermore, this dissertation endeavors to illuminate an efficacious methodology for examining representational shifts over time. This is achieved through a within-subject multi-session design that combines remote learning conditions with in-scanner retrieval, employing a localizer task to scrutinize the representational geometry underpinning inference. While each of these methodological innovations might not be unprecedented in isolation, their confluence within this research area is rare. Such a synthesis holds the potential to augment the current state of sleep and memory research.

However, it’s important to note certain limitations that might influence the interpretation or generalizability of the findings. The scope of this dissertation is defined by its concentrated focus on theories of generalization discussed in *Chapter 1*, particularly emphasizing reactivation as the primary underlying mechanism. In terms of tasks, the research is primarily centered on the phenomena of associative inference, with a specific emphasis on transitive inference throughout. Additionally, while the main emphasis is on offline consolidation, this research does not include any direct physiological measures of reactivation during the offline period. The analysis is anchored solely in recall performance, with only minimal measures following immediate recall. Regrettably, none of the experiments undertaken involve measuring immediate post-learning rest or sleep physiology, which could provide deeper insights into the process of offline consolidation.

For this research, a multifaceted methodological approach was adopted to delve into the intricacies of reactivation-dependent generalization. A vector-based memory models were crafted using both MATLAB and Python, facilitating the simulation of specific experiments. These simulations were useful in shedding light on reactivation-dependent generalization, resonating with the overarching goals of this dissertation. On the behavioral analysis front, the methodology predominantly hinges on logistic mixed-models and employs remote web-based experimentation to dissect time-dependent consolidation. Furthermore, univariate random-effects models have been employed for a meta-analysis of the transitive inference findings, as well as a meta-regression analysis of moderator variables. To enhance the depth of the research, a follow-up experiment was analyzed using model-based representational similarity analysis, complemented by fMRI data.

The dissertation is structured to provide a rigorous exploration of the topic at hand. *Chapter 1* commences with a theoretical exposition, delineating both the classical and more recent paradigms that inform our understanding of human generalization. As the discourse advances, attention is directed towards cognitive models of generalization, with an emphasis on those that can produce offline generalization phenomena. In *Chapter 2*, a methodical comparative analysis is undertaken, juxtaposing two salient cognitive models: REMERGE and MINERVA2. This chapter underscores the efficacy of model-based methodologies in the study of generalization, positing MINERVA2 as an exemplar baseline model. Its value lies in its capacity to account for an array of time-dependent findings with a parsimonious set of parameters. *Chapter 3* presents a triad of empirical investigations centered on the details of time and sleep-dependent generalization as they manifest following the transitive inference task. Of these, one is a reanalysis of an extant dataset, while the subsequent two collected as part of this dissertation, adopt divergent design paradigms: the former adhering to the conventional between-subject design and the latter a novel within-subject approach. Transitioning to *Chapter 4*, the narrative engages in a reevaluation of two secondary datasets published in the realm of transitive inference. This chapter ends with a meta-analytical synthesis, amalgamating insights from our three primary experiments, the reinterpreted datasets, and additional published research, thereby facilitating a comprehensive examination of time and sleep-dependent effects. Concluding the dissertation, *Chapter 5* introduces the final empirical endeavor, which harnesses neuroimaging techniques to probe the representational geometry underpinning successful inference at delayed test, building upon the previously piloted within-subject study of transitive inference.

Chapter 1

Overview of literature

1.1 Proposed mechanisms of time and sleep-dependent consolidation

Over a century ago, Müller & Pilzecker (1900) embarked on a pioneering journey into the world of post-encoding memory stabilization. Through their empirical studies, they shed light on a phenomenon: memories, immediately after they are formed, are susceptible to disruption by new incoming information. However, as time progresses, these memories fortify against such retroactive interference. Drawing a conclusion from their observations, they introduced the concept of memory consolidation, portraying it as a slow process set into motion soon after a memory's inception.

This notion of memory consolidation wasn't limited to their findings. Indeed, the effects of brain lesions on human memory soon corroborated their theory. Studies from Korsakoff (1889) , and Ribot (1887) consistently illustrated a temporal gradient in memory loss post-brain damage; newer memories were more susceptible to loss than older, more distant ones. As the century progressed, Hebb (1949) offered a fresh, neurobiological perspective. He theorized that memories reside within neuronal cell assemblies. With time, these neuronal clusters earned the designation 'memory engrams,' a term rooted in Semon's work (Semon, 1921), which aimed to describe the persistent physiological changes accompanying experiences.

Building on these insights, Brenda Milner's groundbreaking investigations into the renowned amnesic patient H.M. brought memory consolidation again into sharp focus (Scoville & Milner, 1957) . After undergoing surgery to alleviate severe epilepsy, which involved bilateral lesions to his hippocampus, H.M. manifested profound anterograde amnesia, rendering him unable to form new declarative memories. Furthermore, he also exhibited retrograde amnesia, albeit to a lesser extent, affecting memories shortly before his surgery. Milner's observations underscored the pivotal role of the medial temporal lobe, and especially the hippocampus, in the consolidation of new memories (but for more nuanced interpretation see Augustinack et al. (2014)). This also highlighted the notion that while newer memories are anchored within structures like the hippocampus, older ones seem to be localized elsewhere in the brain, signifying a distributed memory storage system.

Further expanding on this foundational understanding, additional research on amnesic patients illustrated the intricacies of 'systems consolidation' (Moscovitch, 2012; Squire, 2009). While contrasting in duration to cellular consolidation, systems consolidation can range from mere hours to decades. It offers a deeper exploration into the formation, structuring, and interaction of engrams across expansive neural landscapes, such as the neocortex, and their interconnected networks. In summary, "systems consolidation" refers to the transformative journey of memory engrams. These memory traces, which initially are heavily dependent on specific brain regions, notably the hippocampus, gradually branch out and assimilate into broader neural circuits, reducing their dependence on their original formation sites.

1.1.1 Standard Systems Consolidation Theory

How might one explain the phenomena of severe anterograde amnesia coupled with time-limited retrograde amnesia? Initial perspectives posited that consolidation unfolds over an extended period, emphasizing the enhanced stability and accessibility of newly formed memories. Declarative memory acquisition, as prominently cited by Squire, Genzel, Wixted, & Morris (2015) , was postulated to hinge significantly on the medial temporal lobe. Within this framework, the MTL memory system undertook a transient role, housing new declarative memories temporarily. In contrast, older or more distant memories were thought to be retrievable without the MTL's intervention.

As memories mature, a uni-directional process, predominantly guided and reinforced by the hippocampus, comes into play. This process, potentially through the replay of hippocampal-neocortical ensembles as delineated by M. A. Wilson & McNaughton (1994), serves to fortify connections amongst neocortical ensemble elements. Over time, these connections are bolstered to such an extent that memory retrieval becomes feasible without the need for hippocampal activation.

This culmination point signifies the consolidation process's termination, marking a phase where memories are fetched directly from the neocortex, bypassing the hippocampus entirely. It's posited within system consolidation theory that memory traces are effectively "fixated" as a result of this process. Lending credence to SCT are several further observations, such as the typically observed decline in forgetting rates over extended periods, particularly evident when sleep forms part of the retention interval, as documented by Jenkins & Dallenbach (1924). This can be rationalized by postulating that newly-formed memories are inherently more fragile and hence more susceptible to forgetting, contingent on an active consolidation phase. Aligning with this notion is the evidence highlighting a temporally graded vulnerability to retroactive interference, as reported by Wixted (2004).

Challenges

Despite the significant empirical support for the Standard Systems Consolidation Theory (SCT), several challenges have arisen, throwing into question some of the central tenets of the theory. Firstly, current research on reconsolidation has unveiled that even memories deemed to be consolidated are not set in stone. Instead, they can undergo a form of "destabilization" upon reactivation. This reactivation-dependent instability contradicts the static nature proposed by SCT (Lee, Nader, & Schiller, 2017). Second, contrary to the notion of a diminishing hippocampal role over time, specific hippocampal lesion studies have unearthed findings that contradict this perspective. Notably, some studies have observed the loss of remote episodic memories, implying that the role of the hippocampus (HPC) might be persistant rather than merely transitional (Sanders & Warrington, 1971; Sekeres, Winocur, & Moscovitch, 2018a). Third, parallel to the aforementioned challenges, the Cognitive Map Theory, introduced by O'Keefe & Nadel (1978), underscores the continuous importance of the hippocampus. According to this theory, which stands as one of the seminal views on hippocampal function, allocentric spatial representations remain perpetually dependent on the hippocampus. For readers interested in delving deeper into the critiques and challenges facing SCT, several comprehensive reviews provide invaluable insights: (Kandel, Dudai, & Mayford, 2014; Moscovitch & Gilboa, 2021; Squire et al., 2015).

1.1.2 Multiple Trace Theory

To address the extensive retrograde amnesia observed for episodic memory and the relatively preserved nature of semantic memory, Nadel & Moscovitch (1997) proposed a framework which argued that the Medial Temporal Lobe, with a primary emphasis on the hippocampus, is essential for the retention and retrieval of contextually rich episodic memories, commonly referred to as autobiographical memories. This concept aligns with the differentiation established by Tulving (1972) , where semantic memory encapsulates general knowledge about oneself and the world, and episodic memory pertains to the recollection of personally experienced events.

In this framework, Nadel and Moscovitch introduced the notion that each instance of memory retrieval leads to the re-encoding of that memory as a separate, sparsely-distributed trace within the hippocampus. This trace is accompanied by the integration of new contextual information associated with the retrieval event, giving rise to the term "multiple trace theory" (MTT). These multiple traces become more abundant and broadly distributed within the hippocampus for older memories. This phenomenon is attributed to the increased opportunities for retrieval that older memories have undergone, as opposed to their newer counterparts. While this principle holds true for episodic memories, the resilience of semantic memories to hippocampal damage aligns with both MTT and SCT. Consistent with SCT, MTT also acknowledges that semantic memories, unlike episodic memories, eventually become independent of the hippocampus over time. However this recognition emphasizes that the consolidation process is not uniform for all declarative memories, which contradicts SCT's assumption (Moscovitch & Gilboa, 2021).

At the core of the ongoing debate between SCT and MTT lies the question of the hippocampus' role in temporal gradients of memory loss. Typically, damage or disruption to the hippocampus, whether due to surgical interventions, pharmacological treatments, or advanced techniques such as optogenetics and transgenic methodologies, has resulted in retrograde memory loss. Notably, this memory loss lacks a discernible temporal gradient, as long as context-specificity remains pivotal for performance (Sekeres, Winocur, & Moscovitch (2018a) , but see J. J. Kim & Fanselow (1992)). Since vividness of episodic memories is commonly reduced as a function of retention interval, until the emergence of studies using more sensitive memory assessments and higher anatomical resolution recordings, cognitive neuroscientists must rely on more indirect forms of evidence to discern between consolidation theories.

Functional neuroimaging studies focused on tasks assessing the differential engagement of the hippocampus in retrieving recent and remote episodic memories in neurologically intact individuals (Svoboda, McKinnon, & Levine, 2006) have provided valuable insights. In line with MTT's predictions, investigations into autobiographical memory consistently demonstrate hippocampal activation during the recollection of vivid, context-specific memories spanning various life stages. Notably, the extent of hippocampal activation is influenced by the richness of episodic details rather than the chronological age of the memory. This pattern is supported by research conducted by Moscovitch, Cabeza, Winocur, & Nadel (2016) and Sekeres, Winocur, & Moscovitch (2018a) . However, when the richness of episodic details is not considered, the anticipated reduction in hippocampal activation with the aging of the memory becomes evident (Boccia, Teghil, & Guariglia, 2019; Gilboa, Winocur, Grady, Hevenor, & Moscovitch, 2004).

For non-autobiographical memories, findings are more inconclusive. Some studies indicate reduced hippocampal activity during recall (Dandolo & Schwabe, 2018; Du et al., 2019; Furman, Mendelsohn, & Dudai, 2012), while others report the opposite (Bosshardt et al., 2005; Gais et al., 2007; J. F. Smith et al., 2010), or even a lack of a univariate relationship between hippocampal activation and time since encoding (Tallman, Clark, & Smith, 2022). A recent analysis of the Natural Scenes Dataset, involving weekly 7-Tesla fMRI scans of eight subjects over 300 days during a recognition memory task, revealed a consistent but modest activation in the anterior hippocampus over time. This finding aligns with MTT's principles (Vanasse et al., 2022). Moreover, multivariate pattern analyses (MVPA) demonstrated that the combined activity of the Medial Temporal Lobe and visual cortex regions achieved superior classification accuracy compared to the visual cortex alone. This suggests a lack of representational transfer from the MTL to the neocortex within the examined timeframe, which contrasts with predictions by SCT.

Furthermore, fMRI not only reveals brain regions activated during the processing of diverse information but also offers insights into the functional connectivity between nodes within neural networks. This capability has facilitated the exploration of functional network connectivity associated with memory transformation and consolidation. Longitudinal studies aimed at tracking representations over periods ranging from days to several weeks have been pivotal in this pursuit. Within this body of research, certain investigations have indicated an persistent connectivity between the medial prefrontal cortex (mPFC) and the hippocampus, a phenomenon predictive of memory retention (Sheldon & Levine, 2013; Söderlund, Moscovitch, Kumar, Mandic, & Levine, 2012). Conversely, other studies (Kesteren, Fernández, Norris, & Hermans, 2010; Kesteren & Meeter, 2020; Squire et al., 2010; Sterpenich et al., 2009) have proposed that interactions among cortical regions are more adept at predicting memory recall, more in line with SCT. A potential explanation for this disparity lies in the nature of the learned task information. Variations between tasks that are either arbitrary or grounded in pre-existing knowledge may account for the divergent findings (Gilboa & Moscovitch, 2021). Note however, in the realm of retrospective autobiographical memory, connectivity studies have largely revealed a shared pattern: hippocampal connectivity with cortical structures remains a common feature, regardless of memory age. This observation persists when memories are equated in terms of qualities such as vividness, rehearsal, and personal significance. These consistent findings provide additional support for a persistent interplay between the hippocampus and cortical regions in the context of memory consolidation (Gilboa & Moscovitch, 2021).

Challanges

While the Multiple Trace Theory (MTT) offers valuable insights into memory consolidation, it is not immune to challenges that warrant careful consideration. A key prediction suggests that hippocampal (HPC) traces should progressively expand with the passage of time or the occurrence of repeated reactivations (Sutherland, Lee, McDonald, & Lehmann, 2020). The physiological recording, calcium imaging, and Immediate Early Genes (IEGs) studies, present a series of findings that run contrary to a fundamental prediction of MTT. This discordance between empirical evidence and MTT's core postulate prompts a reexamination of the theory's framework and implications.

A particularly intriguing challenge emerges from the examination of the relationship between the extent of retrograde amnesia for episodic memory and the size of hippocampal lesions. Contrary to MTT's expectations, instances are observed where the severity of hippocampal damage does not correlate with the degree of retrograde amnesia. In some memory tasks, it becomes evident that substantial HPC damage is requisite to disrupt memories that are only a few days old. This counterintuitive finding raises questions about the precision and specificity of MTT's predictions, particularly those pertaining to the vulnerability of recent memory traces to hippocampal disruption (Sutherland et al., 2020).

Additionally, MTT's predictive scope faces scrutiny when it comes to the interplay between episodic and semantic memories over time. Early consolidation theories largely centered on the extent of hippocampal involvement during the remote retrieval of these memory types. However, the less explored domain of qualitative changes affecting episodic memories over time, particularly the influence of repeated online and offline retrievals, brings with it a fresh set of challenges. Addressing the effects of these retrievals, both online and offline, and understanding how they influence memory representations opens avenues for refining our understanding of memory dynamics (Diekelmann & Born, 2010; Dudai, Karni, & Born, 2015).

Furthermore, MTT, alongside some computational extensions of the Standard Consolidation Theory (SCT), such as the Complementary Learning Systems (CLS), offers mechanisms to extract statistical regularities among events. However, the narrow lens of these theories, characterized by a dichotomous view of declarative memory, hampers a comprehensive grasp of the continuous evolution of memory traces when focusing on their time-dependent evolution. Notably, both MTT and SCT share a limitation in their conceptualization of memory consolidation as a unidirectional process. This perspective overlooks the intricate interactions and interdependencies between episodic and semantic memory, hindering a holistic understanding of memory dynamics. Moreover, both theories tend to treat the hippocampus as a unified entity, disregarding its intrinsic anatomical differentiation. Recent investigations have illuminated the functional specialization along the longitudinal axis of the hippocampus, spanning from its anterior to posterior regions. This functional diversity, coupled with emerging insights into the roles of hippocampal subfields, accentuates the need for a nuanced perspective that acknowledges the complexity of the hippocampal region and its contributions to memory consolidation (Maguire & Mullally, 2013; Robin & Moscovitch, 2017; Sekeres, Winocur, & Moscovitch, 2018a; Zeidman & Maguire, 2016).

1.1.3 Modern theories of system consolidation

Trace Transformation Theory

In response to the challenges and complexities observed within memory consolidation, more recent theories have emerged to offer nuanced perspectives and address the aforementioned issues. Among these theories, the Trace Transformation Theory (TTT), as proposed by Robin & Moscovitch (2017), presents a notable extension of the Multiple Trace Theory (MTT) framework. TTT attempts to account for systems consolidation by introducing a multidirectional and dynamic model that accounts for the coexistence and interaction of diverse forms of memory.

Unlike the conventional notion of unidirectional time-dependent consolidation, TTT posits that memory consolidation involves a multidirectional and ongoing process characterized by interactions between different memory representations. Specifically, TTT delineates four distinct forms of memory: details and gist, which are subcategories within episodic memory that vary in the extent of episodic details, and schema and semantics, subcategories of semantic memory. Schema refers to an integrated schematic representation of commonalities across similar events, while semantics pertain to conceptual knowledge unrelated to specific episodes. The strength or dominance of each representation, and consequently the form of memory expressed, is influenced by temporal progression and task demands during encoding and retrieval.

Crucially, each of these four psychological representations are underpinned by unique neurobiological substrates and processes, driving the dynamics of memory. For instance, posterior hippocampus (pHPC)-posterior cortical neural ensembles underpin detailed contextual representations, while anterior hippocampus (aHPC) interactions with ventromedial prefrontal cortex (vmPFC) give rise to coarser-level event-specific gist representations. Schematic representations mediated by vmPFC, interacting with posterior and lateral cortical regions, encapsulate statistical regularities. Meanwhile, the extraction and integration of decontextualized semantic information involve interactions between the anterior temporal lobe (ATL) and posterior cortical regions.

It is noteworthy that, within the TTT framework, systems consolidation encompasses more than a mere transition of hippocampal involvement to the neocortex over time. Instead, it is conceptualized as a continuous process marked by hippocampal-neocortical interactions that orchestrate memory organization and expression, commencing even before memory acquisition and persisting throughout an individual's lifespan. Therefore, the authors argue that the term "memory systems reorganization" more aptly captures the essence of systems consolidation as portrayed by TTT. For an exhaustive overview of empirical support, encompassing univariate, multivariate, and functional connectivity findings that both substantiate and contrast TTT, refer to Moscovitch & Gilboa (2021).

However, TTT is also not without its vulnerabilities. Some empirical studies have reported an increase in posterior hippocampus (pHPC) activation over time. Notably, Bonnici et al. (Bonnici et al., 2012; Bonnici & Maguire, 2018) utilized a multivoxel pattern analysis approach to discern memory-specific activity during retrieval. Intriguingly, memories spanning a range of 2 to 12 years were differentially identified by clusters of voxels in the posterior HPC, often including the same memories from 2 years prior, exhibiting similar levels of detail and vividness (Bonnici & Maguire, 2018).

Furthermore, a significant observation within the domain of hippocampal involvement relates to the activation of the hippocampus (HPC) during the mental simulation of events that were not directly experienced. This phenomenon, as evidenced by the study conducted by Maguire & Hassabis (2011), prompted subsequent researchers Barry & Maguire (2019), to draw compelling conclusions regarding the nature of hippocampal activations in relation to remote memories, specifically those aged at least 2 years or more. Barry & Maguire (2019) assert that the instances of hippocampal activations witnessed in the context of recalling remote memories do not necessarily signify the reactivation of specific hippocampal circuits associated with those memories. Rather, they propose an alternative interpretation – that these activations are attributable to a broader and more general involvement of the hippocampus in the construction of three-dimensional (3D) scenes, which contrast with MTT/TTT interpretations.

Finally both MTT and TTT introduce perspectives wherein the extent of consolidation is influenced by the nature of memory content as a function of episodic detail, but do not provide quantitative criteria that would dictate what content will consolidate, nor do they expound upon why this content-specificity is behaviorally advantageous (Sun, Advani, Spruston, Saxe, & Fitzgerald, 2023).

Complementary Learning Systems

Arguably one of the most impactful computational models in the realm of systems consolidation is the Complementary Learning Systems (CLS) model, as formulated by McClelland, McNaughton, & O'Reilly (1995) . This model proposes a dynamic process through which memories are initially encoded swiftly within the hippocampus to prevent catastrophic interference with older memories and to safeguard newly-acquired memories from loss (McCloskey & Cohen, 1989). The fundamental premise of the CLS model is predicated on the idea that the rapid acquisition and retention of new information is a task suited to the hippocampus, which then facilitates the gradual integration of these memory traces into neocortical representations over extended periods. Such integration ultimately grants independence upon the newly-acquired memories from the hippocampus.

However, the CLS model is not devoid of limitations. One notable limitation pertains to its adherence to a dual-systems perspective of memory consolidation. While this perspective has been instrumental in catalyzing computational investigations into memory consolidation and has significantly influenced biologically-inspired artificial intelligence models of continual learning (Parisi, Kemker, Part, Kanan, & Wermter, 2019), it imparts a constriction when attempting to comprehend the roles of diverse cortical areas in the consolidation process. The CLS model's focus on semantic memories is another noteworthy limitation. It largely neglects episodic memories, which in turn hampers its capacity to account for challenges associated with retrieving remote episodic memories following hippocampal damage (although see (Kumaran &

McClelland, 2012a; Schapiro, Turk-Browne, Botvinick, & Norman, 2016)).

Moreover, the model's characterization of "slow" neocortical learning mediated by the medial prefrontal cortex (mPFC) has been deemed limited in scope. It has been argued that even when not intrinsically tied to prior knowledge, neocortical learning can also be rapid, often occurring concurrently with hippocampal learning. In certain contexts, neocortical learning is essential to facilitate the swift acquisition of memories grounded in hippocampal processes (Takehara-Nishiuchi, 2020 ; S. Brodt et al., 2018; Hebscher, Wing, Ryan, & Gilboa, 2019; Tompary & Davachi, 2017).

Significantly, some of these limitations have been addressed in subsequent updates to the CLS model. These include proposals for bidirectional connections between the medial temporal lobe and the neocortex, the delineation of distinct functional roles within MTL subregions, explorations into hippocampal generalization, and the proposition of rapid schema-dependent consolidation (Kumaran, Hassabis, & McClelland, 2016a). These refinements underscore the model's adaptive nature and its ongoing evolution in response to emerging insights and empirical observations.

Contextual Binding Theory

The Context Binding Model (CBT) presents an alternative perspective on anterograde amnesia in individuals with hippocampal damage, contrasting with the tenets of the Standard Consolidation Theory (SCT). According to CBT, the hippocampus plays a pivotal role in the integration of items and context into episodic memory, with memory deterioration stemming from contextual interference (Yonelinas, Ranganath, Ekstrom, & Wiltgen, 2019a).

The CBT framework capitalizes on the phenomenon of contextual drift – the gradual transformation of environmental, cognitive, and emotional context over time – along with its corresponding neural representations, to conduct systems consolidation (DuBrow, Rouhani, Niv, & Norman, 2017) . Given that the retrieval of episodic memories hinges upon reestablishing the encoding context during recall, contextual drift serves as a contributing factor to memory interference, facilitation, and temporal organization. Notably, recent memories bear greater contextual overlap with new learning, rendering them more susceptible to interference and subsequent forgetting.

Resonating with MTT, the CBT theory posits that the repetition of an item (via re-study or retrieval) leads to its re-encoding alongside novel context information. Also both MTT and CBT align in their contention that the hippocampus is continuously instrumental in the retention of contextually rich episodic memories, a facet distinct from cortical memory (Moscovitch & Gilboa, 2021).

Regarding the influence of sleep, CBT argues that the swift emergence of certain sleep-induced benefits (e.g., <60-minute naps) finds more cogent explanations within its framework compared to SCT (Yonelinas et al., 2019a). CBT suggests that if these benefits were indicative of active consolidation, a substantial proportion of memories that would otherwise have been forgotten would need to be rapidly (within an hour) consolidated into the cortex – a notion at odds with the prevailing understanding of system consolidation as a gradual process spanning years. Moreover, sleep's role in diminishing proactive interference also aligns more harmoniously with CBT. SCT solely addresses the consolidation of recent events, sidestepping considerations related to proactive, future events, again favor CBT in predictive terms.

Furthermore, SCT posits that individuals with hippocampal damage should exhibit accelerated forgetting across delays that encompass periods of sleep. However, empirical evidence suggests that even extensive damage to the medial temporal lobe does not lead to accelerated forgetting during sleep-containing intervals, implying that the hippocampus may not play a causal role in decelerating episodic memory loss (Yonelinas et al., 2019a).

Yet, the CBT model is also not immune to criticism. Notably, it falls short in delineating the roles of schemas and the intricate interactions between the medial prefrontal cortex (mPFC) and the hippocampus during their acquisition and maintenance (Moscovitch & Gilboa, 2021). Furthermore, CBT assigns sleep the role of experiential quiescence, providing an interference-free interval for memory consolidation and posits that replay may reflect primarily context-related residual activity. Additionally, its been argued that observed adaptive replay dynamics related to prioritization of weak, infrequent events or never-experienced trajectories cannot be easily explained by residual activity (Antony & Schapiro, 2019). Although studies on targeted memory activation during sleep suggest that replay is pivotal rather than the mere absence of interference (Hu, Cheng, Chiu, & Paller, 2020), CBT posits that reactivating hippocampal memory traces during an offline period might influence subsequent memory performance, even if the hippocampus does not directly train the cortex as SCT proposes (Yonelinas, Ranganath, Ekstrom, & Wiltgen, 2019b). The existing literature on sleep's effects on memory presents a paradox. Numerous studies have highlighted the role of sleep in decontextualizing memories and promoting the formation of semantic representations (Lewis & Durrant, 2011) . However, conflicting evidence suggests that context-specific memories may be preserved during sleep, as some studies have reported no negative interaction between context memory and sleep on memory performance (R. Cox, Tijdens, Meeter, Sweegers, & Talamini, 2014; Groch et al., 2011;

Jurewicz, Cordi, Staudigl, & Rasch, 2016). In fact, there are even studies indicating that associations between contextual information and items may be strengthened or forgotten less during nap or sleep (R. Cox et al., 2018; Helm, Gujar, Nishida, & Walker, 2011; Kurinec, Whitney, Hinson, Hansen, & Van Dongen, 2021; Lewis, Manning, & Critchley, 2011; Wang, Weber, Zinke, Noack, & Born, 2017). Notably, the exploration of how contexts influence replay dynamics in terms of sleep-dependent benefits has been limited to only a few studies (X. L. Liu, O'Reilly, & Ranganath, 2021; Schechtman, Heilberg, & Paller, 2023b, 2023a; Z. Zhou, Kahana, & Schapiro, 2023) . A deeper understanding of how spatiotemporal and even semantic context impact both spontaneous and triggered reactivation dynamics could be instrumental in adjudicating between the competing theories of Consolidation-Based Theory (CBT) and Synaptic Consolidation Theory (SCT). Furthermore, shedding light on the specific circumstances in which sleep maintains or strengthens the preservation of contextual information, as opposed to fostering the formation of decontextualized memories, would be a significant contribution to resolving the ongoing debate between active and passive theories of sleep-dependent consolidation (Ellenbogen, Payne, & Stickgold, 2006).

Having an all-encompassing theory of systems consolidation is almost tantamount to having a unified theory of memory. In essence, most consolidation theories typically portray memory as changing from an unstable but dynamic state to a stable but inflexible final state. However, such a fixed goal state does not take advantege of the full potential of a memory trace, especially in a changing world where maximal utility requires adaptation. Although considerable advancements have been made in the mathematical modeling of memory consolidation, and biologically inspired artificial intelligence models for continual learning show potential for developing memory systems capable of addressing the stability-plasticity dilemma, the field has not yet fully realized this goal. What remains to be done is to create computational instantiations of each promising theory, fit these models to behavioural and imaging data and test the predictions of each theory that survived scrutiny (Turner, Forstmann, & Steyvers, 2019). Testing these models in ecologically valid paradigms that mirror real-world conditions, with tasks that necessitate some form of generalization, will be particularly beneficial. Such an approach will help to shed light on the mechanisms of memory consolidation that are responsible for our ability to continually adapt to the dynamic environment in which we reside.

1.2 Proposed mechanism for generalization

The phenomenon of generalization, wherein individuals extrapolate information from previous experiences to navigate novel, related situations, exhibits varying degrees of complexity in its inputs, outputs, and underlying mechanisms. At one extremity of this spectrum, we see stimulus generalization during classical conditioning, where an organism conditioned to respond to a specific stimulus also exhibits responses to similar stimuli, even without direct conditioning. This phenomenon is evident even in organisms with simple nervous systems, such as flatworms, where light paired with a shock elicits conditioned responses (Thompson & McConnell, 1955). Conversely, at the other end of the spectrum, humans engage in sophisticated informal reasoning, a semi-structured cognitive process relying on intuition, experiential knowledge, and common sense to arrive at conclusions and resolve problems in daily life. For instance, consider a seasoned therapist providing guidance to a client seeking to feel happy again. The therapist draws upon an intricate model of the client's mental state, prior reactions to challenges, and the depth of trust established between them to formulate a pertinent response (Tannen, Labov, & Fanshel, 1981). However, the realm of investigation typically lies between these two extremes when explored by experimental psychologists studying generalization, with the primary focus remaining on lower levels of complexity to ensure precise experimental control, reproducibility, and potential cross-species comparisons. This allows for the identification of underlying psychological and neural mechanisms (Ghirlanda & Enquist, 2003).

In this discussion, my objective is to examine the various mechanisms possibly underlying generalization, embracing a process-oriented perspective over a problem-centric one (Feeney & Heit, 2010). While my primary focus will be on mechanisms linked to episodic inference tasks—a suite of tasks designed to assess the ability to synthesize discrete episodes or fragments of information in order to infer or deduce novel insights that were not explicitly provided. Consequently, rather than concentrating on the diversity of generalization problems, the process view emphasizes the characterization of psychological processes underpinning observed phenomena (for an alternative approach see Kemp & Jern (2013)). I will not focus on domains of generalization related to vision, language, or motor control, rather on mechanisms that have been argued to lie at the intersection of episodic and semantic memory.

1.2.1 Complementary learning systems

In the realm of memory generalization, theories encompassing encoding processes are compelled to grapple with the intricate balance between fostering generalized knowledge and facilitating the retrieval of individual memories. This quandary parallels the plasticity-stability dilemma within statistical learning theory, wherein plasticity denotes the capacity to assimilate novel information, while stability preserves prior knowledge (Mermillod, Bugaiska, & Bonin, 2013). During the early 1970s, David Marr introduced a dual-system model of the hippocampus to tackle this challenge (Michael E. Hasselmo & Hinman, 2017). His proposition posited a swift encoding mechanism by the hippocampus followed by a gradual transference of these experiences to the neocortex for enduring storage.

A biologically grounded and formal framework, the Complementary Learning Systems theory (CLS), was subsequently introduced as a potential realization of these distinct memory functions within the brain (McClelland et al., 1995). This perspective advocates for the existence of complementary neural learning systems with distinct functions: the hippocampus, serving as the locus of fast-paced learning, is responsible for encoding specific events, whereas other brain regions undergo gradual and incremental synaptic modifications. These widespread neural adaptations contribute to the formation of generalized memory representations that capture the statistical regularities observed across various experiences.

This dichotomous division emerges due to the likelihood of hippocampal memory representations being sparser and less overlapping than their cortical counterparts (Becker, 2005). These representations are engendered through pattern separation, a mechanism that imparts orthogonality to similar input patterns. This process enables the episodic memory system to assimilate akin events through one-trial learning while circumventing interference between them (Yassa & Stark, 2011).

Initial formulations postulated that the gradual learning systems of the neocortex lacked the capacity for swift consolidation, as new associative connections could supersede prior connections, causing catastrophic interference. However, subsequent simulations demonstrated that even distributed networks can swiftly incorporate novel information by altering weights consistently with “prior knowledge” (McClelland, 2013). The notion of “consistency” was subsequently refined, contending that integration speed depends on whether item features align with established dimensions. Rapid learning in the absence of interleaving occurs when projections align with known dimensions, while integrating new dimensions mandates gradual interleaved learning. This shift underscores a feature-level discourse on integration, diverging from an item-level discourse on the timing of integration (McClelland, McNaughton, & Lampinen, 2020). Significantly, this theory posits that to avert interference, old memories impeding new learning must be replayed during subsequent memory consolidation phases.

Furthermore, a recent update and expansion of the theory pertains to the role of the hippocampus in generalization, which was initially attributed to swift encoding, pattern separation, and bolstering new information for robust training of the gradual learning neocortex (Kumaran et al., 2016a) . However, contemporary evidence suggests that hippocampal replay can assume various forms: (i) it can take on an imaginative nature, evidenced by robust hippocampal firing patterns corresponding to diverse alternatives to actual experiences (Comrie, Frank, & Kay, 2022); (ii) it can manifest as reorganization, characterized by the creation of novel shortcut paths that stitch together trajectory components (Gupta, Meer, Touretzky, & Redish, 2010); (iii) it can facilitate look-ahead online planning during goal-directed behaviors (H. Ólafsdóttir, Bush, & Barry, 2018); and (iv) it can exhibit bias to mirror trajectories through rewarded locations in the environment (Bendor & Wilson, 2012). These findings collectively advocate for a pervasive role of hippocampal replay in constructing, revising, and deploying environment representations, thereby prompting revisions in the understanding of the hippocampal circuit's functioning.

Regarding anatomical considerations, intriguingly, the neocortex might not be the sole repository of slow-learning systems within the brain. Emerging evidence implies that competitive interactions occur between the basal ganglia and hippocampal memory systems (Poldrack et al., 2001; Poldrack & Packard, 2003). Deactivation of the hippocampus seems to amplify engagement of the implicit basal ganglia system in various scenarios such as transitive inference (Frank, O'Reilly, & Curran, 2006), passive spatial learning (Poulter et al., 2019), and motor skill acquisition (Brown & Robertson, 2007).

In the realm of modeling, three notable strands of the Complementary Learning Systems (CLS) framework bear relevance to the study of generalization: First, work centered around "simple" feed-forward, deep linear neural networks has supplied mathematically informed boundaries for understanding semantic cognition formation (Saxe, McClelland, & Ganguli, 2019). Additionally, through experiments employing a comparable artificial neural network to the neocortex in a teacher-student setup, it has been demonstrated that unchecked replay within traditional CLS can lead to overfitting. This highlights the notion that memory consolidation should only transpire when it contributes to generalization rather than mere memorization as in the original CLS. This research offers a normative framework for assessing the conditions under which "regulated" systems consolidation proves advantageous, assuming a predictability monitoring signal is available. This approach argues that moderate signal-to-noise ratios, where learning samples are limited and environmental predictability is moderate, stand to benefit the most from controlled consolidation (i.e.: controlled replay) to avoid overfitting (Sun, Advani, Spruston, Saxe, & Fitzgerald, 2021).

Second, modeling endeavors employing the Emergent neural network simulation environment, which implements the Leabra ("Local, Error-driven and Associative, Biologically Realistic Algorithm") algorithm, have been deeply influenced by CLS. A specific modeling study within this framework demonstrated how sleep can integrate information from temporally separated learning episodes. This study leveraged behavioral outcomes to showcase that sleep can mitigate the detrimental effects of retrieval practice on learning episodes separated by time. It was proposed that during sleep, semantic connections in the neocortex enable the cortex to train the hippocampus by reactivating associations and bypassing temporal context (X. Liu, Ranganath, & O'Reilly, 2022).

Thirdly, within the realm of continual learning in machine learning research, where successive tasks are learned sequentially without catastrophic forgetting, CLS-inspired dual-memory systems have shown competitiveness on various benchmarks (Parisi et al., 2019) . These systems have also provided valuable insights for crafting mathematical theories of generative replay when systems seek to optimally learn in environments approaching real-world complexity (Hayes et al., 2021a; Wittkuhn, Chien, Hall-mcmaster, & Schuck, 2021a).

1.2.2 Memory integration theory

Memory Integration Theory (MIT) proposes a comprehensive framework for understanding how memories are integrated and knowledge is extracted from various events (Eichenbaum, 2003). Unlike earlier theories, MIT emphasizes the rapid involvement of the hippocampus in generalization, a contrast to the role attributed to the hippocampus by the Complementary Learning Systems theory (CLS) ((McClelland et al., 1995), although alternative viewpoints have been presented as well (Kumaran & McClelland, 2012a)).

MIT suggests that the process of extracting associative structures across multiple events, which is essential for the development of schemas, involves reactivating memories of events that share common features and binding novel information with these reactivated memories. Research involving lesion studies in both rodents and humans has demonstrated the critical role of the hippocampus in extracting commonalities across distinct events. While damage to the hippocampus spares memory for individual associations, it impairs the ability to make inferences that necessitate knowledge of the complete hierarchy (Bayley, Gold, Hopkins, & Squire, 2005; Dusek & Eichenbaum, 1997b).

Recent neuroimaging investigations have provided insight into the mechanisms underlying the extraction of knowledge across episodes. One step in forming knowledge schemas involves reactivating previously experienced events that share features with current experiences. When new experiences overlap with prior episodes, hippocampal pattern completion mechanisms reactivate previous memories, facilitating the inference of unobserved relationships (Kuhl, Shah, DuBrow, & Wagner, 2010; Molitor, Sherrill, Morton, Miller, & Preston, 2020; Margaret L. Schlichting & Preston, 2015; Zeithamova, Dominick, & Preston, 2012). Furthermore, human neuroimaging studies utilizing multivariate analysis have revealed that the hippocampus represents indirectly related elements from overlapping events as more similar to one another after learning, signifying cross-episode hippocampal integration (Collin, Milivojevic, & Doeller, 2015; L. Mack & Preston, 2016; Margaret L. Schlichting, Mumford, & Preston, 2015a). Additionally, hippocampal activity patterns encode information about the temporal, spatial, and conceptual distances between multiple objects (Park, Miller, Nili, Ranganath, & Boorman, 2020a; Theves, Fernandez, & Doeller, 2019; Theves, Fernández, & Doeller, 2020; Viganò & Piazza, 2020).

In terms of hippocampus's role in binding items with their contextual information during the encoding process as argued by CBT, recent advances in the field of statistical learning have revealed that hippocampal representations tend to become more similar for items that consistently follow each other in a continuous sequence of events, reflecting their binding in memory (Schapiro, Kustner, & Turk-Browne, 2012). This phenomenon is also observed in spatial navigation, where hippocampal representations become more similar for objects perceived as closer in proximity within a recently learned virtual environment (Deuker, Bellmund, Navarro Schröder, & Doeller, 2016). These forms of representational binding are believed to facilitate predictive reinstatement during memory retrieval, aiding effective decision-making through predictive pattern completion mechanisms that operate in associative, temporal, or spatial domains (Ranganath, 2010), supported by distinct cortical retrieval networks (J. S. Kim & Lee, 2023).

In situations where associative relationships are less reliable, hippocampal representations tend to diverge for items that co-occur intermittently (Schapiro et al., 2012), hinting at a potential mechanism underlying content-specific differentiation. While these observations primarily focus on the representation of directly observed associations, the fundamental mechanisms of binding and reinstatement outlined here lay the groundwork for constructing more intricate associative schemas that capture structural information beyond direct experience. Interestingly, the hippocampus can also assign distinct representations to overlapping associations, as seen in studies by Chanales, Oza, Favila, & Kuhl (2017) and Favila, Chanales, & Kuhl (2016). Moreover, Koster et al. (2018a) discovered indications suggesting that the hippocampus maintains separate event representations during memory encoding, while aiding the integration of overlapping associations during memory retrieval, a perspective that aligns with the REMERGE computational model (Kumaran & McClelland, 2012a) .

Contemporary models should aim to probe the circumstances under which hippocampus-dependent memory integration or differentiation occurs, some contending that the outcome is influenced non-monotonically by competition (Ritvo, Turk-Browne, & Norman, 2019). Lower competition may lead to similar representations of overlapping events in the hippocampus, while higher competition might result in orthogonalized representations (Dimsdale-Zucker et al., 2022). It has been posited that episodic memory and statistical learning can function in parallel within the hippocampus to support either differentiated or integrated memory representations, respectively. For instance, CA1 (MSP) exhibits heightened pattern similarity for related items, indicative of integration, while CA2/3/DG (TSP) concurrently displays distinct pattern differentiations for related items (Dimsdale-Zucker, Ritchey, Ekstrom, Yonelinas, & Ranganath, 2018; Molitor, Sherril, Morton, Miller, & Preston, 2021) .

The growing evidence regarding the hippocampus' role in relational integration presents challenges in determining whether these diverse inference tasks involve an explicit cognitive map or simply the reactivation of associated experiences. Recent support for the notion that cognitive maps underlie "structural inferences" comes from a study demonstrating that a 2D social hierarchy, defined by independent social dimensions of popularity and competence, could be reconstructed from binary comparisons between pairs differing by one rank level on each dimension. Despite not needing to combine the dimensions behaviorally, analyses of BOLD patterns in the hippocampus (HC), entorhinal cortex (EC), and orbitofrontal cortex (OFC) favored a 2D cognitive map, where closer individuals in true 2D social space were represented more similarly (Park, Miller, & Boorman, 2021).

The boundary conditions of this behavioral flexibility, where learning in one context is transferred to plan actions in a novel environment, remain unclear. Two hypotheses have emerged, both centered around the notion that grid cells in the medial entorhinal cortex (MEC) encode the environment's structure and facilitate generalization, but with differing proposed mechanisms (Yu, Park, Sweigart, Boorman, & Nassar, 2021). One hypothesis suggests that grid codes, supported by grid cells in MEC, encode eigenvectors of the online-learned successor representation (SR) during a task (Momennejad, 2020; Stachenfeld, Botvinick, & Gershman, 2017). The SR represents the probability of future occupancy over states based on past experiences, offering a middle ground between model-free and model-based reinforcement learning in terms of both computational efficiency and sample efficiency (Gershman, 2018). The other hypothesis posits that grid codes reflect the inferred global task structure, providing greater flexibility than the SR hypothesis. This perspective assumes that grid codes represent a structural code independent of specific sensory content, enabling the transfer of structural information across tasks. Consequently, the brain could make one-shot inferences without further experience. If the hippocampal-entorhinal cortex (HC-EC) system constructs a cognitive map based not only on experiences in one context but also on separate experiences from multiple contexts, it may develop a combined representation in a shared multidimensional space, assuming a common latent structure exists. This concept is fundamental in contemporary memory integration models like the Tolman-Eichenbaum machine (TEM) (Whittington et al., 2019b), which asserts that the HC encodes both sensory (lateral EC) and structural codes (medial EC), enabling the rapid connection of sensory observations to specific locations and facilitating the transfer of structural knowledge between environments.

In summary, along with TEM, non-formal (Eichenbaum, 2017; Shohamy & Wagner, 2008; Zeithamova & Preston, 2010) and formal models (M. W. Howard, Fotedar, Datey, & Hasselmo, 2005b), along with related perspectives such as the autoencoder model of the hippocampus (Gluck, Meeter, & Myers, 2003), propose that the overlap of hippocampal neural codes for related experiences is crucial for generalization.

Temporal Context Model

The Temporal Context Model (TCM), which originated to explain recency and contiguity effects in free recall tasks, posits that items are linked to a slowly changing temporal context signal (Estes, 1950). This temporal context signal at encoding serves as a cue for retrieval. In TCM, behavioral associations between items are not attributed to direct connections formed between item representations during association. Instead, functional associations in TCM are mediated by the influence of items on temporal context and the temporal context's capacity to cue item retrieval. Simulations of TCM by M. W. Howard, Fotedar, Datey, & Hasselmo (2005c) demonstrated that memory integration arises due to the hippocampus' role in encoding item information relative to the temporal context of the encounter. The model suggests that when learning an association between items A and B, the hippocampus encodes the link between A, B, and a representation of the temporal context. When B is later encountered with item C, the presentation of B triggers reactivation of the temporal context representation, enabling retrieval of A. Thus, the hippocampus forms an "intermediate representation" linking A, B, and C. This process creates a new temporal context in which A becomes associated with both B and C. TCM implies that the hippocampus can link representations of separate experiences by retrieving and associating the representation of temporal context from a previous experience with a new overlapping one.

TCMs have been closely connected to work in cognitive neuroscience. A key prediction is the presence of a temporal context vector "c," displaying temporal autocorrelation spanning macroscopic time periods. Rodent studies show robust signals with autocorrelation over seconds, minutes, and even hours or days in brain regions like the hippocampus and prefrontal cortex (Cai et al., 2016). Another significant TCM prediction is the retrieval of the state of temporal context during episodic memory recall, akin to mental travel. Invasive human recordings have provided evidence of this phenomenon in various memory paradigms (Folkerts, Rutishauser, & Howard, 2018; Umbach et al., 2020), fMRI studies of free recall (Chan, Applegate, Morton, Polyn, & Norman, 2016), and real-world memory over extended periods (Nielson, Smith, Sreekumar, Dennis, & Sederberg, 2015). More recently the model has also been updated to work a wider range of context information (Polyn, Norman, & Kahana, 2009).

In, summary these findings and successful descriptive model simulations suggest that reactivating related memories during learning promotes binding of the of temporally, spatially or associatively neighboring items, leading to formation of integrated memory networks that represent associations beyond direct experience. Such knowledge formation through integration may support prediction and inference in novel situations and likely represents the first step in schema formation (Alison R. Preston & Eichenbaum, 2013). Moreover, some have argued that hippocampus can locally rapidly form cognitive maps, albeit with limited flexibility, whereas others have argued it is the HC-EC together that can slowly learn the map but, when learned, can immediately generalize it. While these are not mutually exclusive possibilities, future experiment will be necessary to arbitrate between them and specify the boundary conditions of when each dominates (Whittington, McCaffary, Bakermans, & Behrens, 2022; Yu et al., 2021) as well as clarify to role of the PFC when interacting with HC-EC circuit (Hok, Save, Lenck-Santini, & Poucet, 2005; Jacobs et al., 2013).

1.2.3 Retrieval-based generalization

While we've discussed potential mechanisms underlying the rapid formation and updating of hippocampal schemas (MIT) and incremental formation and updating of neocortical schemas (CLS), both these models built on seminal work involving schemas which are not directly observable. Retrieval-based generalization asks the question whether we even need to assume that such latent variables even exist in the brain and relatedly whether we even need multiple memory systems when trying to understand generalization? The existence of multiple memory systems, forming distinct kinds of memory representations to serve distinct memory functions, has been the dominant view for some time. The single-system view posits however that multiple memory decisions could be made based on separate representations of individual experiences, without the need for stored generalized representations (Jamieson, Johns, Vokey, & Jones, 2022).

This challenges the conventional notion of distinct memory systems forming various memory representations for specific functions. The single-system view suggests that individuals form separate memories of individual events and subsequently make generalization decisions based on those specific memories depending on task demands (Curtis & Jamieson, 2018; Kinder & Shanks, 2003; Zaki, Nosofsky, Jessup, & Unverzagt, 2003). This view simplifies the role of the hippocampus in both memory specificity and generalization by proposing that the hippocampus encodes individual events that can be flexibly accessed for making judgments of either type.

Formalizing the single-system view for generalization has been primarily done in the study of categorization (Ashby & O'Brien, 2005). Here exemplar models propose that concepts can be represented by specific category examples, allowing generalization to new items by considering exemplars from relevant categories (Nosofsky & Kruschke, 1992; Zaki & Nosofsky, 2001b). This contrasts with prototype models that suggest individuals extract the central tendency across category exemplars to guide categorization decisions (Minda & Smith, 2001; J. D. Smith & Minda, 2001).

Distinguishing between integrated representations and on-demand inference from separate memories poses a challenge in identifying the underlying representations from a behavioral perspective. Some studies suggest that generalization judgments based on integrated representations tend to be more accurate and faster than those relying on demand-driven inference from separate representations (Margaret L. Schlichting, Zeithamova, & Preston, 2014; Shohamy & Wagner, 2008). However, others argue that strong source memory leads to better rather than worse generalization, aligning with exemplar-based models (Banino, Koster, Hassabis, & Kumaran, 2016).

While exemplar models fit behavioral data well, only a couple of studies have provided neural evidence for the single-system view (Ashby & Rosedahl, 2017; Michael L. Mack, Preston, & Love, 2013; Nosofsky, Little, & James, 2011). Recent research suggests that the hippocampus can support memory abstraction and generalization in line with exemplar models, potentially through a separate circuitry within the hippocampus (Schapiro et al., 2016).

MINERVA2

Two prominent exemplar models of generalization are particularly relevant to episodic inference tasks. One of these models is MINERVA2, which falls within a category of "global" memory models, encompassing a diverse spectrum of paradigms including recognition, cued recall, serial recall, and free recall, while making minimal assumptions (Surprenant & Neath, 2013). MINERVA2 operates under the premise of integrating both episodic and semantic memory within a unified framework. This model relies on cues to evoke responses and operates as a distributed memory model, positing that memories are represented as vectors in a high-dimensional space.

Inspired by the Multiple-trace theory, which posits that each memory retains a distinct code or memory trace and maintains a substantial hippocampal involvement (Moscovitch, Nadel, Winocur, Gilboa, & Rosenbaum, 2006; Nadel, Samsonovich, Ryan, & Moscovitch, 2000), MINERVA2 adopts a similar conceptual framework. In the context of this model, recognition and recall processes hinge upon a similarity computation between the given cue and all stored memory traces. The greater the similarity between the current cue and a stored memory trace, the stronger the activation of that particular trace. This mechanism is akin to a high-level instantiation of the Hebbian principle "what fires together wires together." The computed similarities across all traces are then aggregated, resulting in a "summed similarity" that signifies memory strength (aka "echo intensity") in the model. Decision-making in this framework involves comparing the summed similarity to a predetermined threshold to determine the familiarity/recognition of the probe.

The concept of "summed similarity," often referred to as "global matching," encapsulates the notion that recognition judgments are founded upon the cumulative evidence garnered from all traces in memory, rather than being contingent on a solitary, specific trace. This conceptual framework effectively accounts for several phenomena observed in recognition memory research (Clark & Gronlund, 1996). For instance, the frequency effect can be explained as the consequence of items encountered multiple times generating multiple memory traces, thereby contributing to a higher summed similarity and echo intensity (Osth & Dennis, 2020a). Furthermore, the model exhibits noise robustness, as it can successfully recognize items even when certain traces are noisy or the probe exhibits slight alterations. This resilience is attributed to the ability to rely on the collective evidence provided by more accurate traces (Osth & Dennis, 2020a). Moreover, the model effectively addresses the concept of generalization, wherein probes similar but not identical to stored traces can still elicit a recognition decision tilted toward "familiar." This again is made possible through the cumulative effect of the summed similarity derived from multiple traces, enabling the model to generalize from prior experiences. Additionally, incorporating recurrence into the retrieval process allows the model to iteratively utilize its own output (Hintzman, 1986a). For example in a model of an associative inference paradigm, this can result in the retrieval of D item when cued with A after learning AB, BC and CD associations and stepwise reactivating the necessary intermediate representations.

REMERGE

The REMERGE model (Recurrency and Episodic MEmory Results in GEneralization) presents an extended framework to explain the emergence of generalization through neural systems that are optimized for orthogonalizing memories (Kumaran & McClelland, 2012b). Distinct from Minerva2, the REMERGE model posits a unique mechanism for generating on-the-fly generalization. Memories of individual events are discreetly stored within the hippocampus as separate representations. However, unlike Minerva2, where generalization is attributed to a single similarity computation during retrieval, REMERGE's generalization process occurs dynamically in response to task demands through the co-activation or sequential activation of multiple individual memories (Kumaran & McClelland, 2012b).

The architectural foundation of REMERGE combines principles from interactive activation competitive (IAC) networks (McClelland & Rumelhart, 1981) and earlier exemplar memory models, like MINERVA2 (Hintzman, 1986b, 1988). It simplifies the mutlicomponent hippocampal circuitry into two fundamental layers: a feature layer and a conjunctive layer, which correspond broadly to the entorhinal cortex (ERC) and the hippocampus proper, respectively. The pivotal concept in REMERGE is the “recurrent similarity computation”. This computational process enables the model to adequately account for a plethora of rapid generalization phenomena, including recognition memory, categorization, transitive inference, paired associate inference, and acquired equivalence (Kumaran & McClelland, 2012b). Unlike traditional perspectives of the hippocampal system as a unidirectional feedforward circuit, REMERGE introduces the principle of “big-loop” recurrence. This concept involves reciprocal interactions between the hippocampus proper (e.g., DG/CA3) and neocortical regions like the ERC. This interplay permits the cyclic reintegration of output as successive input into the system. This incorporation of recurrence in REMERGE is inspired by anatomical connections existing between the superficial and deep layers of the ERC (Strien, Cappaert, & Witter, 2009), facilitating the dynamic reprocessing of information within the neural circuitry.

Contemporary models related to retrieval-based generalization exhibit significant connections with both the Global matching model and the REMERGE model, presenting valuable insights into the intricacies of memory phenomena. Global matching models have demonstrated remarkable success in accounting for a wide array of memory effects, including list-length, list-strength, and similarity effects, as well as serial position effects and prototype formation. Notably, these models have effectively accounted for the differential rates of forgetting observed for exemplars and prototypes, while also satisfying benchmarks in scenarios involving cued and free recall, thus solidifying their status as comprehensive models of episodic memory (Osth & Dennis, 2020b). Recent advancements have further expanded their applicability by successfully modeling independent sources of interference that impact long-term recognition memory retrieval, encompassing item-noise, context-noise, and background-noise (Fox, Dennis, & Osth, 2020). Moreover, these models have displayed improved fitting with regard to reaction time data (G. Cox & Shiffrin, 2017), and they have incorporated semantic item representations through the integration of word embeddings derived from distributed semantic models (Johns, Jamieson, & Jones, 2023).

Turning to the REMERGE model, the trisynaptic pathway of the C-HORSE hippocampal model showcases striking parallels with classical exemplar models, as it stores distinct traces of individual exemplars. Notably, a reimplementation of REMERGE has revealed how a neural circuit could effectively embody exemplar-style representations (Schapiro, Turk-Browne, Botvinick, & Norman, 2017a). The sparsity-inducing mechanisms that give rise to pattern separation of individual episodic memories within the Trisynaptic pathway correspondingly facilitate pattern separation across diverse exemplars (Schapiro, Turk-Browne, et al., 2017a; Sučević & Schapiro, 2022). Furthermore, the model's monosynaptic pathway has been suggested to function analogously to a prototype model to a certain extent. This notion finds support in initial neuroimaging findings, indicating coexisting neural representations in the hippocampus that possess both prototype-like and exemplar-like features (Zeithamova & Bowman, 2020a). Such a conceptualization is reminiscent of that proposed by the Complementary Learning Systems theory in terms of HPC-neocortex, which outlines the functional roles of these representations.

Relationship to Memory Integration Theory

In the context of the Memory Integration Theory, which asserts that the hippocampal role in generalization and inference extends beyond storing specific memories, several findings have been introduced that appear to contradict a purely "exemplar and retrieval-based" account of generalization.

A plausible resolution to this contradiction is the hypothesis of "stored generalizations." Under this premise, once a generalized representation emerges as an output of the exemplar network, it becomes available as an input to downstream brain regions. Depending on the connectivity between regions, the generalized representation might also be encoded back into the original exemplar network itself. This integration would lead to the generalized representation becoming a more robust component of the network of individual traces. The idea that cueing reactivates prior memories that are then re-encoded as a part of a new memory was promoted by Hintzman himself to account for behavioral phenomena, such as recognition confidence and judgments of frequency (Hintzman, 2004). Koster et al. (2018b) have recently described a neural model that embodies this concept. Notably, a loop recurrence is present within the medial temporal lobe, connecting entorhinal layers that receive hippocampal output with those that provide input to the hippocampus. This arrangement offers a mechanism by which hippocampal output can be recirculated as new input back into the hippocampus.

The question arises as to when these stored generalizations emerge. Do integrated and separated representations evolve concurrently to support distinct functions, or do they develop over time in response to the demands of generalization? Furthermore, is their behavioral relevance a pivotal factor for the mechanistic and representational validity of exemplar-based models of generalization (Zeithamova & Bowman, 2020a)? Behavioral data aligned with the notion that related memories might initially exist as separate representations but become co-activated and interconnected in response to task requirements have been reported within episodic inference paradigms. Instances of such data can be observed in the context of acquired equivalence. Araujo Sanchez & Zeithamova (2023) observed robust generalization learning during testing, even in the absence of feedback on test trials. In the domain of associative inference, Carpenter & Schacter (2017) demonstrated reduced memory for specific contextual details, such as an increased propensity for false attribution of associations, albeit only following successful inference.

To summarize, recent investigations provide compelling evidence supporting the notion that the hippocampus possesses the capacity to facilitate memory abstraction and generalization, thereby aligning with the principles espoused by exemplar models of generalization to a certain degree. This notion finds support in a variety of studies as reviewed by Zeithamova & Bowman (2020a). However, the underlying representational scheme that governs hippocampal generalization remains enigmatic, particularly regarding whether integration takes place during encoding or retrieval processes. While an increasing body of evidence stemming from diverse generalization paradigms suggests that the hippocampal role extends beyond the mere storage of specific memories, encompassing the creation of generalized representations spanning multiple experiences, uncertainties remain concerning the timeline of emergence of these generalized representations within the hippocampus. This uncertainty arises due to conflicting research findings, with some studies supporting MIT and others lending support to the Retrieval-based generalization mechanisms. Furthermore, recent behavioral research has argued that relational semantics associated with task structure (due e.g., instructions) can influence the whether episodic inference corresponds to encoding or retrival-based mechanisms (Hecker, Müller, Kirian Dill, & Christoph Klauer, 2023), raising the possibility that task demands and prior knowledge can determine which of these two processes is active in a given instance.

However, it is important to note that not all instances of generalization are contingent upon hippocampal involvement. Consistent with the tenets of systems consolidation theories, which propose the presence of cortical representations for generalized knowledge, a number of studies indicate that individuals with hippocampal damage are still capable of performing generalization tasks (Knowlton & Squire, 1993; Nieuwenhuis & Takashima, 2011; Winocur & Moscovitch, 2011). This observation suggests the existence of diverse mechanisms for generalizing across experiences, each reliant on distinct brain systems.

1.2.4 Offline generalization, replay, and reactivation

Historically, the domains of neuroscience and psychology have concentrated on studying the human brain during intervals of “online” attention, where participants are actively immersed in the processing of sensory inputs. This is particularly true of generalization research supporting MIT and Retrieval-based generalization models. Interestingly, humans allocate a minimal portion of their time to actively attending to their environment. A significant fraction, approximately one-third of an individual’s life, is reserved for sleep and nearly half of the waking hours are expended in other ‘offline’ modes, marked by activities such as daydreaming, episodes of mind wandering, or phases where an individual is unresponsive to the immediate environment. While the underlying reasons for these offline states remain a topic of discussion, a prevailing hypothesis suggests that the brain requires an offline state for effective consolidation (Wamsley & Summer, 2020) . Pivotal research spanning decades provides evidence that post-learning sleep substantially enhances memory retention in contrast to an equivalent time spent in wakefulness (Berres & Erdfelder, 2021). Recent findings introduce the perspective that even a momentary lapse into an eyes-closed resting state can produce similar behavioral outcomes as sleep. This underscores the significance of viewing states of arousal and attention as a continuum in consolidation studies (Humiston, Tucker, Summer, & Wamsley, 2019; Wamsley, 2022; Wamsley, Arora, Gibson, Powell, & Collins, 2023).

An analogous differentiation can be implemented in the exploration of generalization. Memory representations may undergo a transformation from being specific to adopting a more generalized or integrated format. This metamorphosis can transpire during the actual learning process, denoted as online, or it can manifest spontaneously during intervals of rest or sleep, termed as offline. Furthermore, a synergy or interleaving of both online and offline processes may also contribute to this transformation (Roscow, Chua, Costa, Jones, & Lepora, 2021; Singh, Norman, & Schapiro, 2022) . Reactivation emerges as a dominant mechanism underpinning offline generalization, aligning with previously discussed models of generalization. However, it is crucial to acknowledge that this is not the sole mechanism responsible for such generalizations (Svenja Brodt, Inostroza, Niethard, & Born, 2023). Reactivation can be categorized based on its occurrence: while an experience is ongoing (online replay) or during subsequent periods of rest or sleep (offline replay).

SPW-R

A salient neural hallmark of replay, predominantly observed during Non-Rapid Eye Movement (NREM) sleep is the sharp-wave ripples (SPW-R). Research reveals that throughout sleep, CA3 pyramidal neurons situated in the hippocampus are spontaneously activated in synchronous bursts. This spontaneous activity culminates in a pronounced activation of CA1 pyramidal cells. Within the stratum radiatum, the CA3 input acting on pyramidal cell dendrites engenders the sharp wave. Conversely, in the CA1 pyramidal cell layer, the interaction between the triggered pyramidal cells and the interneurons produces the ripple, which manifests as a fast oscillation ranging from 100 to 250 Hz (Buzsáki, 2015). Elucidating this mechanism further, the two-stage model presented by Buzsáki (1989) hypothesizes an initial phase where a subset of CA3 and CA1 cells are synchronized by theta oscillations during an experiential event. This results in the formation of cell assemblies which encode the pertinent new information. Subsequent to this, during following sleep intervals, these CA3 assemblies spontaneously initiate SPW-R events. This action reactivates the associated CA1 ensembles, reinforcing their connections, thereby facilitating memory consolidation.

Lending credence to this theory, pairs of CA1 pyramidal cells, observed to cofire during open-field exploration, retain this correlation during the ensuing sleep SPW-Rs (M. A. Wilson & McNaughton, 1994). This enduring correlation of activity from wakefulness to subsequent sleep phases is conventionally termed offline reactivation. Drawing upon an array of methodologies (Tingley & Peyrache, 2020), subsequent investigations have determined that both cofiring patterns and entire sequences of place cells activated during wakefulness are recapitulated during the SPW-Rs of the subsequent sleep cycle. This phenomenon is colloquially referred to as replay (Pfeiffer, 2018).

Early research endeavors delving into SPW-Rs predominantly targeted non-human animals. Nevertheless, the occurrence of hippocampal SPW-Rs has been documented in humans, especially during memory encoding and retrieval of diverse stimuli (Bragin, Engel, Wilson, Fried, & Buzsáki, 1999; Le Van Quyen et al., 2008). Significantly, in the context of episodic and semantic memory tasks, there is a higher likelihood of hippocampal ripples preceding successful retrieval as opposed to failed attempts (Ryan, Cox, Hayes, & Nadel, 2008). It remains imperative, however, to acknowledge the substantial discrepancies observed in the detection and temporal dynamics of SPW-R across different studies and model systems. Such variations pose challenges in deriving clear translational interpretations (A. Liu et al., 2022). Contemporary research underscores that awake 'online' SPW-Rs—those which support immediate behavior—and 'offline' SPW-Rs—which underpin memory consolidation during restful state—can be demarcated based on the time interval subsequent to the preceding task engagement (F. Ólafsdóttir, Carpenter, & Barry, 2017). This delineation might epitomize a transition from a brain state oriented externally to one that is internally-focused.

Additionally, studies have illuminated a pivotal role for sharp-wave ripples (SWR) in mediating generalization (Karlsson & Frank, 2009; Roumis & Frank, 2015). They appear not to be restricted merely to replaying directly encountered information. Instead, their potential extends to forecasting impending experiences (Pfeiffer & Foster, 2013) synthesizing discrete yet interconnected experiences (Jadhav, Rothschild, Roumis, & Frank, 2016), and possibly in rearranging events in congruence with an acquired rule (Buzsáki & Fernández-Ruiz, 2019). Consequently, the spiking activity in the hippocampus during SPW-Rs could serve as a streamlined mechanism facilitating the integration of memories for generalization ((Diba & Buzsáki, 2007; Gupta et al., 2010; Pfeiffer & Foster, 2013), but see (Duszkiewicz et al., 2023; Gillespie et al., 2021). This integrative process might be particularly pronounced when it confers a biological advantage.

A conjectured signature for memory consolidation lies in the synchronized activity between the hippocampus and the neocortex during sleep and rest intervals. Notably, hippocampal SPW-Rs exhibit a correlation with sleep spindles observed in the neocortex during slow-wave sleep (Siapas & Wilson, 1998). Although this hippocampal-neocortical coupling during restful states might predominantly strengthen representations of learned data, it is conceivable that hippocampal spiking activity simultaneously augments the establishment of generalized representations within the neocortex. Hence, periods of rest and sleep may therefore provide an opportunity to integrate memories across neural circuits, to support flexible and adaptive behavior in the future (Buzsáki, 2015).

While the fundamental function of SPW-Rs ostensibly relates to wake consolidation and retrieval (Joo & Frank, 2018a), empirical findings assert that such retrievals these retrievals can serve a plethora of functions spanning both wakefulness and sleep (Roumis & Frank, 2015). The ambiguity in distinguishing SPW-Rs specific to consolidation from those dedicated to retrieval has culminated in recent speculations. These postulate that each SPW-R concurrently retrieves a memory while catalyzing its consolidation. Such a hypothesis aligns seamlessly with recent investigations highlighting overlapping molecular mechanisms between memory retrieval and consolidation

(Joo & Frank, 2018b). This intertwined relationship ensures an adaptive mechanism: memories that are consistently retrieved and employed undergo reinforcement and transformation through consolidation. This conceptualization resonates with the multiple-trace theory and the retrieval-based generalization proposition. Within this framework, every act of retrieval engenders a fresh, composite memory, combining the retrieved data with its prevailing context, thereby facilitating the extraction of generalities (Hintzman, 1984, 1986a; Nadel & Moscovitch, 1998).

Mechanisms Underpinning Memory Transformation Leading to Generalization

Within the domain of cognitive neuroscience, the phenomenon of memory transformation, which eventually manifests as generalization behavior, is attributed to a constellation of mechanisms. While the primary focus of this chapter centers around reactivation-centric perspectives on generalization, it is imperative to acknowledge the multifaceted nature of this domain. Below is a non-comprehensive list of the diverse mechanisms discussed in the literature:

1. **Synaptic normalization** postulates that strengthening of certain synapses is counterbalanced by the weakening of others in the neural circuitry, thus facilitating memory transformation (C.-H. Wu, Ramos, Katz, & Turrigiano, 2021).

2. **Reconsolidation** hypothesizes that once memories are retrieved, they become labile and undergo a process of stabilization once again (Robertson, 2018).

3. The **decay** mechanism underscores the natural attenuation of memories over time (Baddeley & Scott, 1971).

4. **Interference** postulates that the formation or retrieval of certain memories can be obfuscated by other memories (Yassa & Reagh, 2013; Yonelinas et al., 2019a).

5. The role of **task structure** (C. M. Wu, Schulz, & Gershman, 2019).

6. **Pre-existing knowledge structures** have historically been implicated in memory transformation (Gilboa & Marlatte, 2017).

7. **Spaced retrieval** implies that memory transformation might be contingent upon the frequency and spacing of memory retrieval (Vlach, 2014).

8. The **level of awareness during retrival and cueing** influences memory transformation (Tal, Schechtman, Caughran, Paller, & Davachi, 2023).

9. The **reconstruction** mechanism underscores that memories, when retrieved, are influenced by antecedent knowledgem, prevailing goals and motivations (Schacter, 2012).

10. **Episodic simulation, imagination, and dreaming** serve as avenues through which memories might be reshaped (Hoel, 2021).

11. **Functional connectivity within and between brain networks** offers a neurobiological perspective on global dynamics underlying memory transformation (Webler et al., 2021).

It is crucial to note that these mechanisms are not mutually exclusive. Additionally, certain mechanisms may be more predisposed to depend on memory reactivation, while others might particularly benefit from offline neural states.

Active consolidation In terms of sleep and reactivation, there has been a burgeoning interest in understanding the processes underpinning memory consolidation and generalization. Central to this investigation is the concept of "active consolidation," which underscores the significant role that offline states, especially sleep, play in these processes. This perspective suggests that the intrinsic characteristics of these offline periods facilitate both consolidation and generalization. Notably, offline reactivation emerges as one of the primary mechanisms in this paradigm, although it is by no means the only one (Rasch & Born, 2013).

Among the many theories that have been posited, the "active systems consolidation hypothesis" stands out, not just for its comprehensive nature but for its deep-rooted empirical backing (Svenja Brodt et al., 2023; Inostroza & Born, 2013; Klinzing, Niethard, & Born, 2019; Rasch & Born, 2013). Indeed, Klinzing et al. have hailed this hypothesis as "the [...] most integrative account of sleep-dependent memory consolidation" (2019, p. 1598).

The hypothesis offers an intricate understanding of the processes at play. It postulates that memory representations, which take shape during wakeful periods, are not localized but rather dispersed across different neural regions. The pivotal role of the hippocampus in this scenario is to bind these disparate components, thus forging a singular, cohesive episodic memory representation. As the individual transitions into sleep—slow-wave sleep (SWS) to be specific—there is a subsequent replay of a select subset of these newly minted hippocampal representations. Complementing this hippocampal replay, the various components of this memory representation undergo reactivation within their respective neural loci. This results in qualitative alterations, a process likened to synaptic upscaling. However, these upscaling processes do not operate in isolation. They are firmly nestled within the broader framework of global synaptic downscaling, a mechanism that curbs neuronal firing rates during sleep, ensuring a recalibration of synaptic connection strengths. Hence, the essence of the active systems consolidation hypothesis is a duality—it argues for the co-existence and interplay of both synaptic and systems consolidation in the realm of sleep-dependent memory processing. This conceptual model harks back to, and indeed, builds upon the foundational two-stage model of memory (see McClelland (1995) and Squire (2015)).

However, while the active systems consolidation hypothesis provides a compelling narrative, certain nuances remain enigmatic. The specific temporal dynamics governing synaptic and systems consolidation, as well as the factors that may modulate them, remain subjects of intense inquiry. Preliminary evidence does suggest that sleep plays a cardinal role in tethering various components, thus enabling the crystallization of a unified episodic memory. This process seems particularly pronounced for shorter retention intervals, such as 10 hours (Weber, Wang, Born, & Inostroza, 2014). Conversely, for more extended periods, like three days, sleep appears to foster the melding of memories into pre-existing cognitive networks, resulting in more gist-like memory representations for longer retention intervals (e.g., 3 days; Klinzing et al., 2019).

Permissive and passive consolidation The “active” perspective posits that offline states inherently contribute to memory consolidation and generalization. In contrast, the “permissive” perspective suggests that offline states simply create conditions conducive to these processes. It acknowledges that offline states, especially sleep, can promote recall, but it questions whether these states provide any unique benefits beyond those of online states. The primary advantage, according to this view, might be the reduced potential for retroactive interference, which is more common during wakeful periods when new experiences might interfere with the consolidation of prior memories.

A more minimalist perspective, the “passive” view, posits that offline states primarily protect memories from interference. Unlike the “permissive” stance, the “passive” perspective asserts that no active consolidation occurs during these offline states.

Therefore, any observed memory benefits following offline periods are temporary and will diminish as soon as the level of interference matches that observed in a control group, a notion supported by Ellenbogen et al. (2006).

These views are in line with the previously mentioned Contextual-binding theory (Yonelinas et al., 2019a) . As mentioned earlier, in this framework, ‘context’ is broadly defined, encompassing various elements of a learning scenario, be it cognitive, emotional, spatial, or temporal components. The theory suggests that the retrieval of a specific memory can be hindered by subsequent or prior learning, but only if the new information’s content or context mirrors that of the original memory, leading to retroactive interference. This interference can be exacerbated by factors such as temporal proximity or similarity in content or context. However, because new learning is essentially absent during sleep, there is a reduction in retroactive interference. Consequently, post-sleep periods might be characterized by improved memory recall, especially when contrasted with wakeful periods, as outlined by Yonelinas et al. (2019a).

Models of offline generalization The study of offline generalization has seen a proliferation of models and theories, with a prominent focus on the role of sleep in the generalization process. Several theories and hypotheses have been put forth to account for the underlying mechanisms. Lewis, Knoblich, & Poe (2018a) highlights the integral role of sleep in enhancing procedural skill learning and augmenting creative problem-solving. This has paved the way for models such as the ‘information overlap to abstract’ (iOtA) model. As proposed by Lewis et al. (2018a), the iOtA model suggests that during non-rapid eye movement (NREM) sleep, memory consolidation and gist extraction transpire. This is followed by an integration of these consolidated memories into existing cognitive schemas during rapid eye movement (REM) sleep. The model posits that replay-induced strengthening of correlated memories, juxtaposed with synaptic downscaling, results in a prominent amplification of shared features of memories, while other less arbitrary details are eliminated. This process fosters the abstraction of central information, thereby catalyzing the formation of cognitive schemata.

However, the realm of sleep-dependent generalization isn’t devoid of contention. Cordi & Rasch (2021), among others, have expressed reservations, debating the veracity of sleep-dependent generalization effects. Questions permeate regarding the specific conditions facilitating these effects, the neural underpinnings driving them, and the true robustness of the observed outcomes. Lerner & Gluck (2019b) further enrich this discourse through their systematic review on the role of sleep in tasks geared towards the extraction of hidden regularities. They discerned two dominant patterns: Firstly, the effects are considerably task-dependent; both behavioral and electrophysiological findings are often replicated across analogous tasks. Secondly, sleep-dependent explicit detection of patterns predominantly surfaces when the underlying regularity exhibits a temporal dimension and this is especially associated with NREM sleep metrics.

Drawing from these observations, Lerner (2017) propose the Temporal Scaffolding hypothesis. This hypothesis contends that sleep plays a pivotal role in the explicit generalization of memories exhibiting temporal patterns. This claim is rooted in the nature of memory replay in the hippocampus during slow-wave sleep (SWS), which is characteristically time-compressed. As a result, temporal regularities, which might have originally been segregated by several seconds, become compressed within the confines of typical Hebbian learning timescales. This compression potentially enables the offline discernment of these temporal patterns, akin to the detection of non-temporal patterns during wakefulness.

In advancing our understanding of offline generalization, computational modeling emerges as a valuable tool. As emphasized by Guest & Martin (2021), computational modeling can further constrain these existing arguments, thus advancing scientific inference over and above experimental research, systematic reviews and phenomenological theories. Numerous neural network models rooted in the cortical architecture have been advanced. These models seek to provide a mathematical understanding of the processes by which the brain might leverage sleep to facilitate the extraction of semantic patterns.

One seminal framework that has been pivotal in this realm are stochastic hierarchical models. The primary objective of such models is to optimize the likelihood of observed data under a generative construct. Notable exemplars include the Helmholtz machine proposed by Dayan, Hinton, Neal, & Zemel (1995) and the Wake–Sleep algorithm developed by G. E. Hinton, Dayan, Frey, & Neal (1995). These models have showcased the feasibility of fusing online (awake) and offline (sleep) states to foster the learning of semantic representations. Essentially, the models endeavor to decode observed data, employing hierarchical processing to infer latent variables.

A distinct but related proposal was put forth by Káli & Dayan (2004) . They postulated a model focusing on the interplay between neocortical and hippocampal regions. Central to their hypothesis is the role of offline hippocampal replay. The model posits that such replay acts as a linchpin in preserving and recalling declarative memories within the cortical domain. An intriguing corollary of their hypothesis suggests that in the absence of such hippocampal replay, consolidated memories become susceptible to fragility, especially when cortical plasticity is at play (i.e.: catastrophic interference). However, contrary to what one might presume, these models do not harness offline states to refine their generative constructs. Rather, their training is explicitly oriented towards reconstructing sensory inputs encountered during wakeful periods. This approach underscores a critical area of exploration in the domain of neural network models - the true potential of offline states in enhancing the fidelity of generative constructs and whether future iterations of these models can be developed to leverage sleep states more effectively.

Oscillating Learning Rule in REM One such proposal states that REM sleep employs a so called “Oscillating Learning Rule” (Norman, Newman, & Perotte, 2005), which offers a fresh perspective on the traditional CLS model. Rather than solely focusing on the idea that offline interleaved replay revolves around an immutable, stationary cortical signal, this approach posits two distinct offline learning mechanisms, thus enriching the concept of memory consolidation during sleep. The unique feature of this algorithm is its dynamic strength adjustment. It enhances memories when their strength falls short of supporting recall and attenuates those memories which might obstruct the successful recall of a target memory. The underlying aim of this oscillation is to counteract interference, thereby accelerating the process of cortical learning. Building upon this, it’s theorized that the activities during wakefulness, coupled with hippocampal reactivations during Slow-Wave Sleep (SWS), might inadvertently interfere with or “damage” associated cortical memories. To mitigate this, spontaneous cortical reactivations during REM phase are believed to streamline these memories, reducing interference through an adaptive pattern separation mechanism (McDevitt, Duggan, & Mednick, 2015) .

Taking this a step further, it was recently extended with a more detailed hippocampal model (Schapiro, Turk-Browne, Botvinick, & Norman, 2017b), more biologically plausible learning dynamics and a larger focus on the NREM stage and benefits of NREM/REM autonomous interactions (Singh et al., 2022). A central tenet of this revised model is the differentiation in communication between the hippocampus and neocortex during NREM and REM sleep. Specifically, the proposed sleep algorithm employs a combination of high-fidelity reactivation of recent memories in NREM sleep (likened to spindles) with lower-fidelity reactivation due to a “high inhibition” state, and binding with more remote memories during a “low inhibition” state (likened to theta oscillations) in REM. Highlighting the efficacy of this model, it not only elucidates the mechanisms behind sleep-dependent categorization but also illustrates how the alternating patterns of sleep stages can bolster continual learning. This is especially evident when navigating challenges like the classical AB–AC interference paradigm.

Model of Perturbed and Adversarial Dreaming Similar to OLR the Model of Perturbed and Adversarial Dreaming (PAD) focuses on interaction of NREM and REM, but draws inspiration from the principles of generative adversarial networks (Deperrois, Petrovici, Senn, & Jordan, 2022). It postulates a mechanism by which learning during active wakefulness and during the distinct sleep phases collaboratively contribute to the formation of object category representations. Furthermore, this model embodies features that mirror the phenomena of replay and dream states observed in both NREM and REM sleep.

In our waking state, the cortical feedforward routes become adept at discerning that low-level neuronal activity stems from external stimuli. Concurrently, feedback mechanisms are refined to recreate these activities from higher-order neuronal representations. These advanced representations then find a repository within the hippocampus. Transitioning to the NREM phase, the feedforward routes endeavor to reconstruct these high-level activities, which are replayed from the hippocampus but are now subject to minor low-level disturbances. This phase has been metaphorically described as ‘perturbed dreaming’.

As we delve deeper into the REM phase, an intriguing dynamic ensues between feedforward and feedback pathways, operating in what can be described as an ‘adversarial’ mode, coining the term ‘adversarial dreaming’. Here, feedback routes are responsible for producing virtual low-level activities. These activities are conceived by fusing multiple memories stored in the hippocampus with spontaneous cortical activities, the latter being represented as Gaussian noise in the model. While the feedforward mechanisms strive to identify these low-level activities as internally orchestrated, the feedback mechanisms attempt to mislead the feedforward routes.

Significantly, this model is not just a standalone theory but is rooted in cognitive interpretations of REM dream states. Such states are often punctuated by dreams that meld creativity with realism, generating lifelike sensory experiences. This unique characteristic of REM dreams can be attributed to the interplay between episodic memories and a notable reduction in the synchronization between the hippocampus and the cortex. This connection is supported by a foundation of research, prominently including works by Fosse, Fosse, Hobson, & Stickgold (2003), Wierzynski, Lubenov, Gu, & Siapas (2009), and Lewis et al. (2018a).

Reinforcement learning and replay In the vast realm of neural network modeling, a significant gap pertains to the question of how certain memories are tagged for replay. Distinctly, memories that carry salient attributes such as emotional valence, reward potential, or future relevance demand prioritization (Mattar & Daw, 2018) .

Reinforcement Learning (RL) delves into the dynamics of how agents should strategize their actions within an environment to optimize cumulative rewards. As the agent interacts with its environment—observing, acting, and subsequently receiving rewards—it endeavors to refine its actions to accrue maximum expected long-term rewards. An inherent challenge within this process lies in maintaining an equilibrium between exploiting acquired knowledge and embarking on exploration to reduce uncertainties (Sutton & Barto, 2018). A long-standing proposition in the RL sphere is the potential for agents to strategize these exploratory actions during periods of inactivity, analogously during quiet wakefulness or sleep.

A cornerstone concept in modern RL, especially pivotal in the backdrop of complex environments is the notion of "experience replay" or offline replay. This concept emerges as a solution to the challenges that surface when neural networks intersect with RL and rewards based learning needs to operate on and update a distributed learning system. These challenges span stabilizing the learning process, streamlining the utilization of past experiences via a designated replay buffer, and enhancing generalization through diverse replay tactics (Momennejad, Otto, Daw, & Norman, 2018). The burgeoning interest in replay-based RL converges on its applicability as a paradigm for modelling latent or structure learning tasks. Herein, optimal actions are contingent on cognizing the relational structure of the environment, including episodic inference tasks featuring discrete state spaces (e.g., A, B, C). The adoption of RL in these scenarios is rooted in its biological feasibility. It not only postulates testable neural theories related to structure learning but also delves into the specifics of replay content, frequency, and their behavioral implications (Cazé, Khamassi, Aubin, & Girard, 2018; Roscow et al., 2021). Moreover, RL principles, even devoid of rewards, can adeptly assimilate and predict the intricacies of learned structures (Gershman & Daw, 2017). Empirically grounded in a suite of studies, both computational and experimental, is the hypothesis that relational maps in memory are architected in alignment with RL-specific prioritization and replay principles (Momennejad, 2020).

The presence of different representational schemes carries significant implications for behavior, which holds considerable importance for both the fundamental comprehension of memory processes and their practical applications in domains such as aging, education, and mental health.

Moving forward, testing episodic inference paradigms for model-based signs of time-dependent consolidation and employing cognitive models that encompass the measurement of both specificity and generalization within the same study, while also harnessing neural indicators to address cognitive inquiries, will assume increasing significance in fully characterizing the memory representations. This integrative approach will prove instrumental in describing mechanisms through which generalization occurs across diverse experiences.

Chapter 2

A simulation-based toy model of reactivation-dependent memory consolidation

2.1 Abstract

Recent qualitative reviews indicate that sleep-dependent memory consolidation (SDMC) effects largely hinge on the specific task at hand (Chatburn, Lushington, & Kohler, 2014; Lerner & Gluck, 2019a). A growing body of research posits that both encoding-related spontaneous reactivation (Bendor & Wilson, 2012) and reactivation due to memory cueing during sleep (Hu et al., 2020) play pivotal roles in SDMC, especially in the context of associative information and gist abstraction (Lewis, Knoblich, & Poe, 2018b). To comprehensively unpack the interplay between task-dependency, reactivation, and rapid generalization, a formal framework is indispensable. While existing work has explored connectionist memory models to analyze these relationships, we posit that a vector-symbolic approach is complementary, considering the diversity of published sleep and memory effects. Specifically, we focus on the MINERVA2 global matching memory model (Hintzman, 1984) and highlight how it is complementary to existing connectionist computational models of reactivation and rapid generalization (Kumaran & McClelland, 2012a) . By conceptualizing offline reactivation as internally generated cued recall, these approaches adeptly account for a wide spectrum of behavioral SDMC findings, including paired association, episodic inference tasks, and motor memory tasks. We attempt to elucidate how exemplar models align with the empirical and theoretical

SDMC work and highlight testable hypotheses it offers pertaining to task variables, such as interference, memory strength, complexity, and context.

2.2 Sleep and reactivation

Current theories highlight an active, multifaceted role for sleep in which memories undergo a process of systems consolidation, with reactivation being but one of many proposed mechanisms facilitating memory consolidation (Rasch & Born, 2013). There is supporting neurophysiological evidence for spontaneous reactivation of memories during sleep in animals and humans (Foster, 2017), as well as behavioural evidence in the form of cued reactivation studies, with a recent meta-analysis showing reactivating memories during non-rapid eye movement sleep (NREM) improved later memory performance with a small to medium effect size (Hu et al., 2020). Artificial learning systems also seem to benefit from reactivation when it comes to avoiding catastrophic forgetting of continually acquired knowledge (Hayes et al., 2021b; Wittkuhn, Chien, Hall-McMaster, & Schuck, 2021).

While reactivation's causal role in memory retention is well documented, less is known about its role in forming inferences, whether those processes denote some form of inductive, statistical generalization – e.g., pattern extraction (Lerner & Gluck, 2019a) – or more higher order analogical generalization (Lewis et al., 2018b). Mechanistic underpinnings of reactivation's role in associatively chaining "item" or "location" memory during neural activity not solely determined by prior experience is emerging (Buzsáki, 2015; Gupta et al., 2010), with more recent studies showing that future planning (F. Ólafsdóttir, Barry, Saleem, Hassabis, & Spiers, 2015) and abstract trained rules (Y. Liu, Dolan, Kurth-Nelson, & Behrens, 2019) also influence replay content.

Standard consolidation theory (SCT, Dudai (2004)) and synaptic homeostasis theory (SHY, Tononi & Cirelli (2014)) are two of the main theories that aim to explain the mechanisms behind sleep-based benefits to memory. In short, SCT argues that memories for events are only temporarily dependent on the hippocampus and will decay unless they go through a consolidation process that redistributes the content of these memories to the neocortex, thus becoming independent of the hippocampus; SHY argues that the role of sleep is to downscale synaptic strength to a baseline level that has been elevated during wakefulness. Early hypotheses of sleep and memory have also focused on distinctions between memory systems and sleep stages; an emergent notion from research with paradigms looking at dissociable non-REM

(NREM) and rapid eye movement sleep (REM) effects was that declarative memories preferentially benefit from NREM sleep, typically slow wave sleep (SWS), and non-declarative, especially procedural, memories preferentially benefit from REM sleep (Gais & Born, 2004; Peigneux, Laureys, Delbeuck, & Maquet, 2001; Plihal & Born, 1997). However, a dissociation of NREM and REM sleep serving different memory systems is incongruent with many current findings (Chatburn et al., 2014; Lerner & Gluck, 2019a).

Alternatively, hypothesis considering the sequential nature of sleep cycles led to the development of the sequential hypothesis, which emphasizes the importance of interleaving NREM and REM sleep cycles in memory processing. The sequential hypothesis proposes that NREM sleep contains selective processes that weaken non-adaptive memories before REM sleep stores the surviving memories and integrates them with pre-existing memories (Giuditta et al., 1995), an idea that is supported by studies relating overnight retention of words to the integrity of NREM-REM sleep cycles (Ficca, Lombardo, Rossi, & Salzarulo, 2000). The active systems consolidation hypothesis, similarly, assumes cooperation between the cycles, stating that events of wake are encoded across cortical networks and bound together by the medial temporal lobe; then, during SWS, spontaneous neuronal replay (Ji & Wilson, 2007) originating from the hippocampus consolidates memory via hippocampal-neocortical information transfer and the strengthening of cortico-cortical connections (Born & Wilhelm, 2012). This memory reactivation during sleep is assumed to have a two-fold role: stabilizing/strengthening labile memories making them less susceptible interference (Ellenbogen, Hulbert, Jiang, & Stickgold, 2009), and secondly reorganizing and integrating new memories into the pre-existing networks of long-term memories through system consolidation (Stickgold & Walker, 2013). This can lead to structural changes on a representational level: the extraction of statistical regularities (Cousins, El-Deredy, Parkes, Hennies, & Lewis, 2016; Lewis & Durrant, 2011), the abstraction of gist information (Lutz, Diekelmann, Hinse-Stern, Born, & Rauss, 2017) and the emergence of cognitive schemas (Lewis & Durrant, 2011).

Whilst there are significant sleep-dependent memory consolidation findings, recent qualitative reviews struggle to pinpoint: 1) what are the necessary conditions for these effect to emerge (Chatburn et al., 2014; Cordi & Rasch, 2021; Dringenberg, 2019b; Lerner & Gluck, 2019a), and 2) how these conditions might be modulating, or be modulated by, reactivation during sleep. Creating or adapting existing formal cognitive models to understand existing experiments is a necessary step in shifting from exploratory to theory-driven research (Burghardt & Bodansky, 2021; Fried, 2020;

Guest & Martin, 2021).

2.3 Computational models of memory during offline states

The main goal of computational modelling in behavioural science is to use precise mathematical models to make better sense of behavioural data, however, the level of precision and the amount of explanatory value should be tailored to the specific aim of the research project (R. C. Wilson & Collins, 2019). Computational memory models can explain behavioural findings from a variety of experimental paradigms, but research has produced multiple competing models, and as a result, since different models focus on different phenomena, there isn't one single model best suited to explain findings from all paradigms (Kahana, 2012). When considering computational models of reactivation during sleep, one review distinguished between models of (a) dynamics of replay activation, (b) memory consolidation dependent on replay events and (c) complementary learning systems (Rennó-Costa, Silva, Blanco, & Ribeiro, 2019). Models of dynamics of replay activation tend to focus on the mechanistic origin of reactivation-related neuronal activity observed during sleep, particularly hippocampal reactivation (e.g., Michael E. Hasselmo (2008)). The latter two classes of models (b) and c)), involve computational studies that focus on the mechanisms by which offline reactivation of memory traces affects encoding, storage, and retrieval of memories. While this group of studies encompasses a wide range of perspectives, from focusing on consolidation of recent (August & Levy, 1999) to remote memories (Kirkpatrick et al., 2017), from "neuroplausible" (Wei, Krishnan, & Bazhenov, 2016) to highly abstracted memories (Káli & Dayan, 2004), most of these models are not interested in accounting for inferential features of reactivation, instead focusing on the stabilizing, protective role of sleep. A notable exception includes the Complementary Learning System (CLS) (McClelland et al., 1995), and more specifically relevant to my current modelling aims, the model REMERGE, proposed as an extension to the CLS in order to account for rapid, hippocampal-dependent generalization (Kumaran & McClelland, 2012a) .

2.3.1 REMERGE

Presented as a fundamentally two-layer recurrent neural network complemented by a response layer for behavioral analysis (refer to Fig. 1A), this model adeptly accounts for a multitude of episodic inference and categorization task behavioral effects. Instead of relying on the creation of an explicit cognitive map (Behrens et al., 2018), it leverages a retrieval-based search process where inference results from recurrent similarity-based computation (see Fig. 1C, red lines). Furthermore, the authors propose a sleep-dependent memory consolidation mechanism for the model,whereby during sleep (specifically SWS) the network proceeds by strengthening existing association weights further (see Fig. 1B, red lines). Through a more conceptual lens REMERGE labels itself as a synthesis of classical exemplar models of memory (Hintzman, 1984; Medin & Schaffer, 1978; Nosofsky, 1984) and a class of connectionist networks, termed interactive activation and competition (IAC) models (Mcclelland, Mirman, Bolger, & Khaitan, 2014). In relation to contemporary models of generalization, REMERGE is in stark contrast to models that propose that it is the overlap between hippocampal representations for related stimuli that is critical to generalization, such as the Temporal Context Model (M. W. Howard, Fotedar, Datey, & Hasselmo, 2005a) or relational theory (Neal J. Cohen et al., 1997).

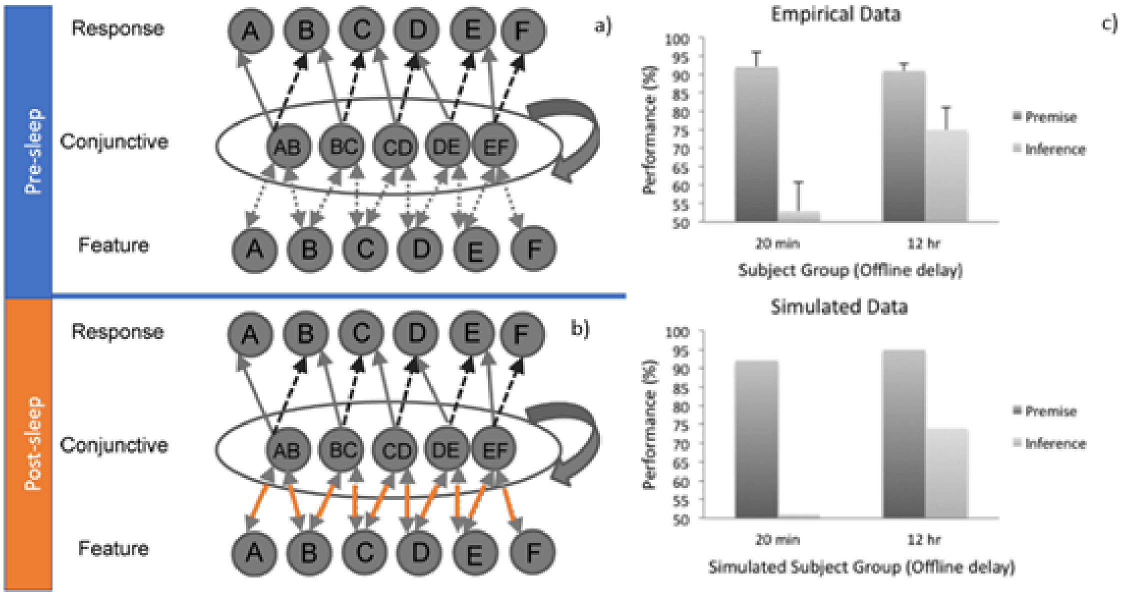


Figure 1: Overview of REMERGE a) REMERGE model of a six item Transitive Inference Task b) Weight strengths for the 20-min (a) and 12-hr groups (b), respectively: 1.29, 1.35. c) Transitive inference task: empirical (upper panel) and simulated (lower panel) data from Ellenbogen et al.'s (2007) study. Performance of 20-min and 12-hr subject group shown (x-axis) with performance on y-axis. Groups differ as a function of the length and nature of the delay period interposed between training and testing, adapted from Kumaran & McClelland (2012)

In recurrent neural network architectures, the mechanism of information flow and activation computation often involves a series of interconnected equations. The REMERGE model, a two-layer recurrent neural network retrieval model, follows this principle. It's fundamental equations offer a comprehensive view of the network's dynamics, from how each neuron processes its input to the final decision or output of the system. Below are the foundational equations governing the behavior and function of the REMERGE model:

$$\text{net}_i(t) = \lambda * \text{cnet}_i(t) + (1 - \lambda) * \text{cnet}_i(t-1) \tag{2.1}$$

Eq. (2.1) models the temporal dynamics of the neural activation. The net input $\text{net}_i(t)$ for neuron i at time t is computed as a blend of its current and previous inputs. The factor λ determines the extent to which the present versus past inputs influence the neuron's net input (fixed at 0.2 in Kumaran & McClelland (2012a)).

$$\text{cnet}_i(t) = \sum_{j=1}^{N} w_{ij} y_j(t) + estr * \text{ext}_i(t) + \varepsilon_\sigma, \tag{2.2}$$

Eq. (2.2) determines the combined input (often termed 'current net input') for neuron i at time t. The input is the weighted sum of activations $y_j(t)$ from all N other connected neurons, influenced by synaptic weights w_{ij}. An external stimulus, represented by $\text{ext}_i(t)$ and scaled by a factor estr (fixed at 0.5 in Kumaran & McClelland (2012a)), is added to this sum. Finally, a gaussian noise term ε_σ is added to introduce variability in certain simulations.

$$y_i = \frac{1}{1 + e^{-net_i/\tau}} \tag{2.3}$$

Eq. (2.3) represents a typical sigmoid activation function. The value of y_i (feature layer neuron's activation) lies between 0 and 1. The rate of change in the sigmoid curve is determined by the parameter τ, temperature parameter, regulating how strongly output y varies with net_i, which represents the net input to the neuron i.

$$y_i = \frac{e^{net_i/\tau}}{C^{1/\tau} + \sum_{i=1}^{N} e^{net_i/\tau}}, \tag{2.4}$$

Eq. (2.4) describes a normalization function for the activation of conjunctive neuron i, ensuring the summed outputs over all N neurons are bounded (called 'hedged softmax'). It is a form of the softmax function adapted for this model, where activations are exponentiated and then normalized by the sum of all activations. C is a constant that may be used to adjust the scale or influence of the normalized output.

$$P(1) = \frac{e^{y_1/\beta}}{e^{y_1/\beta} + e^{y_2/\beta}} \tag{2.5}$$

Eq. (2.5) provides the probability $P(1)$ based on the activations y_1 and y_2. It used to produce a choice in a binary forced-choice scenarios (e.g., old/new recognition or transitive inference B?D) The function can be seen as a specialized softmax operation, where the activations are scaled by a factor β and normalized to generate a probability output.

These equations, when interpreted collectively, provide a holistic view of the REMERGE model's computational strategy, highlighting its capacity for handling complex episodic memory data. While REMERGE has extensive explanatory breadth for the number of free parameters it assumes there are some key limitations, namely:

1. there is no learning algorithm proposed for how the conjunctive layer is formed (and what the feature-conjunctive weights should be after learning);

2. it is not clear how one would model tasks where a priori conjunctive nodes are not obvious from the task structure such as statistical learning tasks

3. because of the recurrent nature of the system, it is highly sensitive to initial conditions, namely the exact tuning of the "external cueing strength", weights, and temperature parameter.

4. it does not propose mechanism for storing/integrating newly formed generalizations

While these could be viewed as surmountable issues, there are strong theoretical assumptions that come with the model that have been briefly mentioned:

1. it assumes that reactivation – for the tasks that it purports to account for – is always 'generalized' – irrespective of cueing one item, all items become activated like in IAC networks in a context-dependent manner

2. it subscribes to a strictly associative learning-based view of memory (both for wake and sleep), although there are good theoretical arguments and empirical evidence that this learning strategy is not optimal for environments that involve a large number of multiple overlapping sequences (see Kalm & Norris (2018)).

2.3.2 MINERVA2

To overcome the limitations listed above we propose the use and extension of a multiple-trace memory model as conceptual model of offline reactivation that is in line with retrival-based generalization theories. MINERVA2 (Hintzman, 1984) was the first formal model to argue that semantic memory traces might not be stored separately from episodic traces (rather be computed at recall) with the primary aim to account for to category learning and recognition memory literature in parallel. As with many abstract memory models of recognition, during encoding, memory traces are individually stored in long-term storage. Due to the notion of "multiple traces", the addition of new memories doesn't compromise the integrity of earlier ones. During recall, the model calculates the similarity between the test prompt and every stored memory, hence the name "global matching models'. By assessing each memory's compatibility, an overall familiarity signal can be determined. In the MINERVA2 system, memory traces are represented as vectors. Each element, or more precisely feature of these vectors can be 1 (signifying the presence of a feature), -1 (denoting the absence of a feature), or 0 (showing that the status of the feature is uncertain). The length of these vector (ϵ) is a free parameter chosen by the modeler. During encoding, a fresh memory trace vector is produced, capturing features from the original item vector based on a probability termed as L, the *learning rate*. If any feature isn't copied into the memory, the trace for that feature is marked as 0. After encoding a collection of items, memory's content is represented by a matrix, M, holding all the study-phase created trace vectors. Further, as time passes, features that were precisely copied might nonetheless change to 0 with a probability δ. This δ, known as the *decay rate*, embodies the effects of the time elapsed between studying and testing. Longer intervals entail greater decay rates.

Global matching is achieved by computing the match between the probe or cue vector P with every trace vector in the memory matrix M. The computed similarity, denoted as S_i, between the probe and the i^{th} trace vector is driven by the formula Eq. (2.6) (see also Fig. 2A).

$$S_i = \sum_{j=1}^{N} \frac{P_j M_{i,j}}{\epsilon_i} \tag{2.6}$$

Here, M_i represents the i^{th} trace vector within M. The factor j signifies the j^{th} feature being compared between the probe and trace vectors. Meanwhile, ϵ_i is the count of features where either the cue or trace is non-zero. An S_i value of 1 indicates that the probe vector is a perfect match with the trace vector, while a value of 0 suggests that the two vectors are orthogonal. The similarity values undergo a transformation, leading to an activation value represented by Eq. (2.7) for each trace.

$$A_i = S_i^3 \tag{2.7}$$

This cubing emphasizes a non-linear linkage between similarity and activation, ensuring maximum activation for trace vectors closely resembling the cue vector. The sum of all these activation values gives us the "echo intensity" I (i.e.: global match), which essentially gauges familiarity (see Eq. (2.8) and Fig. 2B).

$$I = \sum_{i=1}^{N} A_i \tag{2.8}$$

It is important to highlight that it is here that MINERVA2 demonstrates an adeptness in distinguishing between singular and conjunctive episodes. While matches distributed over multiple saved traces are summed linearly, when multiple feature matches correspond to a singular trace corresponding to a conjunctive episode, their aggregation escalates due to the cube rule. For instance, contrast two traces each with a match value, S_i, of 0.5 to a single trace with a match value, S_i, of 0.7. With the cube rule applied, the cumulative match value, I, for the first scenario becomes $0.5^3 + 0.5^3 = 0.25$. Conversely, in the latter case, $I = 0.7^3 = 0.343$. Another important property of the retrieved trace is the "content"; *MINERVA2* simulates recall by determining a weighted total, C, of all items, i, retained in memory. The weight of each item is based on its correspondence to the test cue. The j^{th} part of C is described by Eq. (2.9).

$$\mathrm{C}_j = \sum_{i=1}^{N} A_i M_{i,j} \tag{2.9}$$

Each of these two values can be used in "downstream" processing. The "echo intensity" I is usually further evaluated against a decision threshold Cr. If I surpasses Cr, it signifies a high familiarity, prompting an "old" response in a recognition memory context. Conversely, a value below Cr yields a "new" response. As the echo intensity peaks when the similarity between the probe and traces is at its zenith, I values are generally elevated for targets. This implies a higher expected similarity between a target trace vector and its associated cue than between a trace vector and an unrelated probe. Echo intensity can be viewed as the activation strength of the echo resulting from the cue and can be computed via the cosine of the echo and cue; a proxy measure for familiarity in memory search (Yonelinas, 2002). In terms of "echo content", more a memory trace overlaps with the cue, the more influential it becomes in shaping the echo. However, there's invariably an influence from other interfering traces. Thus, the act of recalling memory always introduces some distortions and recalled information is a blend of various memories. Amplifying this distortion, the returned echo gets stored as a new trace. Via this mechanism, MINERVA2 not only encapsulates the concept that our memories continually evolve based on experiences and even the mere act of retrival, but can account for how schemas can form based on exemplars (Hintzman, 1984). When the system is cued solely with categorical or context encoding cue vector it will produce an echo that is a prototypical composite off all the exemplars related to that category/context. Further, while the process of retrieval continually introduces distortions, it also aids in pinpointing a specific memory trace. By employing the returned echo to refine the memory search, by reintroducing it as a cue, it triggers a progressively limited set of memory traces. This results in individuals achieving a more precise recall than if they solely relied on the initially returned, most generalized, memory trace as can be viewed as model of deliberate search process.

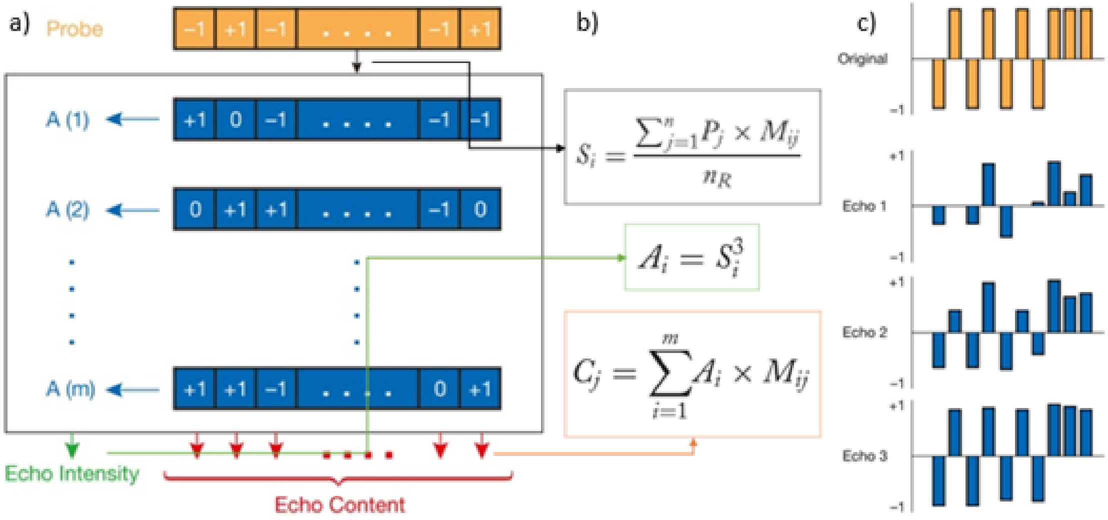


Figure 2: Overview of MINERVA2 a) Echo Intensity (activation strength) and content in MINERVA2 adapted from (Radvansky, 2017) b) Mathematical functions involved in MINERVA2; Si denotes the similarity of trace, i, to the cue, Pj ; Ai denotes the activation of trace as the cube of its similarity to the cue c) Ci refers to the echo computed by weighting each of the 1.. . m traces in memory by their activation's strength then, summing all m traces into a single vector

When we run the system in a "recurrent" manner by reintroducing the echo as a cue, we can see the model gradually performing more accurate retrieval, as it activates a smaller and smaller set of memory traces (Fig. 2C) which evokes the recurrent similarity computation of REMERGE.

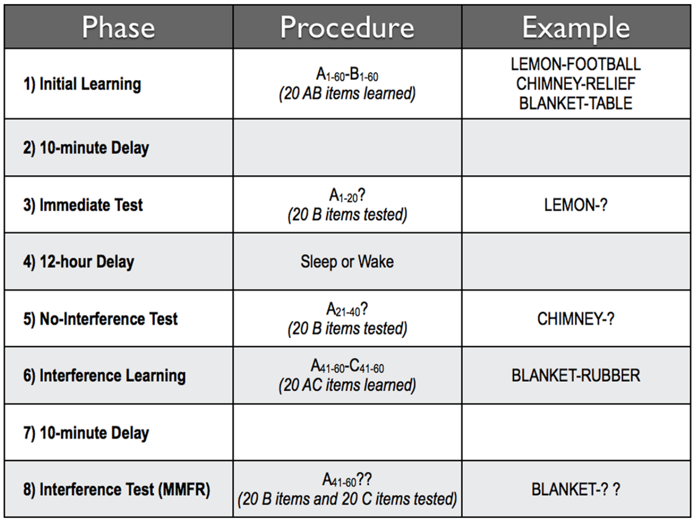
To recap, Hintzman (1984, p. 96) summarizes MINERVA2's key assumptions as follows:

1. only episodic traces are stored in memory,
2. repetition produces multiple traces of an item,
3. a retrieval cue contacts all traces simultaneously,
4. each trace is activated according to similarity to the cue,
5. all traces respond in parallel, and retrieved information reflects their summed output.

MINERVA2 and Sleep

By modelling offline reactivation as spontaneous, internally cued recall combined with recurrent similarity computation, we attempt to conceptually replicate a subset of sleep-dependent memory consolidation findings that highlight the explanatory range of MINERVA2:

Paired associates: interference learning protocol To investigate the notion that sleep safeguards memory from ensuing disturbances, a study utilized the A–B, A–C word-pair interference scheme (see Fig. 3) . In the study's setup, participants initially learn a sequence of $A_{1\ldots60} - B_{1\ldots60}$ word pairs. Subsequently, a subset of these pairs, specifically $A_{41\ldots60}$, is linked with $C_{1\ldots20}$, introducing retroactive interference with the earlier learned $A_{41\ldots60} - B_{41\ldots60}$ pairs. The authors not only found that improved recall the in the sleep group for both the interfered and non-interfered words when compared to a wake group at delayed recall, but found that sleep was most beneficial for the interfering subset. This supported an active interpretation of the role of sleep as it showed that retroactive interference is less disruptive if the memories had been consolidated during sleep (Ellenbogen et al., 2009) .



Phase	Procedure	Example
1) Initial Learning	$A_{1\text{-}60}$-$B_{1\text{-}60}$ *(20 AB items learned)*	LEMON-FOOTBALL CHIMNEY-RELIEF BLANKET-TABLE
2) 10-minute Delay		
3) Immediate Test	$A_{1\text{-}20}$? *(20 B items tested)*	LEMON-?
4) 12-hour Delay	Sleep or Wake	
5) No-Interference Test	$A_{21\text{-}40}$? *(20 B items tested)*	CHIMNEY-?
6) Interference Learning	$A_{41\text{-}60}$-$C_{41\text{-}60}$ *(20 AC items learned)*	BLANKET-RUBBER
7) 10-minute Delay		
8) Interference Test (MMFR)	$A_{41\text{-}60}$?? *(20 B items and 20 C items tested)*	BLANKET-? ?

Figure 3: Experimental design of the Interference learning protocol. Participants began by studying a set of 60 word pairs, illustrated as A–B (examples include Lemon-Football, Chimney-Relief, and Blanket-Table). A test on 20 selected pairs from this set followed after a 10-minute interval (e.g., Lemon-?). After a 12-hour period, which either included sleep or wakefulness, a separate set of 20 pairs from the initial 60 was tested (e.g., Chimney-?). Subsequently, participants learned 20 new A–C pairs (e.g., Blanket-Rubber) where each A-cue word was linked to an untested pair from the initial set. Figure adapted from Ellenbogen et al., 2009.

Transitive inference task The basic concept of the transitive inference task is this: if A surpasses B and B outperforms C, then it can be deduced that A exceeds C, even without a direct experience of this relationship (see Fig. 4). Relational learning involves understanding indirect connections between stimuli based on their direct relationships. In a research by Ellenbogen, Hu, Payne, Titone, & Walker (2007) – later replicated by Werchan & Gómez (2013a) – the transitive inference task was employed to investigate the impact of sleep on recalling paired associations (e.g., how does A relate to B) and making inferences (e.g., how does B relate to D). The study revealed that after a period of sleep, inferential ability was enhanced when compared to an equal awake duration.

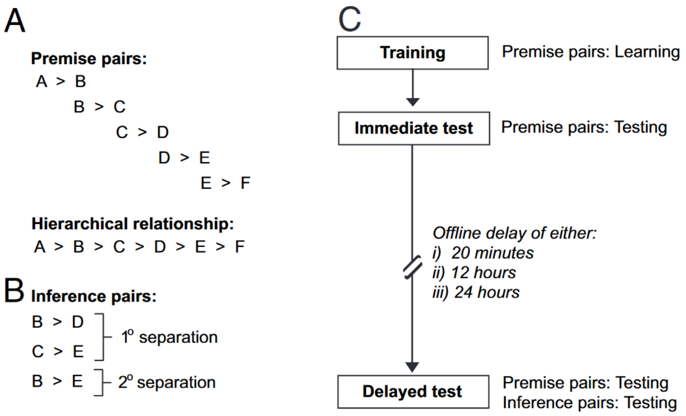


Figure 4: Experimental design of the Transitive inference task. (A) Six object visuals (conceptually represented as A–F) were assembled into five premises. The > indicates a preference, denoting "should be chosen over". Unbeknownst to participants, an inherent order existed: A>B>C>D>E>F. (B) Post-assessment of the hierarchy entailed testing both the initial premises and new "inference" pairs. These untrained inference combinations had either one (B>D, C>E) or two (B>E) levels of separation between items. (C) Participants first familiarized themselves with the premises via a session providing corrective cues. Right after this session, feedback cues were omitted, and participants were assessed on premise recall. After designated breaks of 20 minutes, 12 hours, or 24 hours, participants revisited the premises and also engaged with the inference pairs to measure understanding of the hierarchy. Adapted from Ellenbogen et al., 2007.

Serial reaction time task The Serial Reaction Time Task (SRTT) is a widely used perceptual-motor skill learning task in cognitive psychology to study implicit learning, which refers to the learning of patterns or sequences without the conscious awareness of the learner (Nissen & Bullemer, 1987). The task typically requires participants to respond in a 4-alternate forced choice design as quickly as possible to a series of stimuli presented sequentially in specific locations, where the sequence of cues covertly follows an embedded grammar. Faster reaction times and improved accuracy when the cues follow the grammar sequence compared with conditions where the cues follow a random sequence are interpreted as evidence of implicit learning. In one notable study, researchers found that sleep led to increased reduction in RT for the grammatical sequence as compared to the random sequence and argued that post-learning reactivation observed during sleep reflected the reprocessing of latent information of sequential contingencies (Peigneux et al. (2003) , see Fig. 5).

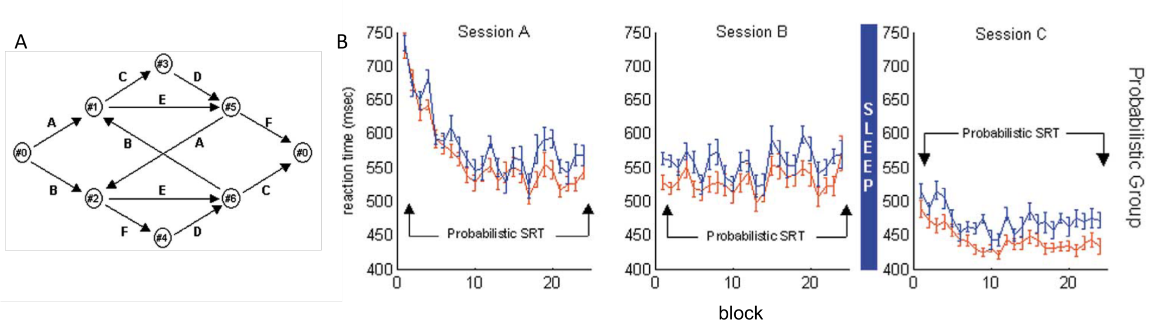


Figure 5: Experimental design of the Serial reaction time task. (A) Depiction of the probabilistic finite-state grammar dictating allowed transitions between consecutive trials, based on Jimenez et al. (1996). Each represented letter aligns with a stimulus position on the display. Upon reaching each node (e.g., #2), a random arrow from the available choices (e.g., #4) is chosen. (B) Line graph showcasing average reaction times, accompanied by standard errors, for each block. Grammatical stimuli are marked by red lines, while random stimuli are indicated by blue lines. This is shown for sessions before and after sleep. Every session encompassed 24 blocks, each with a chain of 205 consecutive stimuli. Modified from Peigneux et al., 2003

MINERVA2 modifications and other models

MINERVA2 traditionally is formulated as shown in Fig. 2B. However, in an attempt to tailor it to the individual tasks we are exploring in this study, we have created new versions of it that either use a different similarity function (S_i) or have a different activation function (A_i). The similarity function serves as the global objective function for the system and by modifying it, MINERVA2 can generate markedly different representations. For example depending on the similarity function we can return more pattern separated (e.g., using Pearson's correlation as a similarity function) or more pattern completed (e.g., using cosine similarity) echoes. Additionally, by modifying the activation function, one can control the magnitude of how differences in similarity between the cue and traces in memory become exaggerated or attenuated over time (e.g., by changing the cubic S_i^3 activation function shown in Fig.~2B (green bounding box) to a quartic function S_i^4, a quantic function S_i^5, or a septimic function S_i^7).

- MINERVA2$_{\text{cosine}}$, where the similarity function operates as cosine similarity (Eq. (2.10), activation function is cubic (S_i^3):

$$S_i = \frac{\sum_{j=1}^{N} P_j M_{i,j}}{\sqrt{\sum_{j=1}^{N} P_j^2}\sqrt{\sum_{j=1}^{N} M_{i,j}^2}} \tag{2.10}$$

For binary vectors, this can be interpreted as the similarity between sets, as the formula for cosine similarity for binary vectors is essentially the count of co-occurrences (both vectors have a value of 1) divided by the geometric mean of the counts of 1s in each vector.

- MINERVA2$_{\text{correlation}}$ (also referred to as MINERVA2$_{\text{AL}}$), where similarity function operates as Pearson's correlation (Eq. (2.11), activation function is cubic (S_i^3):

$$S_i = \frac{\sum_{j=1}^{N}(P_j - \bar{P})(M_{i,j} - \bar{M_i})}{\sqrt{\sum_{j=1}^{N}(P_j - \bar{P})^2 \sum_{j=1}^{N}(M_{i,j} - \bar{M_i})^2}} \tag{2.11}$$

For binary vectors, the correlation can be thought of as capturing the difference in the proportion of times the two vectors co-occur (both are 1) versus when one is 1 and the other is 0.

- MINERVA2$_{\text{Exp+1}}$ which follows the same similarity function as MINERVA2$_{\text{correlation}}$, but with activation function modified from S_i^3 to $S_i^{(3+1)}$

- MINERVA2$_{\text{Exp+2}}$ which follows the same similarity function as MINERVA2$_{\text{correlation}}$, but with activation function modified from S_i^3 to $S_i^{(3+2)}$

- MINERVA2$_{\text{Exp+4}}$ which follows the same similarity function as MINERVA2$_{\text{correlation}}$, but with activation function modified from S_i^3 to $S_i^{(3+4)}$

Additionally, simulations present outcomes from various vector-based memory models, offering alternative memory representations in contrast to MINERVA2 (which employs a multi-vector representation to store memories). While the vector-based memory models were not the primary focus of the hypothesis, they provide a valuable baseline for understanding Minerva models. Further details can be found in Kelly, Mewhort, & West (2017).

- Hologram: This model represents memories using compressed higher-order tensors, such as "3D" or "4D" matrices (Murdock, 1993).

- Tesseract: Memories are represented as auto-associative fourth-order tensors in this model, which has been demonstrated as a higher-order tensor variant of MINERVA2 (Kelly et al., 2017).

- Intersector: This model has similarities with Hologram but incorporates an added lateral inhibition component (Levy & Gayler, 2009).

- Matrix: This model depicts memories using auto-associative second-order tensors, recognized as a higher-order tensor version of MINERVA2 (M. W. Howard & Kahana, 2002).

2.4 Results

2.4.1 Paired associates: interference learning protocol

In the study by Ellenbogen et al. (2009) , researchers observed a higher recall performance in a paired associate task for the Sleep group across sets of words. A notable “retention interval [Wake, Sleep] x interference [interference, no interference]” interaction was also identified, emphasizing a more pronounced difference in recall performance between the Wake and Sleep groups for words with introduced interference pairs (A-C). In our approach, interference between paired associate memories $AB_i - BC_i$ is quantified based on the cosine similarity of their respective traces, as recalled by MINERVA2 variants. A higher cosine similarity indicates increased interference, while a lower value suggests the opposite. Interestingly, by employing $\text{MINERVA2}_{\text{Exp+2}}$ and $\text{MINERVA2}_{\text{Exp+4}}$, we successfully replicated a pattern of results akin to Ellenbogen et al. (2009). This was achieved by modeling the retention intervals (Wake/Sleep) using activation functions with distinct polynomial degrees (S_i^5 vs S_i^7) applied to the same memory set (see Fig. 6 and Fig. 7). The interplay between these activation function polynomials and recurrent recall dynamics qualitatively emulates the sleep-dependent resistance to retroactive interference outlined in Ellenbogen et al. (2009). This outcome stems from the interaction between recurrence and the activation function applied to the correlation-based similarity function. Specifically, the more sensitive, lower-degree polynomial activation function generates more pattern-separated AB-BC echoes during iterative cueing than its higher-degree counterpart, resulting in diminished cosine similarity. Such findings demonstrate that $\text{MINERVA2}_{\text{Exp+2}}$ can manifest representational drift, effectively separating interfering memory traces—a phenomenon intrinsically linked to the active role of sleep-dependent memory consolidation.

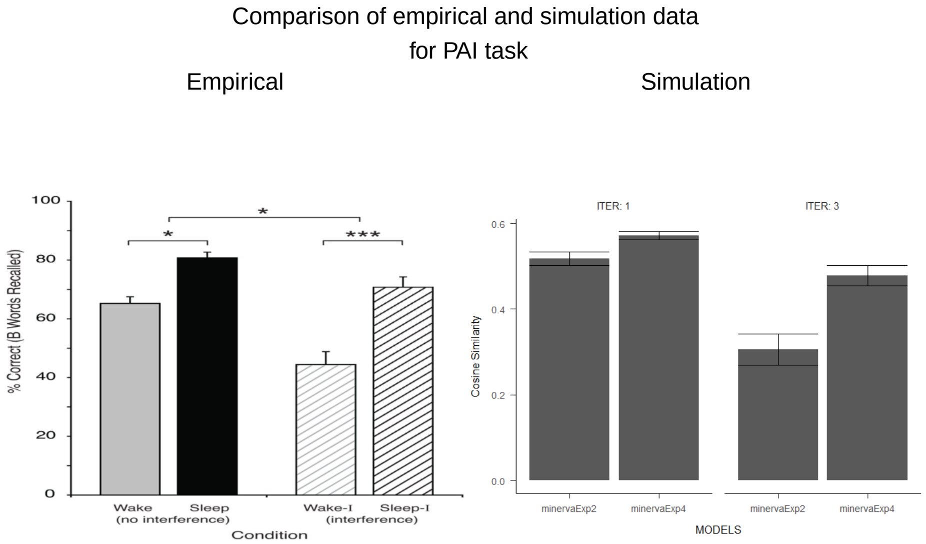


Figure 6: PAI: Comparison of empirical behavioral findings and model simulations. Left panel) Percent correct recall for B words from the original A–B pair following 12-hour retention interval, with no interference and with interference (A–C) prior to testing. Bar indicates one standard error of the mean. Right panel) Performance of MINERVA$2_{\text{Exp+2}}$ and MINERVA$2_{\text{Exp+4}}$ model, left facet shows average cosine similarity between A-B and B-C pairs the 1st iteration of the model being cued, right facet shows 3rd iteration, performance calculated as cosine similarity between each A-B and B-C cued echo at each iteration, simulation detail in the supplement, error bar indicates one standard error of the mean.

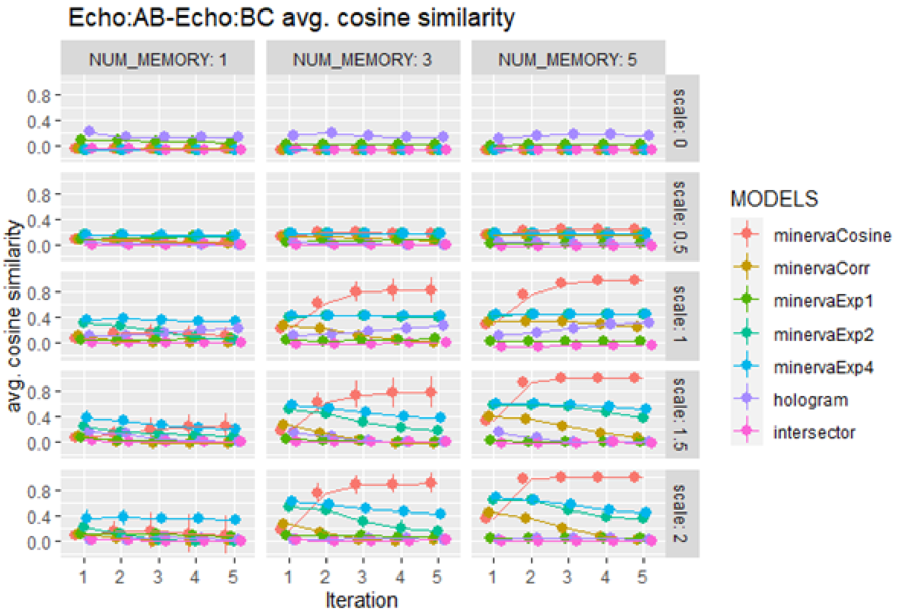


Figure 7: PAI: Detailed model simulations. This figure presents an in-depth simulation analysis considering variations in memory strength and the introduction of uniform noise across multiple recurrent iterations. The columns differentiate the encoding strength associated with each word-pair, while the rows segregate the scale variable that governs the magnitude of uniform noise infused into the feature vectors representing the word-pairs. The x-axis depicts the iterative evolution, highlighting the average cosine similarity (depicted on the y-axis) between the echoes of AB and BC word-pairs. A higher value on this metric suggests increasing similarity between the pairs over iterations, whereas a lower value indicates growing distinctiveness. The error bars encapsulate one standard error of the mean.

2.4.2 Transitive inference task

In the study conducted by Ellenbogen et al. (2007), the researchers observed a notable enhancement in inference performance for novel item pairs, such as B?D , C?E , and B?E , in a transitive inference task after a prolonged delay. Interestingly, the performance for premise pairs, like A?B , remained relatively consistent. One specific analysis delved into the impact of the retention interval duration, comparing a short 20-minute interval to a longer 12-hour gap, between the encoding and recall phases. This was examined for both premise and inference trials using a between-subject design. In our approach, we quantify recall performance based on the cosine similarity between the probe and the traces as recalled by different versions of MINERVA2. For inference probes, intermediate item traces were included; for instance, the probe for B?D incorporated features from items B, C, and D. A higher cosine similarity to the inference prompt indicates a more precise pattern completion based on overlap. By employing the MINERVA2$_{\text{cosine}}$ model, we successfully replicated the findings of Ellenbogen et al. 2007. We modeled the retention interval durations (20 min and 12 h) as the varying number of recurrent iterations, wherein the model continually receives its previous output. Interestingly, after 10 iterations, a pattern-completed trace representing associative inference becomes evident, mirroring the results from the original study (see Fig. 8 and Fig. 9). This phenomenon can be attributed to the synergistic interaction between recurrence and the cosine similarity-based function, which collectively chain together distributed memory traces based on their overlap.

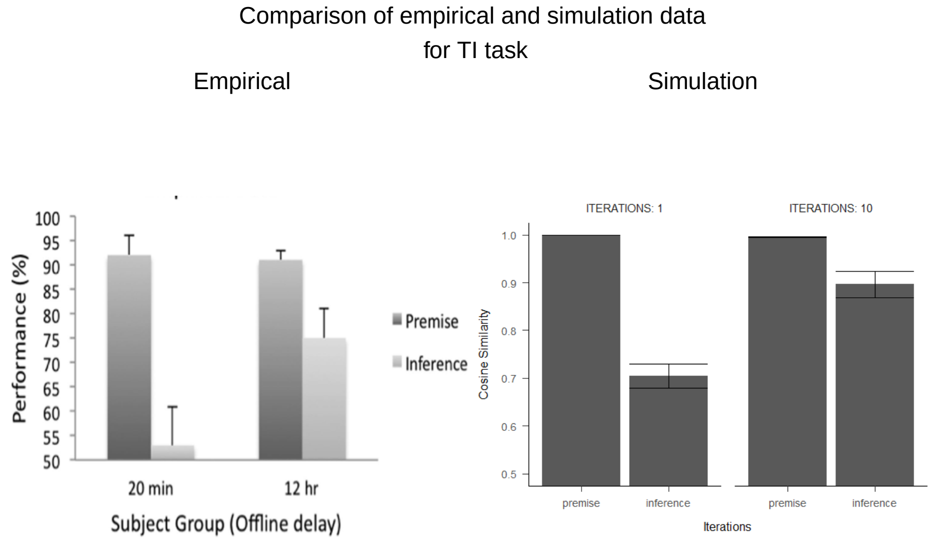


Figure 8: TI: Comparison of empirical behavioral findings and model simulations. Left panel) Empirical results from Ellenbogen et al. (2007), figure adapted from Kumaran & McClleland (2012). The x-axis differentiates between the two experimental groups with 20-minute and 12-hour delays, while the y-axis quantifies performance in percentages. Distinct performances between the groups arise due to variations in the duration and characteristics of the delay between training and evaluation sessions. Performance for premise pairs, represented in dark gray, is averaged across all relevant pair combinations (e.g., A–B, B–C, ... E–F). Inference performance, shown in light gray, is averaged over both close (B–D, C–E) and distant (B–E) inference pairs. Right panel) Results from the MINERVA2$_{\text{cosine}}$ model. The left facet displays premise and inference performance during the model's initial iteration, while the right facet represents the 10th iteration. Performance is measured by the cosine similarity between the initial cue and its echo. Detailed simulation methodologies are provided in the supplementary section. Bars denote one standard error of the mean.

In Ellenbogen et al.'s 2007 study, a particularly intriguing observation was made regarding the impact of sleep on memory retention. The researchers discovered that a retention interval encompassing sleep, as opposed to wakefulness, had a more pronounced positive effect on the performance of distant inference pairs, such as B?E , compared to close inference pairs like B?D and C?E . This differential effect was interpreted by the researchers as evidence of sleep-dependent consolidation processes. They postulated that sleep might enhance individual memory items by facilitating the associative integration of these elements, culminating in the formation of a "metamemory representation". In our analysis, we demonstrate that the MINERVA2 model, akin to other associative chaining models like REMERGE, struggles to replicate this sleep-dependent symbolic distance effect observed in the study. Specifically, while MINERVA2 can successfully achieve distant inference, it invariably exhibits superior performance on close inference pairs compared to distant ones, contrary to the empirical findings.

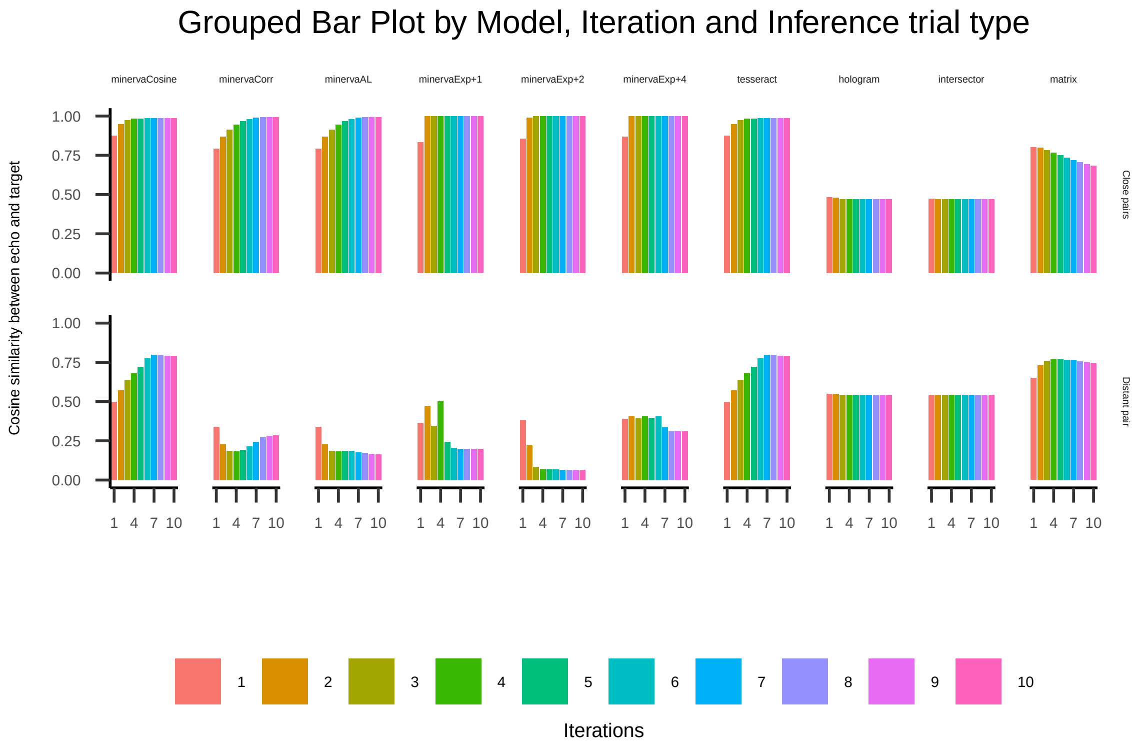


Figure 9: TI: Detailed model simulations. Upper Panel) Performance across iterations for various memory models when tested with Close pair cues (e.g., B-D) that are separated by two degrees. The models, in sequence, are minervaCosine, minervaCorr, minervaAL, minervaExp+1, minervaExp+2, minervaExp+4, tesseract, hologram, intersector, and matrix. The color gradient, ranging from red to pink, represents iterations from the 1st to the 10th. Performance is quantified as the cosine similarity between the initial cue and its echo. Lower Panel) Analogous performance metrics for the same memory models, but when tested with Distant pair cues (e.g., B-E) that are separated by three degrees. The model sequence and color gradient remain consistent with the upper panel.

2.4.3 Serial reaction time task

Peigneux et al. (2003) conducted a study on the robabilistic serial reaction time task (SRTT) that highlighted that after sleep, there was a more pronounced learning of high-order probabilistic rules in participants who had prior training on a probabilistic sequence. Intriguingly, based on neuroimaging results, the authors argued that the neural reactivation in the cuneus during post-training REM sleep wasn't just a result of acquiring basic visuomotor skills. Instead, it signified the reprocessing of detailed information about the sequential contingencies present in the learned material.

With the goal of modelling a simplified deterministic SRTT, we adopt the approach of Jamieson & Mewhort (2009), who adapted MINERVA2 to model sequence learning tasks, incorporating a unique exemplar encoding format. Their adapted model suggests that after each response in the task, participants store an event trace. This trace encompasses the current stimulus, its associated response (location information), and the context given by the preceding response (previous location response). As participants encounter the next stimulus, they utilize this stored trace to retrieve the correct response. With continued practice, the sequence's redundancy offers additional cues pointing to the right response. This extra information subsequently leads to a reduction in response time.

A pivotal aspect of their modeling using MINERVA2 was the incorporation of the iterative resonance model (IRM) concept. This was applied to Minerva's cued recall retrieval mechanism, as detailed by Mewhort & Johns (2005). The IRM, in its functionality, computes a new echo at every time step. The activation function's exponent is initially set to 1 and increases by 1 with each iteration. This increment accentuates the probe's similarity differences to each trace, enhancing the signal-to-noise ratio. As a result, the iterative retrieval process emphasizes traces most similar to the probe. This iterative search is sustained until the echo's response segment closely matches one of the potential responses, determined by a fixed decision criterion (in their simulation, this was defined as $Cr = 0.99$).

Interestingly, the number of these iterations can be perceived as a computational representation of the deliberation time that precedes the motor response in the SRTT by participants. Advancing this methodology, we associate a post-sleep RT trajectory of a previously learned sequence with a lower decision threshold for the iterative retrieval. Our simulations using $\text{MINERVA2}_{\text{sequence}}$ mirrored the results pattern of Peigneux et al. (2003) (see Fig. 10 and Fig. 11). We linked post-sleep structured RT trajectory with a lower decision threshold for iterative retrieval and a random sequence RT trajectory with a higher one. Notably, after ten blocks, distinct qualitative differences became evident between the iteration counts governed by low and high decision thresholds. The lower decision threshold consistently led to quicker correct responses.

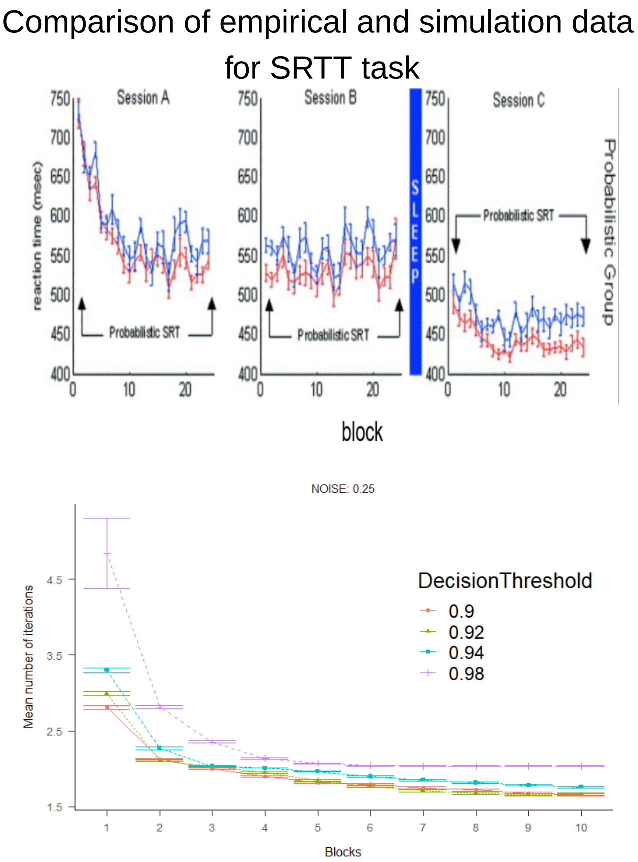


Figure 10: SRTT: Comparison of empirical behavioral findings and model simulations. Top panel) Average reaction times (with standard errors) per block for probabilistic structured (depicted in red) and random (shown in blue) stimuli during both pre- and post-sleep SRTT sessions in the Probabilistic condition. Figure adapted from the Peigneux et al. (2003) study. Bottom panel) Performance metrics of MINERVA2$_{\text{sequence}}$ the model. The x-axis represents block averages of 100 trials, while the y-axis indicates the mean number of recurrent memory retrievals required to identify the correct response, allowing the experiment to advance. Colors represent different levels of decision thresholds.

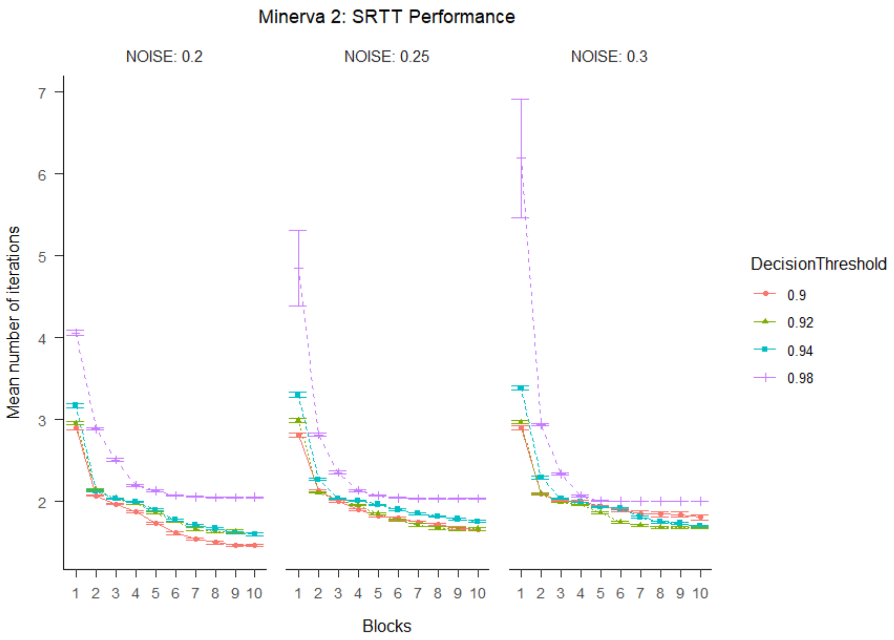


Figure 11: SRTT: Detailed model simulations. Model simulations showcase the impact of varied uniform noise levels across blocks. Each column represents a distinct level of uniform noise. The x-axis displays block averages, each consisting of 100 trials, while the y-axis represents the average number of retrieval iterations required to identify the correct response. The colors in the graph signify different decision thresholds. Error bar provides a visual representation of one standard error of the mean

2.5 Discussion

Qualitative reviews suggest that sleep-dependent memory consolidation (SDMC) doesn't strictly follow the traditional distinctions of memory systems or cognitive domains; instead, it appears to be task-specific (Chatburn et al., 2014; Lerner & Gluck, 2019a). We believe that developing streamlined computational memory models, adaptable to a broad spectrum of existing 'sleep and memory' paradigms, offers a fresh, algorithmic perspective. This approach aids researchers in comprehending tasks at a representational level and in formulating testable predictions that differentiate between competing models of reactivation-dependent effects. Here we show that even a very simple model such as MINERVA2 can account for a range of SDMC effects in the serial reaction time task, transitive inference task and the word-pair learning task (involving interference learning protocol) relying primarily on spontaneous reactivation dynamics.

In the case of 'paired association with interference' this manifested by choosing a similarity function that when applied recurrently could increase separation between AB and BC conjunctions and modelling offline retention interval where this process is facilitated (see Fig. 6 and Fig. 7). When applied to the transitive inference task we saw stepwise associative chaining with each recurrent iteration, where we considered a longer post-encoding retention interval to be akin to longer recurrent sampling during recall (see Fig. 8 and Fig. 9). For the serial reaction time task, we showed that by varying the decision threshold that determines the quality of the stored echo (i.e.: replay content) we could manipulate the learning trajectory (see Fig. 10 and Fig. 11). Interestingly the lower reactivation fidelity granted by the more lenient threshold produced faster learning (i.e.: faster correct responses) on this task which we associated with offline reactivation. An overarching theme is that these results were all produced by manipulating recall dynamics, not the stored representations themselves. While this might seemingly align with permissive and passive views of sleep-dependent memory consolidation (Ellenbogen et al., 2006), MINERVA2 is neutral in this regards as it can store the traces it produces which would lead to the kind of representational drift that could be associated to active system consolidation, whereby reactivated episodes form gist and schemas (Klinzing et al., 2019) . We argue that this afforded flexibility also favors it against many other exemplar models of generalization when modelling time and sleep-dependent consolidation. An additional important question is how can offline consolidation both reduce interference between interfering memories in certain cases (Ellenbogen et al. (2009) although for detailed review see Abel, Nickl, Reßle, Unger, & Bäuml (2023)) while also integrate and associatively bind memories in others (i.e. pattern separation versus pattern completion). We do not provide a formal answer to this question, but instead show that MINERVA2 can exhibit both phenomena given the appropriate similarity function.

An attempt to re-examine the existing literature through these competing (P.-C. Chen, Niknazar, Alaynick, Whitehurst, & Mednick, 2021) or cooperating (Born & Wilhelm, 2012) processes using cognitive models could lead to testable sleep-stage specific hypotheses (Doxey, Hodges, Bodily, Muncy, & Kirwan, 2018; Hanert, Weber, Pedersen, Born, & Bartsch, 2017; Poh & Cousins, 2018) and help better explain null results (Svenja Brodt, Pöhlchen, Täumer, Gais, & Schönauer, 2018; Hołda, Głodek, Dankiewicz-Berger, Skrzypińska, & Szmigielska, 2020; Talamini et al., 2022a). An example can be found in a recent meta-analysis summarizing the effects of sleep in the Deese–Roediger–McDermott (DRM) paradigm (13 studies) that observed no general increase of false memories after retention across sleep when compared to wakefulness (Newbury & Monaghan, 2019a) . However, when conducting a moderator analysis two important factors emerged: one being the performance measure, as sleep effects were greater in studies using recall measures than in recognition tests. Second, the size of the word list, whereby sleep increased false memories when lists of 10 words were used but decreased false memories in lists of 15 words. The authors posited that smaller learning list results in a sparse associative network in long-term memory, for which sleep can better increase the spreading activation leading to more false memories as compared to a densely related network. Similar capacity limitations have been shown in the paired association task (Gordon B. Feld, Weis, & Born, 2016a), and in mnemonic discrimination tasks (Cellini, Mercurio, Vanzetti, Bergamo, & Sarlo, 2020). These kinds of three-way interactions are ideal problems to apply cognitive models to and can be tackled by spreading activation models like REMERGE or similarity-based models like MINERVA2.

When considering list size — whether it's composed of words, images, or sounds — in relation to encoding strength or interference during recall, exemplar models predict a positive correlation between memory strength, inference, and false memory (Newbury & Monaghan, 2019a) when directed to pattern complete. This leads us to ponder: To what extent are offline reactivations (or replay) observed following episodic memory tasks reminiscent of exemplars? Exemplar-based abstraction theories seem to be better suited for modeling spontaneous hippocampal reactivation as a process that initially starts out as singular event reactivations as opposed to one that is generalized from the very start (Kumaran & McClelland, 2012a). If offline reactivation post-encoding is strongly influenced by prior knowledge or highly abstracted schemas (i.e., more cortically driven), these models might not be entirely applicable. However, in the latter case they can still serve as baseline models for such experiments. Further testable predictions arise when considering the differences in assumed representational geometry of source memory between REMERGE and MINERVA2. REMERGE relies on orthogonal conjunctive memories during its recurrent similarity computation which are assumed to stay orthogonal following retrieval. MINERVA2 on the other hand can operate on either orthogonal or overlapping initial traces and account for increased similarity based overlapping over repeated recall. These differences can be approached using model-driven neuroimaging studies involving representational similarity analysis (Kriegeskorte, Mur, & Bandettini, 2008). While a comprehensive, multivariate depiction of the precise temporal dynamics of reactivation content remains elusive for the studies discussed in this chapter, research employing RSA to investigate associative inference (Barron et al., 2020; Erie D. Boorman, Sweigart, & Park, 2021a; Koster et al., 2018a) and statistical learning (Kalm & Norris, 2018) suggests a leaning towards overlap-based models of learning. This contrasts with the associative strengthening-based account proposed by REMERGE.

We contend that MINERVA2 holds value in particular scenarios, such as examining reactivation-dependent shifts in basic episodic memory and statistical learning tasks. However, like all models, it is not without its limitations. As a representation of rapid or retrieval-based generalization, it serves as a useful counterpart to REMERGE (Kumaran & McClelland, 2012). However, when considering sleep-dependent generalization across long time scales and multiple memory systems, MINERVA2 lacks the intricate complexity needed to differentiate between detailed episodes, gist, and schema information (Sekeres, Winocur, & Moscovitch, 2018a; J. E. Taylor et al., 2021) changing across multiple memory systems. These elements appear to possess distinct offline reactivation dynamics (Fenn, Margoliash, & Nusbaum, 2013). Other memory models that leverage its simplicity could potentially be expanded to represent SDMC, integrating interactions between short-term and long-term memory crucial for sleep-driven generalization (Grossberg, 1986; Nelson & Shiffrin, 2013). Another significant limitation to note is the absence of a temporal dimension, which seems essential for accurately modeling sequential memory tasks. Incorporating this dimension will likely be pivotal to capture time-compressed replays coordinated between the hippocampus and cortex during sleep and to comprehend the emergence of limited replay opportunities (M. D. Howard, Skorheim, & Pilly, 2019). Finally, given its reliance on similarity-based reactivations, MINERVA2 may not be apt for capturing the time and sleep-dependent effects of relational generalization. Such generalization doesn't hinge on similarities in stimulus features and contexts but arises from information transfer based on arbitrary learned associations. It's crucial to emphasize this, especially as increasing evidence suggests that reactivation dynamics are influenced by cognitive maps (and graph) structures (Wittkuhn, Chien, Hall-mcmaster, & Schuck, 2021b).

Nonetheless, we contend that, given the uncertainties surrounding the boundary conditions of when time and sleep facilitate generalization, simple models adaptable to various memory paradigms will be crucial. As the decoding of reactivation content improves, whether it's cued or spontaneous, models like REMERGE and MINERVA2 can assist researchers in exploring the interplay between bottom-up and top-down influences governing reactivation. This exploration can be achieved through combined representational and behavioral model-fitting. Furthermore, aligning with theories that underscore the sequential significance of sleep-stage specific differences in reactivation content, exemplar models offer a preliminary look into the effects of interleaving different "types" of replays. For instance, understanding the interplay between pattern-separating and pattern-completing types of replays can shed light on the boundary conditions that produce robust representations (whether at an episodic, gist, or schema level), or lead to no sleep benefit or even catastrophic forgetting. However, we believe that even in its current state, MINERVA2 can contribute meaningfully to the discourse on the active, passive, and permissive roles of sleep in consolidation (Durrant, Taylor, Cairney, & Lewis, 2011; Rasch & Born, 2013; Stickgold & Walker, 2013).

2.6 Methods

2.6.1 Paired associates: interference learning protocol

For the configuration of the MINERVA2 memory matrix, we utilized an overlap-based encoding scheme, specifically for the AB and BC traces that represent word-pairs. The task representation was real-valued. The model was set to run with a total of 10 iterations and was executed over 20 runs. Within the memory, there were 6 items each for AB and BC. The dimensionality of the model was set at 96. To introduce variability, uniformly distributed random noise was added to the vectors representing the word list. This noise was adjusted according to a scaling variable and was mean-centered.

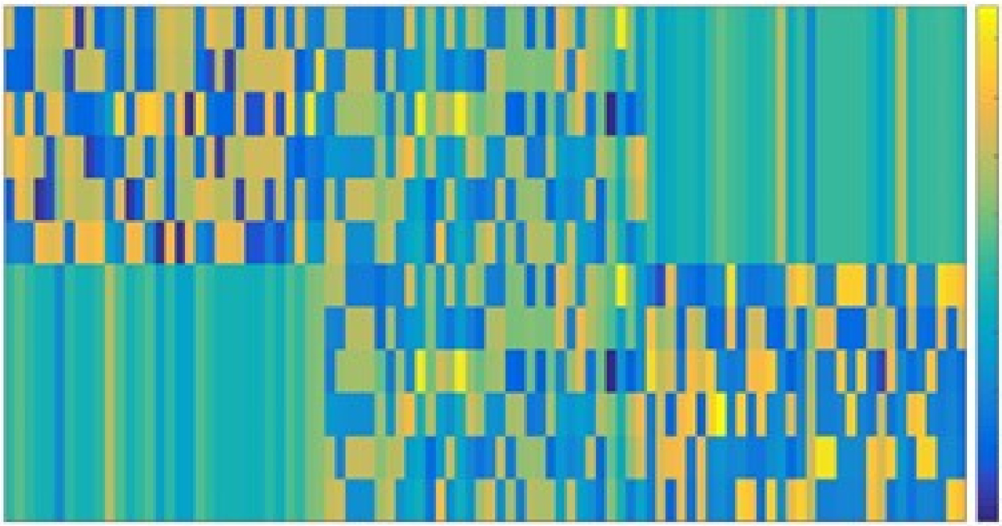


Figure 12: Colormap of the Memory Matrix for the Paired associate task. This matrix visualizes the interference protocol using an overlap-based encoding scheme. Each row delineates a specific paired association (Ai-Bi), followed by padding to maintain uniformity across entries.

2.6.2 Transitive inference task

In configuring the MINERVA2 memory matrix, we adopted an overlap-based encoding scheme for neighboring conjunctions, specifically representing premise pairs such as AB, BC, and so forth. The task representation was set to binary. The model was designed to undergo 10 iterations and was executed across 20 runs. The memory was populated with a total of 8 items, and the dimensionality of the model was established at 9.

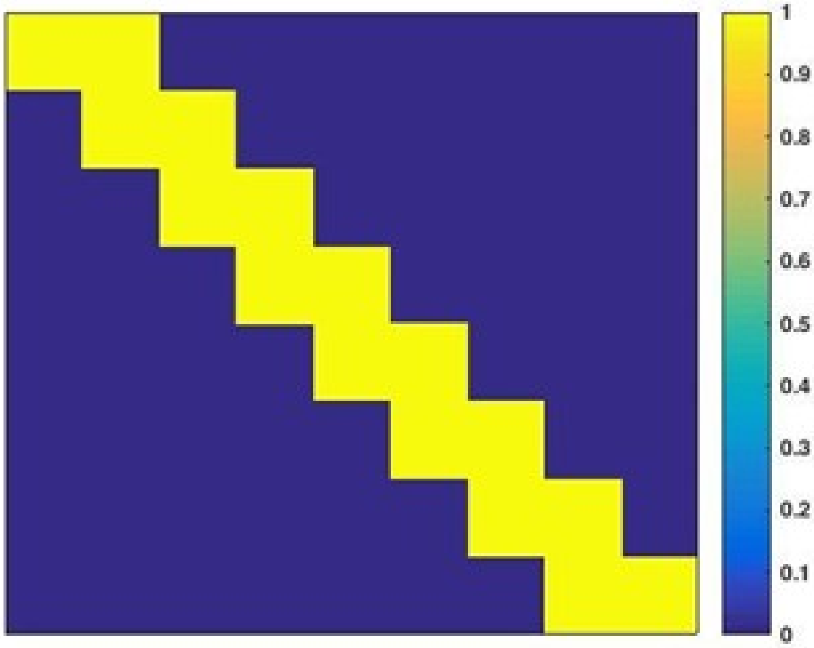


Figure 13: Colormap of the Input Matrix for the Transitive Inference Task: Visualizing the overlap-based encoding scheme, where each row delineates a specific overlapping association, progressing from A-B to B-C and so on.

2.6.3 Serial reaction time task

For the configuration of the MINERVA2 memory matrix, we adhered to the design proposed by Jamieson & Mewhort (2009) . In their approach, to account for the instructions provided to participants before commencing the SRTT experiment, the memory was populated with one example of each S(i)//R(i-1)//R(i) contingency, setting L to 1. The task representation was binary in nature. The model was designed to undergo 10 iterations and was executed over 20 runs. Initially, the memory was equipped with 6 items, and the overall dimensionality was set at 60. The noise parameter, denoted by L, indicates the proportion of features encoded as zero.

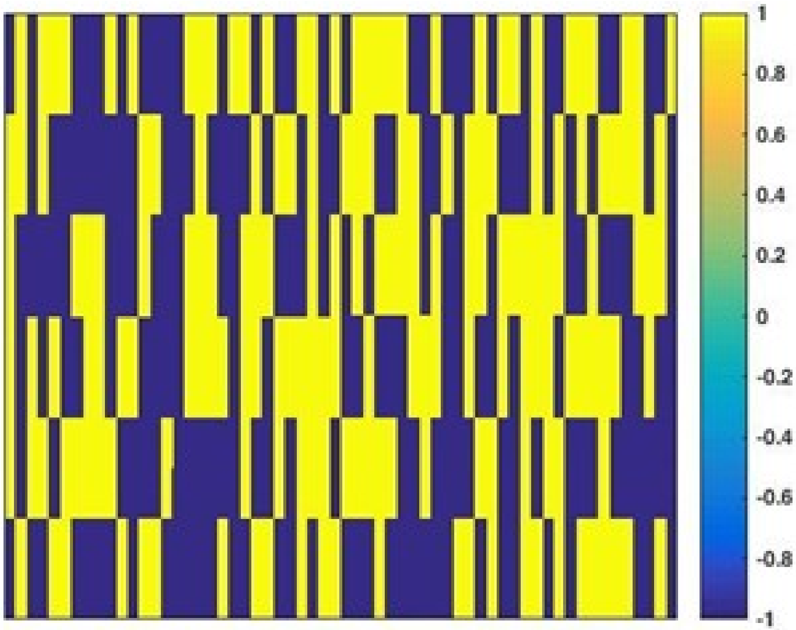
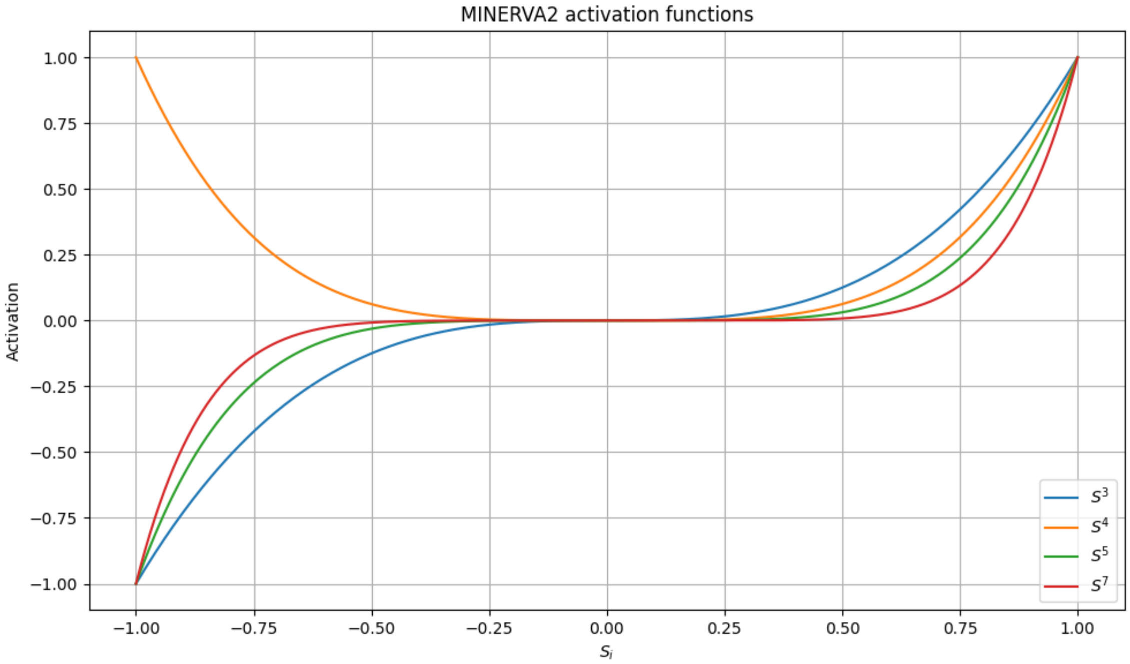


Figure 14: Colormap of the Input Matrix for the Serial Reaction Time Task. Utilizing an overlap-based encoding scheme, each row showcases a distinct binary memory. Indices 1 to 20 represent the unique stimuli encoding the item-location association. Indices 21 to 40 capture the correct response to the stimuli from the preceding trial, while indices 41 to 60 denote the correct response for the current trial, culminating in a Si//R(i-1)//Ri encoding structure

2.7 Acknowledgement

The simulations showcased herein were elaborations upon the foundational code developed by Dr. Mary Alexandria Kelly, which can be accessed at the following repository: `https://github.com/ecphory/Memory-Tesseract`.

2.8 Supplemental materials



Supplemental Figure S1: Graphical representation of the MINERVA2 activation functions. The plot showcases four different activation functions: S^3, S^4, S^5, and S^7, over the range of S_i from -1 to 1. The y-axis represents the activation value, while the x-axis represents the value of S_i. Each function exhibits distinct curvature characteristics within the specified range.

Chapter 3

Sleep-related benefits to transitive inference are modulated by encoding strength and joint rank I.

3.1 Abstract

Transitive inference is a measure of relational learning which has been shown to improve across sleep. Here, we examine this phenomenon further by studying the impact of encoding strength and joint rank. In Experiment 1, participants learned adjacent premise pairs and were then tested on inferential problems derived from those pairs. In line with prior work, we found improved transitive inference performance after retention across a night of Sleep compared to Wake alone. In Experiment 2, we aimed to replicate previously observed sleep-dependent benefits of transitive inference using a larger and more diverse online sample of students participating for course credit; however, our results were mixed, showing no significant effect of sleep on delayed inference performance, contrasting with the findings from Experiment 1. Experiment 3 extended these findings using a within-subject design and found superior transitive inference performance on a hierarchy, consolidated across 27 hours including sleep compared to just 3 hours of wake. In all three experiments, consolidation-related improvement was enhanced when pre-sleep learning (i.e., encoding strength) was stronger. We also explored the interaction of these effects with the joint rank effect, in which items are scored according to their rank in the hierarchy, with more dominant item pairs having the lowest scores. Interestingly, the consolidation-related benefit was greatest for more dominant inference pairs (i.e., those with low joint rank

scores). Overall, our combined findings provide further support for the improvement of transitive inference across a consolidation period which includes sleep. We additionally show that encoding strength and joint rank strongly modulate this effect.

3.2 Introduction

Relational reasoning is the cognitive process of identifying and understanding relationships between stimuli or concepts. It involves the ability to identify patterns, e.g. similarities and differences between stimuli, and to understand how they relate to each other (Halford, Wilson, & Phillips, 2010). Relational reasoning is important for a wide range of cognitive tasks, such as problem-solving (Dumas, Alexander, Baker, Jablansky, & Dunbar, 2014), decision-making (Dumas, 2016), and language acquisition (Gentner & Namy, 2006). Transitive inference is a specific type of relational reasoning that involves making inferences about the relationships between items in a hierarchy. For example, if A is dominant to B, and B is dominant to C, then we can infer that A dominates C when probed (A?C). This type of inference requires the ability to reason about the relationships between stimuli based on their relative positions in a hierarchy. Transitivity, a property of all ordered sets, has been extensively studied in human developmental psychology for over a century (Piaget, 1921) and by comparative cognition research (McGonigle & Chalmers, 1977) for close to half a century. However, the mechanisms underlying transitive inference are not fully understood (Holyoak & Lu, 2021).

The study of time-dependent consolidation in relational learning has mostly been carried out in animal models, such as rats and non-human primates involving extensive training procedures that aim to establish a high level of accuracy in the animals' behavior. However, there is growing interest in investigating this phenomenon in humans, given its potential relevance for education (H. A. Golkashani, Ghorbani, Leong, Ong, & Chee, 2023a; H. Golkashani et al., 2022a), rehabilitation and clinical settings (Avery et al., 2021; Lepage, Hawco, & Bodnar, 2015; Musa, Khan, Mujahid, & El-Gaby, 2022; Nour, Liu, Arumuham, Kurth-Nelson, & Dolan, 2021). To date, several studies have examined the effects of sleep and wakefulness on the consolidation of relational memories in humans, but the results have been mixed (Berens & Bird, 2022a; Ellenbogen et al., 2007; Matorina & Poppenk, 2021a; Werchan & Gómez, 2013b) and the underlying mechanisms are not well understood.

3.2.1 Encoding strength, sleep-related benefits and offline generalization

The process of offline consolidation, which refers to the strengthening and integration of memories during sleep, is influenced by various factors, including encoding strength, prior knowledge and individual differences in cognitive abilities. Encoding strength refers to how well information is initially encoded into memory, and it has been argued to play a crucial role in the offline consolidation process (for review, see Diekelmann, Wilhelm, & Born (2009)). Several studies have demonstrated that sleep-mediated memory benefits are strongly influenced by pre-sleep memory strengths. For instance, when participants had already learned the material to a high degree of accuracy before sleep, sleep did not further enhance memory performance, however, beneficial effects of sleep were observed for weaker memory traces in the same study (Drosopoulos, Schulze, Fischer, & Born, 2007a; Kuriyama, Stickgold, & Walker, 2004; McDevitt et al., 2015; Schapiro, McDevitt, et al., 2017; Schapiro, McDevitt, Rogers, Mednick, & Norman, 2018). On the other hand, when memories are encoded too weakly, sleep-mediated consolidation can be compromised (Walker et al., 2019). The memory intensities attained in these studies were influenced by a variety of factors, including the depth of encoding (Drosopoulos, Schulze, Fischer, & Born, 2007b), salience (Wilhelm et al., 2011), number of encoded items (Gordon B. Feld et al., 2016a; Kolibius, Born, & Feld, 2021a), encoding strategy (Denis, Bottary, Cunningham, Tcheukado, & Payne, 2023) (for an overview see (Berres & Erdfelder, 2021)) . Manipulating memory strength can also be achieved by varying the amount of learning that takes place, such as offering more or fewer immediate cued recall rounds before the retention interval. Immediate testing after learning has been shown to enhance memory formation through the testing effect, which involves more effortful processing and facilitates memory accessibility during later recall (Sutterer & Awh, 2015). Repeated testing of information immediately also increases their future relevance, which has been found to promote the sleep effect (Wilhelm et al., 2011, although see (J. Ashton & Cairney, 2021; Wamsley, Hamilton, Graveline, Manceor, & Parr, 2016)). With regards to neurophysiology, baseline performance during learning has been found to predict the involvement of sleep spindles in memory consolidation. Sleep spindles are brief bursts of brain activity that occur during non-rapid eye movement (NREM) sleep and are believed to play a role in memory consolidation. A study by Wislowska, Heib, Griessenberger, Hoedlmoser, & Schabus (2017) found that individuals with higher baseline performance during learning showed a stronger association between sleep spindles and memory consolidation. Moreover, immediate recall EEG patterns, which reflect the electrical activity of the brain, may serve as markers of the entry into the consolidation process. Heib et al. (2015) conducted a study in which they found that specific EEG patterns during immediate recall were associated with subsequent memory performance, suggesting that these patterns could indicate the initiation of memory consolidation.

Secondly, a related concept to learning strength is the degree of conformity of the learned material to prior schemas (prior knowledge), the direction of this influence, however, is controversial and has been the subject of debate in the literature. Some studies have reported that a higher level of prior knowledge for the to-be-learned information before sleep can lead to a lower sleep-mediated memory benefit. For example, Payne et al. (2012) found that for a set of semantically related and unrelated word-pairs in the retest conducted after 12 hours, the overall memory retention was markedly better following a night's sleep as opposed to a day of wakefulness. Nonetheless, this disparity in performance was largely due to a significant decline in the recall of unrelated word pairs over the period of wakefulness, with no noticeable difference in the recall of related word pairs between sleep and wake states. Similarly, J. E. Ashton, Staresina, & Cairney (2022) the findings of the study revealed that in a noun-colour pairing task ,while plausible pairings were generally better remembered than implausible pairings, the benefits of sleep were stronger for implausible memories compared to plausible memories. The study conducted by Havas et al. (2018) aimed to investigate the influence of semantic and phonological schema on the acquisition and overnight consolidation of new spoken words. The researchers examined the impact of existing knowledge (schema) on word learning by manipulating both semantic familiarity (familiar vs. unfamiliar objects) and phonological familiarity. Spanish-speaking participants learned the spoken forms of novel words in their native language (Spanish) or in an unfamiliar, non-native language (Hungarian). The participants underwent training and testing with a 12-hour intervening period that included either overnight sleep or daytime wakefulness. The findings of the study revealed that sleep improved recognition memory more for words that sounded like they came from the second, non-native language as compared to words that sounded like they came from the person's first language.

Conversely, numerous studies emphasize the critical role sleep plays in assimilating new information, which engages the hippocampus, into pre-existing cortical knowledge networks. This is proposed to be a fundamental function of sleep in memory formation, as per the active systems consolidation hypothesis (Born & Wilhelm, 2012;

Diekelmann & Born, 2010; Klinzing et al., 2019; Lewis & Durrant, 2011). The study conducted by Hennies, Lambon Ralph, Kempkes, Cousins, & Lewis (2016) aimed to investigate the interaction between sleep-related mechanisms and schema-dependent memory consolidation. The researchers examined how sleep spindle activity, a characteristic brain oscillation during sleep, influenced the effect of prior knowledge on memory consolidation. Participants in the study first established a schema over a period of two weeks. They then encoded new facts that were either related to the schema (SR) or unrelated (NR). After a 24-hour retention interval, including a night of sleep, participants encoded a second set of facts. Memory for all facts was tested using fMRI. The findings of the study revealed a improved SR memories as well as a schema × delay interaction, suggesting that the schema effect increased across time. Furthermore the sleep spindle density predicted an increase in the schema benefit to memory across the retention interval. Higher spindle densities were associated with reduced decay of schema-related memories. Functionally, spindle density predicted increased disengagement of the hippocampus, a brain region important for memory, specifically for schema-related memories. These results suggest that sleep spindle activity is associated with the effect of prior knowledge on memory consolidation. In another study, Durrant, Cairney, McDermott, & Lewis (2015) showed that only schema-conformant items, but not non-conformant items, benefited from retention intervals filled with sleep. In a second language learning task, Zion, Nevat, Prior, & Bitan (2019) found that participants with a higher degree of meta-linguistic knowledge benefited more from a retention interval filled with sleep. Groch, Schreiner, Rasch, Huber, & Wilhelm (2017) found that the benefits of targeted memory reactivation during sleep were observed only for familiar word-object associations, suggesting that prior knowledge is a prerequisite for sleep-mediated memory consolidation.

The degree of schema congruence could relate to the question of whether a given learning session get rapidly consolidated or not which thus limiting the role of sleep as stated by ASC hypothesis. This suggests that systems consolidation, the gradual process through which memories are redistributed to long-term stores, could occur much faster than previously believed, possibly reducing sleep's role in this process. For example studies have shown that when the online material is already strongly linked with existing cortical information, the benefits of offline consolidation are diminished, indicating the influence of associative encoding (Himmer, Müller, Gais, & Schönauer, 2017). Indeed, recent research has shown that memory engrams, the physical representations of memory, are formed not just in the hippocampus but also in slow-learning systems like the neocortex, right from the onset of learning (S. Brodt et al., 2018; Hebscher et al., 2019). However, specific studies indicate that sleep retains a crucial role in memory consolidation even amid rapid neocortical plasticity (Pöhlchen & Schönauer, 2020).

A study conducted by Cordi, Schreiner, & Rasch (2023) aimed to investigate the role of prior knowledge in sleep-associated memory benefits and whether additional training opportunities could restore these benefits under conditions of low prior knowledge, thus examining the relative importance these two previously discussed factors. Participants in the study were divided into two groups: a high prior knowledge group (German native speakers) and a low prior knowledge group (French native speakers) in foreign language (Dutch) word learning task. Memory strength was additionally manipulated via varying cued retrival blocks. Both groups underwent a learning phase where they were presented with word-picture associations. The high prior knowledge group received associations that were related to their existing knowledge, while the low prior knowledge group received associations that were unrelated and unfamiliar. After the learning phase, both groups had a 12-hour retention interval that included either sleep or wakefulness. Following the retention interval, participants underwent a recognition memory test to assess their memory performance. The findings of the study revealed - while showing robust main effect of sleep vs wake - that memory strength, rather than prior knowledge, determined whether sleep enhanced memory or not. Strong initial memories benefited from sleep compared to wakefulness, while weak memories did not. In Germans, the necessary memory strength was achieved after two learning trials, while it took three trials for French participants. This suggests that prior knowledge can expedite achieving the necessary memory strength, but this advantage can be compensated by additional learning efforts. The study concluded that initial memory strength, rather than similarity to the native language, determines whether sleep enhances memory retention.

Thirdly, research has demonstrated that sleep-dependent consolidation is influenced by individual differences in skill levels and cognitive abilities (G. B. Feld, Bernard, Rawson, & Spiers, 2022a; Fenn & Hambrick, 2012; Tucker & Fishbein, 2008), suggesting a unique interplay between personal characteristics and sleep's role in memory formation and learning. The study by Tucker & Fishbein (2008) investigated the effects of daytime napping on declarative memory performance and how it is contingent on the strength of initial task acquisition. The researchers proposed a physiological model of sleep-dependent memory consolidation, suggesting that the reactivation of hippocampal and neocortical networks during immediate recall primes these networks for subsequent NREM sleep-related processing. Authors evaluated performance on three declarative tasks after a training-retest interval of 3.5 hours, which either incorporated a daytime NREM nap or did not. To examine how task acquisition factors might influence sleep-associated memory processing,participants were subjected to two types of encoding conditions (test/not tested prior to sleep). Furthermore, for each task, based on their training performance,participants were post hoc categorized into high and low performers (i.e., those ranking in the upper and lower halves of the sample according to a median split). The study found that high performers not only gain more from sleep compared to those who don't nap in each task, but this performance improvement is also linked with the proportion of Slow Wave Sleep (SWS) experienced during the nap.

In terms of moderating effects of cognitive capacity, the study by Fenn & Hambrick (2012) aimed to investigate the relationship between working memory capacity (WMC) and sleep-dependent declarative memory consolidation. The researchers examined whether individual differences in online cognitive processes, specifically WMC, were related to the offline processes of sleep-dependent memory consolidation. The study involved participants learning word pairs and then undergoing either a period of sleep or wakefulness. Memory performance was assessed before and after the sleep or wake period. The results showed that memory for word pairs improved after sleep, but not after wakefulness. Importantly, there was a significant positive correlation between WMC and the increase in memory performance specifically after sleep, suggesting that individuals with higher WMC showed greater improvement in memory after sleep. The correlation between WMC and memory performance was specific to the change in memory due to sleep, as there was no significant correlation between WMC and performance during the initial test, suggesting that WMC is specifically related to the sleep-dependent consolidation process rather than overall memory performance. These findings were recently corroborated by Martini, Marhenke, Martini, Rossi, & Sachse (2020) (but see (Gordon B. Feld et al., 2016a; Kolibius, Born, & Feld, 2020; Scullin, Trotti, Wilson, Greer, & Bliwise, 2012)).

While most research on sleep-related encoding strength focuses on simple declarative memory involved in paired associations, our understanding of the relationship between encoding strength and sleep-related generalization is less developed. The extraction of the gist, or general meaning, from information is expected to be enhanced by sleep, particularly when learning involves a word list centered around a common topic or gist. Some studies supported this claim, demonstrating that both nighttime sleep and napping increased the recall of false memories compared to wakefulness in the Deese-Roediger-McDermott (DRM) paradigm, however, a recent meta-analysis of 13 studies in the DRM paradigm did not find a general increase in false memories after sleep compared to wakefulness (Newbury & Monaghan, 2019a), but see (Mak, O'Hagan, Horner, & Gaskell, 2023a). The size of the word list was however identified as an important factor, with sleep increasing false memories in 10-word lists but decreasing false memories in 15-word lists (Newbury & Monaghan, 2019a). This discrepancy may be due to the presence of a sparse associative network in long-term memory for smaller word lists, which sleep can enhance to produce more false memories.

The effect of sleep on gist extraction and problem-solving insight is likely to be more specific and task-dependent than previously believed. For example in a DRM experiment, Diekelmann, Born, & Wagner (2010) found that the effect of sleep on DRM false recall was modulated by veridical recall, such that sleep (vs. wake) increased false recall, but only among low "veridical" performers (although this did not replicate in a larger sample Mak, O'Hagan, et al. (2023a)). Another DRM study found that sleep increased veridical and gist-based memory when list words were presented in order of associative strength, while veridical memory was enhanced when list words were presented randomly (Wernette & Fenn, 2023a).

It is possible that there is a threshold of memory strength that needs to be reached for consolidation to occur, and sleep may consolidate memories above this threshold but not weaker memories. Deep encoding, which produces stronger memories than shallow encoding, may generate memories that reach this consolidation threshold (Wernette & Fenn, 2023a). Other encoding conditions, such as intentionality, may also influence retrieval processes and strategies. Incidental encoding, where individuals do not anticipate the importance of the information, may lead to a reliance on gist-based representations during testing, particularly when source information or item-specific details are not available or when veridical memory is weak. Taken together, these findings highlight the significance of online factors, such as encoding strength, associative encoding, repeated retrieval, baseline performance, in shaping the effects of sleep on human memory irrespective of whether gist-related generalization processes can be assumed to be involved.

3.2.2 Transitive inference and time/sleep-dependent consolidation

One area of ongoing research in the study of transitive inference is the role of time and sleep-dependent consolidation in the formation and retention of transitive inference abilities. Several studies have proposed that the process of memory consolidation, particularly that which occurs during sleep, plays an integral role in this cognitive phenomenon. This is based on the premise that sleep serves not only to stabilize memories, but to qualitatively transform them and this transformation may result in the abstraction of the gist of a memory or the discovery of latent structure underlying a learned material (Inostroza & Born, 2013; Lerner & Gluck, 2019a; Lewis & Durrant, 2011; Lewis et al., 2018b). In one of the first studies investigating the impact of time and sleep on performance in the transitive inference task Ellenbogen et al. (2007) reported that participants performed transitive inferences significantly above chance if they were allowed a night of sleep between training and testing, whereas those tested immediately after the session were at chance. This study also found that there was a more pronounced improvement in performance on more distant inference pairs, (e.g., B?E, which requires two inferential steps to determine their relationship, C and D), when participants had slept between the training and testing session. The symbolic distance effect (SDE) is a phenomenon observed in various cognitive psychology experiments where reaction times and/or accuracy are affected by the numerical or conceptual distance between stimuli (Moyer & Landauer, 1967). Specifically, response times decrease and accuracy increases as the distance between stimuli increases. It has been observed in various tasks, and in the context of transitive inference has been used to support the idea that participants learn the relative rank of stimuli in the hierarchy. While Werchan & Gómez (2013, 2016) partially support this sleep-dependent symbolic distance effect, their amended results are not technically significant, and other variations also fail to replicate the effect. For instance, a shorter delayed test (3h) either involving a short nap or no nap (Morgan & Stickgold, 2017) and within-subject studies involving a 12-hour delayed test on a 6-item hierarchy (Matorina & Poppenk, 2021b) or 24-hour delayed test on a 7-item hierarchy (Berens & Bird, 2022b) found above chance inference performance at delayed test, but no time-dependent SDE.

Studies have suggested that time, and more specifically sleep, plays a crucial role in the consolidation of newly acquired relational knowledge, potentially helping to strengthen the connections between stimuli in a hierarchy and improve overall inference performance (Alger & Payne, 2016; Lau, Tucker, & Fishbein, 2010). Others have argued that many earlier findings showing sleep-mediated benefits on memory are not robust (Cordi & Rasch, 2021) and that sleep-dependent benefits in statistical learning tasks are highly task-dependent and implicit only (Lerner & Gluck, 2019a). We propose that employing diverse study designs and control variables can facilitate the elucidation of boundary conditions for time and sleep-dependent gist extraction. Although considerable knowledge exists regarding moderators of sleep-related benefits for non-relational episodic memories (Berres & Erdfelder, 2021), it remains unclear how these factors pertain to sleep-dependent generalization studies (Pereira et al., 2023). An essential factor of interest is encoding strength, as studies may establish various learning criteria in their study phase, depending on whether previous research suggests that sleep enhances the retention of weaker memories to a greater extent than stronger memories (Diekelmann et al., 2009) or vice versa (Tucker & Fishbein, 2008). While the use of a learning criterion can ensure that participants have acquired a minimum level of proficiency, individual differences in acquired performance can lead to varying sleep benefits (Denis et al., 2021a). We hypothesize that encoding strength will affect not only overall inference performance but also qualitative measures of hierarchical learning. Thus, we not only expect to replicate earlier findings showing sleep-dependent differences in symbolic distance effect (a measure of relative positional encoding) but also propose an additional measure of absolute positional encoding, based on the summed rank of a given pair's constituent item ranks (e.g., if A>B>C the summed rank of A?B would be 1+2=3 and for B?C it would be 2+3=5), hereafter referred to as joint rank value (Jensen, Alkan, Muñoz, Ferrera, & Terrace, 2017b).

This is motivated by two strands of research. On one hand, a study by Kao, Jensen, Michaelcheck, Ferrera, & Terrace (2020) extended the single hierarchy transitive inference (TI) paradigm by using a derived list (or derived hierarchy) approach, in which participants were instructed to learn the ordinal structure of five hierarchies consisting of five items each (e.g., $A_{H1} > B_{H1}$ $C_{H1} > D_{H1} > E_{H1}, \ldots, A_{H5} > B_{H5} > C_{H5} > D_{H5} > E_{H5}$). Participants were then tested on both adjacent and nonadjacent pairs of items from five different derived hierarchies, and responses were scored such that the hierarchies were mixed, but the ordinal positions of all items on the derived hierarchy were maintained. This means that the positions of items that were learned during training sessions retained their ordinal positions during testing sessions, but were paired with novel items from different hierarchies (e.g., $A_{H1} > B_{H3} > C_{H5} > D_{H2} > E_{H3}, \ldots, A_{H5} > B_{H4} > C_{H2} > D_{H3} > E_{H1}$). The authors found evidence of greater than chance accuracy on these novel pairings during test as well as an SDE and argued that for transitive inference to emerge in derived hierarchies is only possible if paired with an additional representation of absolute position.

Secondly, Ciranka et al. (2022) focused on the behavioral modeling of emergent transitive inference during relational learning. They showed that inference performance is worse for inference trials with higher compared to lower values for joint rank (e.g., accuracy(B?D) « accuracy(C?E)). The authors argue that this observed reduction in the ability to differentiate between the more dominant items (e.g. A?B, pair with low joint rank value) compared to the less dominant items (e.g. E?F, pair with high joint rank value) could be caused by compressed representations of magnitude that can emerge from an asymmetric learning policy. In other words, if participants consistently update their belief only about the winner (or loser) during premise pair learning the model predicts diminishing rank-based discrimination as a function of joint rank value.

In three studies, we aimed to investigate the interaction between encoding strength and time/sleep-dependent consolidation effects on relational learning in humans using a transitive inference task. We sought to replicate and extend the findings of Ellenbogen et al. (2007), who showed that time and more specifically, sleep enhances the consolidation of transitive inference performance in healthy young adults by additionally introducing a lower learning criterion (>66%) to the one they used (>75%). Specifically, we wanted to explore whether sleep-dependent consolidation (Sleep vs Wake) would interact with encoding strength when modelling inference performance. Further, we explored how these factors impacted representational geometry of the mental schema as defined by symbolic distance and the joint rank effect. By shedding light on the encoding and temporal dynamics of relational memory consolidation in humans, these studies have implications for understanding the cognitive and neural mechanisms underlying this process. Additionally we provide a quantitative summary of the literature of time and sleep-dependent consolidation and transitive inference.

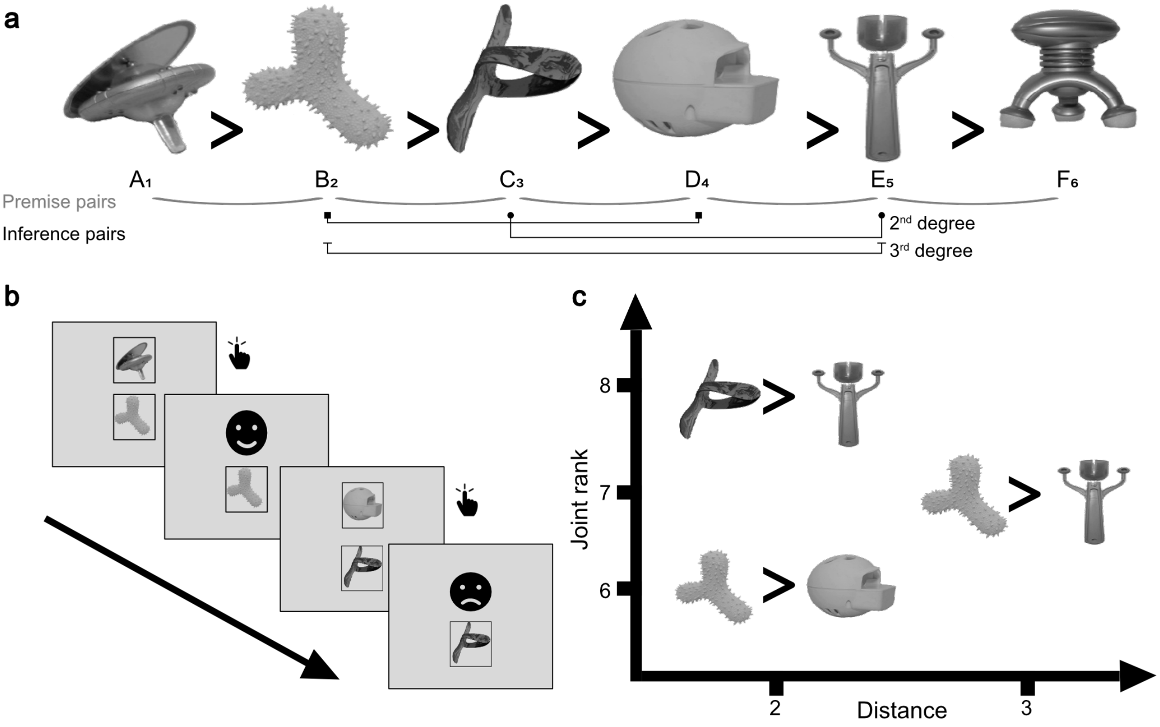


Figure 1: Experimental design a) Example hierarchy and hidden relational rank-order structure. Participants were presented with randomly generated hierarchies from a stimulus set that either involved faces, scenes or objects. Adjacent premise pairs (e.g. A?B) were used during training, non-adjacent inference pairs were used during a delayed test to assess relational learning. b) Example training trials. Participants are asked on each trial to select the item hiding a smiley face and were given feedback after each selection. c) Stimulus pairs can be represented along two orthogonal feature dimensions: symbolic distance (the difference in rank), which is represented on the x-axis, and joint rank (the sum of the ranks), which is represented on the y-axis.

3.3 Experiment 1: Materials and methods

3.3.1 Participants

24 adults (age= 22 ±3.72) with no self-reported history of neurological, psychiatric, sleep or motor disorders participated in the experiment. All participants provided written informed consent and were reimbursed for their time. The experiment was approved by the School of Psychology Ethics Committee at Cardiff University. All participants agreed to abstain from caffeine and alcohol during the study and for 24 hours before it.

3.3.2 Experimental protocol

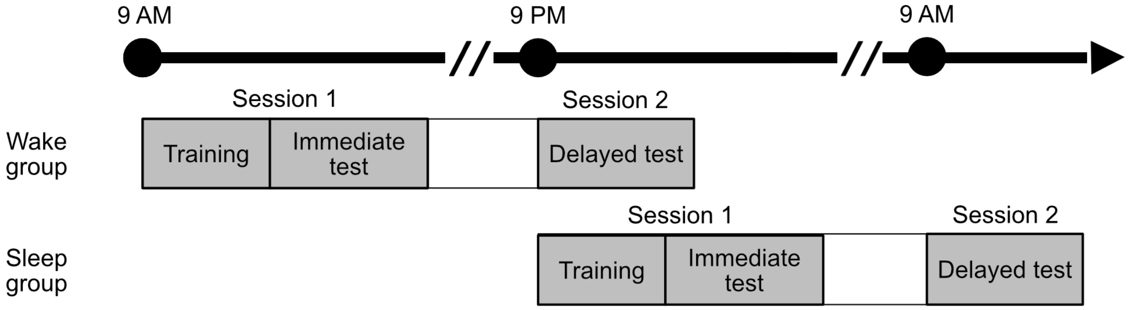


Figure 2: Experimental procedure. Experiment 1. Participants were randomly assigned to either the Wake group or the Sleep group. The Wake group started the experiment at 9 AM (±30min) and the Sleep group started the experiment at 9 PM (±30min). Both groups followed the same procedure. In Session 1 participants learned three separate hierarchies to criterion (premise pairs only). Immediately after training, participants were tested on all three hierarchies (premise pairs only). In Session 2, twelve hours later, participants were tested again on all three hierarchies, as well as on the novel inference pairs.

The data presented as Experiment 1 below were collected by a different experimenter (Lorena Santamaria). Here we are re-analyzing the results from their experiment.

Participants were randomly assigned to one of two groups: Wake or Sleep. Two participants were discarded for not being able to reach criteria (66% accuracy in two consecutive blocks on middle pairs). Both groups participated in two sessions: an initial training with an immediate test session and a delayed test session separated by 12 hours. Wake group participants arrived at the lab at 9 am (±30 min) for the first session and came back at 9 pm (±30 min) for the second one, carrying on with their normal daily routines. Those in the Sleep group arrived at the lab at 9 pm (±30 min) and came back the next morning at 9 am (±30 min).

Premise pairs training

Training involved the presentation and learning of the five-item pairs of each of the stimulus categories in a 6-items hierarchy, hereafter referred to as “premise pairs”. A hierarchy can be schematically represented with letters A>B>C>D>E>F where “>” describes the relationship “choose over” (e.g., “A>B” denotes ‘choose “A” over “B” ’). The order within the hierarchy was randomly selected for each participant at the start of the training phase. Premise pairs were presented on the screen, such that one image was located at the top of the screen and the other one on the bottom. On each trial, after the participant saw one of the five premise pairs (either A?B, B?C, C?D, D?E or E?F), were required to identify the correct item through a process of trial and error. However, with repeated exposure and feedback, participants were able to learn the correct item and make accurate selections. If the participant selected the correct item of the pair, the chosen item was replaced by a smiling-face stimulus on the left side of the screen, and the other item will be presented on the right side. When participants selected the wrong member of the pair, the chosen item was substituted by an angry-face stimulus also on the left side of the screen and the other item on the right. Finally, a purple circle in the middle of the screen was presented to indicate the end of each trial.

Items were organized into blocks, each containing 10 trials of each stimulus category (a total of 30 trials per block). Therefore, each block presented each of the five items of each hierarchy twice, counterbalancing the position from top to bottom or vice versa (e.g. A?B and B?A, where A is the correct selection in both trials). Additionally, all the premise pairs within each hierarchy were presented in a pseudo-random order to minimize the chance of revealing the latent hierarchy (e.g. A?B will never be followed by B?C). At the end of each block, the mean accuracy for that block was shown on the screen to keep participants engaged with the task. Additionally, the order of the stimulus categories was counterbalanced across participants. All subjects underwent a minimum of three blocks of training. After the third block, performance was automatically scored for each stimulus category. If the performance on the "middle premise pairs" (B?C, C?D, D?E) for two of the last three blocks was over 66% for any given hierarchy, the participant stopped receiving feedback for that particular hierarchy. Premise pairs were still displayed for such hierarchies to avoid different number of presentations across hierarchies. When the criteria was reached, or a maximum of 10 blocks, for all three stimulus categories participants, the program automatically stopped. Participants were given a 5-minute break before advancing to the next phase.

Immediate and Delayed test

During the Immediate test similar block protocol was employed as in the Premise pair training with the exception that feedback cues were removed. Participants performed a total of four blocks and in between blocks a series of two easy arithmetic problems had to be solved as a distractor task, to clear the participant's short-term memory (Hecker, Klauer, & Aßfalg, 2019). Following a delay of 12 hours, participants returned to the lab for the Delayed test phase. This phase involved three novel inference pairs (B?D,B?E and C?E) and an anchor pair (A?F), in addition to the five premise pairs. Participants were instructed that they may see novel combinations and, if that happened, to make their best guess on that trial. At the end of each trial, participants were asked how sure they were of their answer on a scale ranging from -2 (guessing) to +2 (completely sure). Similar to the immediate test, participants performed four blocks with two arithmetic problems between each block. After completing this phase participants had to fill out a questionnaire to probe their awareness of the existence of a latent hierarchy underlying the items in each stimulus category.

Apparatus and stimuli

In Experiment 1 a computerized memory task was presented in a quiet room using PsychToolbox (Kleiner, Brainard, & Pelli, 2007). The visual stimuli consisted of three sets of images (hereafter referred to as Stimulus categories), including female faces, unusual objects (Horst & Hout, 2016), and landscapes, each comprising six items. These items were selected randomly from a set of 12 images for each category. All the items were presented in a greyscale and matched for luminescence. Each item was distinguishable from the items within and between stimulus categories. The order of the category presentation was counterbalanced across participants and the order of the stimuli within each relational hierarchy was completely randomized for each subject at the start of the learning phase.

3.3.3 Behavioural data analysis

To examine the relative impact of Encoding strength, Distance and Joint rank on delayed inferential performance, separate hierarchical multiple regression analyses were conducted for Distance and Joint rank. We employed specific transformations on the variables used. First, we mean centered the Distance and Joint rank variables to enhance interpretability and reduce multicollinearity among predictor variables. Additionally, we applied a transformation to the Encoding strength variable by multiplying the original values, which initially ranged between 0 and 1, by a factor of ten, enabling us to assess the impact of Encoding strength in terms of a rate per units of 10%.

To first identify the contributions of Encoding strength, we entered our experiment-specific condition factor into the model in Step 1 (Group), Encoding strength in Step 2, and the interaction of these measures at Step 3 as predictors of memory recall for inferential pairs. Next, we extended the interaction model by adding in the Distance factor as Step 4, an interaction between Distance and the respective condition factor as Step 5 and finally a model that included a full factorial combination between condition, Encoding strength and Distance as Step 6. We also repeated this for Joint rank replacing Distance. An alpha value of p<.05 was employed for all analyses.

We used R (version 4.2.2) and the R-packages *lme4*,*afex* and *emmeans* for all our statistical analyses (Bates, Machler, Bolker, & Walker, 2015; Lenth, 2023; Singmann, Bolker, Westfall, Aust, & Ben-Shachar, 2023) and *sjPlot* for generating the regression tables (Lüdecke, 2023).

3.4 Experiment 1: Results

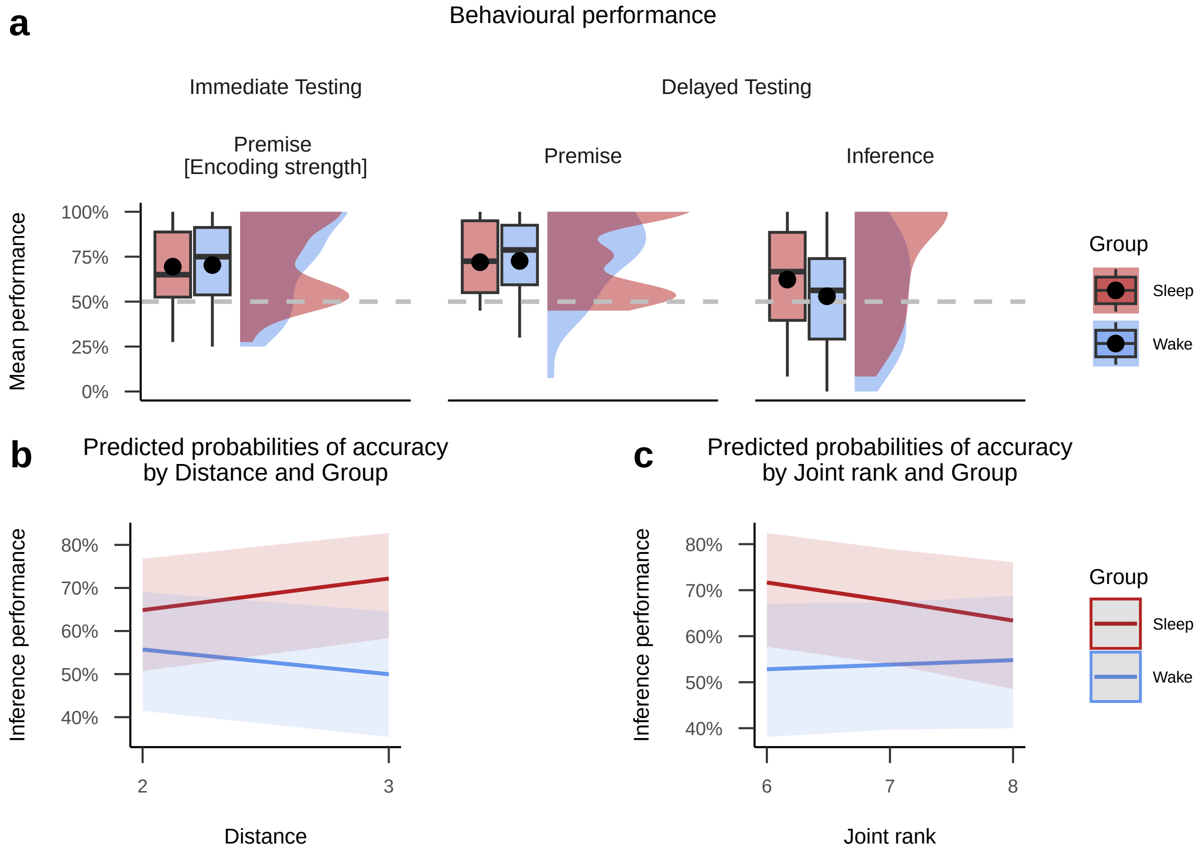


Figure 3: Experiment 1: Behavioral performance and factors predicting inference accuracy at delayed test. a) Raincloud plot with mean and median performance across Wake and Sleep conditions. Dashed line represents chance performance at 50% and dot represents mean values b) Predicted probabilities of accuracy by distance and group, with distance levels of 2 and 3. (c) Predicted probabilities of accuracy by joint rank and group, with joint rank levels ranging from 6 to 8. In panels b) and c), shadowed areas represent 95% confidence intervals

Our primary interest was the inference pairs, which were tested only at session 2 (post-retention interval). One-sample t-tests showed that inference performance was significantly higher than chance for the Sleep group ($M = 0.62$, 95% CI $[0.54, \infty]$, $t(35) = 2.53$, $p = .008$), but not for the Wake group ($M = 0.53$, 95% CI $[0.45, \infty]$,

$t(35) = 0.62$, $p = .271$). For overview of performance see 3. For a table of descriptive table of pair level performance, see Table 4.

To test for baseline differences in premise pair memory which might have confounded the above result, we performed a mixed three-way ANOVA on the mean premise pair accuracy measure with between-subject factor Group and within-subject factor Session and Stimulus category. This revealed no effect for either of the factors (smallest p=0.088) or their interactions (smallest p=0.733, for detailed results, see Table S1).

3.4.1 Encoding strength

To assess the effects of our factors of interest (Encoding strength and Group) on inference performance at test, we next conducted a hierarchical regression with a series of nested mixed effects logistic models. We started with the Group variable, which captured whether the participant was in the Sleep or Wake condition, however, this did not outperform the intercept-only null model ($\chi^2(1) = 2147.49$, p = 0.18). We then added the Encoding strength variable, which measured the mean accuracy of premise pair recall at Immediate testing. This significantly improved the model's ability to predict accuracy compared to the null model, $\chi^2(2) = 2141.41$, p = 0.019. Next, we added an interaction term between Group and Encoding strength, which again significantly improved fit, $\chi^2(1) = 2126.26$, p < 0.001. We took this model as our baseline model for all further analysis. In this baseline model, Group was a significant predictor of the outcome variable ($\hat{\beta} = 1.61$, 95% CI $[0.25, 2.96]$, $z = 2.32$, $p = .020$), as was the interaction between Group and Encoding strength ($\hat{\beta} = -0.31$, 95% CI $[-0.47, -0.15]$, $z = -3.89$, $p < .001$, see Table 1. This indicates that the effect of Encoding strength on accuracy differs significantly between the Wake and Sleep groups. To probe the interaction, simple effects coefficients were computed at 60% ($\mathrm{OR}_{\mathrm{Sleep/Wake}} = 1.30$, SE = 1.52, $p = 0.53$) and 80% ($\mathrm{OR}_{\mathrm{Sleep/Wake}} = 2.42$, SE = 1.52, $p = 0.03$) values of Encoding strength, suggesting increased sleep-dependent benefit with stronger encoding.

Table 1: Experiment 1: LMM Group/Encoding strength. Results of the mixed-effects logistic regression model examining the effects of Group (Wake/Sleep) and Encoding strength on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	1.05	0.38 – 2.91	0.932
Group [Wake]	4.98	1.28 – 19.30	0.020
Encoding strength	1.10	0.97 – 1.25	0.121
Group [Wake] X Encoding strength	0.73	0.63 – 0.86	<0.001
Random Effects			
σ2	3.29		
τ00participant	0.93		
ICC	0.22		
N participant	24		
Observations	1728		
Marginal R2 / Conditional R2	0.050 / 0.258		

Note: ^{a}CI = Confidence Interval. All p-values are two-tailed.

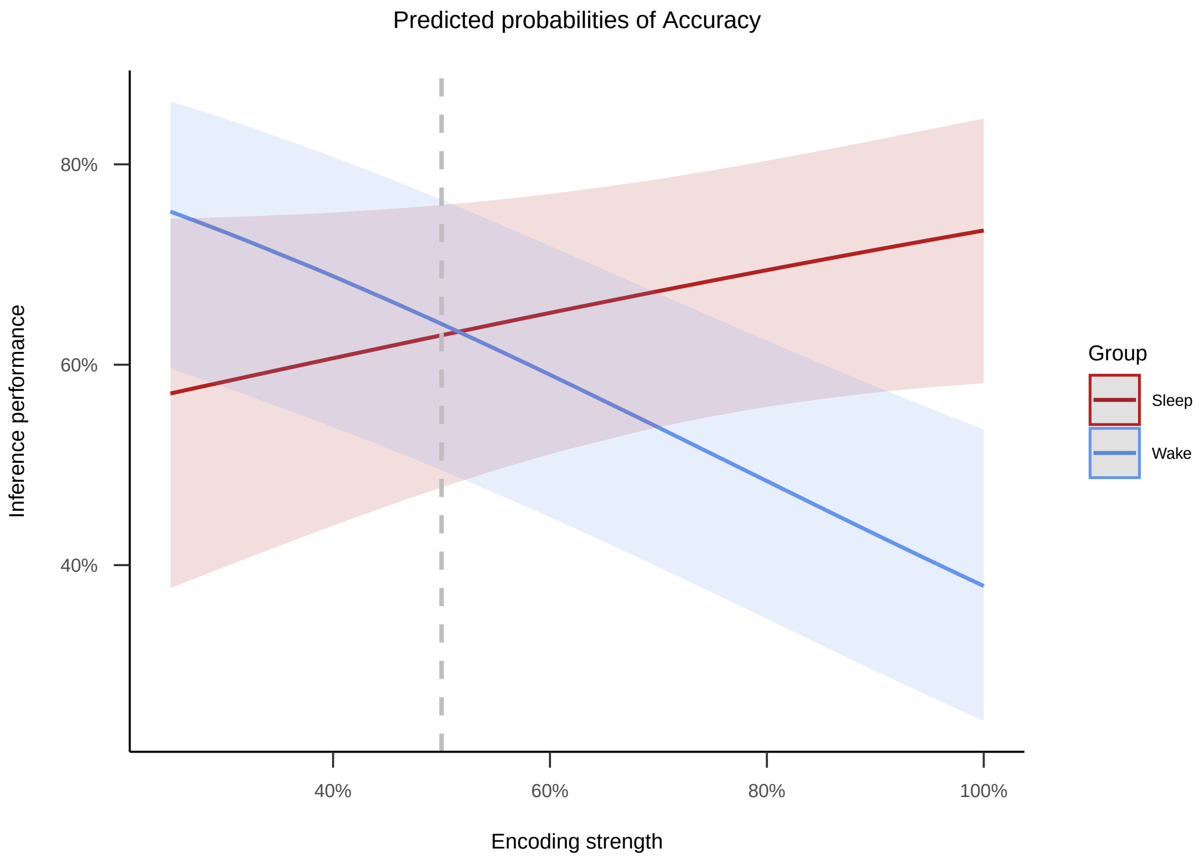


Figure 4: Experiment 1: Group X Encoding strength. Plotting model predicted interaction between Group and Encoding strength. Shadowed areas represent 95% confidence intervals.

3.4.2 Distance

To evaluate the influence of Distance on inference performance at test, we continued the hierarchical regression analysis, extending previous findings of the "Group and Encoding strength interaction" model described above (Table 1, baseline model). We first added the Distance variable, which captured whether a given trial was a "distant" (B?E; degree of separation: 3) or "close" inference trial (B?D or C?E; degree of separation: 2). This did not significantly improve fit model, $\chi^2(1) = 2126.14$, p = 0.74. Next, we added an interaction between Group and Distance, which significantly improved fit, $\chi^2(1) = 2119.73$, p = 0.011. Finally, we included a full factorial combination of the predictor variables Group, Encoding strength, and Distance (comprising of all two-way interactions and a three-way interaction, see Methods). The additional terms did not significantly improve the model,$\chi^2(2) = 2115.80$, p = 0.14. The best-fitting model showed the same effects of Group and Encoding strength as reported by the baseline model. Additionally there was a main effect of

Distance ($\hat{\beta} = 0.34$, 95% CI $[0.02, 0.66]$, $z = 2.07$, $p = .039$) and a Distance-Group interaction (, Table 2). To investigate the interaction, we calculated the coefficients of simple effects at close ($\mathrm{OR}_{\mathrm{Sleep/Wake}} = 1.47$, SE $= 1.57$, $p = 0.39$) and distant ($\mathrm{OR}_{\mathrm{Sleep/Wake}} = 2.58$, SE $= 1.60$, $p = 0.04$) pair values of Distance, suggesting increased sleep-dependent benefit for accuracy on inference pairs of greater symbolic distance. In summary, this analysis showed that the symbolic distance over which participants had to make an inference was a significant predictor of inference performance, and also that the impact of this differed between Sleep and Wake groups.

Table 2: Experiment 1: LMM Group/Encoding strength/Distance. Results of the mixed-effects logistic regression model examining the effects of Group (Wake/Sleep), Encoding strength and Symbolic distance on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	4.02	1.76 – 9.16	0.001
Encoding strength	0.91	0.84 – 0.98	0.014
Distance	1.41	1.02 – 1.94	0.039
Group [Wake]	0.56	0.24 – 1.35	0.199
Distance X Group [Wake]	0.57	0.36 – 0.88	0.012
Random Effects			
σ2	3.29		
τ00participant	1.11		
ICC	0.25		
N participant	24		
Observations	1728		
Marginal R2 / Conditional R2	0.032 / 0.276		

Note: ^{a}CI = Confidence Interval. All p-values are two-tailed.

3.4.3 Joint rank

In parallel to the above, we also assessed the effects of Joint rank on inference performance at test. We again started with our baseline model, first adding the Joint rank variable, which encoded the summed rank of the inference pairs (B?D: 6; B?E: 7 and C?E: 8). This did not significantly improve the model,$\chi^2(1) = 2125.92$, p = 0.56. We next added the interaction between Group and Joint rank, which again did not significantly improve the model, $\chi^2(2) = 2124.32$, p = 0.38. Finally, we included a full factorial combination of the predictor variables Group, Encoding strength, and Joint rank. The additional terms significantly improve the fit, $\chi^2(4) = 2114.12$, p = 0.016. This model showed the same effects of Group and Group x Encoding strength as the baseline model. In addition, it showed an effect of Joint rank ($\hat{\beta} = 0.85$, 95% CI $[0.21, 1.50]$, $z = 2.59$, $p = .010$), Encoding strength x Joint rank interaction ($\hat{\beta} = -0.15$, 95% CI $[-0.24, -0.05]$, $z = -3.09$, $p = .002$) and a three-way interaction between them ($\hat{\beta} = 0.12$, 95% CI $[0.00, 0.24]$, $z = 1.99$, $p = .047$, Table

3). In order to examine the interaction, we computed the coefficients of simple effects for BD (JR:6, $OR_{Sleep/Wake} = 2.26$, SE = 1.55, $p = 0.06$) and CE trials (JR:8, $OR_{Sleep/Wake} = 1.43$, SE = 1.55, $p = 0.41$). From this post-hoc analysis, we can conclude that for BD trials, the odds of success in the Sleep group compared to the Wake group were estimated to be 2.26 times higher, although this difference was not statistically significant (p = 0.06). In summary, this analysis showed that Joint rank was a significant predictor and that it interacted with both baseline encoding strength and whether or not participants consolidated across sleep.

Table 3: Experiment 1: LMM Group/Encoding strength/Joint rank. Results of the mixed-effects logistic regression model examining the effects of Group (Wake/Sleep), Encoding strength and Joint rank on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	1.04	0.37 – 2.93	0.944
Group [Wake]	5.05	1.28 – 19.84	0.021
Encoding strength	1.11	0.98 – 1.25	0.113
Joint rank	2.35	1.23 – 4.48	0.010
Group [Wake] X Encoding strength	0.73	0.62 – 0.85	<0.001
Group [Wake] X Joint rank	0.53	0.23 – 1.25	0.148
Encoding strength X Jointrank	0.86	0.78 – 0.95	0.002
(Group [Wake] X Encoding strength) X Joint rank	1.13	1.00 – 1.28	0.047
Random Effects			
σ2	3.29		
τ00participant	0.96		
ICC	0.23		
N participant	24		
Observations	1728		
Marginal R2 / Conditional R2	0.060 / 0.272		

Note: ^{a}CI = Confidence Interval. All p-values are two-tailed.

3.5 Experiment 1: Discussion

The transitive inference (TI) paradigm evaluates how well individuals can grasp the relational arrangement of a group of stimuli without relying on any overt hierarchical cue while learning about them. To achieve this, the training involves displaying pairs of images selected from a pre-determined list of neighboring items in order of rank and providing incentives for participants to correctly identify the dominant item. Work in both humans and other animals has supported the idea that the ability to assess relational dominance for non-neighboring, inference items rely on some form of gist abstraction and that this is facilitated by time and sleep-dependent memory consolidation in humans (Inostroza & Born, 2013; Lerner & Gluck, 2019a;

Lewis & Durrant, 2011; Lewis et al., 2018b). While initial findings (Ellenbogen et al., 2007; Werchan & Gómez, 2013a, 2016) showed strong time and sleep-dependent benefits in inference performance and sleep-dependent SDE, subsequent studies that implemented modified parameters failed to replicate these findings, suggesting that the time/sleep-inference relationship may be more complex and dependent on specific experimental conditions (Cordi & Rasch, 2021). Here, we evaluated whether we could replicate the original sleep-dependent TI findings by Ellenbogen et al. (2007) in two experiments that involve learning multiple hierarchies, with lower learning criterion and using a broader range of stimuli compared to the original study.

Crucially this data support earlier findings (Ellenbogen et al., 2007; Werchan & Gómez, 2013a, 2016) by showing increased inference performance at delayed test in the experimental Sleep group, Experiment 1 (see Table 1). Additionally, the sleep-benefit effect increases as a function of encoding strength. Concerning the sleep-dependent SDE shown in past work (Ellenbogen et al., 2007; Werchan & Gómez, 2013a), we were also able to replicate this in our AM-PM design in Exp 1 (Fig. 3).

We next explored the sleep-dependent joint rank effect (JRE). JRE, whereby participants - when comparing performance for pairs that have equivalent symbolic distance - exhibit lower accuracy the higher the cumulative sum of the item ranks, has been observed in both primates (Ciranka et al., 2022; Munoz et al., 2020a) and humans (Ciranka et al., 2022; Jensen et al., 2017b). JRE has been suggested as a valuable complementary variable to symbolic distance, as it serves as a measure of absolute positional encoding during serial learning and retrieval (Jensen et al. 2017) and is indicative of a cognitive model supporting TI (Behrens et al., 2018). The reduced ability to choose between more dominant items (e.g. B?D, pair with low joint rank value) compared to less dominant items (e.g. C?E, pair with higher joint rank value) observed in wake-only studies could be caused by compressed representations of magnitude that can emerge from an asymmetric learning policy (Ciranka et al., 2022). Interestingly, this study found that this pattern reversed across a retention interval containing sleep, with participants showing a lesser ability to discriminate during low dominance inference pairs (C?E, JR:8), and a greater ability to discriminate between high dominance items (B?D, JR:6) after consolidation (see Fig. 3). This “inverse joint rank effect” could suggest active sleep-dependent consolidation processes, whereby experience is reorganized based on the learned absolute rank-structure prioritizing high dominance items over low dominance ones (Y. Liu et al., 2019; Y. Liu, Mattar, Behrens, Daw, & Dolan, 2021).

One limitation of this study stems from the fact that we did not record sleep. Although Experiment 1 compared consolidation across 12 hours of wake and 12 hours including an overnight sleep, we cannot assume that the benefit observed in the overnight condition relates specifically to sleep rather than to a combination of wake and sleep. A further limitation of our study is the use of a relatively small sample size in our first experiment. Although we were able to obtain significant results, a larger sample size would have increased the generalizability and robustness of our findings.ndefined Additionally, while we used well-established transitive inference tasks to assess memory consolidation, these tasks do not capture all aspects of relational memory, and other cognitive processes may have played a role in our results (e.g., awareness (Libben & Titone, 2008), strategy (S. Moses, Villate, & Ryan, 2006)).

3.6 Experiment 1: Acknowledgement

Thank you to Dr Lorena Santamaria for sharing the data she collected with me for reanalysis.

3.7 Experiment 1: Supplemental materials

Supplemental Table S1: Experiment 1: ANOVA table of baseline differences

Effect	$\hat{\eta}^2_G$	90% CI	F	df^{GG}	$df^{\mathrm{GG}}_{\mathrm{res}}$	p
Group	.000	[.000, .030]	0.01	1	22	.913
StimCategory	.020	[.000, .043]	1.53	1.95	42.82	.229
Session	.003	[.000, .118]	3.19	1	22	.088
Group × StimCategory	.002	[.000, .000]	0.16	1.95	42.82	.844
Group × Session	.000	[.000, .000]	0.01	1	22	.939
StimCategory × Session	.000	[.000, .000]	0.25	1.60	35.30	.733
Group × StimCategory × Session	.000	[.000, .000]	0.07	1.60	35.30	.902

Table 4: Descriptive Statistics for Delayed testing performance by Sleep and Wake Groups. The table presents the median, mean, and standard deviation (SD) for performance during delayed testing, grouped by Sleep and Wake conditions. Data is further broken down by Pair values.

		Group					
		Sleep			**Wake**		
Session	Pair	Median	Mean	SD	Median	Mean	SD
Delayed testing	AB	0.94	0.67	0.40	1.00	0.73	0.39
Delayed testing	BC	0.88	0.75	0.34	0.88	0.67	0.41
Delayed testing	CD	0.88	0.73	0.30	0.81	0.71	0.34
Delayed testing	DE	0.88	0.71	0.36	0.88	0.72	0.34
Delayed testing	EF	0.88	0.74	0.34	1.00	0.81	0.34
Delayed testing	BD	0.62	0.62	0.35	0.62	0.54	0.39
Delayed testing	CE	0.62	0.58	0.38	0.62	0.56	0.42
Delayed testing	BE	0.75	0.67	0.35	0.50	0.50	0.37
Delayed testing	AF	0.88	0.69	0.39	1.00	0.70	0.40

3.8 Experiment 2: Materials and methods

3.8.1 Purpose

Advantages of conducting cognitive psychology research online for memory research and time-dependent consolidation include increased accessibility and convenience for participants, the ability to collect large amounts of data quickly, and the potential for conducting longitudinal studies. Online research allows researchers to reach a larger and more diverse pool of participants, which can enhance the generalizability of findings (Clifford & Jerit, 2014). Many studies have found similar effects between online and laboratory studies in diverse domains of cognitive psychology such as attention (Crump, McDonnell, & Gureckis, 2013; Germine et al., 2012; Haas et al., 2021). Additionally, conducting research online eliminates the need for participants to travel to a physical location, making it more convenient for them to participate. This convenience factor may also increase participant retention rates and reduce attrition in longitudinal studies (Ashton et al., 2018). Longitudinal studies can provide valuable insights into the long-term effects of memory consolidation and the role of time and sleep in this process. Longitudinal studies involve collecting data from the same participants over an extended period of time, which can help researchers understand how memory consolidation changes over time and how it is influenced by various factors, including sleep (Gómez & Edgin, 2015). Several studies argue that online behavioral studies offer greater ecological validity compared to laboratory studies. For example Enkavi et al. (2019) conducted a large-scale analysis of test-retest reliabilities of self-regulation measures and found that data acquired online exhibited high reliability, suggesting that online studies can capture real-world behaviors more accurately. Online research is particularly beneficial when combined with wearable devices that monitor biophysical parameters (Fazio, Mattei, Al-Naami, Vittorio, and Visconti, 2022).

However, there are also some disadvantages to conducting cognitive psychology research online for memory research. One potential limitation is the lack of control over the research environment. In online studies, participants complete tasks and questionnaires in their own homes or other non-laboratory settings, which may introduce confounding variables that can affect the validity and reliability of the results (Ashton et al., 2018). For example, participants may be more easily distracted or may not adhere to the instructions as closely as they would in a controlled laboratory setting. Another limitation is the potential for selection bias in online research. Online studies often rely on self-selected samples, which may not be representative of the general population. This can limit the generalizability of the findings and may introduce biases that affect the validity of the results (Ashton et al., 2018). Additionally, online research may face challenges in ensuring data quality and participant engagement. Without direct supervision, participants may not fully engage with the tasks or may provide inaccurate or incomplete responses. Researchers must implement strategies to mitigate these issues, such as including attention checks and monitoring participant compliance (Ashton et al., 2018).

One criticism of the sleep and memory literature, as with other areas of research, is that many studies lack statistical power and generalizability, leading to mixed and confusing evidence (Mantua, 2018; Nemeth, Gerbier, & Janacsek, 2019). In recent years however, researchers investigating the impact of sleep on memory have started using web-based tools for conducting online sleep experiments (Ashton & Cairney, 2021, Kroneisen and Kuepper-Tetzel, 2021, Denis et al, 2022). It can be argued that generally such experiments do not limit the capacity to detect the impact of sleep on memory (Mak, Curtis, Rodd, & Gaskell, 2022, 2023; Mak, O'Hagan, Horner, & Gaskell, 2023b). A recent positive example where a well-powered study successfully replicated previous laboratory findings was done by Denis, Sanders, Kensinger, & Payne (2022). The researchers conducted two large-scale replication experiments with over 250 participants from diverse backgrounds and age ranges to confirm the preferential consolidation of emotional memory during sleep using a well-established task. Their findings confirmed that sleep enhances memory for negative emotional objects while impairing memory for their paired neutral backgrounds, supporting their previous work and providing clear evidence for the role of sleep in emotional memory formation. In a second experiment, they examined whether this effect extends to positive emotional memory, but found that sleep did not modulate the memory enhancement for positive objects compared to neutral backgrounds. Unsuccessful online replication studies, include a study by (J. Ashton & Cairney, 2021)Ashton & Cairney (2021) who conducted an online replication study to test the selective strengthening of future-relevant memories during sleep. They found that the memory benefits of overnight consolidation were observed in both the sleep and wake groups, regardless of whether participants expected a future test or not. These findings suggest that the selective sleep-memory effects observed in prior laboratory-based studies did not emerge in an online environment.

3.8.2 Participants

The present study involved 137 participants completing Session 1 (Mean age=20.44± 2.15, female=104). However, due to various reasons, only 116 participants completed Session 2 (Mean age=20.51± 2.18, female=99). The reasons for attrition included participant withdrawal and technical issues. We further excluded participants who didn't reach a minimum of 60% mean premise pair performance at Session 1, resulting in a final sample of 73 (Mean age=20.37± 1.89, female=56).

All participants reported no history of neurological, psychiatric, sleep, or motor disorders based on self-report.

Participants were primarily recruited from two universities, namely Cardiff University, UK and Eötvös Loránd University, Hungary. The recruitment process was facilitated through the respective Experimental Management Systems (EMS) of each university. The EMS platforms served as a centralized hub for managing participant recruitment and scheduling.

To recruit participants, study advertisements were posted on the EMS platforms, where eligible individuals from the university communities were able to view and express their interest in participating. Interested participants were directed to an online screening form to assess their eligibility based on the predetermined inclusion criteria.

All participants provided written informed consent and were awarded course credit for their time. The experiment was approved by the School of Psychology Ethics Committee at Cardiff University. All participants agreed to abstain from caffeine and alcohol during the study and for 24 hours before it.

3.8.3 Experimental protocol

Upon obtaining informed consent, participants were randomly assigned to conditions in a two-dimensional design involving learning rate and wake/sleep conditions. The learning rate condition consisted of two levels: 66% and 75%. The wake/sleep condition also consisted of two levels: wake and sleep. This resulted in a total of four experimental conditions: 66% learning rate/wake, 66% learning rate/sleep, 75% learning rate/wake, and 75% learning rate/sleep. In all other respects it followed the same design as Experiment 1.

Random assignment to the conditions was conducted using a computerized randomization procedure (*Randomizer*, 2022). Each participant was assigned to one of the four conditions with an equal probability of assignment to each condition. The randomization was implemented to ensure that the assignment of participants to the different conditions was unbiased and independent of any potential participant characteristics.

Participants in the Wake group underwent Session 1starting between 9:00 AM and 11:00 AM, followed by Session 2 on the same day between 21:00 PM and 23:00 PM. Participants in the Sleep group, on the other hand, began with Session 1 at approximately starting between 21:00 PM and 23:00 PM, and proceeded to Session 2 on the subsequent day between 9:00 AM and 11:00 AM.

All experimental sessions were conducted remotely using the Qualtrics online platform (Qualtrics, 2020). Participants were provided with comprehensive instructions on how to complete the computerized task, which involved responding to neutral images presented on the computer screen according to predefined rules (Experiment 1: Experimental protocol).

Session 1, lasting approximately 15 minutes, aimed to establish baseline memory performance. During this session, participants completed the computerized task under standard wakefulness conditions. Following a 12 hour predetermined time interval, Session 2 commenced, lasting approximately 10 minutes.

To accommodate participants' preferences and availability, they were given the opportunity to select their desired starting day for the study. The scheduling of the sessions was adjusted according to the chosen starting day and participants' local time zones to minimize any potential confounding factors related to circadian rhythm variations.

Apparatus and stimuli

Participants completed the experiment on their own computer or tablet device using a web browser (e.g., Chrome). Experiment 2 was programmed using PsychJS (Peirce, 2007) and run through the Pavlovia platform (`https://pavlovia.org/`). In the first experiment, the visual stimuli consisted of three sets of images (hereafter referred to as Stimulus categories), including female faces, unusual objects (Horst & Hout, 2016), and landscapes, each comprising six items. These items were selected randomly from a set of 12 images for each category. Participants were randomly assigned to one of the stimulus categories, with both the Group and Wake hierarchy consisting of a six-item subset. All the items were presented in greyscale and matched for luminescence. Each item was distinguishable from the items within and between stimulus categories. The order of the category presentation was counterbalanced across participants and the order of the stimuli within each relational hierarchy was completely randomized for each subject at the start of the learning phase.

3.8.4 Behavioural data analysis

We employed the same analysis as outlined in Experiment 1.

3.9 Experiment 2: Results

Due to a large number of participants performing at or below chance at immediate test we designated a 60% Encoding strength cut-off (see Fig 5).

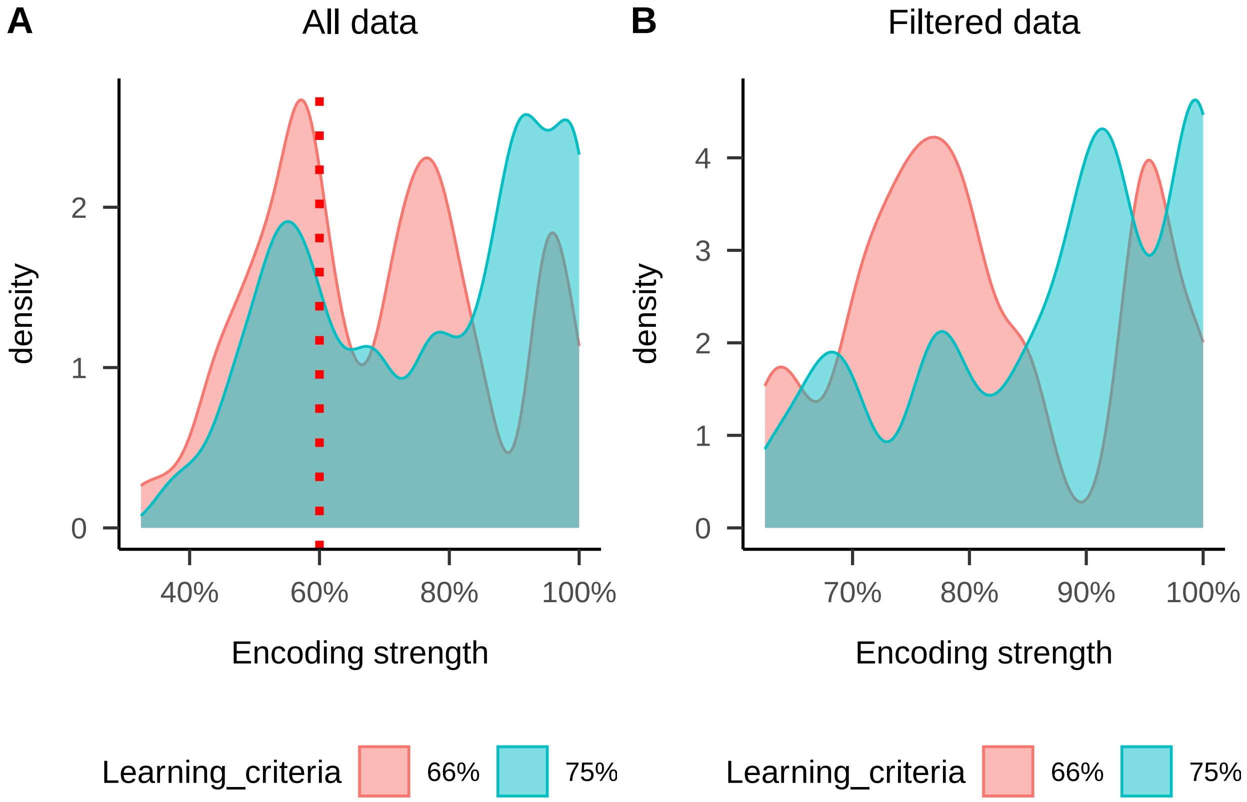


Figure 5: Experiment 2: Density Plot of Encoding Strength. Figure displays the density plot of encoding strength for all data (A) and filtered data (B) using a threshold of 60%. The red dotted line represents the encoding strength threshold.

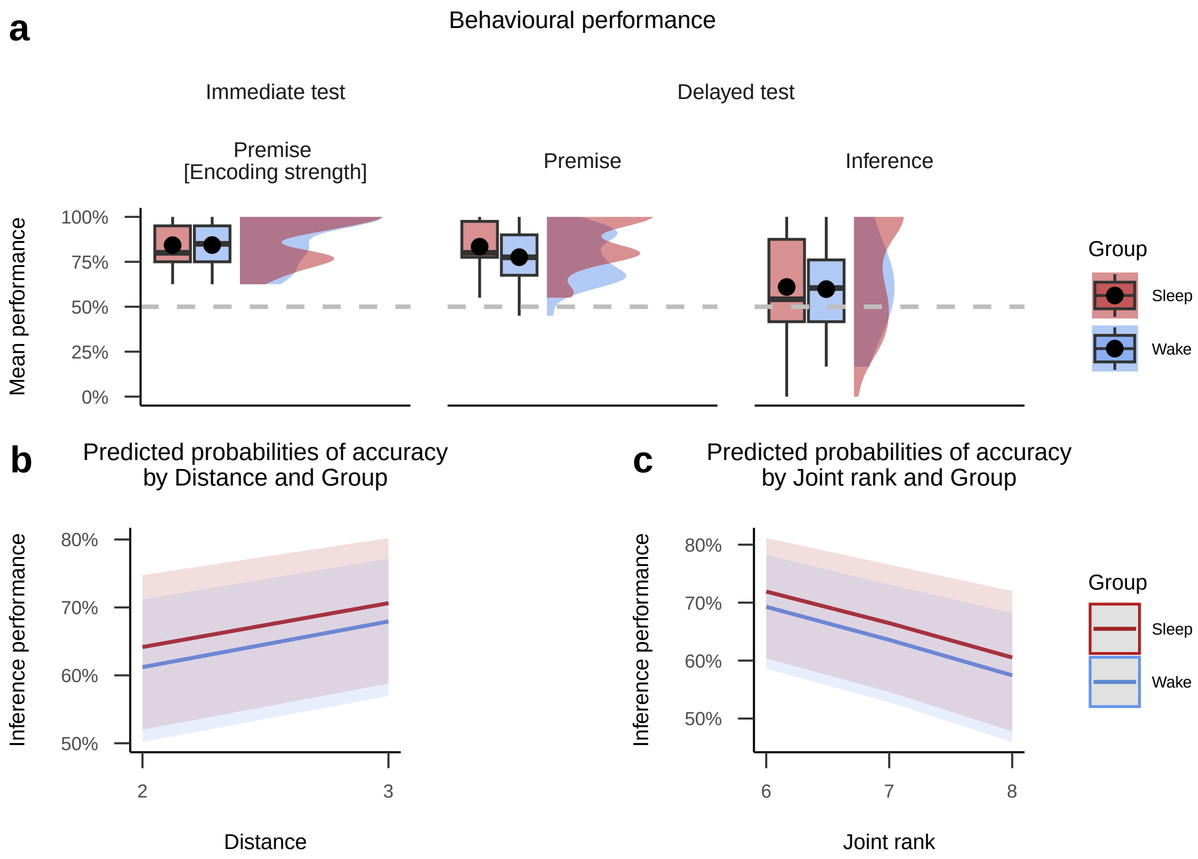


Figure 6: Experiment 2: Behavioral performance and factors predicting inference accuracy at delayed test. a) Raincloud plot with mean and median performance across Wake and Sleep conditions. Dashed line represents chance performance at 50% and dot represents mean values b) Predicted probabilities of accuracy by distance and group, with distance levels of 2 and 3. (c) Predicted probabilities of accuracy by joint rank and group, with joint rank levels ranging from 6 to 8. In panels b) and c), shadowed areas represent 95% confidence intervals

Our primary interest was the inference pairs, which were tested only at session 2 (post-retention interval). One-sample t-tests showed that inference performance was significantly higher than chance for both the Sleep group ($M = 0.61$, 95% CI $[0.53, \infty]$, $t(32) = 2.25$, $p = .016$) and Wake group ($M = 0.60$, 95% CI $[0.54, \infty]$, $t(39) = 2.64$, $p = .006$). For overview of performance see Fig. 6. For a table of descriptive table of pair level performance, see Table 8.

To test for baseline differences in premise pair memory which might have confounded the above result, we performed a mixed three-way ANOVA on the mean premise pair accuracy measure with between-subject factor Group, Stimulus category and within-subject factor Session. This revealed of effect Session ($F(1, 67) = 13.33$, $p < .001$, $\hat{\eta}^2_G = .024$, 90% CI $[.000, .114]$). No other factor or their interaction was significant (smallest p=0.054, for detailed results, see Table S2).

3.9.1 Encoding strength

To assess the effects of our factors of interest (Encoding strength and Group) on inference performance at test, we next conducted a hierarchical regression with a series of nested mixed effects logistic models. We started with the Group variable, which captured whether the participant was in the Sleep or Wake condition, however, this did not outperform the intercept-only null model ($\chi^2(1) = 2038.00$, p = 0.65). We then added the Encoding strength variable, which measured the mean accuracy of premise pair recall at Immediate testing. This did not significantly improve the model's ability to predict accuracy compared to the null model, $\chi^2(2) = 2032.63$, p = 0.061. Next, we added an interaction term between Group and Encoding strength, which significantly improved fit, $\chi^2(3) = 2028.08$, p = 0.017. We took this model as our baseline model for all further analysis. In this baseline model, Group was a significant predictor of the outcome variable ($\hat{\beta} = 5.18$, 95% CI $[0.32, 10.04]$, $z = 2.09$, $p = .037$), as was the interaction between Group and Encoding strength ($\hat{\beta} = -0.63$, 95% CI $[-1.20, -0.06]$, $z = -2.15$, $p = .031$, see Table 5). This indicates that the effect of Encoding strength on accuracy differs significantly between the Wake and Sleep groups. To probe the interaction, simple effects coefficients were computed at 60% ($\text{OR}_{\text{Sleep/Wake}} = 0.25$, SE = 2.18, $p = 0.07$) and 80% ($\text{OR}_{\text{Sleep/Wake}} = 0.88$, SE = 1.43, $p = 0.71$) and 100% ($\text{OR}_{\text{Sleep/Wake}} = 3.09$, SE = 1.78, $p = 0.05$) values of Encoding strength, suggesting increased sleep-dependent benefit with stronger encoding (although no significant differences were found). For a visualisation of the interaction, see Figure 7.

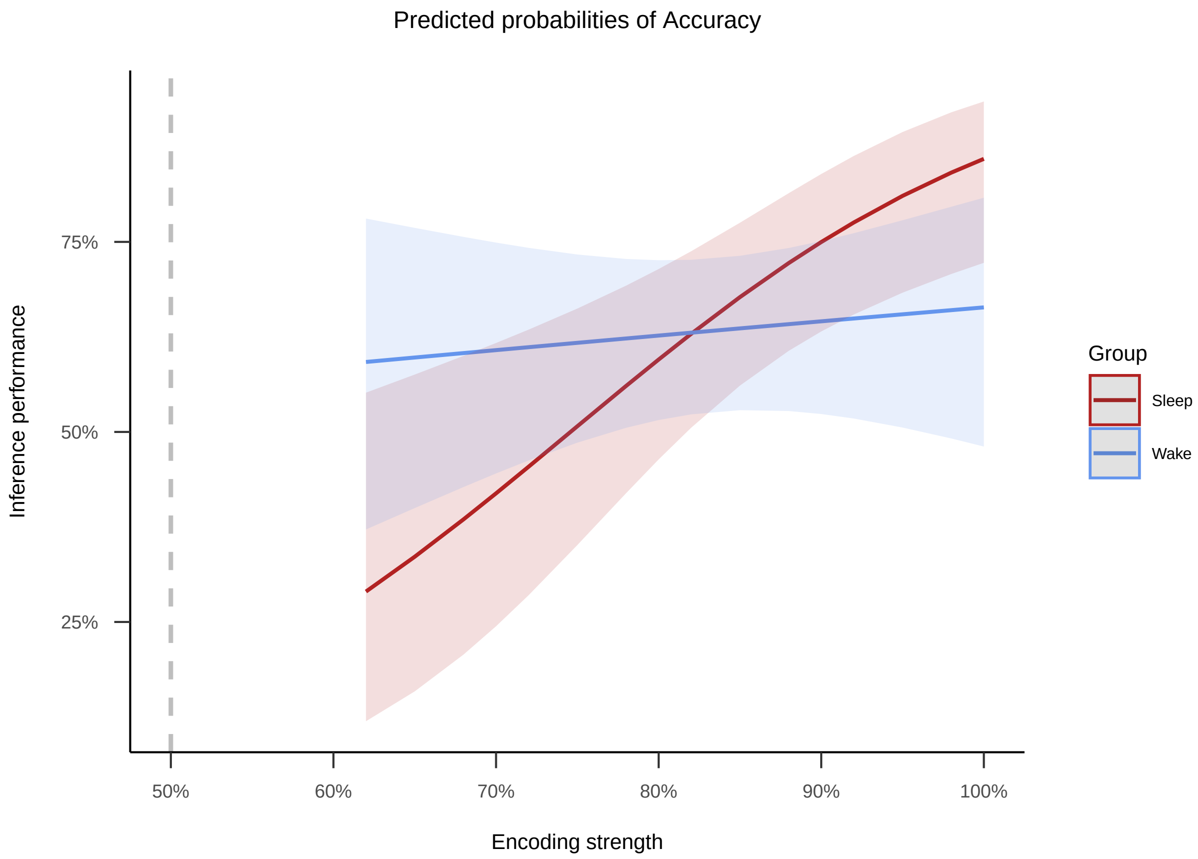


Figure 7: Experiment 2: Group X Encoding strength. Plotting model predicted interaction between Group and Encoding strength. Shadowed areas represent 95% confidence intervals.

Table 5: Experiment 2: LMM Group/Encoding strength. Results of the mixed-effects logistic regression model examining the effects of Group (Wake/Sleep) and Encoding strength on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	0.00	0.00 – 0.21	0.006
Group [Wake]	177.32	1.37 – 22878.49	0.037
Encoding strength	2.04	1.31 – 3.17	0.002
Group [Wake] * Encoding strength	0.53	0.30 – 0.94	0.031
Random Effects			
σ2	3.29		
τ00participant	1.76		
ICC	0.35		
N participant	73		
Observations	1752		
Marginal R2 / Conditional R2	0.058 / 0.386		

Note: ^{a}CI = Confidence Interval. All p-values are two-tailed.

3.9.2 Distance

Table 6: Experiment 2: LMM Group/Encoding strength/Distance. Results of the mixed-effects logistic regression model examining the effects of Group (Wake/Sleep), Encoding strength and Symbolic distance on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	1.98	1.20 – 3.25	0.007
Group [Wake]	0.88	0.45 – 1.71	0.708
Encoding strength	2.30	1.37 – 3.87	0.002
Distance	1.34	1.06 – 1.70	0.013
Group [Wake] X Encoding strength	0.48	0.24 – 0.94	0.032
Random Effects			
$\sigma 2$	3.29		
$\tau 00$participant	1.77		
ICC	0.35		
N participant	73		
Observations	1752		
Marginal R2 / Conditional R2	0.061 / 0.390		

Note: ^{a}CI = Confidence Interval. All p-values are two-tailed.

To evaluate the influence of Distance on inference performance at test, we continued the hierarchical regression analysis, extending previous findings of the “Group and Encoding strength interaction” model described above (Table 5, baseline model). We first added the Distance variable, which captured whether a given trial was a “distant” (B?E; degree of separation: 3) or “close” inference trial (B?D or C?E; degree of separation: 2). This significantly improved fit model, $\chi^2(1) = 2022.01$, p = 0.014. Next, we added an interaction between Group and Distance, which did not significantly improve fit, $\chi^2(1) = 2022.00$, p = 0.95. Finally, we included a full factorial combination of the predictor variables Group, Encoding strength, and Distance (comprising of all two-way interactions and a three-way interaction, see Methods). The additional terms did not significantly improve the model,$\chi^2(3) = 2016.22$, p = 0.12. The best-fitting model showed the same effects of Group and Encoding strength as reported by the baseline model. Additionally there was a main effect of Distance ($\hat{\beta} = 0.29$, 95% CI $[0.06, 0.53]$, $z = 2.47$, $p = .013$), Table 6). In summary, this analysis showed that the symbolic distance over which participants had to make an inference was a significant predictor of inference performance.

3.9.3 Joint rank

In parallel to the above, we also assessed the effects of Joint rank on inference performance at test. We again started with our baseline model, first adding the Joint rank variable, which encoded the summed rank of the inference pairs (B?D: 6; B?E: 7 and C?E: 8). This did significantly improve the model,$\chi^2(1) = 2014.28$, $p < 0.001$. We next added the interaction between Group and Joint rank, which did not significantly improve the model, $\chi^2(1) = 2014.19$, p = 0.76. Finally, we included a full factorial combination of the predictor variables Group, Encoding strength, and Joint rank. The additional terms did not significantly improve the fit, $\chi^2(3) = 2013.73$, p = 0.91. The best-fitting model showed a main effect of Joint rank ($\hat{\beta} = -0.26$, 95% CI $[-0.39, -0.12]$, $z = -3.73$, $p < .001$, Table 7). In summary, this analysis showed that Joint Rank was a significant predictor irrespective of whether or not participants consolidated across sleep.

Table 7: Experiment 2: LMM Group/Encoding strength/Joint rank. Results of the mixed-effects logistic regression model examining the effects of Group (Wake/Sleep), Encoding strength and Joint rank on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	11.92	4.08 – 34.84	<0.001
Group [Wake]	0.88	0.45 – 1.72	0.708
Encoding strength	2.31	1.37 – 3.90	0.002
Jointrank	0.77	0.68 – 0.89	<0.001
Group [Wake] * Encoding strength	0.48	0.24 – 0.94	0.032
Random Effects			
$\sigma 2$	3.29		
$\tau 00$participant	1.79		
ICC	0.35		
N participant	73		
Observations	1752		
Marginal R2 / Conditional R2	0.066 / 0.395		

Note: ^{a}CI = Confidence Interval. All p-values are two-tailed.

3.10 Experiment 2: Discussion

Our attempt to replicate previous findings on sleep-dependent benefits of transitive inference yielded mixed results. The primary goal of the experiment was to validate previous results using a larger, online sample of students participating for course credit. We found no significant effect of sleep on inference performance at delayed recall. In line with the outcomes from the initial experiment (Exp 1), our recent investigation also observed a significant interaction between Encoding strength and whether participants slept during the retention interval. This result re-emphasizes the potential role sleep plays in memory consolidation and transitive inference tasks, specifically how the process of sleep can bolster inferential performance as a function of the encoding strength of acquired information (see Table 5). Additionally, when looking at group effects with respect to symbolic distance and joint rank, the results diverge from those of Experiment 1. Unlike the first experiment, this study did not find any significant differences between the Wake and Sleep groups in these aspects.

There are several potential explanations for this discrepancy. Firstly, the larger sample size and the online nature of the experiment might have introduced variables not present in the initial experiment. Factors such as individual sleep patterns, differences in environment, and lack of in-person control could have influenced the outcome. Secondly, the recruitment process may have introduced some bias, as participants were students who received course credit for participation and possibly were less attentive to the task. Additionally, the complexity of cognitive processes associated with symbolic distance and joint rank might be such that sleep does not exert a consistent influence across all individuals. It is possible that other factors, such as cognitive abilities or learning styles, mediate the relationship between sleep and these specific aspects of transitive inference (S. N. Moses, Villate, Binns, Davidson, & Ryan, 2008; S. Moses et al., 2006).

An important observation was that lack of sleep-benefit was not due to lower than expected inference performance in the Sleep group, but higher than expected performance in the Wake group. There could be multiple reasons for this occurrence, many of which warrant further investigation to fully understand. It could be that the tasks were administered at a time when the Wake group was experiencing peak alertness or cognitive functioning, thereby inflating their performance scores. Another possibility is that participants develop inferential performance irrespective whether retention interval contains sleep or not (Matorina & Poppenk, 2021b; Morgan & Stickgold, 2017), but that sleep provides a beneficial environment for memory consolidation by reducing retroactive interference (Ellenbogen, Payne, & Stickgold, 2006; Mednick et al., 2011), which depending the inference induced by the wake period might not result in significant differences. In terms of reorganization, while we did not observe a sleep-dependent inverse-JRE effect in contrast to Experiment 1, there was an overall inverse-JRE (Fig. S2) which still contrasts the JRE observed in Ciranka et al. (2022), when considering their wake-only data. Consequently, it is possible that the inverse-JRE is primarly time-dependent. However, this is a speculative explanation and further research is needed to understand the role of sleep in the reorganization process.

These mixed results suggest the need for further research to clarify the role of sleep in transitive inference and other forms of cognitive processing. Future studies should consider mitigating potential online experiment constraints, perhaps by implementing stricter controls on participant environment, timing and attention checks. Furthermore, exploring the impact of individual differences in cognitive abilities or learning styles on the relationship between sleep and transitive inference could yield valuable insights. Indeed the observed bimodal distribution of encoding strength observed in Experiment 1 and also in the 66% learning criterion condition of Experiment 2 raises the possibility that studies that ignore encoding strength as an important factor are averaging across different consolidation dynamics. By segregating individuals into those with higher or lower encoding strengths, future studies might be able to observe two distinct consolidation behaviors - a nuance that may have been lost had they relied solely on an overall average. Ignoring the potential influence of encoding strength could indeed lead to an oversimplification of the relationship between sleep, memory, and cognitive performance. This is because a singular average might conceal the diverse effects of sleep on individuals with differing encoding strengths, thereby yielding a homogeneous interpretation of the relationship between sleep and memory consolidation, which, in reality, is likely much more complex. While we attempted to explore the interaction between sleep and encoding strength we only looked at linear effect. Specifically, a hierarchical regression model could be utilized to evaluate whether a polynomial model provides a better fit for the behavioral data. In a hierarchical regression, predictors are entered into the regression model in a series of steps, allowing researchers to evaluate the incremental predictive power of each model. By introducing a quadratic term of encoding strength into the regression, one could test for a potential nonlinear (e.g., U-shaped or inverted U-shaped) relationship between encoding strength and memory consolidation. This could also potentially explain the unexpected high performance of the Wake group, which might be associated with specific encoding strength dynamics.

Despite the partial replication, our study contributes valuable findings to the growing body of research examining the complex interactions between sleep, memory encoding, and transitive inference. It underscores the fact that replication studies often yield nuanced insights that can inform and refine our understanding of cognitive processes.

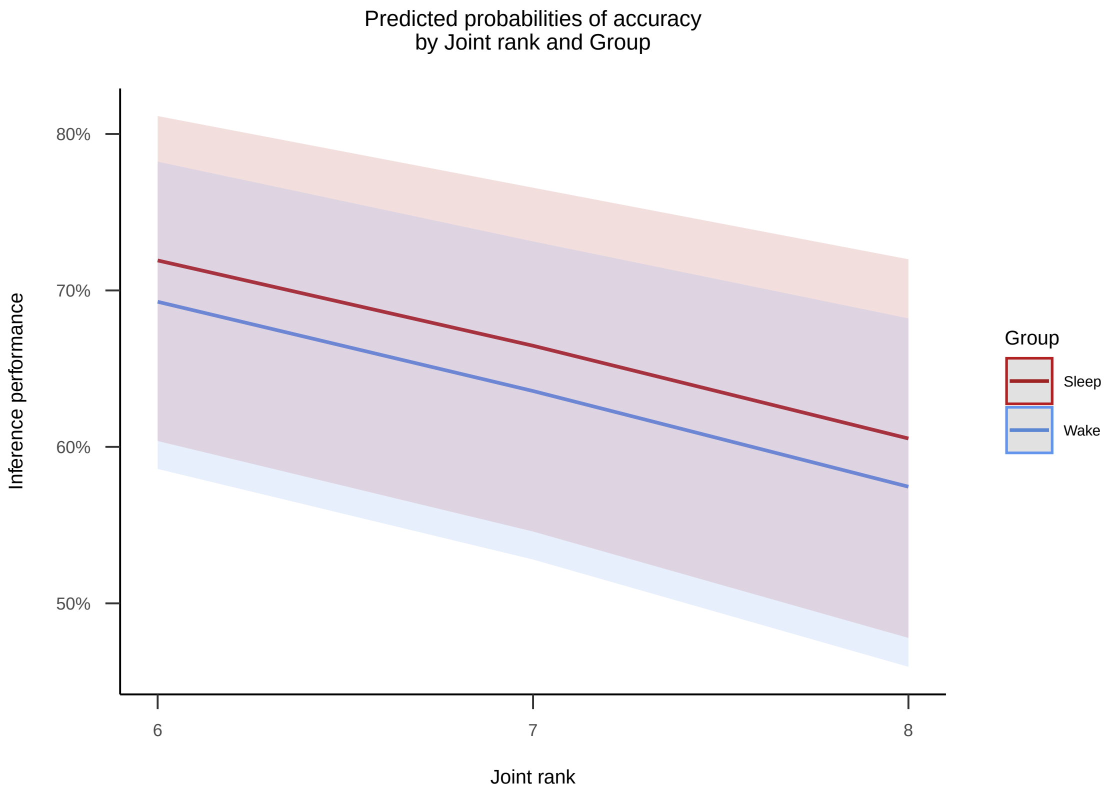
3.11 Experiment 2: Supplemental materials

Supplemental Table S2: Experiment 2: ANOVA table of baseline differences

Effect	$\hat{\eta}^2_G$	90% CI	F	df^{GG}	$df^{\mathrm{GG}}_{\mathrm{res}}$	p
Group	.026	[.000, .118]	2.04	1	67	.157
StimCategory	.002	[.000, .009]	0.08	2	67	.924
Session	.024	[.000, .114]	13.33	1	67	$< .001$
Group × StimCategory	.011	[.000, .062]	0.43	2	67	.651
Group × Session	.003	[.000, .057]	1.52	1	67	.222
StimCategory × Session	.006	[.000, .012]	1.52	2	67	.226
Group × StimCategory × Session	.001	[.000, .000]	0.22	2	67	.803

Table 8: Descriptive Statistics for Delayed testing performance by Sleep and Wake Groups. The table presents the median, mean, and standard deviation (SD) for performance during delayed testing, grouped by Sleep and Wake conditions. Data is further broken down by Pair values.

		Group					
		Sleep			**Wake**		
Session	Pair	Median	Mean	SD	Median	Mean	SD
Delayed test	AB	1.00	0.92	0.14	1.00	0.83	0.26
Delayed test	BC	1.00	0.87	0.28	0.88	0.72	0.33
Delayed test	CD	1.00	0.76	0.36	0.75	0.65	0.35
Delayed test	DE	0.88	0.70	0.36	0.81	0.74	0.30
Delayed test	EF	1.00	0.92	0.20	1.00	0.94	0.16
Delayed test	BD	0.88	0.64	0.37	0.69	0.62	0.32
Delayed test	CE	0.62	0.55	0.41	0.56	0.53	0.38
Delayed test	BE	0.75	0.64	0.39	0.75	0.63	0.36
Delayed test	AF	1.00	0.95	0.12	1.00	0.90	0.21



Supplemental Figure S2: Experiment 2: Group X Joint rank. Plotting model predicted interaction between Group and Joint rank. Shadowed areas represent 95% confidence intervals.

3.12 Experiment 3: Materials and methods

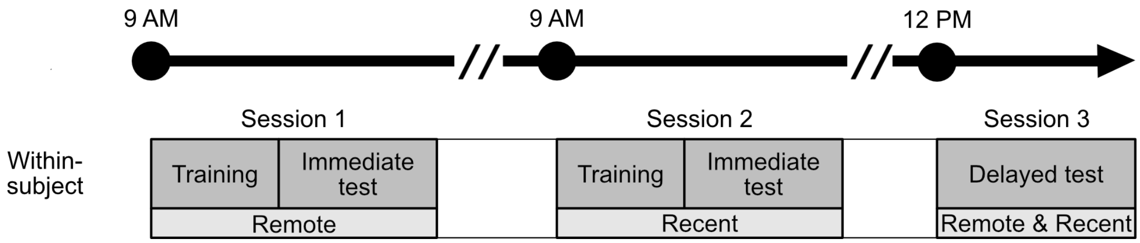


Figure 8: Experimental procedure. Experiment 3. All participants started between 9-11 AM. In Session 1, participants learned one hierarchy (Remote; premise pairs only). Immediately after training, they were tested on these premise pairs. In Session 2, 24 hours later, participants learned a second hierarchy (Recent; premise pairs only). Immediately after training, they were tested on these new premise pairs. In Session 3, three hours later, participants were tested again, this time on premise pairs from both remote and recent hierarchies, as well as their respective novel inference pairs.

3.12.1 Purpose

This study examined the same research question but employed a within-subject, counterbalanced design. Additionally, if our previous hypothesis is again supported utilizing the present within-subject experimental design, then these findings will provide robust evidence of time-dependent consolidation effect reported by Ellenbogen et al. (2007).

Replicating between-subjects findings in within-subject designs is crucial in experimental psychology, particularly in memory research. Memory research often relies on between-subjects designs, where different groups of participants are assigned to different conditions. However, there are limitations to between-subjects designs, such as individual differences and error variance associated with these differences (Greenwald, 1976). Replicating these findings within-subjects helps to address these limitations and provide a more comprehensive understanding of memory processes.

Additionally, within-subject designs provide increased statistical power compared to between-subjects designs (Greenwald, 1976). Within-subject designs involve multiple observations per participant, which increases the sample size and statistical power. This allows for more precise and reliable estimates of the effects being studied. Replicating between-subjects findings within-subjects can help confirm the robustness of the effects and enhance the generalizability of the findings.

Replicating between-subjects findings in within-subject designs also helps to address the complex and interactive nature of memory phenomena (Roediger, III, 2008). Memory research involves various factors, such as types of subjects, events to be remembered, encoding conditions, and test conditions. The use of within-subject designs allows for the manipulation of conditions in a randomized manner, revealing the complex interactions between these factors. By replicating between-subjects findings within-subjects, researchers can explore the robustness of the effects across different conditions and gain a deeper understanding of memory processes.

We again employ a web-based design, but explore the advantages of using Prolific, an online platform that connects researchers with a large pool of participants for their studies (Prolific, 2021). Firstly, Prolific ensures a higher level of participant engagement and motivation compared to students participating for course credit. Participants on Prolific are typically motivated by monetary compensation, which incentivizes them to provide accurate and thoughtful responses (Palan & Schitter, 2018). In contrast, students participating for course credit may be less motivated and may not fully engage with the experiment, potentially compromising the quality of the data collected. Additionally, using Prolific ensures a higher level of data quality and reliability. The platform implements various quality control measures, such as pre-screening participants, monitoring response patterns, and allowing researchers to exclude low-quality or unengaged participants.

3.12.2 Participants

A total of 74 participants completed the study, recruited from Prolific (`https://www.prolific.co/`), an online platform for psychological research. Four participants were excluded from the study due to technical issues related to multiple submissions in either Session 1 or Session 2. Participants (age=23.37 ±4.1) had normal or corrected-to-normal vision, with no-self reported history of neurological, psychiatric, sleep or motor disorders. All participants provided informed consent electronically and were reimbursed for their time. The experiment was approved by the School of Psychology Ethics Committee at Cardiff University. All participants agreed to abstain from caffeine and alcohol during the study and for 24 hours before it.

3.12.3 Experimental protocol

Participants were required to complete three sessions: the first session in the morning (between 9-11 AM local time), the second session 24h later (between 9-11 AM local time) and the third session 3h after completing the second session (or 27h from the first). The first session had an initial training and immediate test part for only one hierarchy (named Remote hierarchy). The second session was structured the same but participants learned a completely novel hierarchy (Recent hierarchy). Finally, for the last session a test involving both hierarchies was performed (Delayed test).

Premise pairs training

Training protocol was identical to Experiment 1 with respect to instructions, stimuli and learning criteria; the only difference being that participants learned only one hierarchy per training session. Participants had to reach learning criteria (66% accuracy in two consecutive blocks on middle pairs) for a given hierarchy within 10 blocks.

Immediate and Delayed test

The Immediate test phase was also identical to Experiment 1 and only tested the hierarchy learned in its respective session. Following a delay of 3 hours after Session 2, participants were required to complete the Delayed test phase. Similar to

Experiment 1, this phase involved three novel inference pairs and an anchor pair, in addition to the five premise pairs for both Remote and Recent hierarchy and lasted 4 blocks. Participants were instructed that they may see novel combinations and, if that happened, to make their best guess on that trial. Similar to the immediate test, participants performed four blocks with two arithmetic problems between them but were not asked to provide confidence ratings after each trial. After completing this phase participants had to fill out a questionnaire to probe their explicit awareness of the existence of a latent hierarchy underlying the items for each hierarchy and finally, complete a hierarchical recall test whereby they were asked to explicitly reconstruct the order of items in the hierarchy to assess implicit awareness.

Apparatus and stimuli

Participants completed the experiment on their own computer or tablet device using a web browser (e.g., Chrome). Experiment 3 was programmed using PsychJS (Peirce, 2007) and run through the Pavlovia platform (`https://pavlovia.org/`). In the first experiment, the visual stimuli consisted of three sets of images (hereafter referred to as Stimulus categories), including female faces, unusual objects (Horst & Hout, 2016), and landscapes, each comprising six items. These items were selected randomly from a set of 12 images for each category. Participants were randomly assigned to one of the stimulus categories, with both the Remote and Recent hierarchy consisting of a six-item subset. All the items were presented in greyscale and matched for luminescence. Each item was distinguishable from the items within and between stimulus categories. The rank-order of the stimuli within each relational hierarchy was completely randomized for each subject at the start of the learning phase.

3.12.4 Behavioural data analysis

To examine the relative impact of Encoding strength, Distance and Joint rank on delayed inferential performance, separate hierarchical multiple regression analyses were conducted for Distance and Joint rank. We employed specific transformations on the variables used. First, we mean centered the Distance and Joint rank variables to enhance interpretability and reduce multicollinearity among predictor variables. Additionally, we applied a transformation to the Encoding strength variable by multiplying the original values, which initially ranged between 0 and 1, by a factor of ten, enabling us to assess the impact of Encoding strength in terms of a rate per units of 10%.

To first identify the contributions of Encoding strength, we entered our experiment-specific condition factor into the model in Step 1 (Hierarchy), Encoding strength in Step 2, and the interaction of these measures at Step 3 as predictors of memory recall for inferential pairs. Next, we extended the interaction model by adding in the Distance factor as Step 4, an interaction between Distance and the respective condition factor as Step 5 and finally a model that included a full factorial combination between condition, Encoding strength and Distance as Step 6. We also repeated this for Joint rank replacing Distance. An alpha value of $p<.05$ was employed for all analyses.

We used R (version 4.2.2) and the R-packages *lme4*,*afex* and *emmeans* for all our statistical analyses (Bates et al., 2015; Lenth, 2023; Singmann et al., 2023) and *sjPlot* for generating the regression tables (Lüdecke, 2023).

3.13 Experiment 3: Results

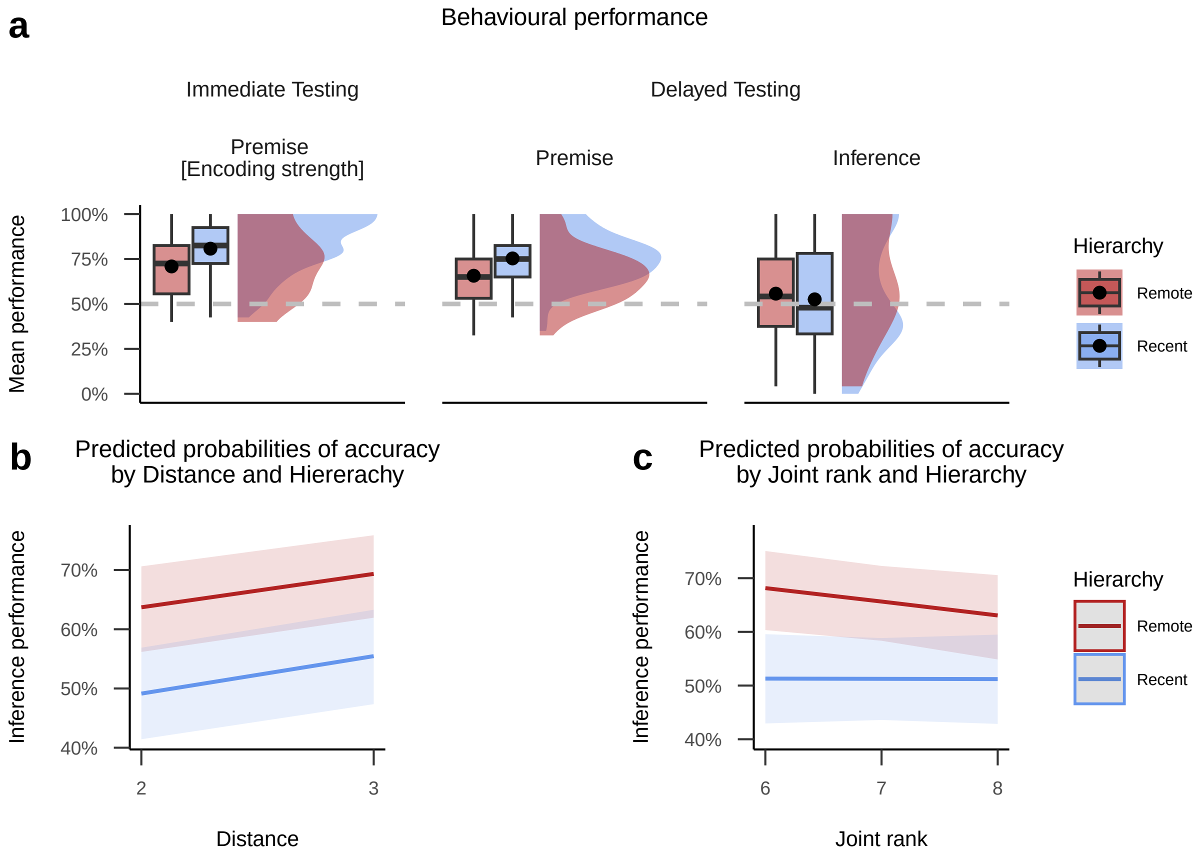


Figure 9: Experiment 3: Behavioural performance and factors predicting inference accuracy at delayed test. (a) Raincloud plot with mean and median performance across Remote and Recent hierarchy conditions. Dashed line represents chance performance at 50% and dot represents mean values (b) Predicted probabilities of accuracy by Distance and Hierarchy, with distance levels of 2 and 3. (c) Predicted probabilities of accuracy by Joint rank and Hierarchy, with joint rank levels ranging from 6 to 8. In panels (b) and (c), shadowed areas represent 95% confidence intervals.

As in Experiment 1, mean inference performance was significantly higher than chance for the Remote hierarchy $M = 0.56$, 95% CI $[0.50, \infty]$, $t(69) = 1.80$, $p = .038$, but not for the Recent hierarchy ($M = 0.53$, 95% CI $[0.47, \infty]$, $t(69) = 0.75$, $p = .228$). For a table of descriptive table of pair level performance, see Table 13.

To test for between Hierarchy baseline differences in premise pair learning, we performed a mixed three-way ANOVA on mean premise pair performance with within-subject factors Hierarchy and Session and between-subject factor Stimulus category. This revealed an effect of Hierarchy ($F(1, 67) = 33.61$, $p < .001$, $\hat{\eta}^2_G = .098$, 90% CI $[.015, .223]$) with higher performance for the Recent hierarchy and Session ($F(1, 67) = 22.33$, $p < .001$, $\hat{\eta}^2_G = .027$, 90% CI $[.000, .120]$) showing decreased performance at Delayed test. There was no main effect of Stimulus category nor interactions (smallest p-value 0.32, for details see Table S3).

These findings suggest that for premise pairs there is a baseline difference in performance between the two hierarchies, whereby participants had higher accuracy for the Recent hierarchy both at Immediate and Delayed test.

3.13.1 Encoding strength

We examined the influence of Encoding strength similar to Experiment 1. We first added the Hierarchy variable to the null model, which captured whether the participants' performance relate to the Remote or Recent hierarchy. The addition of this variable significantly improved the model's ability to predict accuracy, $\chi^2(1)$ = 4211.89, p = 0.045. Next, we added the Encoding strength variable, which significantly further improved fit, $\chi^2(1) = 4104.51$, $p < 0.001$. Finally, we added an interaction term between Hierarchy and Encoding strength, which also significantly improved the model's ability to predict accuracy beyond the previous model, $\chi^2(1)$ = 4079.58, $p < 0.001$. The results of this analysis suggest that both Hierarchy and Encoding strength have significant effects on accuracy and that their interaction also plays an important role in predicting accuracy on inference trials. As can be seen in Table 9, Hierarchy was a significant predictor of the outcome variable ($\hat{\beta} = 1.54$, 95% CI $[0.68, 2.39]$, $z = 3.52$, $p < .001$), as was Encoding strength ($\hat{\beta} = 5.75$, 95% CI $[4.72, 6.79]$, $z = 10.92$, $p < .001$) and the interaction between the two ($\hat{\beta} = -2.81$, 95% CI $[-3.92, -1.71]$, $z = -4.98$, $p < .001$). The interaction indicates that the effect of Encoding strength on accuracy differs significantly between the Remote and Recent conditions. To probe the interaction, simple effects coefficients were computed at 60% ($\mathrm{OR}_{\mathrm{Remote/Recent}} = 1.16$, SE = 1.14, $p = 0.24$) and 80% ($\mathrm{OR}_{\mathrm{Remote/Recent}} = 2.04$, SE = 1.10, $p < 0.001$) values of Encoding strength, suggesting increased time-dependent benefit with stronger encoding.

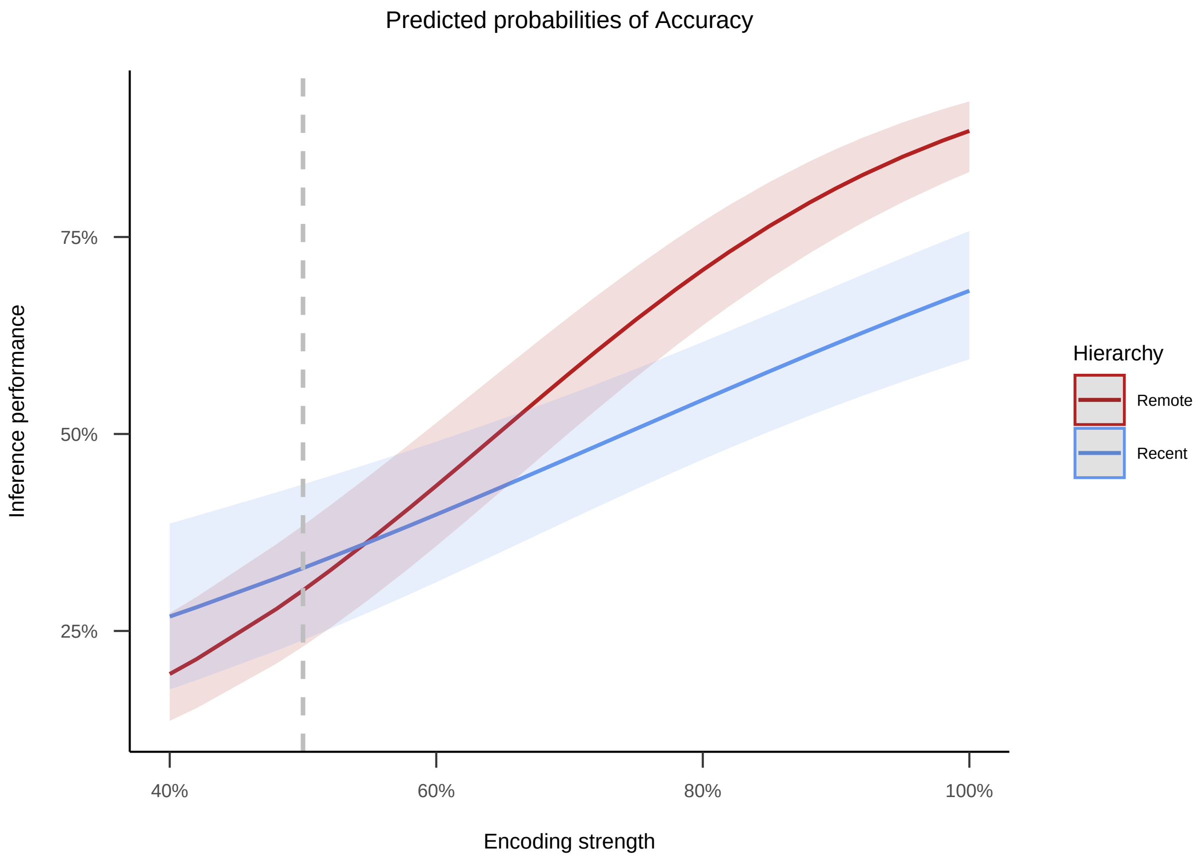


Figure 10: Experiment 3: Hierarchy X Encoding strength. Plotting model predicted interaction between Group and Encoding strength. Shadowed areas represent 95% confidence intervals.

Table 9: Experiment 3: LMM Hierarchy/Encoding strength. Results of the mixed-effects logistic regression model examining the effects of Hierarchy (Remote/Recent) and Encoding strength on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	0.02	0.01 – 0.05	<0.001
Hierarchy [Recent]	4.65	1.97 – 10.97	<0.001
Encoding strength	1.78	1.60 – 1.97	<0.001
Hierarchy [Recent] XEncoding strength	0.75	0.68 – 0.84	<0.001
Random Effects			
σ2	3.29		
τ00participant	1.44		
ICC	0.31		
N participant	70		
Observations	3360		
Marginal R2 / Conditional R2	0.105 / 0.378		

Note: ^{a}CI = Confidence Interval. All p-values are two-tailed.

3.13.2 Distance

In order to assess how Distance affects inference performance during testing, we conducted a hierarchical regression analysis that builds upon the baseline model (Table 9). We first added the Distance variable which significantly improved the model's ability to predict accuracy compared to the baseline model, $\chi^2(1) = 4070.03$, p = 0.002. Next, we added an interaction term between Hierarchy and Distance, which did not significantly improve the model's ability to predict accuracy, $\chi^2(1) = 4069.51$, p = 0.47. Finally, we included the full factorial model (see Methods), but the additional terms did not significantly improve the fit,$\chi^2(3) = 4067.55$, p = 0.48. The best-fitting model showed the same effects of Hierarchy, Encoding strength and Hierarchy x Encoding strength as reported by the baseline model. Additionally, there was a main effect of Distance ($\hat{\beta} = 0.25$, 95% CI $[0.09, 0.41]$, $z = 3.09$, $p = .002$, Table 10). In summary, this analysis showed that the symbolic distance over which participants had to make an inference was a significant predictor of inference performance, but this did not significantly differ between hierarchies.

Table 10: Experiment 3: LMM Hierarchy/Encoding strength/Distance. Results of the mixed-effects logistic regression model examining the effects of Hierarchy (Remote/Recent), Encoding strength and Distance on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	0.02	0.01 – 0.05	<0.001
Hierarchy [Recent]	4.67	1.98 – 11.03	<0.001
Encoding strength	1.78	1.61 – 1.98	<0.001
Distance	1.29	1.10 – 1.51	0.002
Hierarchy [Recent] XEncoding strength	0.75	0.67 – 0.84	<0.001
Random Effects			
σ2	3.29		
τ00participant	1.45		
ICC	0.31		
N participant	70		
Observations	3360		
Marginal R2 / Conditional R2	0.108 / 0.381		

Note: ^{a}CI = Confidence Interval. All p-values are two-tailed.

3.13.3 Joint rank

In parallel, we assessed the effects of Joint rank on inference performance at test. We first added the Joint rank to the baseline model (Table 9) which did not significantly improve fit, $\chi^2(1)$ = 4079.20, p = 0.54. Next, we added an interaction term between Hierarchy and Joint rank, which again did not significantly improve the model's ability to predict accuracy, $\chi^2(2)$ = 4076.11, p = 0.18. Lastly, we compared a full factorial combination of the predictor variables Hierarchy, Encoding strength and Joint rank. The additional terms significantly improved the model's ability to predict accuracy, $\chi^2(4)$ = 4065.05, p = 0.0058. The best-fitting model showed main effect for Joint rank ($\hat{\beta} = -0.12$, 95% CI $[-0.69, 0.46]$, $z = -0.40$, $p = .691$), in addition to effects shown in the baseline model. Furthermore, there was a two-way interaction effect between Hierarchy and Joint rank ($\hat{\beta} = -1.01$, 95% CI $[-1.93, -0.10]$, $z = -2.17$, $p = .030$) and a three-way interaction between Hierarchy, Encoding strength and Joint rank ($\hat{\beta} = 0.15$, 95% CI $[0.03, 0.27]$, $z = 2.44$, $p = .015$, see Table 11). In order to examine the interaction, we computed the coefficients of simple effects for BD ($\mathrm{OR}_{\mathrm{Remote/Recent}} = 2.03$, SE = 1.14, $p < 0.001$) and CE trials ($\mathrm{OR}_{\mathrm{Remote/Recent}} = 1.63$, SE = 1.14, $p < 0.001$). From this post-hoc analysis, we can conclude that for BD trials, the odds of success in the Remote compared to the Recent condition were estimated to be 2.03 times higher, while for CE trials it was only 1.63 times higher.

Table 11: Experiment 3: LMM Hierarchy/Encoding strength/Joint rank. Results of the mixed-effects logistic regression model examining the effects of Hierarchy (Remote/Recent), Encoding strength and Joint rank on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	0.02	0.01 – 0.05	<0.001
Hierarchy [Recent]	4.68	1.98 – 11.07	<0.001
Encoding strength	1.78	1.61 – 1.98	<0.001
Joint rank	0.89	0.50 – 1.58	0.691
Hierarchy [Recent] XEncoding strength	0.75	0.67 – 0.84	<0.001
Hierarchy [Recent] XJoint rank	0.36	0.15 – 0.91	0.030
Encoding strength X Jointrank	1.00	0.92 – 1.08	0.991
(Hierarchy [Recent] XEncoding strength) XJoint rank	1.16	1.03 – 1.31	0.015
Random Effects			
σ2	3.29		
τ00participant	1.46		
ICC	0.31		
N participant	70		
Observations	3360		
Marginal R2 / Conditional R2	0.109 / 0.383		

Note: ^{a}CI = Confidence Interval. All p-values are two-tailed.

3.14 Combined dataset

In order to provide a comprehensive and well-powered overview of the study-level results, we elected to combine the data from three separate experiments. This decision was based on the need to increase statistical power and enhance the reliability of our findings. By integrating these datasets, we were able to draw on a larger sample size, thereby improving our ability to detect significant effects.

3.14.1 Premise pair performance

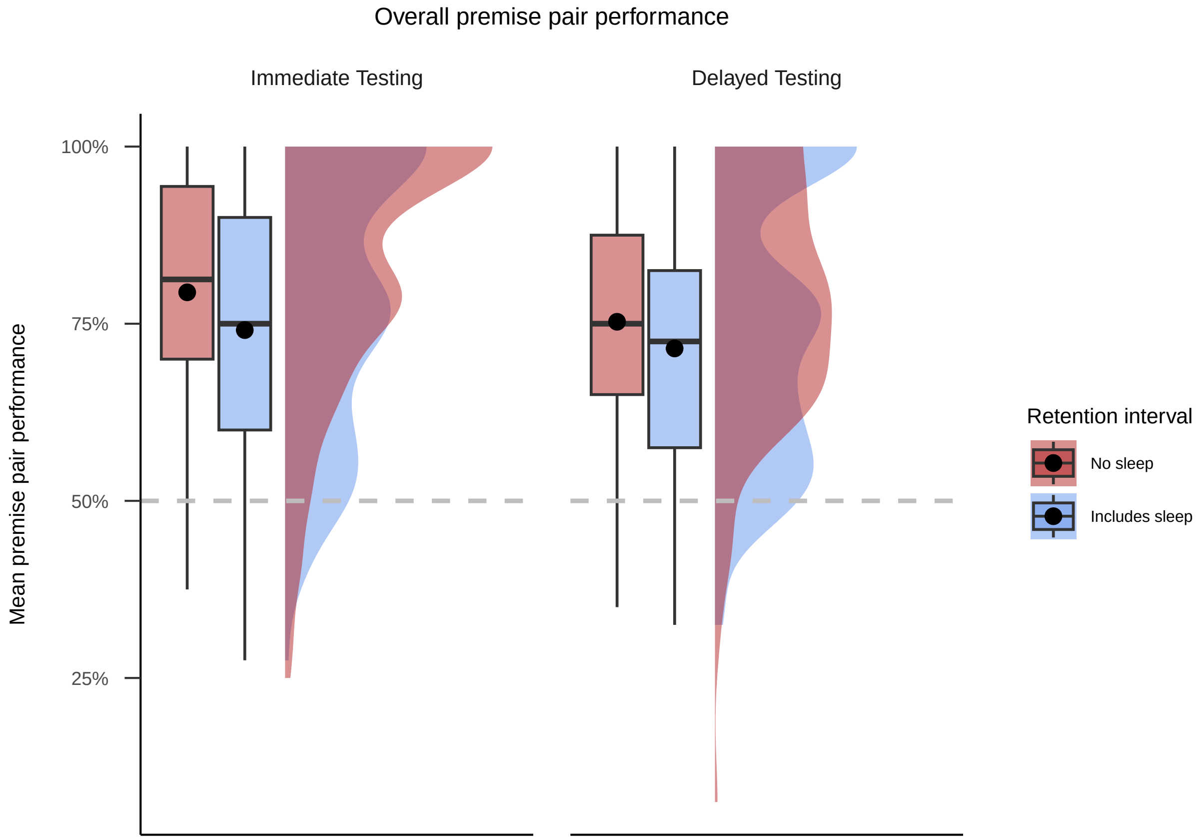


Figure 11: Overall premise pair performance across session and retention interval. Combined (Exp. 1, 2 and 3) behavioral premise pair performance displayed as a raincloud plot, with 'Session' on the x-axis and 'Retention Interval' represented by different fill colors. Each boxplot presents the median and mean premise pair performance, and the central dashed line indicates the 50% performance threshold. Mean performance points are marked as solid circles. The data is organized by Session.

To test for overall baseline difference in premise pair performance we combined immediate and delayed performance recorded in Experiment 1,2 and 3 (see Fig. 11).We fitted a linear mixed model (estimated using REML and nloptwrap optimizer) to predict Mean premise pair accuracy with Retention interval and Session (formula: Mean premise pair accuracy ~ Retention interval * Session). The model included participant as random effect (formula: ~1 | participant). The model's total explanatory power is substantial (conditional R2 = 0.59) and the part related to the fixed effects alone (marginal R2) is of 0.06. The model's intercept, corresponding to Retention interval = No sleep and Session = Immediate Testing, is at 0.82 (95% CI [0.80, 0.85], t(578) = 60.81, p < .001). Within this model the effect of Retention interval [Includes sleep] is statistically significant and negative (beta = 0.09, 95% CI [0.11, 0.06], t(578) = 5.69, p < .001; Std. beta = 0.49, 95% CI [0.67, 0.32]). The effect of Session [Delayed Testing] is statistically significant and negative (beta = 0.04, 95% CI [0.06, 0.01], t(578) = 2.94, p = 0.003; Std. beta = 0.22, 95% CI [0.37, 0.07]). The effect of Retention interval [Includes sleep] × Session [Delayed Testing] is statistically non significant and positive (beta = 0.01, 95% CI [0.02, 0.05], t(578) = 0.72, p = 0.470; Std. beta = 0.08, 95% CI [0.13, 0.29]). The results indicate that a significant decrease in the mean premise pair accuracy is associated with the inclusion of sleep in the retention interval. This suggests that involving a period of sleep in our conditions negatively impacted the accuracy of premise pairs, possibly reflective of decreased alertness due to the AM-PM design. A unit change in the retention interval from ‘No sleep’ to ‘Includes sleep’ is linked to decreasing the accuracy by 0.09 (i.e., ~9%). Furthermore, a shift from immediate to delayed testing is also connected to a significant decline in accuracy, at a smaller degree than the effect of the retention interval. This implies that testing premises after a prolonged period following encoding (>12 hours) shows signs of forgetting compared to immediate testing, evidenced by decreasing accuracy of 0.04 units (i.e., ~4%).

3.14.2 Distance

We fitted a logistic mixed model (estimated using ML and BOBYQA optimizer) to predict Accuracy with Retention interval, Encoding strength and Distance (formula: Accuracy ~ Retention interval * Encoding strength * Distance). The model included participant as random effect (formula: ~1 | participant). The model’s total explanatory power is substantial (conditional R2 = 0.32) and the part related to the fixed effects alone (marginal R2) is of 0.04. The model’s intercept, corresponding to Retention interval = Includes sleep, Encoding strength = 0 and Distance = 0, is at 1.96 (95% CI [2.53, 1.40], p < .001). Within this model the effect of Retention interval [No sleep] is statistically significant and positive (beta = 2.34, 95% CI [1.67, 3.01], p < .001; Std. beta = 0.34, 95% CI [0.49, 0.19]). The effect of Encoding strength is statistically significant and positive (beta = 0.34, 95% CI [0.27, 0.41], p < .001; Std. beta = 0.60, 95% CI [0.47, 0.72]). The effect of Retention interval [No sleep] × Encoding strength is statistically significant and negative (beta = 0.35, 95% CI [0.44, 0.26], p < .001; Std. beta = 0.62, 95% CI [0.77, 0.46]). For detailed results see Table 12.

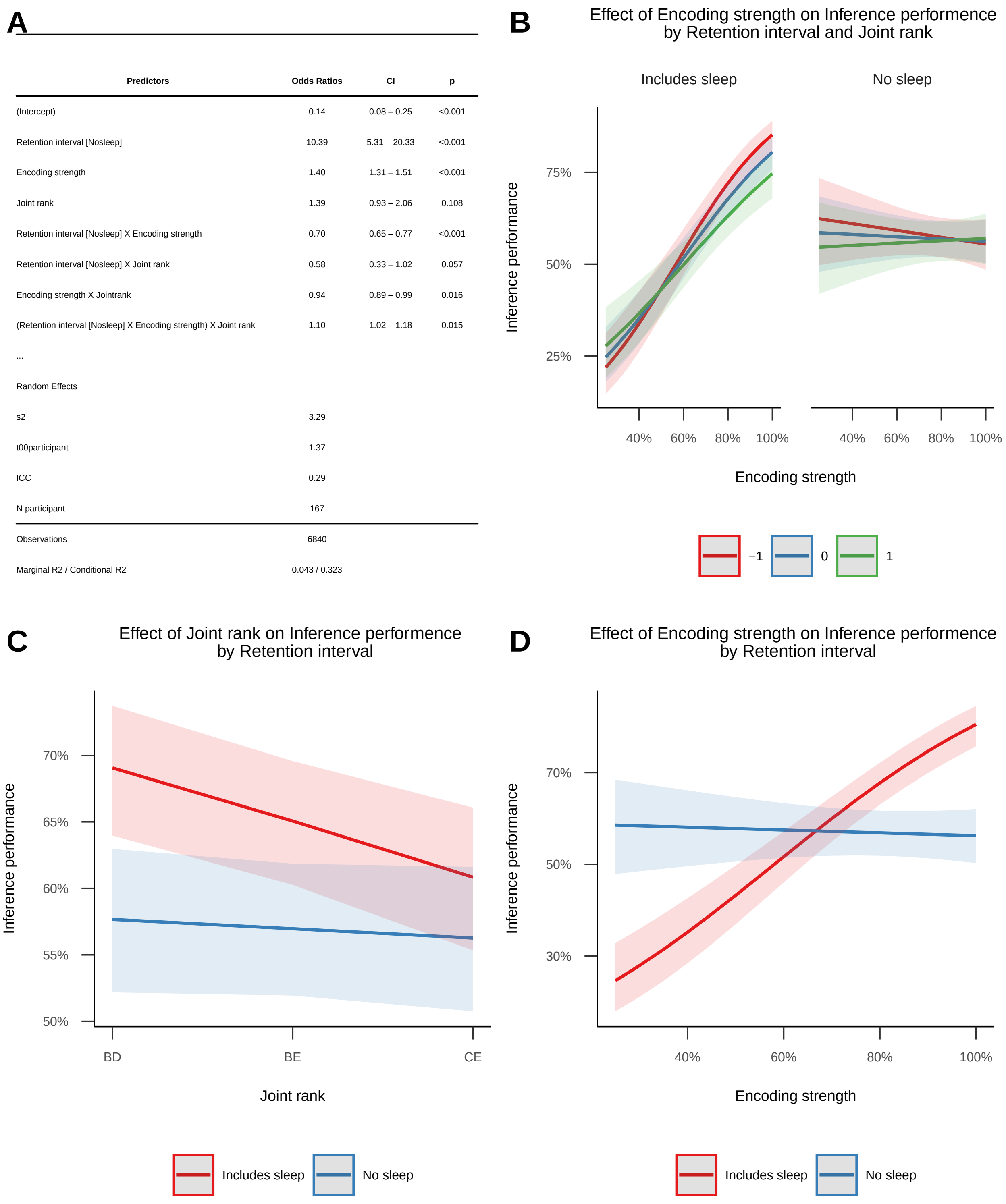
Table 12: Combined dataset: LMM Retention interval X Encoding strength X Distance. Results of the mixed-effects logistic regression model examining the effects of Retention interval (Sleep/No sleep), Encoding strength and Symbolic distance on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	0.14	0.08 – 0.25	<0.001
Retention interval [Nosleep]	10.36	5.29 – 20.28	<0.001
Encoding strength	1.40	1.30 – 1.51	<0.001
Distance	1.72	0.86 – 3.41	0.124
Retention interval [Nosleep] X Encoding strength	0.70	0.65 – 0.77	<0.001
Retention interval [Nosleep] X Distance	0.46	0.17 – 1.23	0.124
Encoding strength XDistance	0.97	0.88 – 1.06	0.509
(Retention interval [Nosleep] X Encoding str..) X D..	1.08	0.95 – 1.22	0.256
Random Effects			
σ2	3.29		
τ00participant	1.36		
ICC	0.29		
N participant	167		
Observations	6840		
Marginal R2 / Conditional R2	0.042 / 0.323		

Note: ^{a}CI = Confidence Interval. All p-values are two-tailed.

3.14.3 Joint rank

We fitted a logistic mixed model (estimated using ML and BOBYQA optimizer) to predict Accuracy with Retention interval, Encoding strength and Joint rank (formula: Accuracy ~ Retention interval * Encoding strength * Joint rank). The model included participant as random effect (formula: ~1 | participant). The model's total explanatory power is substantial (conditional R2 = 0.32) and the part related to the fixed effects alone (marginal R2) is of 0.04. The model's intercept, corresponding to Retention interval = Includes sleep, Encoding strength = 0 and Joint rank = 0, is at 1.96 (95% CI [2.53, 1.40], p < .001). Within this model the effect of Retention interval [No sleep] is statistically significant and positive (beta = 2.34, 95% CI [1.67, 3.01], p < .001; Std. beta = 0.34, 95% CI [0.49, 0.19]). The effect of Encoding strength is statistically significant and positive (beta = 0.34, 95% CI [0.27, 0.41], p < .001; Std. beta = 0.60, 95% CI [0.47, 0.72]). The effect of Retention interval [No sleep] × Encoding strength is statistically significant and negative (beta = 0.35, 95% CI [0.44, 0.26], p < .001; Std. beta = 0.62, 95% CI [0.77, 0.47]). The effect of Encoding strength × Joint rank is statistically significant and negative (beta = 0.07, 95% CI [0.12, 0.01], p = 0.016; Std. beta = 0.10, 95% CI [0.17, 0.02]). The effect of (Retention interval [No sleep] × Encoding strength) × Joint rank is statistically significant and positive (beta = 0.09, 95% CI [0.02, 0.17], p = 0.015; Std. beta = 0.13, 95% CI [0.03, 0.24]). For detailed results see Table 12A. In order to examine the three-way interaction, we computed the coefficients of simple effects for BD (JR:6, $\mathrm{OR}_{\text{Includes sleep/No sleep}} = 1.64$, SE = 1.11, $p < 0.001$) and CE trials (JR:8, $\mathrm{OR}_{\text{Includes sleep/No sleep}} = 1.21$, SE = 1.11, $p = 0.07$). From this post-hoc analysis, we can conclude that for BD trials, the odds of success in the Sleep group compared to the Wake group were estimated to be 1.64 times higher, and this difference was statistically significant. For a visual representation of the three-way interaction, see Figure 12.



Predictors	Odds Ratios	CI	p
(Intercept)	0.14	0.08 – 0.25	<0.001
Retention interval [Nosleep]	10.39	5.31 – 20.33	<0.001
Encoding strength	1.40	1.31 – 1.51	<0.001
Joint rank	1.39	0.93 – 2.06	0.108
Retention interval [Nosleep] X Encoding strength	0.70	0.65 – 0.77	<0.001
Retention interval [Nosleep] X Joint rank	0.58	0.33 – 1.02	0.057
Encoding strength X Jointrank	0.94	0.89 – 0.99	0.016
(Retention interval [Nosleep] X Encoding strength) X Joint rank	1.10	1.02 – 1.18	0.015
...			
Random Effects			
s2	3.29		
t00participant	1.37		
ICC	0.29		
N participant	167		
Observations	6840		
Marginal R2 / Conditional R2	0.043 / 0.323		

Figure 12: Three-way interaction: Encoding strength X Joint rank X Retention interval A) Results of the mixed-effects logistic regression model examining the effects of Retention interval (Includes sleep vs. No sleep), Encoding strength, and Joint rank on inference accuracy. Data from Experiment 1, 2 and 3 have been combined, with condition Sleep (Exp 1, Exp 2) and Remote (Exp 3) being relabeled as 'Includes sleep' and Wake (Exp 1, Exp 2) and Recent (Exp 3) as 'No sleep' B) Predicted probabilities of accuracy by Joint rank, Retention interval and Encoding Strength, with mean centered Joint rank levels ranging from -1 (B?D) to 1 (C?E) C) Predicted probabilities of accuracy by Joint rank and Retention interval, with mean centered Joint rank levels ranging from -1 (B?D) to 1 (C?E). D) Predicted probabilities of accuracy by Encoding Strength and Retention interval. Shadowed areas represent 95% confidence intervals.

3.15 Experiment 3 & Combined data: Discussion

The transitive inference (TI) paradigm evaluates how well individuals can grasp the relational arrangement of a group of stimuli without relying on any overt hierarchical cue while learning about them. To achieve this, the training involves displaying pairs of images selected from a pre-determined list of neighboring items in order of rank and providing incentives for participants to correctly identify the dominant item. Work in both humans and other animals has supported the idea that the ability to assess relational dominance for non-neighboring, inference items rely on some form of gist abstraction and that this is facilitated by time and sleep-dependent memory consolidation in humans (Inostroza & Born, 2013; Lerner & Gluck, 2019a; Lewis & Durrant, 2011; Lewis et al., 2018b). While initial findings (Ellenbogen et al., 2007; Werchan & Gómez, 2013a, 2016) showed strong time and sleep-dependent benefits in inference performance and sleep-dependent SDE, subsequent studies that implemented modified parameters failed to replicate these findings, suggesting that the time/sleep-inference relationship may be more complex and dependent on specific experimental conditions (Cordi & Rasch, 2021). Here, we evaluated whether we could partially replicate the original time and sleep-dependent TI findings by Ellenbogen et al. (2007) in two experiments that involve learning multiple hierarchies, with lower learning criterion and using a broader range of stimuli compared to the original study.

Crucially our Experiment 1 and 3 support earlier findings (Ellenbogen et al., 2007; Werchan & Gómez, 2013a, 2016) by showing increased inference performance at delayed test in the experimental Sleep group, Experiment 1 (see Table 1) and in the Remote condition, Experiment 3 (see Table 9) compared to control. Interestingly, both of these effects showed increase as a function of encoding strength (Exp 1. Fig. 4, Exp 3. Fig. 10. Our third experiment followed a similar design to Berens & Bird (2022b), which found no time-dependent benefit in TI. We speculate that this could be either due to differences in difficulty (we employed 6 items vs their 7-item hierarchy) or differences in encoding strength due to training duration (our 66% learning criterion vs their large number of fixed trials for each hierarchy). While training to ceiling can eliminate baseline, learning-related differences in premise pair encoding strength, which we observed in Exp 3. (Fig. 9), it can also exacerbate practice effects due to increased familiarity with task demands during the Recent hierarchy condition leading to altered consolidation dynamics (Denis et al., 2021b, 2020; Petzka, Charest, Balanos, & Staresina, 2021). In Exp 3, we observed a time-dependent benefit in TI despite weaker encoding strength in the Remote compared to

Recent condition, even though we found encoding strength to be a strong predictor of delayed inference performance overall, which highlights the importance of retention interval-encoding strength interaction.

Concerning the sleep-dependent SDE shown in past work (Ellenbogen et al., 2007; Werchan & Gómez, 2013a), we were able to replicate this in our AM-PM design in Exp 1 (Fig. 3), but could not replicate it in the within-subject design of Exp 3 (nor in Experiment 2). The latter results are a common pattern in studies using a longitudinal design (Berens & Bird, 2022b; Matorina & Poppenk, 2021b). This, again, could be caused by practice effects whereby SDE might take less time to emerge in the Recent hierarchy condition due to familiarity with task demands, leading to smaller condition-specific differences.

To the best of our knowledge, we are the first to explore the time and sleep-dependent joint rank effect (JRE). JRE, whereby participants - when comparing performance for pairs that have equivalent symbolic distance - exhibit lower accuracy the higher the cumulative sum of the item ranks, has been observed in both primates (Ciranka et al., 2022; Munoz et al., 2020a) and humans (Ciranka et al., 2022; Jensen et al., 2017b). JRE has been suggested as a valuable complementary variable to symbolic distance, as it serves as a measure of absolute positional encoding during serial learning and retrieval (Jensen et al. 2017) and is indicative of a cognitive model supporting TI (Behrens et al., 2018). The reduced ability to choose between more dominant items (e.g. B?D, pair with low joint rank value) compared to less dominant items (e.g. C?E, pair with higher joint rank value) observed in wake-only studies could be caused by compressed representations of magnitude that can emerge from an asymmetric learning policy (Ciranka et al., 2022). Interestingly, our two studies found that this pattern reversed across a retention interval containing sleep, with participants showing a lesser ability to discriminate during low dominance inference pairs (C?E, JR:8), and a greater ability to discriminate between high dominance items (B?D, JR:6) after consolidation (see Fig. 3, Fig. 9). This was further supported in the combined analysis (see Fig. 12) This "inverse joint rank effect" could suggest active time and/or sleep-dependent consolidation processes, whereby experience is reorganized based on the learned absolute rank-structure prioritizing high dominance items over low dominance ones (Y. Liu et al., 2019; Y. Liu, Mattar, et al., 2021).

While the current study cannot provide a mechanistic explanation for this phenomenon, one explanation could relate to differences in wake/sleep replay dynamics as it pertains to how veridical or "structured" online vs "disjointed" offline (in sleep) sequential reactivations can affect TI performance (Svenja Brodt et al., 2023). Moreover, while replay can be forward and reverse in both wake and sleep, and whether their relative frequency differs across these states is still unclear, a dominance of forward replay in sleep could potentially explain our observation of inverse joint rank after consolidation across sleep (Findlay, Tononi, & Cirelli, 2020). A further explanation could involve offline replay prioritizing dominant, high-valued items at the expense of lower-ranked items (Momennejad, 2020). Extending existing computational models of transitive inference with biologically inspired replay and fitting to delayed test data could help adjudicate between these possibilities (Ciranka et al., 2021a; Jensen, Terrace, & Ferrera, 2019a; Mattar & Daw, 2018; Roscow et al., 2021).

One limitation of this study stems from the fact that we did not record sleep. Although Experiment 1 compared consolidation across 12 hours of wake and 12 hours including an overnight sleep, we cannot assume that the benefit observed in the overnight condition relates specifically to sleep rather than to a combination of wake and sleep. Our second study suffered from the same problem, but to an even greater extent, as we now compared consolidation across 27 hours, including a night of sleep, to consolidation across 3 hours. Future studies could determine the specific importance of sleep for these effects more definitively by recording polysomnography.

A further limitation of our study is the use of a relatively small sample size in our first experiment. While our study yielded significant results, increasing the sample size would have enhanced the generalizability and robustness of our findings, allowing for a more comprehensive understanding of the investigated phenomena and strengthening the validity of our conclusions. Additionally, while we used well-established transitive inference tasks to assess memory consolidation, these tasks do not capture all aspects of relational memory, and other cognitive processes may have played a role in our results. In Experiment 2, which was web-based, we encountered a potential limitation in the form of reduced participant motivation due to the incentive structure that revolved around course credit, as opposed to the financial compensation provided in Experiment 3. Despite this potential influence on participant engagement, the findings of Experiment 2 offered valuable insights into the nature of cognitive consolidation. Interestingly, this pointed towards the possibility of consolidation-independent generalization occurring during periods of wakefulness, paralleling the traditionally accepted process thought to primarily occur during sleep. This does raise important concerns over using the current transitive inference design and consequent interpretation, as participants could already be performing above chance shortly after encoding (Matorina & Poppenk, 2021a; Morgan & Stickgold, 2017), with inference performance difference be driven by interference, not active consolidation (Yonelinas et al., 2019a).

The results of our study have important implications for understanding the role of time and/or sleep in memory consolidation and suggest that both play a key role in the formation of the relational memories underlying transitive inference. Given that transitive inference is a fundamental cognitive process that is involved in a wide range of daily activities, our findings may have implications for developing computational models of consolidation and improving learning and memory in educational and clinical settings. Future studies may explore whether manipulating the online/offline replay or duration of sleep/wakeful rest can enhance transitive inference performance to better understand how rank-order effects emerge during learning and evolve over time, potentially leading to the development of interventions that can improve memory consolidation. Additionally, conducting sleep and memory studies using a derived hierarchies approach like the one used by Kao et al. (2020) where participants need to generalize across hierarchies would provide a better chance at disentangling sleep-dependent effects on relative vs absolute encoding as well as investigating more flexible forms of generalization than single hierarchy studies can.

In summary, we corroborate earlier findings showing that consolidation across a period of time including sleep benefits inference performance. We also show that this benefit increases as a function of encoding strength, and therefore argue that TI studies that employ a “learn to criterion” approach during training should consider encoding strength as a relevant predictor since there can be significant variation in premise pair performance using standard approaches. Our data, in combination with prior reports, suggests that the emergence of the sleep-dependent symbolic distance effect may be influenced by experimental design. Specifically, studies that employ between-subjects designs have found such effects, whereas those using within-subjects designs have not. The joint rank effect provides a complementary dimension of the mental model which participants use in solving this task which is understudied in the transitive inference literature. We find that it is a strong predictor of inference performance and that it is a sensitive measure of time and sleep-dependent memory consolidation, affected by encoding strength.

3.16 Experiment 3: Supplemental materials

Supplemental Table S3: Experiment 3: ANOVA table of baseline differences

Effect	$\hat{\eta}^2_G$	90% CI	F	df^{GG}	$df^{\mathrm{GG}}_{\mathrm{res}}$	p
StimCategory	.012	[.000, .064]	0.61	2	67	.544
Hierarchy	.098	[.015, .223]	33.61	1	67	$< .001$
Session	.027	[.000, .120]	22.33	1	67	$< .001$
StimCategory × Hierarchy	.007	[.000, .047]	1.13	2	67	.329
StimCategory × Session	.001	[.000, .000]	0.36	2	67	.698
Hierarchy × Session	.000	[.000, .000]	0.03	1	67	.865
StimCategory × Hierarchy × Session	.002	[.000, .000]	0.97	2	67	.384

Table 13: Experiment 3: Descriptive Statistics for Delayed testing performance by Remote and Recent conditions The table presents the median, mean, and standard deviation (SD) for performance during delayed testing, grouped by Remote and Recent conditions. Data is further broken down by Pair values.

		Hierarchy					
		Remote			**Recent**		
Session	Pair	Median	Mean	SD	Median	Mean	SD
Delayed testing	AB	0.88	0.75	0.30	1.00	0.86	0.24
Delayed testing	BC	0.50	0.52	0.39	0.88	0.67	0.36
Delayed testing	CD	0.75	0.65	0.34	0.88	0.73	0.32
Delayed testing	DE	0.62	0.57	0.38	0.75	0.68	0.32
Delayed testing	EF	1.00	0.79	0.32	1.00	0.83	0.28
Delayed testing	BD	0.62	0.56	0.36	0.50	0.50	0.38
Delayed testing	CE	0.50	0.51	0.36	0.50	0.52	0.37
Delayed testing	BE	0.62	0.60	0.33	0.56	0.55	0.36
Delayed testing	AF	1.00	0.85	0.27	1.00	0.85	0.27

Chapter 4

Sleep-related benefits to transitive inference are modulated by encoding strength and joint rank II.

4.1 Abstract

Reanalyzing secondary datasets, offers a valuable opportunity to uncover overlooked insights, validate prior findings, and ensure a comprehensive understanding of the topic at hand. Moreover, conducting small meta-analyses provides an integrative view, enabling the synthesis of disparate findings to discern broader trends and assess the robustness of effects across studies. In this light, our re-analysis of two published datasets on transitive inference highlighted contrasting effects in time-dependent symbolic distance consolidation and provided corroborative evidence for time-dependent inverse joint rank effects. A mini meta-analysis synthesized findings across 6 studies, showing that overall inference performance was significantly above chance and showing a symbolic distance effect at delayed test, albeit with small effect size. Moderator analyses indicated descriptive evidence for increased encoding strength amplifying sleep benefits. These results underscore the significance of encoding strength and joint rank in predicting offline generalization and provoke queries about the necessity of sleep for basic inference emergence. Pertinent methodological considerations are addressed, including the potential advantages of nonlinear modeling approaches in capturing intricate temporal consolidation effects. Overall, this chapter augments

our comprehension of the intricacies in gleaning latent relational structures from experiences over time.

4.2 Experiment BB2022: Reanalysis of Berens and Bird (2022)

4.2.1 Purpose

In the paper titled "Hippocampal and medial prefrontal cortices encode structural task representations following progressive and interleaved training schedules" the authors investigate the influence of different training schedules on generalization in a transitive inference task (Berens & Bird, 2022b). The study aims to understand whether generalization is influenced by progressive training or randomly interleaved training, as well as the effect of overnight consolidation. Participants in the study learned pairwise discriminations from two transitive hierarchies on consecutive days. They were then tested using fMRI to assess their inference performance. The results showed that inference performance was consistently better following progressive training and for pairs that were further apart in the transitive hierarchy (SDE).

Importantly, for the focus of the present chapter, the interleaved condition involved an experimental design similar to what was used in Experiment 3 (see Experimental protocol), with notable differences being a) the size of the hierarchy (6 item vs 7 item hierarchy) and b) encoding strength (dynamic exit criterion involving 66% percent premise pair performance in two consecutive blocks vs fixed number of trials; $\approx$ 128 vs 360 trials) c) the Recent condition was immediate (as opposed to delayed by 3 hours like in Experiment 3).

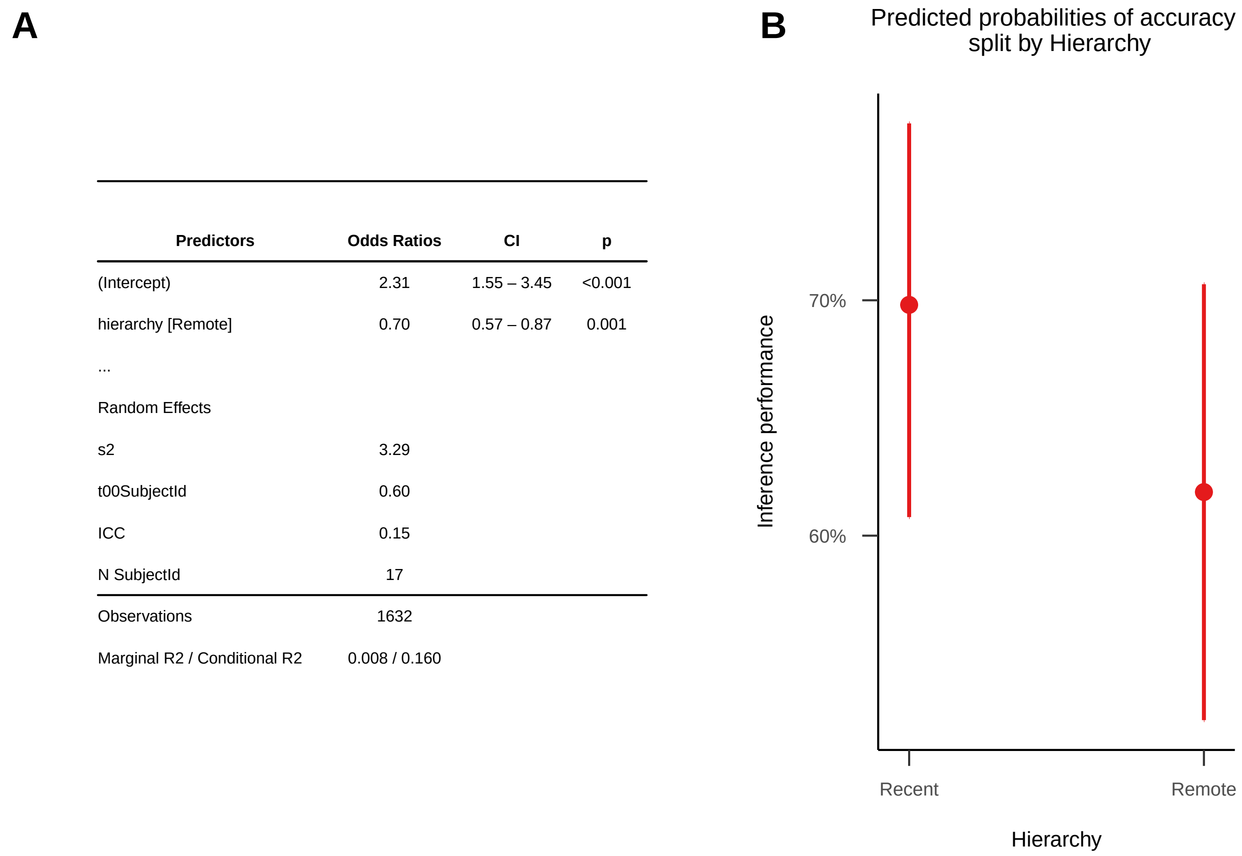
We wished to reanalyze this dataset as it relates to our question regarding time-dependent consolidation and the time-dependent joint rank effect to gain a better understanding of the boundary conditions involving its presentation.

4.2.2 Acknowledgement

Thank you to the authors for sharing the data and analysis code required to reproduce each figure and hypothesis test by uploading it to the Open Science Framework at `https://osf.io/tvk43/`.

4.2.3 Results

Berens & Bird (2022b), when analyzing the interleaved condition, found a main effect of Hierarchy, whereby participants showed greater inference performance in the Recent hierarchy as compared to Remote (see Fig 1 and Fig. 1A). This diverges from the previously discussed time-dependent generalization benefits, suggesting a potential training effect. It is plausible that participants, upon second exposure during the Recent condition, more rapidly acquire and stabilize rank-ordered representations.



Predictors	Odds Ratios	CI	p
(Intercept)	2.31	1.55 – 3.45	<0.001
hierarchy [Remote]	0.70	0.57 – 0.87	0.001
...			
Random Effects			
s2	3.29		
t00SubjectId	0.60		
ICC	0.15		
N SubjectId	17		
Observations	1632		
Marginal R2 / Conditional R2	0.008 / 0.160		

Figure 1: Experiment BB2022: Main effect of Hierarchy A) Results of the mixed-effects logistic regression model examining the effects of Hierarchy on inference accuracy. B) Predicted probability of inference performance split by Hierarchy. Error bars correspond to 95% CI

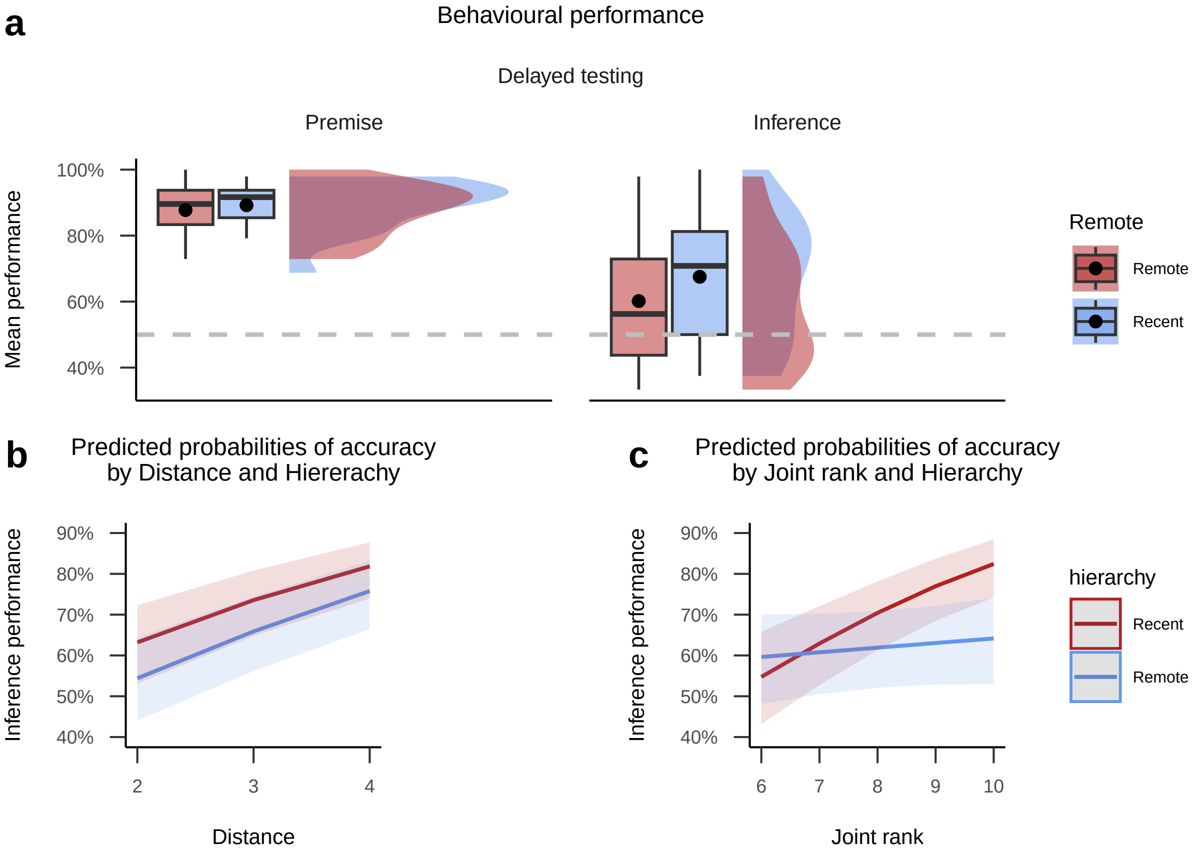


Figure 2: Experiment BB2022: Behavioural performance and factors predicting inference accuracy at delayed test. (a) Raincloud plot with mean and median performance across Remote and Recent hierarchy conditions. Dashed line represents chance performance at 50% and dot represents mean values (b) Predicted probabilities of accuracy by Distance and Hierarchy, with distance levels of 2 and 3. (c) Predicted probabilities of accuracy by Joint rank and Hierarchy, with joint rank levels ranging from 6 to 10. In panels (b) and (c), shadowed areas represent 95% confidence intervals.

Distance

Table 1: Experiment BB2022: LMM Hierarchy/Distance. Results of the mixed-effects logistic regression model examining the effects of Hierarchy (Remote/Recent) and Symbolic distance on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	0.65	0.37 – 1.15	0.141
Hierarchy [Remote]	0.69	0.56 – 0.86	0.001
Distance	1.62	1.39 – 1.88	<0.001
Random Effects			
σ2	3.29		
τ00SubjectId	0.63		
ICC	0.16		
N SubjectId	17		
Observations	1632		
Marginal R2 / Conditional R2	0.040 / 0.193		

Note: ^{a}CI = Confidence Interval. All p-values are two-tailed.

We fitted a logistic mixed model (estimated using ML and BOBYQA optimizer) to predict accuracy with hierarchy and distance (formula: accuracy ~ hierarchy + distance). The model included SubjectId as random effect (formula: ~1 | SubjectId). The model's total explanatory power is moderate (conditional R2 = 0.19) and the part related to the fixed effects alone (marginal R2) is of 0.04. The model's intercept, corresponding to hierarchy = Recent and distance = 0, is at -0.42 (95% CI [-0.99, 0.14], p = 0.141). Within this model:

- The effect of hierarchy [Remote] is statistically significant and negative (beta = -0.36, 95% CI [-0.58, -0.15], p < .001; Std. beta = -0.36, 95% CI [-0.58, -0.15])
- The effect of distance is statistically significant and positive (beta = 0.48, 95% CI [0.33, 0.63], p < .001; Std. beta = 0.36, 95% CI [0.25, 0.47])

Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using a Wald z-distribution approximation.. For detailed odds ratios see Table 1.

The model comparison analysis (Table 3) together with the main effect of Distance suggest that there is overall SDE, irrespective of Hierarchy.

Joint rank

Table 2: Experiment BB2022: LMM Hierarchy/Joint rank. Results of the mixed-effects logistic regression model examining the effects of Hierarchy (Remote/Recent) and Joint rank on inference accuracy.

Predictors	Odds Ratios	CI	p-value
(Intercept)	2.38	1.59 – 3.58	<0.001
Hierarchy [Remote]	0.68	0.55 – 0.85	0.001
Joint rank	1.40	1.24 – 1.59	<0.001
Hierarchy [Remote] XJoint rank	0.75	0.63 – 0.89	0.001
Random Effects			
σ2	3.29		
τ00SubjectId	0.62		
ICC	0.16		
N SubjectId	17		
Observations	1632		
Marginal R2 / Conditional R2	0.033 / 0.186		

Note: [a]CI = Confidence Interval. All p-values are two-tailed.

We fitted a logistic mixed model (estimated using ML and BOBYQA optimizer) to predict accuracy with hierarchy and jointrank (formula: accuracy ~ hierarchy * jointrank). The model included SubjectId as random effect (formula: ~1 | SubjectId). The model's total explanatory power is moderate (conditional R2 = 0.19) and the part related to the fixed effects alone (marginal R2) is of 0.03. The model's intercept, corresponding to hierarchy = Recent and jointrank = 0, is at 0.87 (95% CI [0.46, 1.28], p < .001). Within this model:

- The effect of hierarchy [Remote] is statistically significant and negative (beta = -0.38, 95% CI [-0.60, -0.17], p < .001; Std. beta = -0.38, 95% CI [-0.60, -0.17])
- The effect of jointrank is statistically significant and positive (beta = 0.34, 95% CI [0.21, 0.46], p < .001; Std. beta = 0.44, 95% CI [0.28, 0.60])
- The effect of hierarchy [Remote] × jointrank is statistically significant and negative (beta = -0.29, 95% CI [-0.46, -0.12], p < .001; Std. beta = -0.38, 95% CI [-0.59, -0.16])

Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using a Wald z-distribution approximation.. For detailed odds ratios see Table 2.

In order to examine the interaction, we computed the coefficients of simple effects for BD (JR:6, $\mathrm{OR}_{\mathrm{Recent/Remote}} = 0.82$, SE = 1.22, $p = 0.32$) and CE trials (JR:8, $\mathrm{OR}_{\mathrm{Recent/Remote}} = 1.10$, SE = 1.14, $p = 0.50$). From this post-hoc analysis, we can conclude that for BD trials, the odds of success in the Recent condition compared to the Remote condition were estimated to be 0.82 times lower, although this difference was not statistically significant (p = 0.32).

The model comparison analysis (Table 4) together with the interaction effect of Hierarchy and Joint rank suggest that there is an overall time-dependent JRE, but it appears to be primarily driven by a bias toward recalling low dominant items correctly in the Recent condition.

4.2.4 Discussion

In summary, the findings from the study conducted by Berens & Bird (2022b) reveal an inverse time-dependent consolidation effect. Specifically, they observe a stronger inference performance in the Recent condition compared to the Remote condition, as depicted in Figure 1. Notably, this pattern contrasts with the results reported in Experiment 3, where the opposite pattern was observed, as illustrated in Figure 9 and Table 9. These contrasting findings underscore the complex nature of memory consolidation and suggest that the temporal dynamics of memory retrieval can vary across different experimental conditions.

The observed difference in the consolidation effect between the Recent and Remote conditions could be attributed to several factors. Firstly, one potential factor is the size of the hierarchy employed in the two conditions (6 vs 7 item). The variation in hierarchy size could influence the encoding and retrieval processes, potentially contributing to the differential inference performance observed (Newbury & Monaghan, 2019a). Secondly, encoding strength could also play a role in the observed differences. Experiment 3 exit criterion for encoding strength in the both conditions involved achieving a 66% percent performance on premise pairs in two consecutive blocks, which aimed to ensure that participants could form strong representations for the Remote hierarchy even though they were already more familiar with task demands, due to their Session 1 exposure. On the other hand, in the BB2022 study participants had fix number of learning trials that aimed to train participants to ceiling, which combined with familiarity (or explicit awareness post-remote learning) could result in higher inference performance at Session 2. The disparity in encoding strength requirements between the two experiments could impact the sleep-dependent consolidation of memory representations and subsequently affect inference performance. While our earlier combined findings suggested stronger sleep-benefit as a function of encoding strength in transitive inference, it could be that instead of a linear effect, there is more of an inverted U-shaped effect, with diminishing generalization benefits as one approaches ceiling levels for premise pair performance.

Lastly, the temporal delay between encoding and retrieval might be a contributing factor. In BB2022, retrieval occurred shortly after encoding of the Recent hierarchy, whereas in Experiment 3, there was a delay of three hours between encoding and retrieval. The temporal dynamics of memory consolidation and retrieval processes could influence the accessibility and stability of memory traces, potentially leading to different patterns of inference performance due to altered proactive/retroactive interference effects (Moyano et al., 2022).

In regards to the time-dependent joint rank effect (JRE), we find evidence of a Hierarchy X Joint rank interaction in the same direction as reported in Experiment 3, however it appears to be more driven by JRE of the Recent condition (Fig. 2) as opposed to the "Inverse JRE" of the Remote condition (Fig. 9). The JRE of the Recent condition is in line to that reported in wake-only studies when participants reach high levels of premise pair performance in partial feedback design (no feedback for inference trials, see Ciranka et al., 2022). While this similarity, together with the added possibility of participants being more aware of task demands in the BB2022 Recent condition compared to Experiment 3 Recent condition would argue in favor encoding strength being the strongest contributing factor for authors observe a Remote<Recent inference performance at delayed test, this also cannot fully account for the different findings between the studies. The present findings contribute valuable insights that enrich our understanding of previous research (Experiment 1, 2, and 3) which consistently supported the notion that pre-sleep encoding strength serves as a positive predictor of time and sleep-dependent inference benefit, as illustrated in Figure 12D. However, the current study (BB2022) deviates from these expectations by demonstrating that within-subject design of this paradigm can overshadow time-dependent consolidation when participants reach a ceiling performance on both hierarchies. These novel observations highlight the complex interplay between encoding strength, time-dependent consolidation, and ceiling effects, further refining our knowledge in this field of inquiry.

4.2.5 Supplemental materials

Table 3: Experiment BB2022: Table of model comparisons. Investigating the Effects of Hierarchy and Distance.

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m0.d	3	2001.81	2018.01	-997.91	1995.81	NA	NA	NA
m1.d	4	1962.96	1984.55	-977.48	1954.96	40.85	1	<0.001
m2.d	5	1963.25	1990.24	-976.63	1953.25	1.71	1	0.19

Table 4: Experiment BB2022: Table of model comparisons. Investigating the Effects of Hierarchy and Joint rank

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m0.j	3	2001.81	2018.01	-997.91	1995.81	NA	NA	NA
m1.j	4	1984.84	2006.43	-988.42	1976.84	18.97	1	<0.001
m2.j	5	1975.41	2002.40	-982.71	1965.41	11.43	1	<0.001

4.3 Experiment MP2021: Reanalysis of Matorina and Poppenk (2021)

4.3.1 Purpose

The paper titled “Memory decay distinguishes subtypes of gist” by Matorina & Poppenk (2021b) investigates the role of memory decay in distinguishing different subtypes of gist memory. The study aims to understand how different types of information are retained and forgotten over time. The experimental sessions consisted of an initial study phase followed by a Pre-Sleep test, sleep at home while wearing the Sleep Profiler, and then a Post-Sleep test the next morning. A final memory test was conducted on Day 7. The study included four memory tasks: Deese, Roediger and McDermott (DRM), Transitive Inference (TI), Associative Memory (AM), and Visual Statistical Learning (VSL). Each task had specific procedures and timing details, which are described in the study.

We sought to re-evaluate the Transitive inference dataset for several compelling reasons. Firstly, in the context of time and sleep-dependent transitive inference studies, this particular study boasts a robust sample size, placing it well above the average in terms of statistical power. Secondly, the incorporation of a 7-day follow-up within the experimental design distinguishes it from most other investigations in the field, as such extended follow-ups are infrequently observed in the existing literature.

4.3.2 Acknowledgement

Thank you to authors for sharing their transitive inference data with me following our email correspondence.

4.3.3 Results

Matorina & Poppenk (2021b), when analyzing their TI condition, found no main effect of Session ($\hat{\beta} = 0.02$, 95% CI $[-0.01, 0.05]$, $z = 1.30$, $p = .192$), as participants did not show significantly greater inference performance over time (Post-sleep, Day 7) (see Fig. 1A). This diverges from the previously discussed time-dependent generalization benefits.

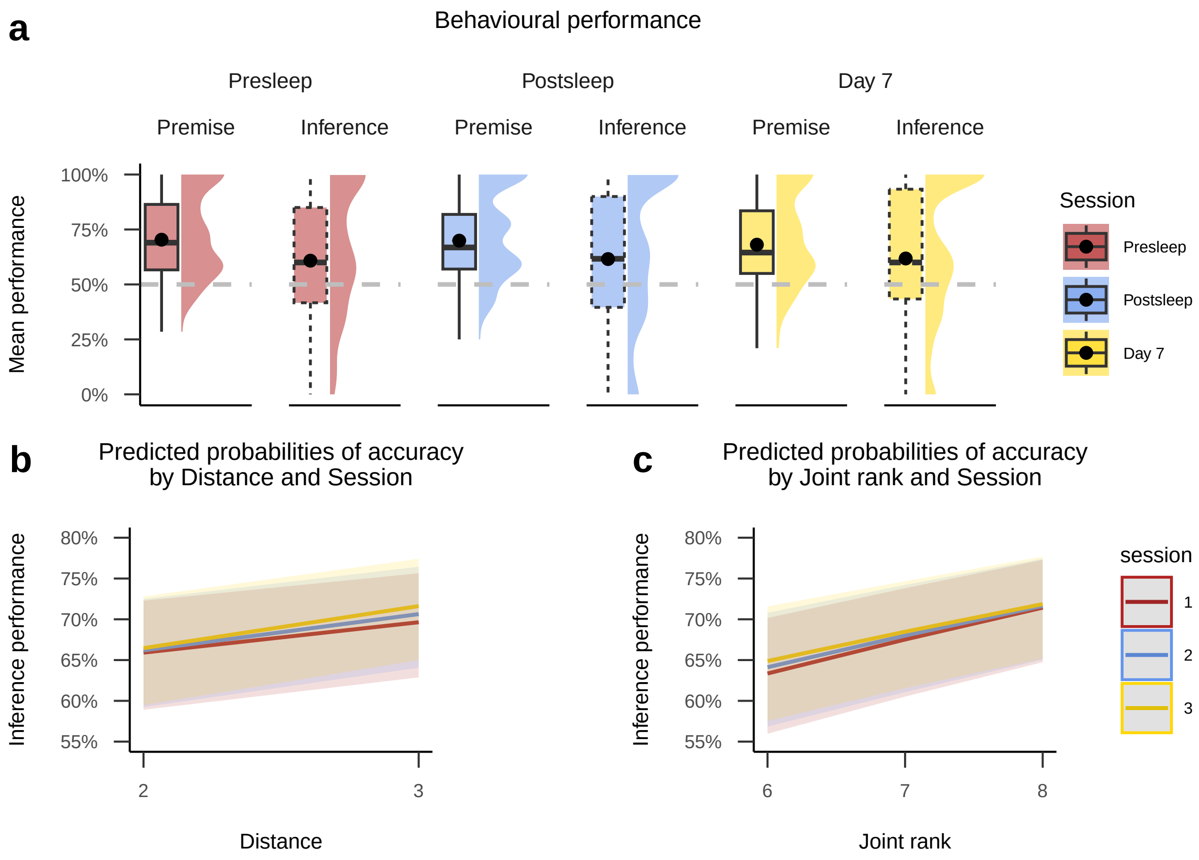
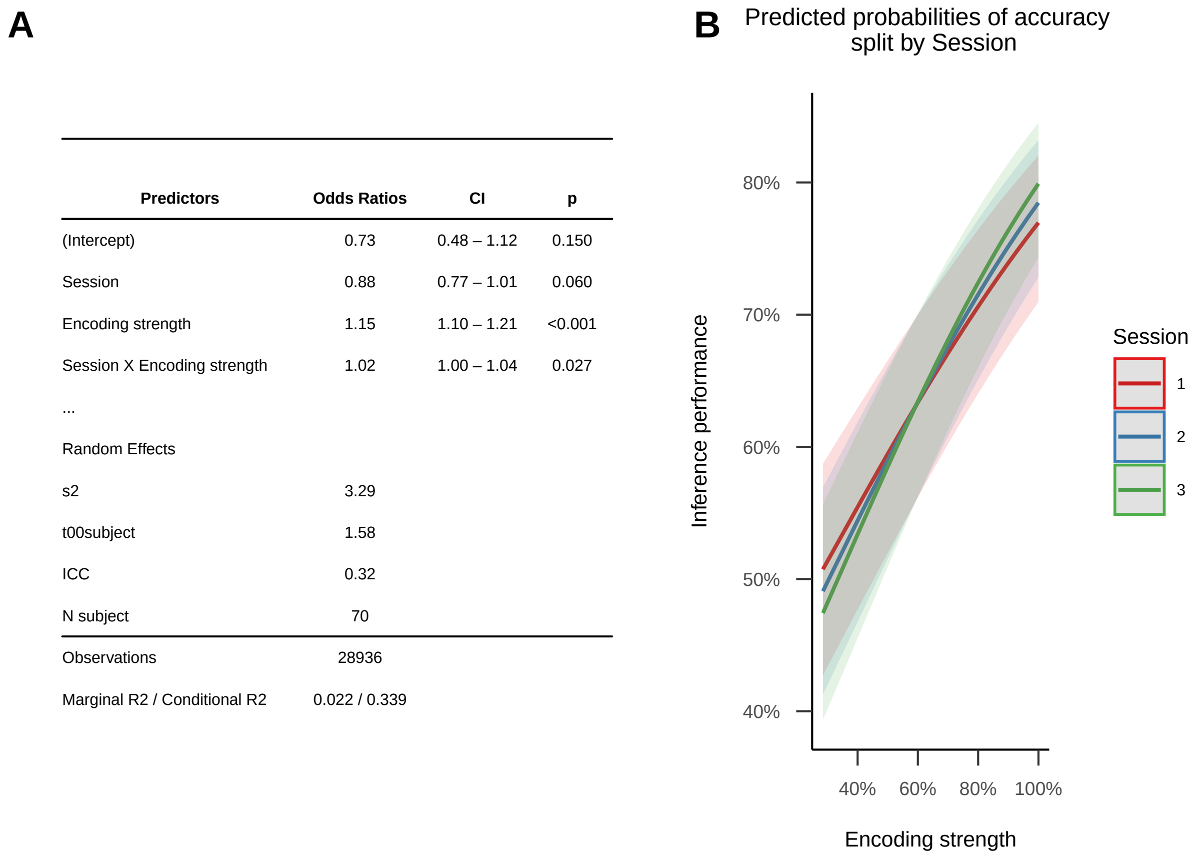


Figure 3: Experiment MP2021: Behavioural performance and factors predicting inference accuracy at delayed test. (a) Raincloud plot with mean and median performance across Pre-sleep/Post-sleep and Day7 conditions. Dashed line represents chance performance at 50% and dot represents mean values (b) Predicted probabilities of accuracy by Distance and Hierarchy, with distance levels of 2 and 3 (rescaled). (c) Predicted probabilities of accuracy by Joint rank and Hierarchy, with joint rank levels ranging from 6 to 8 (rescaled). In panels (b) and (c), shadowed areas represent 95% confidence intervals.

Encoding strength



A

Predictors	Odds Ratios	CI	p
(Intercept)	0.73	0.48 – 1.12	0.150
Session	0.88	0.77 – 1.01	0.060
Encoding strength	1.15	1.10 – 1.21	<0.001
Session X Encoding strength	1.02	1.00 – 1.04	0.027
...			
Random Effects			
s2	3.29		
t00subject	1.58		
ICC	0.32		
N subject	70		
Observations	28936		
Marginal R2 / Conditional R2	0.022 / 0.339		

Figure 4: Experiment MJ2021: Interaction effect Encoding strength X Session A) Results of the mixed-effects logistic regression model examining the effects of encoding strength X Session on inference accuracy. B) Predicted probability of inference performance split by Session.

We examined the interaction of Encoding strength and Session. We fitted a logistic mixed model (estimated using ML and BOBYQA optimizer) to predict correct with Session and Encoding_strength (formula: correct ~ Session * Encoding_strength). The model included subject as random effect (formula: ~1 | subject). The model's total explanatory power is substantial (conditional R2 = 0.34) and the part related to the fixed effects alone (marginal R2) is of 0.02. The model's intercept, corresponding to Session = 0 and Encoding_strength = 0, is at -0.31 (95% CI [-0.74, 0.11], p = 0.150). Within this model:

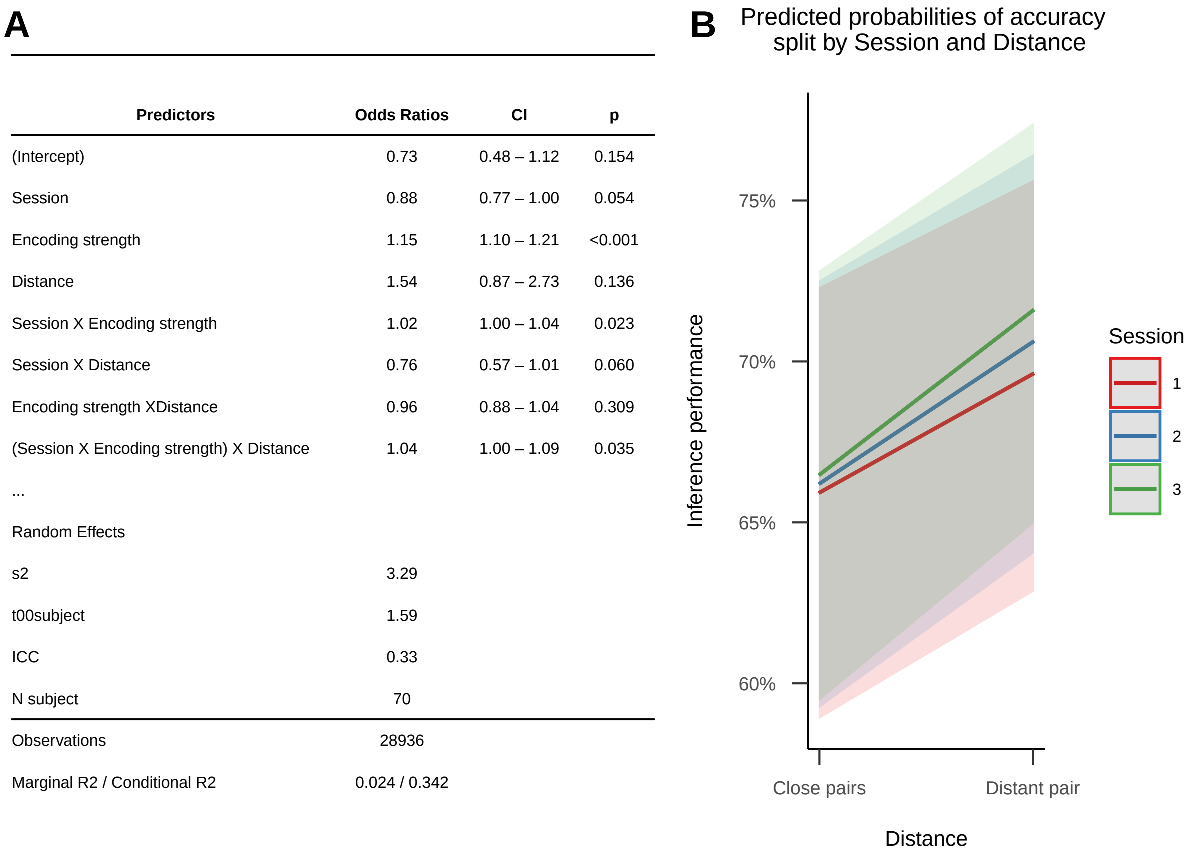
- The effect of Session is statistically non-significant and negative (beta = -0.13, 95% CI [-0.26, 5.27e-03], p = 0.060; Std. beta = 0.02, 95% CI [-7.22e-03, 0.05])
- The effect of Encoding strength is statistically significant and positive (beta = 0.14, 95% CI [0.10, 0.19], p < .001; Std. beta = 0.33, 95% CI [0.28, 0.37])

- The effect of Session × Encoding strength is statistically significant and positive (beta = 0.02, 95% CI [2.39e-03, 0.04], p = 0.027; Std. beta = 0.03, 95% CI [3.55e-03, 0.06])

Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using a Wald z-distribution approximation. For detailed odds ratios see Figure 4A.

To probe the Encoding strength X Session interaction, simple effects coefficients were computed at 60% ($OR_{\text{Session3/Session1}} = 0.80$, SE $= 1.13$, $p = 0.06$) and 80% ($OR_{\text{Session3/Session1}} = 0.80$, SE $= 1.13$, $p = 0.07$) and 100% ($OR_{\text{Session3/Session1}} = 0.81$, SE $= 1.12$, $p = 0.07$) values of Encoding strength, looking at changes between Pre-sleep and Day 7. The results are indicative of increased time-dependent benefit with stronger encoding.

Distance



Predictors	Odds Ratios	CI	p
(Intercept)	0.73	0.48 – 1.12	0.154
Session	0.88	0.77 – 1.00	0.054
Encoding strength	1.15	1.10 – 1.21	<0.001
Distance	1.54	0.87 – 2.73	0.136
Session X Encoding strength	1.02	1.00 – 1.04	0.023
Session X Distance	0.76	0.57 – 1.01	0.060
Encoding strength XDistance	0.96	0.88 – 1.04	0.309
(Session X Encoding strength) X Distance	1.04	1.00 – 1.09	0.035
...			
Random Effects			
s2	3.29		
t00subject	1.59		
ICC	0.33		
N subject	70		
Observations	28936		
Marginal R2 / Conditional R2	0.024 / 0.342		

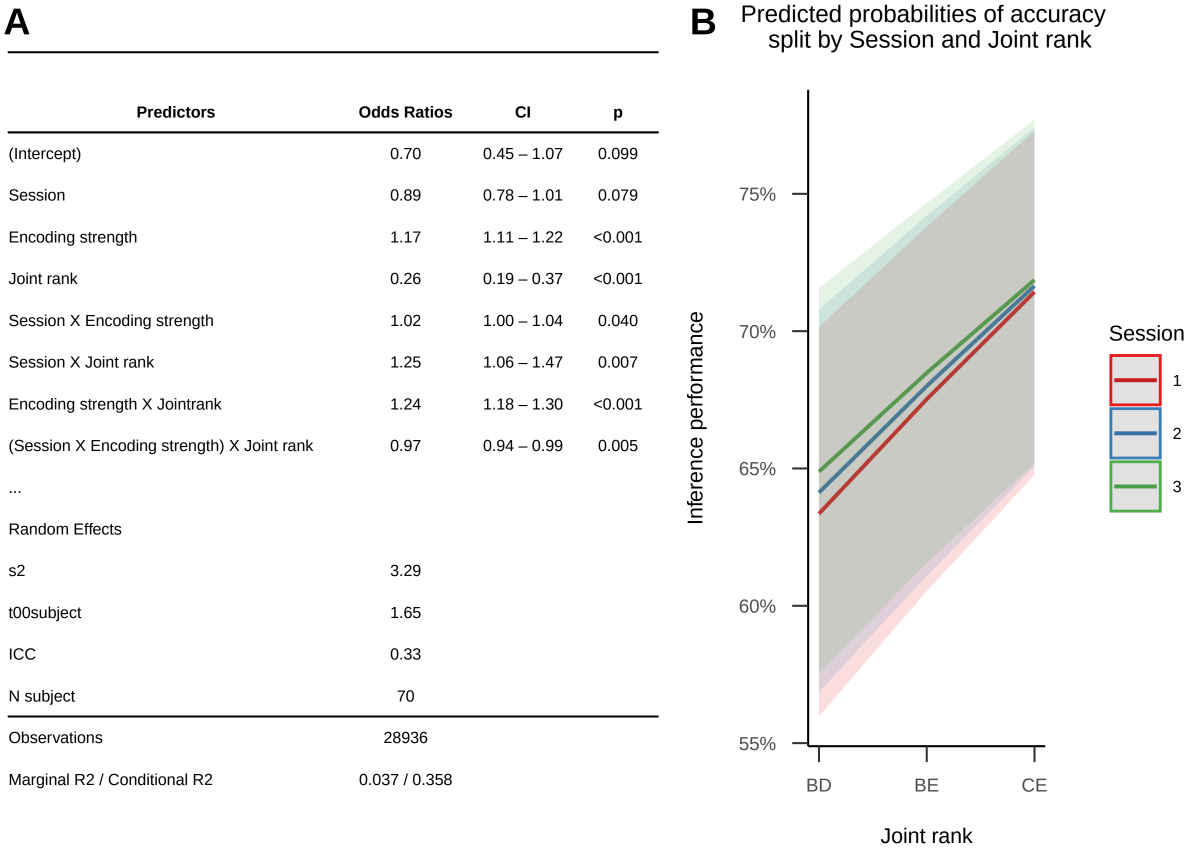
Figure 5: Experiment MJ2021: Interaction effect Distance X Session A) Results of the mixed-effects logistic regression model examining the effects of Distance X Session on inference accuracy. B) Predicted probability of inference performance split by Distance and Session. Shadowed areas represent 95% confidence intervals.

We examined the interaction of Distance, Session and Encoding strength. We fitted a logistic mixed model (estimated using ML and BOBYQA optimizer) to predict accuracy with Session, Encoding_strength and Distance (formula: accuracy ~ Session * Encoding_strength * Distance). The model included subject as random effect (formula: ~1 | subject). The model's total explanatory power is substantial (conditional R2 = 0.34) and the part related to the fixed effects alone (marginal R2) is of 0.02. The model's intercept, corresponding to Session = 0, Encoding_strength = 0 and Distance = 0, is at -0.31 (95% CI [-0.74, 0.12], p = 0.154). Within this model:

- The effect of Session is statistically non-significant and negative (beta = -0.13, 95% CI [-0.26, 2.03e-03], p = 0.054; Std. beta = 0.02, 95% CI [-6.51e-03, 0.05])
- The effect of Encoding strength is statistically significant and positive (beta = 0.14, 95% CI [0.10, 0.19], p < .001; Std. beta = 0.33, 95% CI [0.28, 0.38])
- The effect of Distance is statistically non-significant and positive (beta = 0.43, 95% CI [-0.14, 1.01], p = 0.136; Std. beta = 0.09, 95% CI [0.07, 0.12])
- The effect of Session × Encoding strength is statistically significant and positive (beta = 0.02, 95% CI [2.98e-03, 0.04], p = 0.023; Std. beta = 0.03, 95% CI [4.43e-03, 0.06])
- The effect of Session × Distance is statistically non-significant and negative (beta = -0.27, 95% CI [-0.56, 0.01], p = 0.060; Std. beta = 0.01, 95% CI [-0.01, 0.04])
- The effect of Encoding strength × Distance is statistically non-significant and negative (beta = -0.04, 95% CI [-0.12, 0.04], p = 0.309; Std. beta = 0.03, 95% CI [4.12e-03, 0.06])
- The effect of (Session × Encoding strength) × Distance is statistically significant and positive (beta = 0.04, 95% CI [2.98e-03, 0.08], p = 0.035; Std. beta = 0.03, 95% CI [2.17e-03, 0.06])

Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using a Wald z-distribution approximation. To investigate the interaction between Distance and Session, we calculated the coefficients of simple effects at close ($\text{OR}_{\text{Session3/Session1}} = 1.03$, SE $= 1.04$, $p = 0.53$) and distant ($\text{OR}_{\text{Session3/Session1}} = 1.10$, SE $= 1.06$, $p = 0.09$) pair values of Distance, comparing performance at Session 3 and Session 1. The higher odds of accurate response for distant pairs suggest increased time-dependent benefit for accuracy on inference pairs of greater symbolic distance, although this was not significant. Relatedly, we found a three-way interaction between Encoding strength, Session and Distance suggestive of stronger time-dependent SDE as a function of encoding strength (Fig. S3). In summary, this analysis showed that the influence of symbolic distance over which participants had to make inference only descriptively differed between Session 1 and Session 3, broadly in line with a time-dependent SDE. For detailed odds ratios see Figure 5A.

Joint rank



Predictors	Odds Ratios	CI	p
(Intercept)	0.70	0.45 – 1.07	0.099
Session	0.89	0.78 – 1.01	0.079
Encoding strength	1.17	1.11 – 1.22	<0.001
Joint rank	0.26	0.19 – 0.37	<0.001
Session X Encoding strength	1.02	1.00 – 1.04	0.040
Session X Joint rank	1.25	1.06 – 1.47	0.007
Encoding strength X Jointrank	1.24	1.18 – 1.30	<0.001
(Session X Encoding strength) X Joint rank	0.97	0.94 – 0.99	0.005
...			
Random Effects			
s2	3.29		
t00subject	1.65		
ICC	0.33		
N subject	70		
Observations	28936		
Marginal R2 / Conditional R2	0.037 / 0.358		

Figure 6: Experiment MJ2021: Interaction effect Joint rank X Session A) Results of the mixed-effects logistic regression model examining the effects of Joint rank X Session on inference accuracy. B) Predicted probability of inference performance split by Joint rank and Session. Shadowed areas represent 95% confidence intervals.

We examined the interaction of Joint rank, Session and Encoding strength. We fitted a logistic mixed model (estimated using ML and BOBYQA optimizer) to predict accuracy with Session, Encoding_strength and Joint_rank (formula: accuracy ~ Session * Encoding_strength * Joint_rank). The model included subject as random effect (formula: ~1 | subject). The model's total explanatory power is substantial (conditional R2 = 0.36) and the part related to the fixed effects alone (marginal R2) is of 0.04. The model's intercept, corresponding to Session = 0, Encoding_strength = 0 and Joint_rank = 0, is at -0.36 (95% CI [-0.79, 0.07], p = 0.099). Within this model:

- The effect of Session is statistically non-significant and negative (beta = -0.12, 95% CI [-0.25, 0.01], p = 0.079; Std. beta = 0.02, 95% CI [-8.44e-03, 0.05])
- The effect of Encoding strength is statistically significant and positive (beta = 0.15, 95% CI [0.11, 0.20], p < .001; Std. beta = 0.34, 95% CI [0.29, 0.39])
- The effect of Joint rank is statistically significant and negative (beta = -1.33, 95% CI [-1.66, -1.00], p < .001; Std. beta = 0.14, 95% CI [0.12, 0.17])
- The effect of Session × Encoding strength is statistically significant and positive (beta = 0.02, 95% CI [9.37e-04, 0.04], p = 0.040; Std. beta = 0.03, 95% CI [1.37e-03, 0.06])
- The effect of Session × Joint rank is statistically significant and positive (beta = 0.22, 95% CI [0.06, 0.39], p = 0.007; Std. beta = -7.64e-03, 95% CI [-0.03, 0.02])
- The effect of Encoding strength × Joint rank is statistically significant and positive (beta = 0.22, 95% CI [0.17, 0.27], p < .001; Std. beta = 0.23, 95% CI [0.20, 0.26])
- The effect of (Session × Encoding strength) × Joint rank is statistically significant and negative (beta = -0.03, 95% CI [-0.06, -9.96e-03], p = 0.005; Std. beta = -0.04, 95% CI [-0.07, -0.01])

Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using a Wald z-distribution approximation.

To investigate the two-way interaction between Joint rank and Session, we calculated the coefficients of simple effects at low (BD) ($\text{OR}_{\text{Session3/Session1}}$ = 1.07, SE = 1.05, p = 0.19) and high (CE) ($\text{OR}_{\text{Session3/Session1}}$ = 1.02, SE = 1.05, p = 0.68) pair values of Joint rank, comparing performance at Session 3 and Session 1. The descriptively higher odds of accurate response for the dominant pair (i.e.: BD) suggest increased time-dependent benefit for accuracy on inference pairs of lower joint rank, however this was not significant. Relatedly, we found a three-way interaction between Encoding strength, Session and Joint rank suggestive of stronger time-dependent JRE as a function encoding strength (Fig. S4).

In summary, this analysis showed that the impact of joint rank over inference differed between Session 1 and Session 3, in line with an inverse time-dependent JRE. For detailed odds ratios see Figure 6A.

4.3.4 Discussion

In their study, Matorina & Poppenk (2021b) initially identified a subtle yet consistent improvement in inference performance over time. Our reanalysis, which focused on the interplay between Encoding strength and subsequent time-dependent consolidation, unveiled an interaction with Session. Specifically, the strength of participants' initial encoding of premise pairs appeared to amplify the benefits derived from sleep—and eventually, from the passage of time. Such findings are in line with research that examined the relationship between encoding strength and retention intervals, which either included or excluded sleep (Abel et al., 2023; Gordon B. Feld, Weis, & Born, 2016b; Newbury & Monaghan, 2019b; Petzka et al., 2021; Wernette & Fenn, 2023b). While we didn't identify a definitive interaction between Session and Distance (p=0.06), our post-hoc analysis yielded results that descriptively align with a time-dependent SDE, as illustrated in Fig. 5. We did find a three-way interaction between Encoding strength, Session and Distance that suggested that time-dependent SDE was larger when initial encoding was stronger (see Fig. S3). These underscore the potential value of employing longer, post-1-day follow-ups when studying transitive inference in longitudinal designs. Furthermore, the data did indicate a pronounced interaction between Session and Joint rank. This suggests that the time-dependent advantage was more pronounced for dominant (BD) than for non- dominant (CE) pairs, as illustrated in Fig. 6, in accordance with an inverse time-dependent JRE.

Upon reflection on all three experiments and the two secondary datasets, several key patterns emerge. First and foremost, there is a recurring interaction between Encoding strength and Retention interval across studies. This pattern underscores that as encoding strength increases, there's a concomitant boost in the benefit seen across retention intervals. Notably, the sole exception to this pattern was observed in the BB2022 dataset, where participants were trained to the point of ceiling performance and, additionally, we lacked access to encoding strength data to explore this relationship. In our earlier discussions, we posited that within-subject studies on transitive inference typically do not exhibit time-dependent symbolic distance effects. However, the findings from the MJ2021 dataset challenge this notion, indicating that such an effect can indeed manifest over more extended time intervals, particularly when initial learning is robust. Lastly, in assessing the time-dependent inverse joint rank effect, both secondary datasets exhibited congruent trends, albeit the observed effects were more subdued in magnitude.

4.3.5 Supplemental materials

Table 5: Experiment BB2022: Table of model comparisons. Investigating the Effects of Hierarchy and Distance.

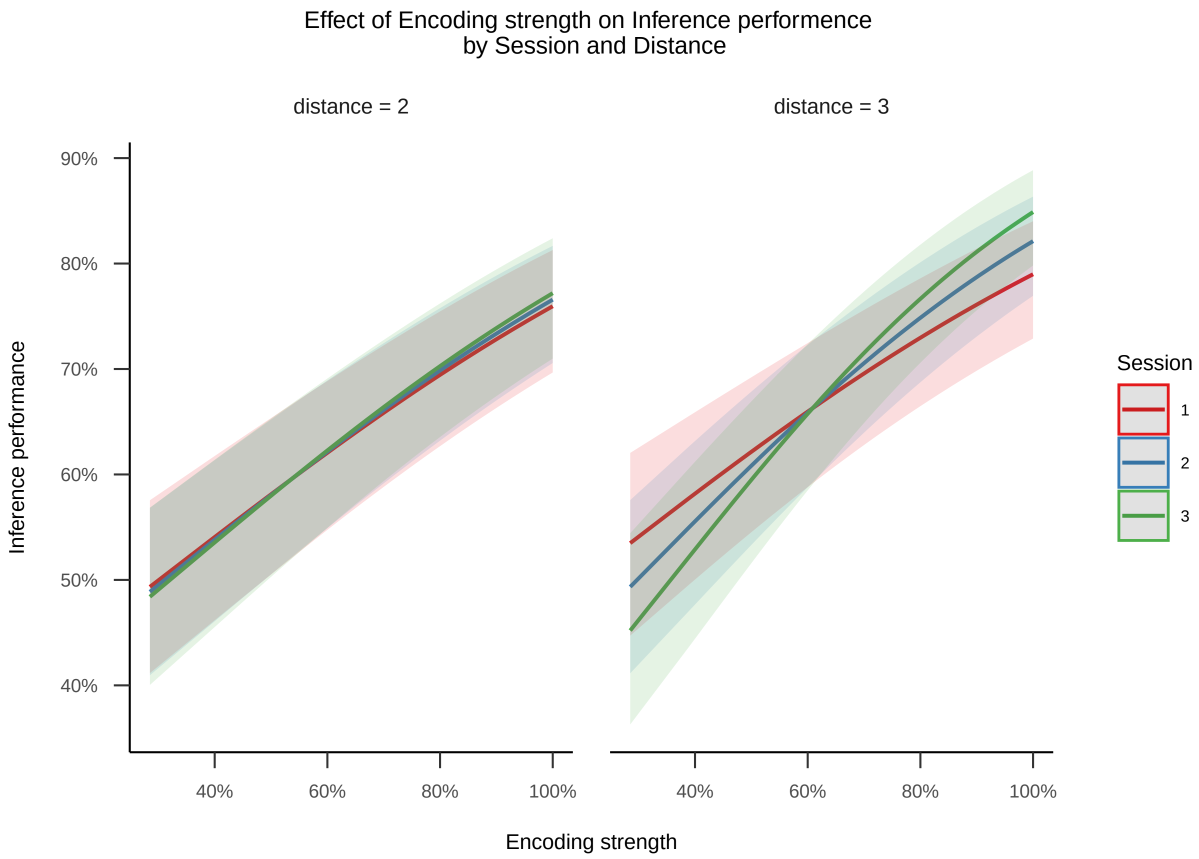
	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m0.e	3	33239.04	33263.86	-16616.52	33233.04	NA	NA	NA
m1.e	4	33045.67	33078.76	-16518.83	33037.67	195.38	1	<0.001
m2.e	5	33042.81	33084.18	-16516.41	33032.81	4.86	1	0.028

Table 6: Experiment BB2022: Table of model comparisons. Investigating the Effects of Hierarchy and Distance.

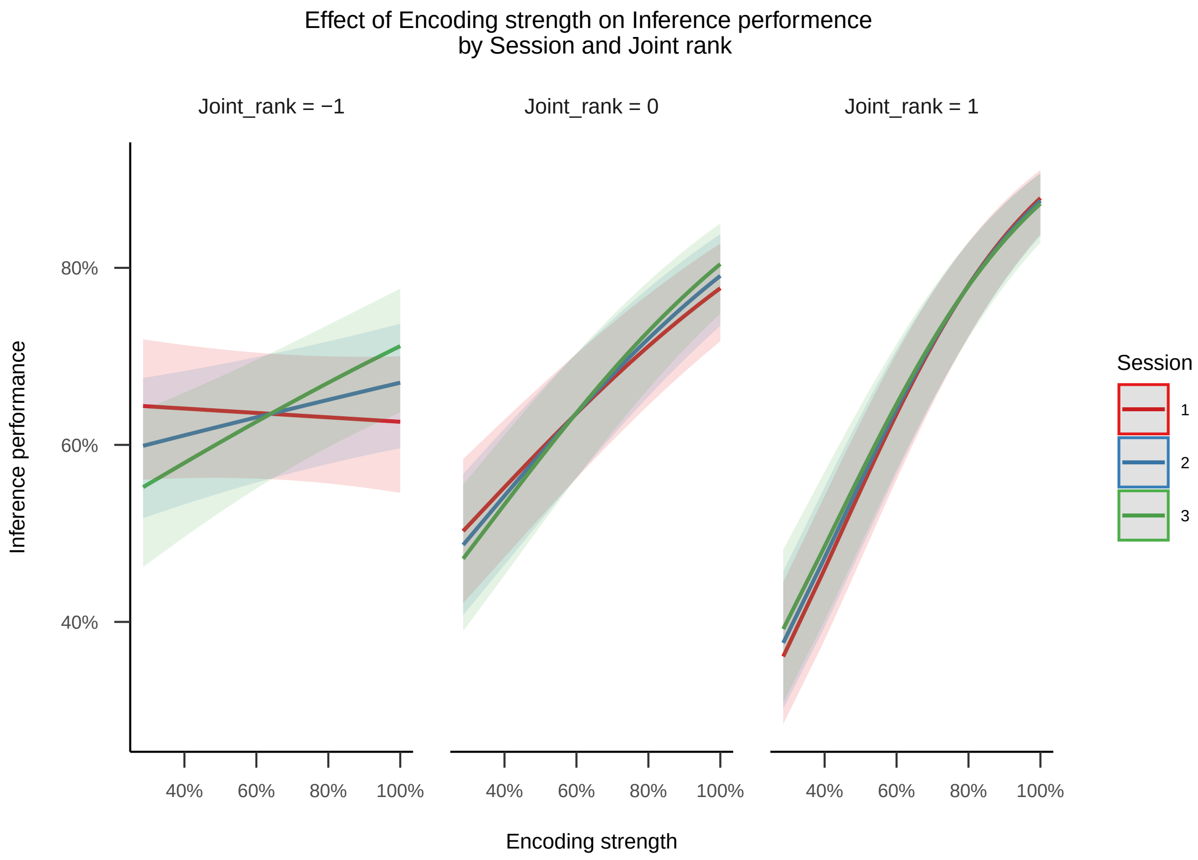
	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m0.d	5	33042.81	33084.18	-16516.41	33032.81	NA	NA	NA
m1.d	6	32999.74	33049.38	-16493.87	32987.74	45.07	1	<0.001
m2.d	7	33001.27	33059.18	-16493.64	32987.27	0.47	1	0.4937
m3.d	9	32996.02	33070.48	-16489.01	32978.02	9.25	2	0.0098

Table 7: Experiment BB2022: Table of model comparisons. Investigating the Effects of Hierarchy and Joint rank

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m0.j	5	33042.81	33084.18	-16516.41	33032.81	NA	NA	NA
m1.j	6	32979.33	33028.97	-16483.67	32967.33	65.48	1	<0.001
m2.j	7	32981.33	33039.24	-16483.67	32967.33	0.00	1	0.98
m3.j	9	32736.49	32810.94	-16359.24	32718.49	248.85	2	<0.001



Supplemental Figure S3: Experiment MP2021: Encoding strength X Session X Distance. Predicted probabilities of accuracy by Encoding strength, Session and Distance, with Distance levels ranging from 2 (close pairs) to 3 (distant pairs). Shadowed areas represent 95% confidence intervals.



Supplemental Figure S4: Experiment MP2021: Encoding strength X Session X Joint rank. Predicted probabilities of accuracy by Encoding strength, Session and Joint rank, with mean centered Joint rank levels ranging from -1 (BD) to 1 (CE). Shadowed areas represent 95% confidence intervals.

4.4 Mini meta-analysis

4.4.1 Purpose

Meta-analysis is a powerful statistical technique that allows researchers to synthesize the results of multiple studies on a particular topic. It involves pooling data from individual studies and analyzing them collectively to obtain an overall estimate of the effect size. While meta-analysis can be a valuable tool in experimental psychology, there are both benefits and limitations to conducting a meta-analysis on a small number of studies. One of the main benefits of conducting a meta-analysis on a small number of studies is that it allows for the inclusion of all available evidence on a particular topic. This can be particularly useful when there are only a few studies conducted on a specific research question. By combining the results of these studies, researchers can obtain a more comprehensive and reliable estimate of the effect size (Goh, Hall, & Rosenthal, 2016). Another benefit of conducting a meta-analysis on a small number of studies is that it can help identify potential moderators or factors that may influence the relationship between variables. Meta-analyses often examine the effects of various moderators, such as sample characteristics or study design, to determine if they influence the overall effect size (Newbury, Crowley, Rastle, & Tamminen, 2021)

However, there are also limitations to conducting a meta-analysis on a small number of studies. One limitation is the potential lack of generalizability. When there are only a few studies available, the results may not be representative of the broader population or may be influenced by specific characteristics of the included studies. This can limit the external validity of the findings. Additionally, the small number of studies may limit the statistical power of the meta-analysis, making it difficult to detect small effect sizes or to conduct subgroup analyses (Jackson & Turner, 2017). Furthermore, the quality of the included studies can also impact the validity of the meta-analysis. If the studies included in the meta-analysis are of low quality or have methodological limitations, the overall estimate of the effect size may be biased or unreliable. It is important to carefully assess the quality of the included studies and consider potential sources of bias (Feeley, 2020).

Goh et al. (2016) discuss the importance and process of conducting a meta-analysis on one's own studies within a manuscript. The authors acknowledge that this practice is still relatively rare and slow to be adopted, possibly due to two misconceptions. First, researchers may believe that conducting a meta-analysis on a small number of studies is not legitimate. Second, researchers may perceive meta-analysis as too complex to be done without expert knowledge or guidance. The article introduces both fixed effects and random effects approaches to meta-analysis. Fixed effects are typically used when the author believes there is one true population effect size, which is more likely when studies are methodologically similar. Random effect modelling is used when the assumption is that the studies are estimating different, yet related, population effect sizes. This heterogeneity in effect sizes can occur due to variations in study design, sample characteristics, or other context-specific factors. Random effects models account for this variability by incorporating both within-study sampling error and between-study variance into the estimation process. This method allows for generalization beyond the set of studies included in the meta-analysis to a broader population of studies (Fleiss, 1993). The authors also mention the Q test, which assesses heterogeneity in the meta-analysis; significant result indicates more heterogeneity than expected from normal sampling variation.

In the present mini meta-analysis we attempt to provide an overview of time and sleep-depedent consolidation studies looking at transitive inference. We began by examining studies that cited Ellenbogen et al.'s seminal work from 2007, which served as an initial point of reference. This step ensured we captured research that directly built on or responded to this influential study. We then scrutinized the methods employed by these studies, specifically seeking those that used random interleaved training during the training of premise pairs. Importantly, we targeted studies that did not explicitly cue participants about any underlying rank-order. This was crucial in eliminating potential bias and ensuring that the effects identified were genuine. Additionally we also set an exclusion criteria. For instance, studies where participants had previously undergone testing in a given condition were excluded. This was done to control for potential testing confounds that might inadvertently cue participants about the underlying rank-order. This eliminated the Recent condition from both Experiment 3 and BB2022 and Session 3 from MJ2021. Finally, we focused on studies that provided a mean and standard deviation (SD) or standard error (SE) for a retention interval that was shorter than two days for inference trials and premise pair performance. This led to the exclusion of two studies (H. A. Golkashani, Ghorbani, Leong, Ong, & Chee, 2023b; H. Golkashani et al., 2022b). This criterion was set to ensure uniformity in the data collection process and minimize heterogeneity that could arise due to longer retention intervals.

4.4.2 Results

In this study, a random-effects meta-analysis was conducted to synthesize the findings of 15 conditions of 6 individual studies using the metafor package in R (Viechtbauer, 2010). The analysis employed the restricted maximum likelihood (REML) method to estimate the between-study variance (τ^2). The meta-analysis results showed a significant overall effect (estimate = 0.46, SE = 0.08, z = 6.07, p <0.001, 95% CI [0.31, 0.6]).

The test for heterogeneity was not significant (Q(df = 14) = 14.83, p = 0.39), indicating homogeneity among the studies. The amount of total heterogeneity (τ^2) was estimated to be <0.001 (SE = 0.03). The I^2 value, representing the proportion of total variability due to heterogeneity, was 0%, suggesting little variability among the studies. The H^2 value, representing the proportion of total variability due to sampling variability, was 1.00.

Overall, the findings indicate a significant effect in favor of the inference performance being higher than chance, with a small effect size. However, it is important to interpret these results with caution due to the limited number of studies included in the analysis (k = 15).

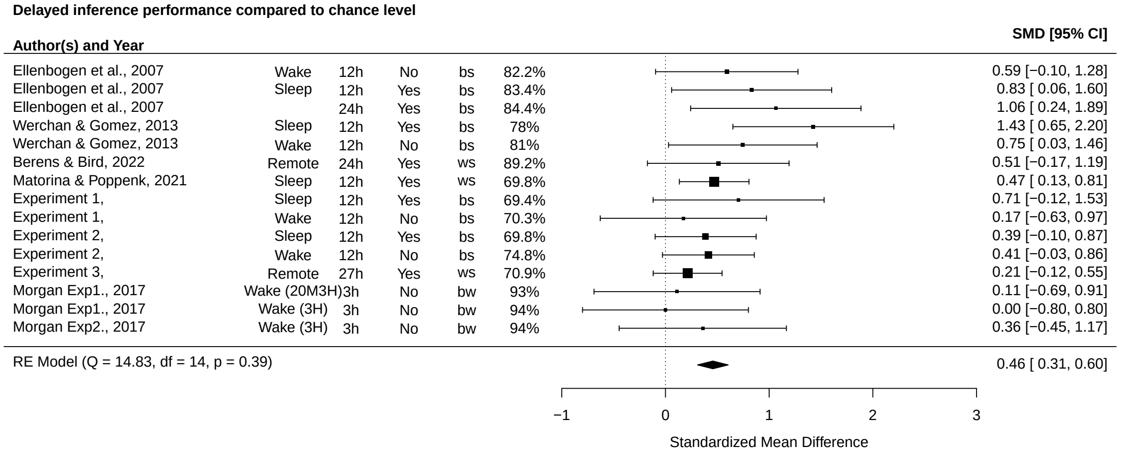


Figure 7: Forest plot displaying the results of the meta-analysis for delayed inference performance compared to chance level. Each study is represented by a square marker, and the size of the marker corresponds to the study's weight in the analysis. The horizontal lines represent the 95% confidence intervals (CI) for each study. The overall effect size estimate is indicated by the diamond at the bottom, with its width representing the CI. The forest plot is based on a random-effects model. The label on the x-axis shows the condition name, time delay (in hours), inclusion of sleep, experimental design, and encoding strength as a percentage. The label on the y-axis shows the author and year of each study.

Table 8: Results of the meta-regression examining the effect of sleep inclusion. Standard errors, z-values, p-values, and 95% confidence intervals are reported for each coefficient.

	Estimate	Std. Error	z value	Pr(>|z|)	CI Lower	CI Upper
Intercept	0.37	0.14	2.68	0.01	0.10	0.65
Included_sleep[Yes]	0.16	0.18	0.89	0.37	-0.19	0.50

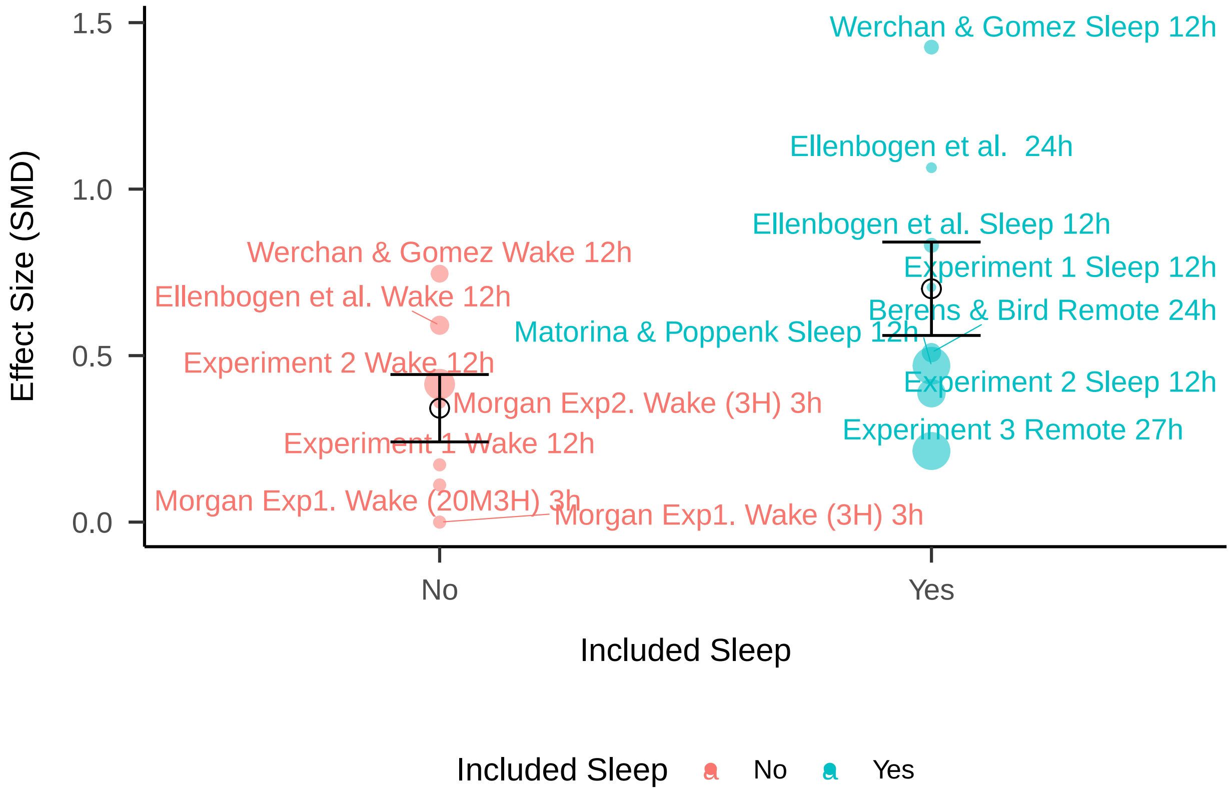


Figure 8: Scatterplot of effect sizes (standardized mean differences) as a function of sleep inclusion. Each point represents a single study and its size corresponds to the study's weight in the meta-analysis. The color of the points indicates whether the study included sleep (blue) or not (red). The black circle represent the mean effect size for each group. Error bars show SE.

To examine the effect of whether the retention interval including sleep affected Inference performance and delayed test, we conducted a meta-regression analysis using the *rma* function from the *metafor* package in R. Our predictor variable was 'Included sleep' (coding for whether the retention interval included sleep or not). We find no significant effect retention interval (for details see Table 8), but note the descriptively we find increased inference performance when the interval includes sleep.

Encoding strength

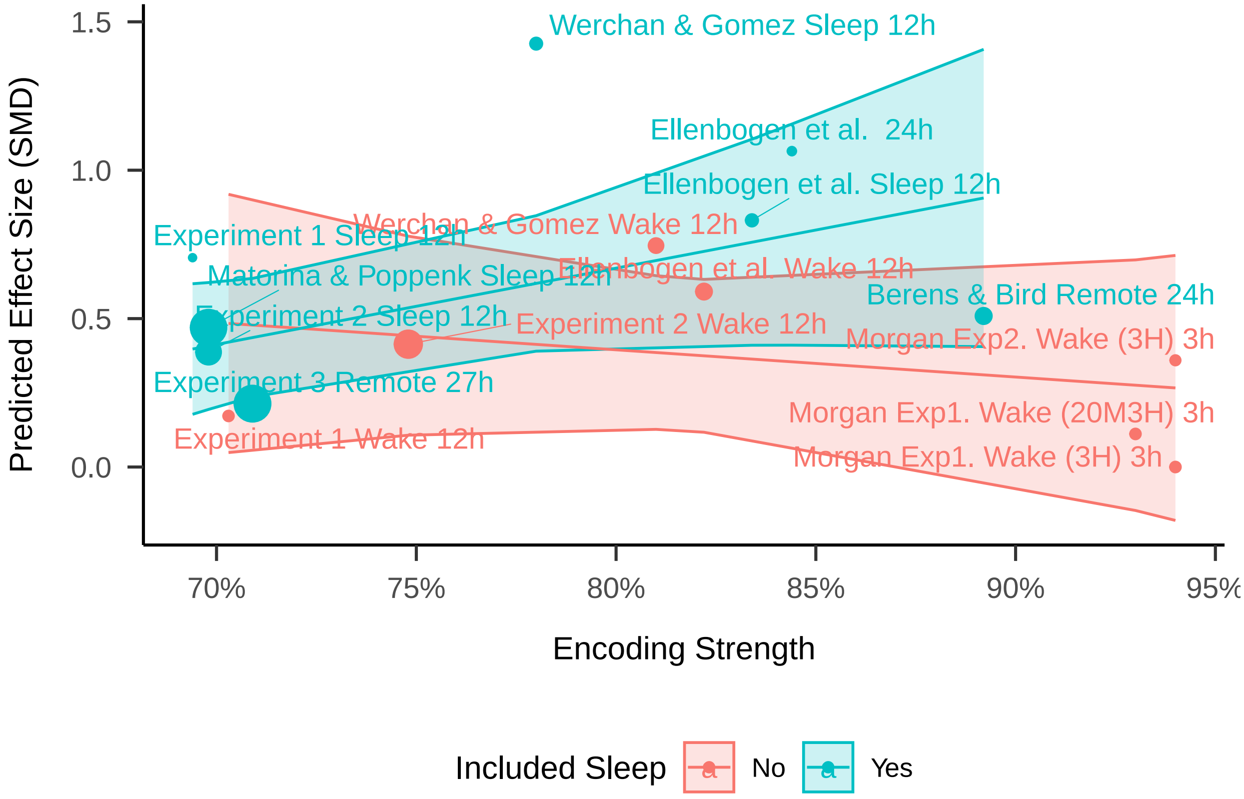


Figure 9: Predicted Effect Sizes (Standardized Mean Differences) as a Function of Encoding Strength and Sleep Inclusion. Effect sizes (SMD) for delayed inference performance as a function of encoding strength and sleep inclusion. The x-axis represents encoding strength, while the y-axis represents the effect size. The data points are color-coded based on the inclusion of sleep. The figure includes ribbons indicating the 95% confidence intervals. The size of the data points corresponds to the weights from the meta-regression analysis. The figure also includes text labels for each data point, showing the author, condition name, and delay duration in hours.

To examine the relationships among our variables of interest, we conducted a meta-regression analysis using the *rma* function from the *metafor* package in R. Our predictor variables were Encoding strength and 'Included sleep' (coding for whether the retention interval included sleep or not), and we were particularly interested in their interaction.

The log-likelihood (logLik) of the model was -3.46 and the deviance was 6.91. The respective AIC, BIC values were 16.91, 18.9.

The estimated amount of residual heterogeneity (τ^2) was 0.0017, with a standard error (SE) of 0.03. The proportion of unaccounted variability in effect sizes attributable to heterogeneity rather than sampling error (I^2) was 1.83%. The ratio of unaccounted variability to sampling variability (H^2) was 1.02, and no heterogeneity

Table 9: Results of the meta-regression examining the interaction between encoding strength and sleep inclusion. Standard errors, z-values, p-values, and 95% confidence intervals are reported for each coefficient.

	Estimate	Std. Error	z value	Pr(>|z|)	CI Lower	CI Upper
Intercept	1.13	1.27	0.89	0.37	-1.36	3.61
Encoding strength	-0.01	0.02	-0.60	0.55	-0.04	0.02
Included_sleep[Yes]	-2.51	1.69	-1.49	0.14	-5.82	0.79
Encoding strength : included_sleep[Yes]	0.03	0.02	1.62	0.11	-0.01	0.08

was accounted for (Rˆ2 = 0).

The test for residual heterogeneity was not significant, QE(df = 11) = 10.93, p = 0.449, indicating that there was no evidence of unaccounted variability in the effect sizes. Similarly, the test of moderators was not significant, QM(df = 3) = 3.85, p = 0.278, suggesting that the included moderator variables did not significantly predict effect size.

We find no significant Encoding strength-'Included sleep' interaction (for details see Table 9), but note that descriptively it seems that increased encoding strength has an increased effect on sleep-related consolidation.

Distance

The examine whether there was a main effect of symbolic distance we again employed the restricted maximum likelihood (REML) method to estimate the between-study variance (τ^2). The meta-analysis results showed a significant overall effect (estimate = -0.16, SE = 0.08, z = -2.06, p=0.039, 95% CI [-0.3, -0.01]).

The test for heterogeneity was not significant (Q(df = 13) = 8.23, p = 0.83), indicating homogeneity among the studies. The amount of total heterogeneity (τ^2) was estimated to be <0.001 (SE = 0.03). The I^2 value, representing the proportion of total variability due to heterogeneity, was 0%, suggesting little variability among the studies. The H^2 value, representing the proportion of total variability due to sampling variability, was 1.00.

Overall, the findings indicate a significant effect showing participants perform better on Distant pairs compared to Close pairs, with a small effect size. However, it is important to interpret these results with caution due to the limited number of studies included in the analysis (k = 14).

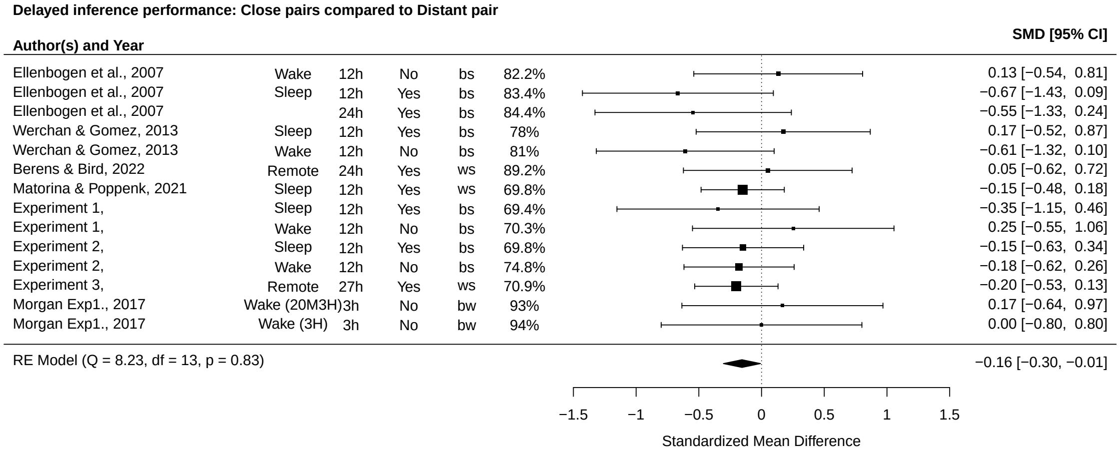


Figure 10: Forest plot displaying the results of the meta-analysis for delayed inference performance Close pairs compared to Distant pair. Each study is represented by a square marker, and the size of the marker corresponds to the study's weight in the analysis. The horizontal lines represent the 95% confidence intervals (CI) for each study. The overall effect size estimate (Close pair - Distant pair performance) is indicated by the diamond at the bottom, with its width representing the CI. The forest plot is based on a random-effects model. The label on the x-axis shows the condition name, time delay (in hours), inclusion of sleep, experimental design, and encoding strength as a percentage. The label on the y-axis shows the author and year of each study.

Table 10: Results of the meta-regression examining the interaction between distance and sleep inclusion. Standard errors, z-values, p-values, and 95% confidence intervals are reported for each coefficient.

	Estimate	Std. Error	z value	Pr(>|z|)	CI Lower	CI Upper
Intercept	-0.09	0.14	-0.63	0.53	-0.35	0.18
Included_sleep[Yes]	-0.10	0.16	-0.62	0.53	-0.42	0.22

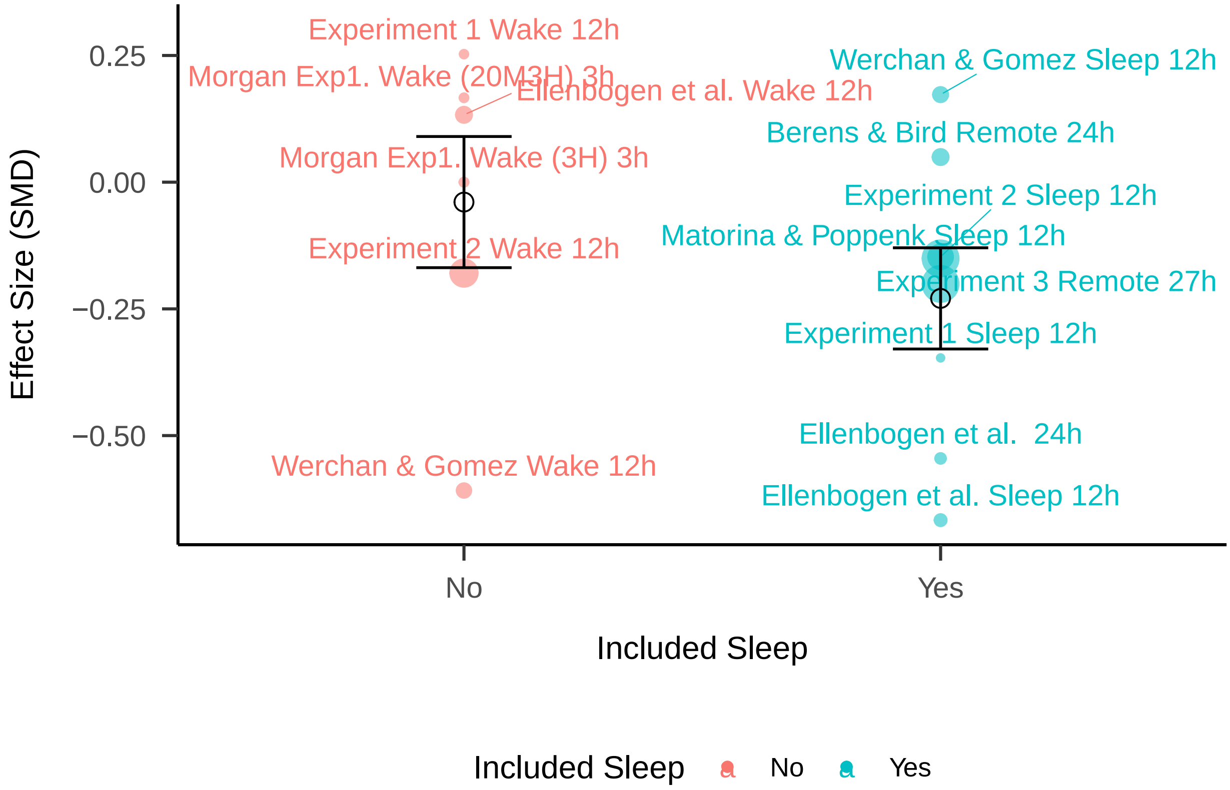


Figure 11: Scatterplot of effect sizes (standardized mean differences) as a function of sleep inclusion. Each point represents a single study and its size corresponds to the study's weight in the meta-analysis. The color of the points indicates whether the study included sleep (blue) or not (red). The black circle represent the mean effect size for each group. Error bars represent SE.

To examine the effect of whether the retention interval including sleep affected Inference performance on delayed test, we conducted a meta-regression analysis, our predictor variable being again 'Included sleep'. We find no significant effect retention interval (for details see Table 10), but not that descriptively the difference between close pairs and distant pairs is larger for 'Included sleep'[Yes].

Joint rank

The examine whether there was a main effect of joint rank we again employed the restricted maximum likelihood (REML) method to estimate the between-study variance (τ^2). The meta-analysis results showed no overall effect (estimate = 0.09, SE = 0.08, z = 1.18, p=0.24, 95% CI [-0.06, 0.25]).

The test for heterogeneity was not significant (Q(df = 11) = 4.7, p = 0.95), indicating homogeneity among the studies. The amount of total heterogeneity (τ^2) was estimated to be <0.001 (SE = 0.03). The I^2 value, representing the proportion of total variability due to heterogeneity, was 0%, suggesting little variability among the studies. The H^2 value, representing the proportion of total variability due to sampling variability, was 1.00.

Overall, the findings do not indicate a main effect of joint rank, whereby participants performance differs between BD and CE trials.

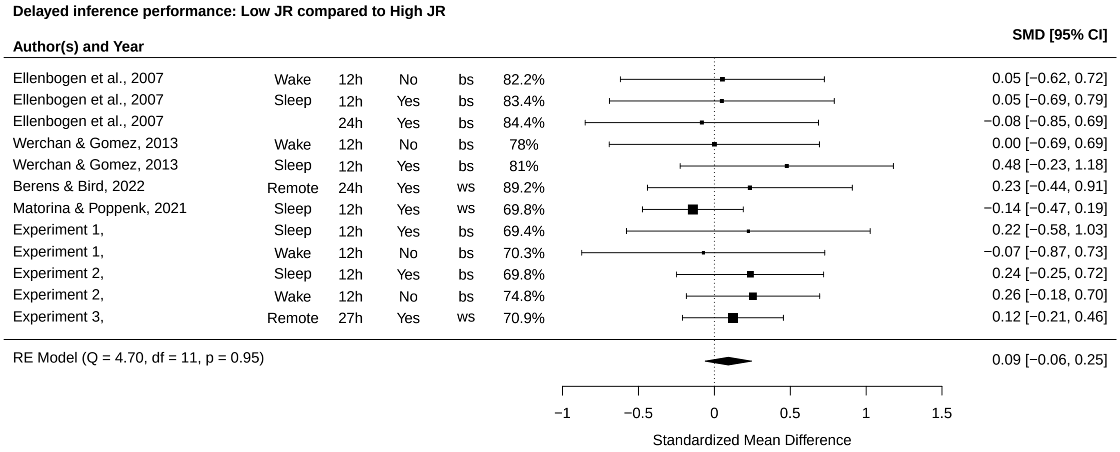


Figure 12: Forest plot displaying the results of the meta-analysis for delayed inference performance Close pairs compared to Distant pair. Each study is represented by a square marker, and the size of the marker corresponds to the study's weight in the analysis. The horizontal lines represent the 95% confidence intervals (CI) for each study. The overall effect size estimate is indicated by the diamond at the bottom, with its width representing the CI. The forest plot is based on a random-effects model. The label on the x-axis shows the condition name, time delay (in hours), inclusion of sleep, experimental design, and encoding strength as a percentage. The label on the y-axis shows the author and year of each study.

Table 11: Results of the meta-regression examining the interaction between joint rank and sleep inclusion. Standard errors, z-values, p-values, and 95% confidence intervals are reported for each coefficient.

	Estimate	Std. Error	z value	Pr(>|z|)	CI Lower	CI Upper
Intercept	0.12	0.15	0.78	0.44	-0.18	0.42
Included_sleep[Yes]	-0.04	0.18	-0.21	0.84	-0.39	0.31

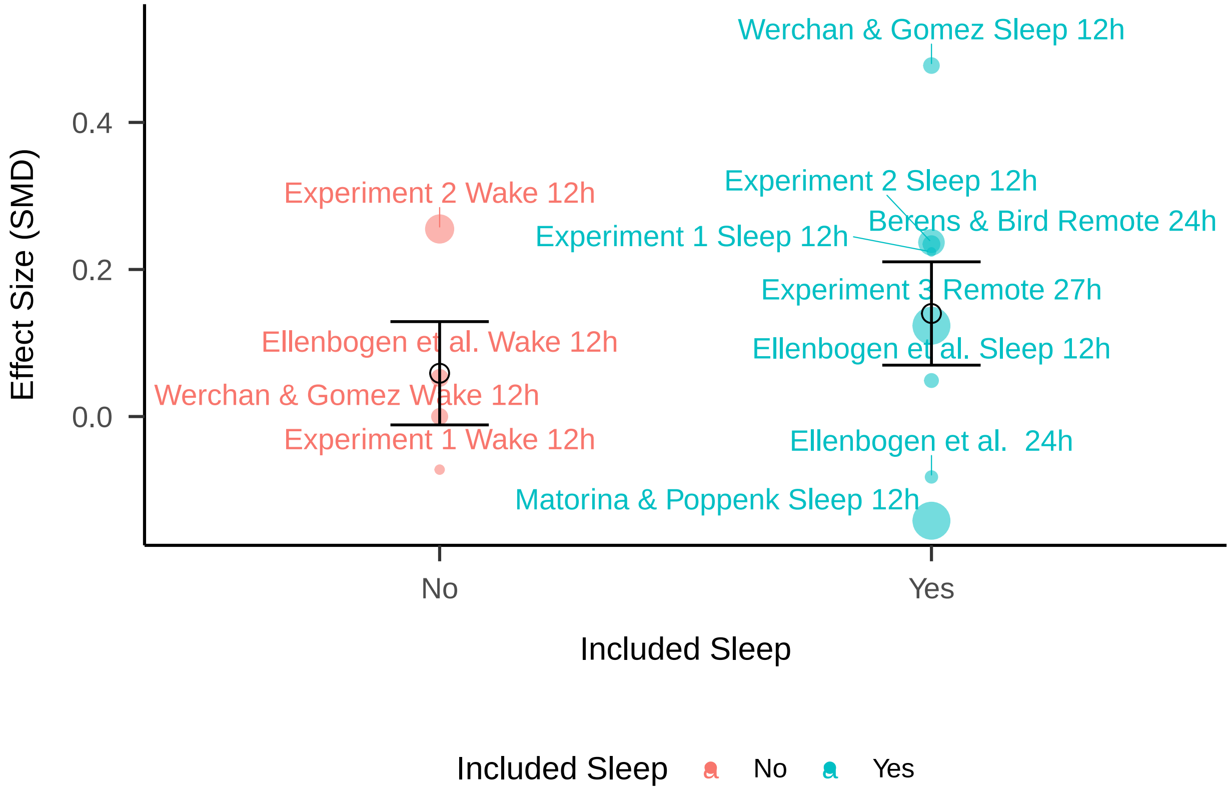


Figure 13: Scatterplot of effect sizes (standardized mean differences) as a function of sleep inclusion. Each point represents a single study and its size corresponds to the study's weight in the meta-analysis. The color of the points indicates whether the study included sleep (blue) or not (red). The black circle represent the mean effect size for each group. Error bars represent SE.

To examine the effect of whether the retention interval including sleep affected Inference performance on delayed test, we conducted a meta-regression analysis, our predictor variable being again 'Included sleep'. We find no significant effect retention interval (for details see Table 11).

4.4.3 Discussion

A mini meta-analysis was conducted on 15 conditions from 6 individual studies. The main analysis showed a significant overall effect, indicating inference performance was significantly higher than chance, albeit with a small effect size. It also showed a high level of homogeneity across the studies, with little variability. Next, two factors, sleep inclusion and encoding strength, were explored as predictors. The inclusion of sleep during the retention interval did not significantly affect inference performance and delayed test. While no significant effect was found for the interaction between encoding strength and sleep inclusion, descriptively however, the results suggested that increased encoding strength might boost sleep-related consolidation in TI.

When examining the effect of symbolic distance, the analysis showed participants performed better on Distant pairs compared to Close pairs, again with a small effect size. The difference between close pairs and distant pairs seemed larger when sleep was included, although this was not statistically significant. Finally, there was no significant main effect found for joint rank, suggesting participants' performance did not differ between BD and CE trials. The inclusion of sleep during the retention interval also did not significantly affect inference performance and delayed test in this case.

In contrast to fixed-effect meta-analysis, where it is assumed that there is one true effect size which underlies all the studies in the analysis, and that all differences in observed effects are due to sampling error, we conducted a random-effects meta-analysis, which allows the true effect size to vary from study to study. This model provides a more conservative (i.e., wider) confidence interval when there is between-study variability. Given additional AM-PM studies of TI a fixed-effect meta-analysis could provide a more precise estimate. However, in the present context, we considered a random-effects meta-analysis provides a more realistic approach, allowing for inherent variability across studies, thereby yielding a more comprehensive and generalized view of the effect size given differences in stimulus category, timing and average encoding strength.

It is important to note that while the results sleep-related results did not show significant overall effects, small meta-analysis such as this cannot provide strong confirmatory or disconfirmatory evidence of the kind of moderators we were primarily interested in, but can serve as a good primer on discussing different experimental designs assessing similar questions. Mini meta-analysis have been used to integrate experiments with differing results within a study (Hughes & Thomas, 2022; Plancher, Massol, Dorel, & Chainay, 2018), and to obtain more precise size estimates of the effects observed across experiments (Gregory, Winstone, Ridout, & Nash, 2019).

Figure 13 shows little difference between low and high JR, we can also conclude that there were only 11 conditions in this analysis compared to the 15 conditions for the encoding strength and distance moderator analyses, there is a lack of wake retention-interval conditions and lastly, the study by Matorina and Poppenk (2021) appears to contrast the small positive differences of the other studies. While its difficult to speculate what could be cause for this - assuming the inverse joint rank effect is a true population effect - its important to note, that when analyzed via using mixed logistic-regression (as opposed to looking at the mean difference via SMD) we also find that the largest time-dependent improved was predicted in the model to be for the BD type inferences, in line with the prediction of the inverse JRE (see Figure. 6). Additionally, we highlight that we've included the 12h/Session 2 data from this study for comparability into the meta-analysis, but 7 day study shows an increased change in this regard.

4.5 Conclusion

In this dissertation, we conducted three studies to investigate the effects of time (Experiment 3) and sleep-dependent benefits (Experiments 1 and 2) on inference performance, as well as exploring moderating effects of encoding strength. Additionally, we re-analyzed two external datasets by Berens & Bird (2022b) and Matorina & Poppenk (2021b) to gain a broader understanding of the boundary conditions related to offline generalization and its long-term trajectory.

Across all studies, except for BB2022, we observed positive interactions between either time or sleep and encoding strength, indicating that strong representation during consolidation leads to inference-related generalization benefits. Interestingly, even when participants were trained to ceiling as in BB2022, we found inference performance broadly in line with a multiplicative effect of retention interval (included sleep or not) and encoding strength (Figure 9). There also appear to be signs of plateau for inference performance in line with an inverted U-shape type relationship between encoding strength and sleep-related inferential benefits (Kolibius, Born, & Feld, 2021b; Stickgold, 2009), suggesting diminishing returns in generalization performance along with a potential risk of explicit awareness emerging which can confound results and make interpreting results more challenging. While the positive sleep-related encoding strength effect could be further evidence of the importance of an online factor in later recall an alternative possibility is that encoding strength is serving as a measure of prior knowledge, awareness of task demands or working memory related individual differences that have been shown to impact subsequent memory consolidation (Fenn & Hambrick, 2010; Groch et al., 2017; Kumaran & Ludwig, 2013). Future research should endeavor to more distinctly unravel these contributing elements.

Furthermore, we examined two important features that provide insights into the mental models participants likely rely on during transitive inference. Firstly, we explored symbolic distance, which has been considered evidence of spatial or quasi-spatial models supporting inference (E. C. Hinton, Dymond, Hecker, & Evans, 2010a). While early studies and our small laboratory sample (Exp 1) of sleep-dependent SDE provided tentative evidence for "symbolic distance x retention interval" interaction (whereby sleep increases the SDE compared to wake), we failed to replicate this in a moderately well-powered web-based experimental design in Exp 2. and Exp 3. In Exp 2. unexpectedly, we observed high inference performance in the wake group of Exp 2 compared to Exp 1, emphasizing the importance of studying consolidation during wakefulness and assessing the criticality of sleep or longer delays in the emergence of inference when minimal training is employed. Future research should investigate this process unfolding over time by testing inference pairs of multiple hierarchies within-subject continuously at multiple time-points with an added focus on short interval following post-encoding. For example, having participants respond to inference pairs of only one of two hierarchies at immediate test (without feedback) can provide an estimate of pre-sleep inference testing effects on consolidation. The null-finding in Exp 3. regarding time-dependent SDE is in line with the findings of Berens & Bird (2022b) and could be due increased awareness of task demands inherent in within-subject design of statistical learning.

Secondly, we introduced the measure of joint rank as a relevant predictor of time and sleep-dependent consolidation. Joint rank assumes that the sum of item ranks (as opposed to the difference as in symbolic distance) moderates inference performance, suggesting the involvement of absolute rank-based ordering representations in addition to relative ranks (Kao, Jensen, Michaelcheck, Ferrera, & Terrace, 2019). Joint rank showed consistent retention interval dependent interaction effects across all reported studies (albeit with varying directions), except in Experiment 2 where only a main effect was observed. Findings from Experiment 1, 3 and Matorina & Poppenk (2021b) suggested a time and/or sleep-related component to the Joint Rank effect (JRE), wherein the more dominant inference trial (BD, low JR) benefited more compared to the less dominant trial (CE, high JR), which we refer to as an inverse JRE, to differentiate it from the JRE effect observed in wake-only studies (Ciranka et al., 2022). Future studies should investigate whether this effect extends to $\geq$ 7-item hierarchies, that provide a larger sample of inference pairs (with respect to both distance and joint rank) and combine a 'lower than ceiling' training criterion with the experimental design of Berens and Bird (2021) and Experiment 3.

The examination of joint rank as a predictor of inference performance has received relatively less attention compared to symbolic distance. This discrepancy can be partially attributed to the fact that the effect of joint rank has been typically reported within the context of symbolic distance, indicating that the predictive power of symbolic distance higher (Munoz et al., 2020b). Nonetheless, it is worth noting that these studies have predominantly focused on extensive premise pair training protocols and investigated inference performance exclusively in wake conditions. Consequently, there remains an unanswered inquiry regarding the emergence of inference performance after minimal training and the potential temporal dynamics of symbolic distance and joint rank, including the possibility of their convergence or even a switch in prominence over time. It is important to note that although we analyzed Distance and Joint Rank separately, they are not necessarily separate mechanisms. A comprehensive investigation should involve simultaneously modeling both factors when examining time and sleep-dependent consolidation, as suggested by Munoz et al. (2020b).

In all of the reported studies, a consistent effect of joint rank condition was observed to varying extents, indicating a notable relationship between joint rank and inference performance. However, in Experiment 2, only a main effect of joint rank was found, suggesting that the sleep-related component of the Joint Rank Effect (JRE) may be implicated. Specifically, it was observed that the more dominant inference trial (~lower joint rank) of the BD trial exhibited greater benefits compared to the less dominant inference trial (higher joint rank) of the CE trial. In Experiment 3's discussion section, we engaged in speculative deliberation concerning the mechanistic underpinnings of this phenomenon. We contend that these results hold implications for comprehending the consolidation mechanisms that drive the temporal progression of statistical learning concerning latent structures. These mechanisms could potentially broaden the application of the transitive inference paradigm—which necessitates that participants intuitively discern the order of a series of items without external spatial or temporal indicators—to more complex domains such as two-dimensional grids and graphs, which may encompass deterministic as well as probabilistic associations (G. B. Feld, Bernard, Rawson, & Spiers, 2022b; Kern et al., 2023). Studies probing sleep or reactivation-dependent benefits for more complex latent structure can better assess boundary conditions. Furthermore, it is important to note that the separate analysis of symbolic distance and joint rank does not imply their independent nature. Rather, a comprehensive examination of time and sleep-dependent consolidation necessitates the simultaneous modeling of both factors. By considering these interconnected mechanisms, a more thorough understanding of the dynamics underlying inference performance can be achieved (Jensen, Alkan, Muñoz, Ferrera, & Terrace, 2017a; Jensen, Munoz, Meaney, Terrace, & Ferrera, 2021; Munoz et al., 2020c).

A meta-analysis was conducted on a subset of studies that examined the original findings reported by Ellenbogen et al. (2007). These studies utilized a similar interleaved training and delayed test design. The meta-analysis focused on extracting mean and standard errors from conditions of interest, allowing for a quantitative overview of the literature. It is important to note that the sample size within the meta-analysis was limited, particularly for the moderator analyses. Unsurprisingly, the findings revealed no significant interactions between retention interval and the predictors of interest, namely encoding strength, symbolic distance, and joint rank. Despite the small sample size, this mini meta-analysis provides valuable insights and effect size estimates that can inform power analysis for future studies in this area of research.

This chapter raises important points that contribute to the understanding of inference performance and its analysis in transitive inference tasks. Apart from underscoring the modulatory effects of encoding strength and joint rank on inference performance at delayed test, the findings from Experiment 2 reveal above chance performance in wake conditions, suggesting the possibility that sleep may not be essential for the emergence of this type of generalization (Talamini et al., 2022b). It raises the question of whether inferential capacity can develop prior to sleep or larger time intervals that include sleep, and whether observed group differences can be primarily attributed to interference differences between Wake/Recent and Sleep/Remote periods take place between encoding and retrieval. This assertion is supported by the contrasting findings of Matorina & Poppenk (2021b) and Morgan & Stickgold (2017) , which show above chance transitive inference performance at immediate tests. Future studies should consider measuring inference performance at immediate test intervals and explore potential interference effects. Methodological solutions, such as employing larger hierarchies and use of post-experimental awareness questionnaires, can be crucial in limiting and monitoring participants' explicit awareness of the underlying relational structure.

Secondly, traditional analysis of transitive inference tasks involves averaging trial-level performance for each participant and conducting ANOVA to compare mean performances. However, this approach assumes normality, which may not hold for inference behavior at delayed tests, particularly when data is split by moderators like symbolic distance. Additionally, treating inference performance as a continuous variable by calculating mean proportions assumes a linear relationship between predictors and outcomes. In contrast, modeling inference performance on the logit scale transforms it into log odds or logit values, assuming a more complex and nonlinear relationship between predictors and outcomes. The logit scale approach can be more sensitive in detecting nuanced changes in probabilities and providing accurate effect estimates, especially when data are sparse (Dixon, 2008). However, it is crucial to note that both linear and logistic regression assume a monotonic relationship between the expected value of inference performance and predictors. If the true relationship follows a U-shaped pattern, standard regression approaches may fail to capture this non-monotonic behavior unless additional predictors, such as squared terms, are included in the model. Future research should explore non-linear relationships between time, sleep-dependent consolidation, and inference performance to gain a more comprehensive understanding of their associations.

Thirdly, this chapter highlights the necessity of accounting for encoding strength, even in studies that use a learning criterion to control for the level of learning. Owing to individual differences in prior knowledge and cognitive abilities, a learning criterion cannot ensure minimal variance in memory strength. We demonstrate in 3 experiments that such variance significantly influences sleep-related improvements in generalization. As an alternative approach, sleep researchers might consider implementing a secondary learning criterion related to the retrieval-based training and testing components of their experiments.

Chapter 5

Time-dependent consolidation in transitive inference: an RSA approach

5.1 Abstract

Research has shown that time including a night of sleep is beneficial for the development of inferential performance. According to theories of memory consolidation, memories get reorganized over time and become supported by distributed cortical network. However, the effects of time on the organization of memories underlying transitive inference remains less clear. In a 2 day study, participants encoded two separate hierarchies of images differing in their opportunity for sleep-dependent consolidation. Adjacent pairs were initially studied either 27 or 3 hours before recall in a functional magnetic resonance imaging (fMRI) scan session. We used model-based representational similarity analysis to examine time-dependent consolidation effects on memory representations in the in the hippocampus, mOFC, PHC and IFG. Our results provide evidence for differences in rank-order representations encoded at separate times, with stronger distance-related pattern similarity for the remote hierarchy in our cortical ROIs. Furthermore we find that the extent of rank-order pattern similarity observed in the left HPC and right PPC correlated with time-dependent benefits in inference performance at delayed test. Together, these results demonstrate reorganization of relational memory traces over time.

5.2 Introduction

5.2.1 Latent structure learning

Throughout each day, we encounter a staggering volume of information, yet our memory systems adeptly encode and subsequently retrieve this data. Without a structured — i.e. organized across dimensions like space, time, semantic and perceived value — storage of information, this task would be immensely challenging, if not outright impossible. Viewing the use of structure to enhance retrieval efficiency as a general mnemonic activity can apply both to the underlying processes that systematize the memory system and to deliberate strategies we employ to aid recall of specific information (Bower, 1970). Consequently, the structure of memory has long captivated cognitive scientists studying memory, as they endeavor to understand how we arrange information and experiences to enable future retrieval (De Brigard & Sinnott-Armstrong, 2022).

Model-based theories suggest that this capability originates from structure learning, typically defined as the process of deducing the underlying model that generates observed data (Peer, Brunec, Newcombe, & Epstein, 2021a). While model-free methods are potent in explaining simple learning and decision-making behaviors, they can fall short in complex real-world scenarios involving intricate stimuli and actions (Decker, Otto, Daw, & Hartley, 2016). It's posited that both humans and animals utilize the inherent structure in real-world tasks to simplify the learning process, relying on inductive model-based priors (Gershman & Niv, 2010). By acknowledging the hidden variables that produce sensory inputs and reinforcements, the task of predicting future reinforcement can be enhanced through inference over these hidden variables.

Within the sphere of cognitive research, the concept of 'cognitive map' has been widely employed to denote the mental framework that captures such latent variables (Peer et al., 2021a). Initially, in the realm of spatial cognition, the term 'cognitive map' was employed to represent our mental conceptualization of physical space and the intricate interconnections between distinct locations or objects within that space. These cognitive maps serve as essential navigational tools, enabling agents to understand, interact with, and orient themselves within their surroundings. The construction of these spatial cognitive maps is a complex process, shaped by a multitude of factors, drawing upon sensory perceptions, movement experiences, and spatial learning. As such, these maps play an integral role in numerous aspects of spatial behavior - from navigation to decision making (Epstein, Patai, Julian, & Spiers, 2017).

In recent years, the conceptual understanding of the cognitive map has expanded beyond its original spatial confines. The scope now encompasses higher-order cognitive processes, representing a more holistic view of cognition. This broader view includes aspects such as generalization, inference, planning and imagination (Erie D. Boorman, Sweigart, & Park, 2021b; Park, Miller, Nili, Ranganath, & Boorman, 2020b; Peer, Brunec, Newcombe, & Epstein, 2021b). Additionally, cognitive maps have also been associated with social cognition, encapsulating our understanding of social relationships and memory, illustrating their role in our recall and understanding of past experiences (Tavares et al., 2015). Crucially, these nonspatial, conceptual cognitive maps, still assume that the mental representations they encode, store and retrieve are structured in a way that reflects the features of real space.

Neural representation of abstract latent states

The brain's ability to effectively manage learning across latent states relies on its capacity to represent these states, their possible values and interrelations. Studies report the neuronal representations of these acquired latent states during or after animals learn behavioral tasks in the entorhinal cortex, hippocampus, posteromedial and parahippocampal, frontal cortex (Erie D. Boorman, Sweigart, & Park, 2021c) and also the parietal cortex (Harvey, Fracasso, Petridou, & Dumoulin, 2015) during or after animals learn behavioral tasks. Although studies vary in the terminology with which they refer to these representations (e.g., cognitive map, schema, abstract task structure, or categorization), it has been argued that these findings likely point to a common mechanism of knowledge representation. Specifically, the entorhinal cortex, hippocampus, and frontal cortex circuits are interconnected and have been demonstrated to be essential for structured learning (Igarashi, Lee, & Jun, 2022).

In terms of cognitive mapping theories of hippocampus, while spatial theories have dominated research on hippocampal function over the past quarter century (Moser, Moser, & McNaughton, 2017), particularly influential
ideas in cognitive science of the past 25 years, argue that HPC implements a learning process that involves the formation of
conjunctive representations (O'Reilly & Rudy, 2001) and that it is involved in the relational processing of stimuli that enables flexible retrieval (Neal J. Cohen & Eichenbaum, 1993a). The Relational Processing Theory suggests that the hippocampus processes relationships between items during memory encoding. Individual items are stored in the perirhinal and parahippocampal cortex. The hippocampus's role in memory is temporary, and while the processing for flexible relations occurs at encod-

ing, it is only evident during retrieval in different contexts (Andersen, 2007).

> Relational representations . . . are created by and can be used for comparing and contrasting individual items in memory, and weaving new items into the existing organization of memories. This form of representation maintains the "compositionality" of the items, that is, the encoding of items both as perceptually distinct "objects" and as parts of larger scale "scenes" and "events" that capture the relevant relations between them. [They] . . . support the flexible use of memories by permitting access to items from various sources and by permitting the expression of memories in various, even novel, situations. (Neal J. Cohen & Eichenbaum, 1993b)

An illustrative example: "comparing and contrasting individual items in memory" is how the hippocampus has been argued to create one-dimensional representations of relationships between items is through a process called transitive inference. In this process, subjects are shown pairs of items arranged in a hierarchy (like A > B > C > D > E) and must learn to pick the correct item from each neighboring pair. Rats with damage to the hippocampus prior to learning were able to learn the trained associations (e.g., B > C), but they were unable to make the transitive inference between B and D (Dusek & Eichenbaum, 1997a), but this has not been replicated in non-human primates (Basile, Templer, Gazes, & Hampton, 2020). Rodent findings might suggest that they did not grasp the hierarchical structure and that hippocampus helps create a cognitive map that encodes the relational distance between items (although see Kumaran & McClelland (2012a) and O'Reilly & Rudy (2001) for an alternative explanation). While work in some animals revealed support for the MTL relational processing idea (Buckmaster, Eichenbaum, Amaral, Suzuki, & Rapp, 2004; Lazareva, Kandray, & Acerbo, 2014), small sample sizes, varying hierarchy lengths and alternate computational mechanisms to solve the TI task limit a clear understanding of the role it plays in episodic inference. Importantly, nearly all species examined demonstrate performance patterns, under certain circumstances, that align with ordinal processing when solving transitive inference (TI) tasks, despite the capability of animals to utilize both associative and ordinal strategies (Gazes, Templer, & Lazareva, 2022). These observations suggest a pivotal role for relational processing theory, which posits that the understanding and manipulation of relationships between stimuli underpin complex cognition and decision-making.

Further consistent with relational theory, human fMRI neuroimaging work show graph-like encoding of statistical transitions of discrete state-spaces involved in serial learning even when subjects are unaware of the underlying transition structure

(Garvert, Dolan, & Behrens, 2017; Peer et al., 2021a; Schapiro et al., 2012; Schapiro, Rogers, Cordova, Turk-Browne, & Botvinick, 2013; Schapiro, Turk-Browne, Norman, & Botvinick, 2015). This findings suggest that the hippocampus is not just a storehouse of memories, but that it also plays a role in generalizing from experiences (Zeithamova & Bowman, 2020a), learning statistical patterns (Covington, Brown-Schmidt, & Duff, 2018; Henin et al., 2019; Brynn E. Sherman, Graves, & Turk-Browne, 2020), and linking related memories together (M. Schlichting, Guarino, Schapiro, Turk-Browne, & Preston, 2017; Wimmer & Shohamy, 2012). These functions require that memories be formed more gradually and in an overlap-based manner, in order to avoid interference between similar memories, possibly via a dedicated processing pathway, at the level of hippocampal subfields or along a gradient along the long hippocampal axis (Collin et al., 2015; Schapiro et al., 2013; Brynn E. Sherman, Turk-Browne, & Goldfarb, 2023).

In terms of cognitive mapping theories of orbitofrontal cortex, these propose that the OFC tracks associations between cues and relevant events, and consequently uses derivative internal model to predict future events and guide behavior. This has led to the idea that the OFC is essential for monitoring ones own location within a cognitive space (Nicolas W. Schuck, Cai, Wilson, & Niv, 2016; J. Zhou et al., 2019). More specifically, the OFC is believed to track the current location within a state space that represents relevant relationships for any task a subject might be engaged in, allowing the OFC to distinguish hidden states not easily identified by external input, using non-observable information (e.g., inductive priors) to fully understand the current situation (Nicolas W. Schuck et al., 2016). Interestingly, these features of the OFC activity have been related to key computational properties of entorhinal cells in spatial domains (Behrens et al., 2018; Whittington et al., 2019a), both argued to play a role in generalization and relational reasoning (Alon B. Baram, Muller, Nili, Garvert, & Behrens, 2019; Bowman & Zeithamova, 2018; Morton & Preston, 2021; Zeithamova & Preston, 2010). For discussion of alternative theories regarding role of OFC during structured learning see Gardner & Schoenbaum (2021).

While considerable number of studies propose that the hippocampus (HPC) and orbitofrontal cortex (OFC) represent abstract task structure (Wikenheiser & Schoenbaum, 2016), their unique role and relationship is less well understood. One popular theory suggests that the HPC provides specific and detailed environmental data. This information is then processed and refined by the OFC to convey details that are more pertinent to the agent's present goals (Lisman et al., 2017; Rudebeck & Rich, 2018). The specifics of the hippocampal "data" is still a matter of debate. Some researchers believe that the hippocampus directly integrates memories, while others believe that it primarily represents separate memories that are then integrated by the mPFC (Zeithamova & Bowman, 2020b).

1D latent structure learning: Associative inference, Transitive inference

Human beings regularly deduce hidden relational patterns through experiential comparisons. For instance, when selecting a meal at a restaurant, if we know that we think that pizza goes well with omelet and an omelet goes well with a spinach, we can "infer" that pizza goes well with spinach. Associative inference is a type of reasoning that is based on the statistical relationships between objects or events; if we know that A is associated with B, and B is associated with C, then we can infer that A is associated with C (Barron et al., 2020; Alison R. Preston, Shrager, Dudukovic, & Gabrieli, 2004; Zeithamova et al., 2012). Other times we may go further and consider that we like pizza more than omelette and omlette more than spinach and consequently order pizza. Interestingly, this inference can be made even if we have never directly compared pizza to spinach. Transitive inference is a type of reasoning that is based on the order of the relationships between objects or events; if we know that A is greater than B, and B is greater than C, then we can infer that A is greater than C (DeVito, Kanter, & Eichenbaum, 2009; Heckers, Zalesak, Weiss, Ditman, & Titone, 2004; S. N. Moses, Brown, Ryan, & Randal McIntosh, 2010; Zalesak & Heckers, 2009).

Researchers have developed experimental tasks to induce these types of episodic inferences involving a set of sequential (1D) distinct latent states in humans, primates and rodents. Subjects learn a large number of overlapping associations by observing pairs of stimuli on a computer screen and in the case of transitive inference, using binary forced-choice paradigm, deterministic feedback and without any explicit spatial or temporal cues. This helps to equate some of the task features across species and facilitates comparison (Gazes et al., 2022).

Underlying mechanims Various theories have been proposed to describe how observers accomplish transitive inferences of non-adjacent relations (e.g., $B > D$) in such settings. One class of models posits that observers learn implicit value representations for each individual element, which then enables judgments of arbitrary pairings (Wynne, 1995). Relatedly, encoding-based models suggest that the hippocampal and medial prefrontal cortex systems learn unified representations that directly establish inferred structured relationships between task features forming an abstract map of task states (Behrens et al., 2018; Eichenbaum, 1999; Shohamy & Wagner, 2008; Zeithamova et al., 2012). Alternatively, associative and transitive inference could be accomplished through more retrieval-based explicit hippocampal memory processes (Curtis & Jamieson, 2018; Hintzman & Ludlam, 1980; Kumaran & McClelland, 2012d; Zaki & Nosofsky, 2001a). Researchers have formalized these theories as computational models that attempt to assess the degree to which model-based (more in line with the encoding-based approach) or model-free (more akin to the retrieval-based approach) reinforcement learning algorithms can account for patterns in performance in multiple TI experimental designs (Ciranka et al., 2021b; Jensen, Terrace, & Ferrera, 2019b) as well as creating formally testable theories at the representational level. Broadly speaking, retrieval-based models argue that the hippocampus encodes pattern separated representations that together with the entorhinal cortex express elemental relationships between adjacent items (“premise pairs”) (Kumaran, Hassabis, & McClelland, 2016b; Kumaran & McClelland, 2012c) and argue that generalisation is supported by recurrent neural dynamics between the two regions, achieving rapid integration over distinct memories at recall. In contrast, encoding-based models suggest that cognitive maps formed mainly by the hippocampus and the MPFC are sufficient to support inference without the need of a dedicated inference mechanism like recurrence, as the relationships between events has been abstracted and stored. In support of encoding-based models, the hippocampus, entorhinal cortex, and medial prefrontal cortex have been found to encode generalised relationships between item representations (Morton, Schlichting, & Preston, 2020). More specifically, the entorhinal cortex and mPFC have been associated with forming abstract cognitive maps (Alon Boaz Baram, Muller, Nili, Garvert, & Behrens, 2021; Barron et al., 2020). One recent study showed that both behaviour and neural activity in EC and vmPFC/mOFC reflected the Euclidean distance between stimuli in a 2D hierarchical reasoning task (Park et al., 2021) . These “maps” are thought to facilitate inference and knowledge transfer across related tasks, although their relationship to behavioural performance is unclear. Consistent with the retrieval-based approach one high resolution fMRI study showed that the recirculation of hippocampal output as a new input mediates the process of integrating information across multiple episodes (Koster et al., 2018a) and a more recent study showing the activation of intermediary cues (“C” when reasoning about B?D) in the medial temporal lobe at the time of choice, also using 7T MRI (Barron et al., 2020).

The two models make competing behavioural predictions as well when it comes to the symbolic distance effect stating that the cognitive capacity (evidenced by time and/or accuracy) needed to “compare two symbols varies inversely with the distance between their referents on the judged dimension" (Moyer & Bayer (1976), p. 229). Retrieval-based models predict that generalisation performance decreases when inferences require integrating over longer distance. In contrast, encoding-based models often predict the opposite relationship, e.g., generalisations are easier when comparing stimuli that are separated by larger distances in an inferred hierarchy in line with the original SDE, due to distance related saliency differences of the compared items. While there is evidence of negative SDE (Banino et al., 2016) the majority of transitive inference paradigms find positive effects in primates (Jensen, Alkan, Ferrera, & Terrace, 2019; Jensen, Altschul, Danly, & Terrace, 2013; Jensen, Muñoz, Alkan, Ferrera, & Terrace, 2015; Lazareva, Paxton Gazes, Elkins, & Hampton, 2020) and in human neuroimaging studies that identified neural correlates in prefrontal and parietal areas (Acuna, Eliassen, Donoghue, & Sanes, 2002; E. C. Hinton, Dymond, Hecker, & Evans, 2010b; Zalesak & Heckers, 2009). Studies that use shorter item list (5 item) do not find SDE, note however that these studies have symbolic distances confounded with terminal item effects (Wright & Howells, 2008).

While it is tempting to pit these two classes of models against each other, it is important to remember that they may not be mutually exclusive (J. E. Taylor et al., 2021; Zeithamova & Bowman, 2020a). In other words, it is possible that both integrative encoding and retrieval-based on-the-fly inference can contribute to generalization behavior. This is supported by the fact that studies have found evidence for both mechanisms in the episodic memory tasks (Margaret L. Schlichting & Preston, 2016a; Zeithamova & Preston, 2010, 2017).

Temporal trajectory of generalization Relatedly, an often under appreciated aspect of computational models of generalization and episodic inference, is time. Notably, the formation of relevant knowledge structures or basic associations might not exclusively occur during encoding or retrival. It's plausible that they surface following a consolidation phase. Both rodent (Wiltgen & Silva, 2007) and human studies (Audrain & McAndrews, 2022; Sekeres, Winocur, Moscovitch, Anderson, et al., 2018; Tompary & Davachi, 2017) have shown that memory traces can shift from being detailed to more abstract. Work by Tompary & Davachi (2017) showed that following a consolidation period, the mPFC and the HPC began reflecting shared features across memories. Notably, during the recall of overlapping memories, the neural patterns manifested were more alike compared to those elicited by non-overlapping memories memories. This phenomenon was distinctly noticeable only when memories were retrieved a week post-encoding, supporting system-level consolidation theories. Later studies found complementary evidence of for a dissociation along the long axis of the hippocampus that emerges with consolidation, whereby representational patterns for paired-associate memories initially formed prior to sleep become differentiated in anterior hippocampus and overlapping in posterior hippocampus overnight (Cowan et al., 2021).

Systems-level consolidation theories propose that memories, initially anchored in the hippocampus, become redistributed to the cortex through consistent reactivation, notably during periods of rest (Alvarez & Squire, 1994; Nadel et al., 2000). This reactivation refines cortical memory traces, leading to broadly distributed patterns that guard against severe memory disruption (i.e.: catastrophic forgetting, McClelland et al. (1995)). Another model, the Trace Transformation Theory (TTT; Sekeres, Winocur, & Moscovitch (2018a)), proposes that as memories get ingrained in the cortex, they undergo a psychological transformation, that makes them more abstract, capturing shared features across various events. These two theories highlight two primary neural mechanisms for the time-sensitive transformation of relational memories: (1) synchronized reactivation of memories within the hippocampus and cortex post-learning, and (2) the evolution of cortical memories capturing commonalities across various interconnected events.

One proposed mechanism that could drive these two processes is replay during wakeful rest or sleep intervals (Lewis & Durrant, 2011; Lewis et al., 2018b). Sharp-wave ripples observed in the hippocampus (SWR) during these states, as argued by Buzsáki (2015), could signify memory replay. This replay might further represent the active exploration and/or protection from interference of task-associated cognitive maps (Gupta et al., 2010; Y. Liu et al., 2019). Previously discussed work has already pointed to mPFC as a key cortical region supporting integrated memories that emerge during structured learning. While multiple studies show increasing involvement of mPFC in memory retrieval over time (Bonnici et al., 2012; Frankland & Bontempi, 2005; Sterpenich et al., 2009; Sweegers, Takashima, Fernández, & Talamini, 2014; Atsuko Takashima et al., 2009; A. Takashima et al., 2006; Takehara-Nishiuchi & McNaughton, 2008), there is much less evidence in humans that these neural representations are predictive of behavioral measures of memory integration over the course of systems-level consolidation.

An additional complicating factor in the study of time-dependent generalizations is that, in juxtaposition to the gradual abstraction portrayed by system-level consolidation theories, swift relational generalization can be also be seen across species even with minimal training or in few-shot situations (McKenzie & Eichenbaum, 2011;

Nelli, Braun, Dumbalska, Saxe, & Summerfield, 2023; Shohamy & Wagner, 2008; Tse et al., 2007; Zeithamova & Bowman, 2020c). The intricacies of how time-influenced generalization varies, depending on initial encoding environments like prior knowledge or task guidelines, retrieval-based recall (often used in transitive inference tasks), and the period and nature of the subsequent retention interval, remain unclear.

Most research on time and sleep-dependent consolidation of declarative memories emphasizes the coordinated activity between the hippocampus and the mPFC. However, the transformation of relational memory essential for episodic inference may encompass more diverse processes, hinging on context, task instruction, and difficulty (Gazes & Lazareva, 2021b).

Inference performance A meta-analysis conducted by X. Zhang et al. (2022) investigated the neural correlates associated with episodic inference (i.e.: spatial, associative and transitive inference) involving 32 univariate, whole-brain corrected fMRI results on healthy adults summarizing contrasts between inference and control conditions. A conjunction analysis between associative inference and hierarchical inference identified the left hippocampus, left medial superior frontal gyrus, right anterior cingulate cortex, left angular gyrus, bilateral precuneus, and left precentral gyrus. While univariate studies were crucial in identifying a set of possible regions contributing to indirect associations, multivariate pattern similarity analyses enable researchers to take a representational geometric perspective on brain and cognition, and allowes them to study the properties of neural representations that support successful episodic inference at a more granular level.

In one such study using a neurocomputational approach, participants learned image pairs (AB, BC) from distinct triads (ABC) that shared the same internal structure. The participants were then tested on their ability to make inferences based on the learned common structure (Morton et al., 2020). The results showed that abstract information about commonalities among events is represented in the hippocampus, lateral parietal cortex (LPC), parahippocampal cortex (PHC), frontopolar cortex and dorsolateral prefrontal cortex (DLPFC). By modelling response time during inference trials as a variable versus fixed-speed process, related to retrival versus encoding-based inference the showed that representational geometry in hippocampus is related to efficient pattern completion, while geometric structure in LPC and PHC support vector-based retrieval.

In terms of transitive inference, model-based RSA has been commonly used to study the representations underlying recall in episodic inference in tasks involving larger number of latent states (>3). For example, in a study of deductive reasoning Alfred, Connolly, & Kraemer (2018) had participants recall the "estimated" rank of each item in a transitive hierarchy used whole-brain searchlight RSA to discover brain regions encoding the true rank-order structure. This revealed that three brain regions in the right hemisphere encode the newly-learned representational space. These regions are located within the intraparietal sulcus, precuneus, inferior frontal gyrus and RLPFC. In another study, using three hierarchies Alfred, Connolly, Cetron, & Kraemer (2020) reported finding pattern similarity representative of symbolic distance-based mental models in the intraparietal sulcus, inferior frontal gyrus, and anterior prefrontal cortex in a transitive inference task, replicating earlier findings. RSA has also been used for studying the representational changes involved in linking two distinct ordered hierarchies into a combined hierarchy (Nelli et al., 2023). The authors found symbolic distance related pattern similarity with peaks in dmPFC and inferior parietal lobe, after minimal trials involved in linking the two hierarchies observed BOLD patterns rapidly reconfigure into a novel geometry that reflects this merger in the PPC and dmPFC.

Using a transitive inference paradigm, Berens & Bird (2022b) investigated how different training schedules influence generalization mechanisms in memory, focusing again on encoding-based and retrieval-based generalization, in remotely vs recently learned hierarchies. The authors compared the effects of progressive training (presenting related information in ordered blocks) and interleaved training (presenting related information randomly) on generalization, whilst also examining the impact of overnight consolidation on generalization. Both training conditions showed better alignment with an encoding-based model in terms of behavioral data, univariate results, and RSA. BOLD pattern similarity correlated with hierarchical distances in the left HPC and superior mPFC, suggesting map-like structural representations. It is interesting to note that the mPFC only appeared to represent structural codes for recently learned stimuli. Additionally, while the authors identified mPFC representations in most participants, they tended to be most strongly expressed in participants who did not achieve high levels of inferential performance.

To examine how neural representations underlying transitive inference task change over time, we collected fMRI data in an implicit TI paradigm. On two consecutive days we presented participants with paired neighboring images sampled from two separate linear hierarchical structures (Fig. 1). We then used fMRI to measure neural similarity of the items to see whether they resembled a rank-order distance model, hypothesizing that this measure would be greater for a remotely learned hierarchy compared to a recently learned one. While prior work measured time-dependent pattern similarity changes in a non-relational memory tasks and distance representations underlying an explicitly cued transitive inference task at recall, here we combine both aims in one study in order to find multivariate representational markers of integration over time in a non-cued TI task. We examined stimulus pattern similarity in hippocampus, mOFC as well as parahippocampal cortex and the inferior frontal gyrus to test whether difference in how items of each hierarchy were represented could explain time-dependant benefits in inferential performance.

5.3 Methods

5.3.1 Participants

32 right handed participants took part in the experiment (25 females, aged 19-36, mean age=23.5 years) after giving informed consent in accordance with the protocol approved by the Cardiff University School of Psychology Research Ethics Committee. Participants were required to achieve 66% accuracy on middle pairs (B?C, C?D, D?E) in two consecutive learning blocks for each hierarchy (Remote/Recent) within 10 blocks. Those who did not meet this criterion were excluded from the study. After implementing this exclusion criterion, the final sample comprised of 26 participants (mean age of 23.8 years, range 19-36 years, with 21 identifying as female).

All participants had normal or corrected-to-normal vision, no history of major neurological or psychiatric illness and reported having a regular sleep-wake cycle. They were requested to abstain from psychostimulants and intense physical activity for the 48h preceding the experiment. Participants received 30£ in total for their participation in all three sessions.

5.3.2 Task Design

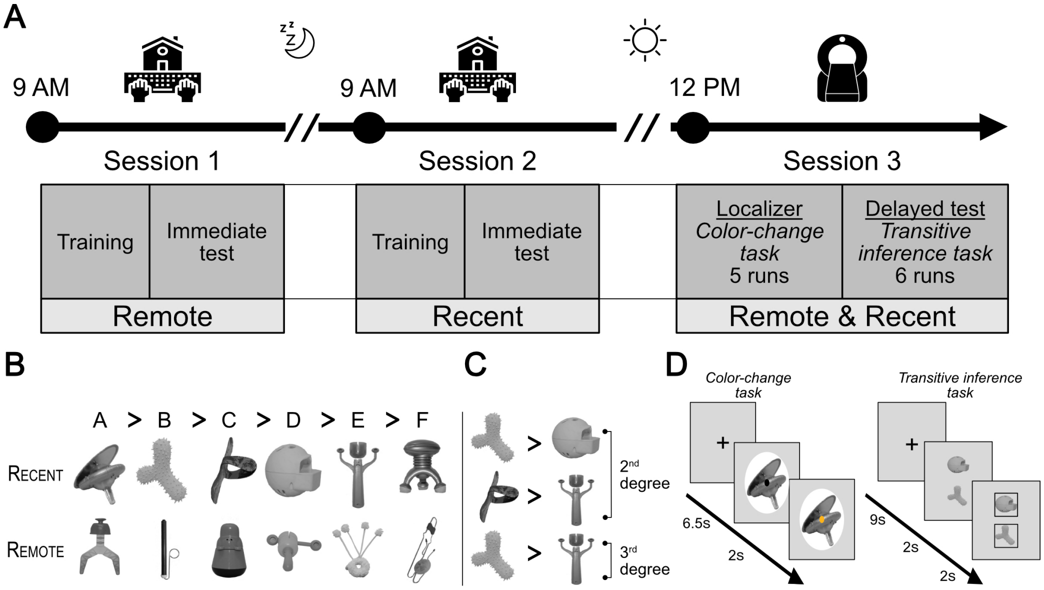


Figure 1: Experimental paradigm. A) Timeline of the procedure. Participants were sent scheduled emails at pre-agreed times containing a link to the experiment. 24 hours after they completed Session 1 they were sent an email containing a link to Session 2, with the MRI session starting approximately 3 hours later. B) Example of the two randomized hierarchies used in the experiment. C) Example of resultant inference pairs from (B) for Recent D) Example of a Color-change and Transitive inference task trial in Session 3. In the Colour-change task a cross appearing on the screen indicates the start of the trial (fixed 6.5s), then one of the images of the hierarchies appear on the screen with a small black dot in the middle that changed to either blue or orange. In the Transitive inference task similarly a cross appearing on the screen indicates the start of the trial (fixed 9s), then two images of one of the hierarchies appear on the screen for a planning period of 2s. Next rectangles appear around the images signalling the start of the response window for 2s.

Twelve pictures were used, all depicting unusual objects in greyscale (Horst & Hout, 2016). The images were randomly divided into two 6 item sets (H1 and H2) and were randomly assigned a hierarchical structure (referred to as A>B>C>D>E>F or R1>R2>R3>R4>R5>R6) for each participant. The experimental design [2 (hierarchy: within-subject) x 3 (session: within-subject)] used was a similar to that used in Experiment 3 (see Experiment 3 protocol). The experiment started in the morning, between 9 AM and 1 PM and was conducted online ("Session 1"). Participants were trained on a two-alternative forced choice paired association task where the adjacent item pairs ("premise pairs": A?B, B?C, C?D, D?E, E?F) belonged to a 6-item linear hierarchy. The pairs were presented on the screen, vertically stacked in the center. Subjects were instructed to select the item "hiding" a smiley face which served a visual feedback for correct choices (e.g., choosing "B" when shown "B" vs "C"). If the participant selected the lower ranked item ("C" when shown "B" vs "C") an angry cartoon face replaced the stimuli. In addition participants were told they had to reach a certain performance to advance to the next session. The training session

(“Recent” hierarchy) was organized in blocks containing 10 trials each for a maximum of 8 blocks (“Training phase”). Each block presented the five adjacent pairs twice with their vertical position counterbalanced and pseudorandomized sequential order to ensure that pairs with overlapping items do not consecutively appear, potentially revealing a latent hierarchical structure. At the end of each block, the overall performance for that block was shown on the screen to keep participants engaged with the task. Participants needed to achieve a minimum of 66% accuracy in two consecutive blocks on the inner premise pairs (B?C,C?D,D?E) in order to successfully complete the training phase. If criteria was reached before the 8th block feedback was hidden to balance out item exposure among participants during training, whilst controlling for learning performance. After finishing the learning phase, participants took a 5 minute break and were subsequently tested on the same pairs they learned for 4 blocks with no feedback (“Test phase”). If participants performed above 50% in the test phase they advanced to Session 2 otherwise they were informed of not reaching criteria and paid for their involvement up to this point. Twenty-four hours after finishing Session 1 participants were again asked to learn adjacent item pairs for a second hierarchy (“Remote”) online (“Session 2”) with the same task structure as in Session 1. Both training sessions lasted approximately 15 minutes and had to be completed within 2 hours of the pre-agreed start time to ensure comparability within the sample with regards to time elapsed between the two sessions. Approximately 3 hours after the Session 2 was completed, the fMRI scanning session took place (“Session 3”, see Fig. 1). This session lasted 90 minutes, and participants completed two tasks while in the scanner. First participants completed a stimulus viewing task that allowed us to estimate the BOLD activity response for each stimulus (“Localizer task”). Participants were presented with images of the 12 objects, while completing an incidental color-change detection task. There were five runs of the Localizer task; each run contained three blocks of 12 item presentations totaling three presentations per item per run. Presentation order was randomized within each block of each run. Each stimulus was presented for 2s, with a fixed interstimulus interval of 6.5 s. Participants performed a change-detection task during the presentation of the stimuli to ensure sustained attention. Participants pressed one button if a small black circle at the center of the image turned blue and another if the circle turned orange; the color change occurred between 0.25-0.75s after picture onset. The number of missed responses were shown to participants at the end of each run.[ˆ1]

In the transitive inference task (“TI task”) participants were shown a similar task to the test phase of Session 1 and Session 2, but crucially were presented with 4 additional pairs, leading a total of 9 pairs per hierarchy. Three were so-called inference pairs (B>D, B>E and C>E), where correct choice can be inferred by generalizing over the premise pairs. Additionally, we also added a so-called anchor pair (A>F) as a control, since terminal items are associated with perfect information (100 and 0% positive feedback, respectively) and correct choice can be inferred without relational reasoning. Prior to starting the task participants were instructed that they may see novel item pairs and they should try to make their best guess on that trial, given the time limit. There were six runs of the task; each run contained 18 item presentations containing one item pair from each hierarchy. Presentation order was randomized within each run and consecutive runs counterbalanced vertical position of each item. Each stimulus was presented for 4s with a fixed interstimulus interval of 9 s. Item presentation was split into a 2s “planning” and 2s “response” window to be better able to separate memory processes from motor response, as participants were asked to not respond during planning.

5.3.3 fMRI data

Preprocessing was performed using *fMRIPrep* 20.2.7 (Esteban, Markiewicz, et al. (2018); Esteban, Blair, et al. (2018); RRID:SCR_016216), which is based on *Nipype* 1.7.0 (K. Gorgolewski et al. (2011); K. J. Gorgolewski et al. (2018); RRID:SCR_002502).

5.3.4 Data collection

MRI was performed on a Siemens PRISMA 3T MRI Scanner at Cardiff University Brain Research Imaging Centre (CUBRIC). Earplugs were used to attenuate scanner noise, and head movement was restricted using foam pillows. Stimuli were generated using Psychopy (Peirce, 2009) and displayed on a screen, which participants viewed through a tilted mirror placed in front of them. Subjects responded pressing one of two buttons on a button box they held in their right hand. There was a total of eleven functional scans acquired using a T2*-weighted multiband-accelerated EPI pulse sequence (repetition time [TR]: 2000 msec, echo time [TE]: 30 msec, flip angle: 77°, field of view: 220 mm, X slices, 1.7 × 1.7 × 1.7 mm voxels, multiband factor: 3, generalized autocalibrating partially parallel acquisitions (GRAPPA) factor: 2, phase partial Fourier: 7/8). A single-band reference (SBRef) image was also collected at the start of each EPI scan to aid in alignment to anatomical scans. A field map (TR = 668 msec, TE = 5 msec/ 7.46 msec, flip angle = 5°, 67 slices, 1.7 × 1.7 × 2 mm voxels) was collected before the first functional scan to correct for distortions in the magnetic field in functional scans. Two oblique coronal T2-weighted (T2w) structural images were acquired perpendicular to the main axis of the hippocampus (TR: 13150 msec, TE: 82 msec, flip angle: 150°, 60 slices, 0.4 × 0.4 mm in-plane resolution, 1.5-mm through-plane resolution) to facilitate segmentation of hippocampal anatomy. These images were coregistered and averaged to generate a mean coronal image for each participant. A T1-weighted 3-D magnetization prepared rapid gradient echo volume was collected (TR: 1900 msec, TE: 2.43 msec, flip angle: 9°, field of view: 256 mm, 192 slices, 1 × 1 × 1 mm voxels) to facilitate alignment and normalization of the functional data.

5.3.5 Preprocessing

A total of 1 T1-weighted (T1w) images were found within the input BIDS dataset. The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with `N4BiasFieldCorrection` (Tustison et al., 2010), distributed with ANTs 2.3.3 (Avants, Epstein, Grossman, & Gee, 2008, RRID:SCR_004757), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a *Nipype* implementation of the `antsBrainExtraction.sh` workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using `fast` (FSL 5.0.9, RRID:SCR_002823, Y. Zhang, Brady, & Smith, 2001). Brain surfaces were reconstructed using `recon-all` (FreeSurfer 6.0.1, RRID:SCR_001847, Dale, Fischl, & Sereno, 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (RRID:SCR_002438, Klein et al., 2017). Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with `antsRegistration` (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: *ICBM 152 Nonlinear Asymmetrical template version 2009c* [Fonov, Evans, McKinstry, Almli, & Collins (2009), RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym],

Functional data preprocessing

For each of the 11 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated by aligning and averaging 1 single-band references (SBRefs). A B0-nonuniformity map (or *fieldmap*) was estimated based on a phase-difference map calculated with a dual-echo GRE (gradient-recall echo) sequence, processed with a custom workflow of *SDCFlows* inspired by the `epidewarp.fsl` script and further improvements in HCP Pipelines (Glasser et al., 2013). The *fieldmap* was then co-registered to the target EPI (echo-planar imaging) reference run and converted to a displacements field map (amenable to registration tools such as ANTs) with FSL's `fugue` and other *SDCflows* tools. Based on the estimated susceptibility distortion, a corrected EPI (echo-planar imaging) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using `bbregister` (FreeSurfer) which implements boundary-based registration (Greve & Fischl, 2009). Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using `mcflirt` (FSL 5.0.9, Jenkinson, Bannister, Brady, & Smith, 2002). BOLD runs were slice-time corrected to 0.952s (0.5 of slice acquisition range 0s-1.91s) using `3dTshift` from AFNI 20160207 (R. W. Cox & Hyde, 1997, RRID:SCR_005927). First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. The BOLD time-series were resampled onto the following surfaces (FreeSurfer reconstruction nomenclature): *fsaverage*. The BOLD time-series (including slice-timing correction) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as *preprocessed BOLD in original space*, or just *preprocessed BOLD*. The BOLD time-series were resampled into standard space, generating a *preprocessed BOLD run in MNI152NLin2009cAsym space*. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al. (2014)) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al. (2002)). FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* (following the definitions by Power et al., 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks.

The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using `mri_vol2surf` (FreeSurfer). First, a reference volume and its skull-stripped version were generated by aligning and averaging 1 single-band references (SBRefs). A B0-nonuniformity map (or *fieldmap*) was estimated based on a phase-difference map calculated with a dual-echo GRE (gradient-recall echo) sequence, processed with a custom workflow of *SDCFlows* inspired by the `epidewarp.fsl` script and further improvements in HCP Pipelines (Glasser et al., 2013). The *fieldmap* was then co-registered to the target EPI (echo-planar imaging) reference run and converted to a displacements field map (amenable to registration tools such as ANTs) with FSL's `fugue` and other *SDCflows* tools. Based on the estimated susceptibility distortion, a corrected EPI (echo-planar imaging) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using `bbregister` (FreeSurfer) which implements boundary-based registration (Greve & Fischl, 2009). Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using `mcflirt` (FSL 5.0.9, Jenkinson et al., 2002). BOLD runs were slice-time corrected to 0.951s (0.5 of slice acquisition range 0s-1.9s) using `3dTshift` from AFNI 20160207 (R. W. Cox & Hyde, 1997, RRID:SCR_005927). First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. The BOLD time-series were resampled onto the following surfaces (FreeSurfer reconstruction nomenclature): *fsaverage*. The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as *preprocessed BOLD in original space*, or just *preprocessed BOLD*. The BOLD time-series were resampled into standard space, generating a *preprocessed BOLD run in MNI152NLin2009cAsym space*. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al. (2014)) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al. (2002)). FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* (following the definitions by Power et al., 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (*CompCor*, Behzadi, Restom, Liau, & Liu, 2007). Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's *aseg* segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or

1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using `mri_vol2surf` (FreeSurfer).

Many internal operations of *fMRIPrep* use *Nilearn* 0.6.2 (Abraham et al., 2014, RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIPrep*'s documentation.

Quality control

Given that the data was already in BIDS-compliant format, no modifications or adaptations were needed prior to deploying MRIQC (Esteban et al., 2017). MRIQC, or Magnetic Resonance Imaging Quality Control, is a software tool that is used to assess the quality of MRI data. It is designed to provide objective measures of image quality and to identify potential issues or artifacts that may affect the accuracy and reliability of the MRI results. MRIQC uses a combination of image processing algorithms and statistical analysis to evaluate various aspects of the MRI data and can assess the overall image quality, including factors such as signal-to-noise ratio, contrast-to-noise ratio, and spatial resolution. It can also detect specific artifacts, such as motion artifacts, ghosting, and susceptibility artifacts. Each T1w image and BOLD scan generated a unique visual report through MRIQC, which was evaluated by the author.

To assess motion confounds in our functional data we used framewise displacement as calculated in MRIQC (version: 22.0.0rc1, (Esteban et al., 2017)). The framewise displacement at each time point is calculated as the sum of the absolute values of the differential realignment parameters (i.e., derivatives of the six motion parameters). These realignment parameters represent translation in the x, y, and z directions (measured in mm) and rotations about x, y, and z axes (measured in degrees). However, for calculating FD, the rotational displacements are also converted to mm. To convert rotations to translations (so they can be summed), the rotations are typically multiplied by the assumed radius of the head (which can vary, but a common value is 50 mm (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012)). The end result is a single time series (i.e., a single value for each frame/volume) of framewise displacements in mm that can serve as composite measure of overall movement. We used the following metrics in specific to assess signal quality: *fd mean* refers to the average framewise displacement over the course of the scan in mm, *fd num* refers to the number of frames (or volumes) that have a framewise displacement above a certain threshold (0.5mm; the unit is count/number of frames) and *fd perc* would be the percentage of frames (or volumes) that have a framewise displacement above a certain threshold (unit: percent/%). These results are illustrated in Fig. 2.

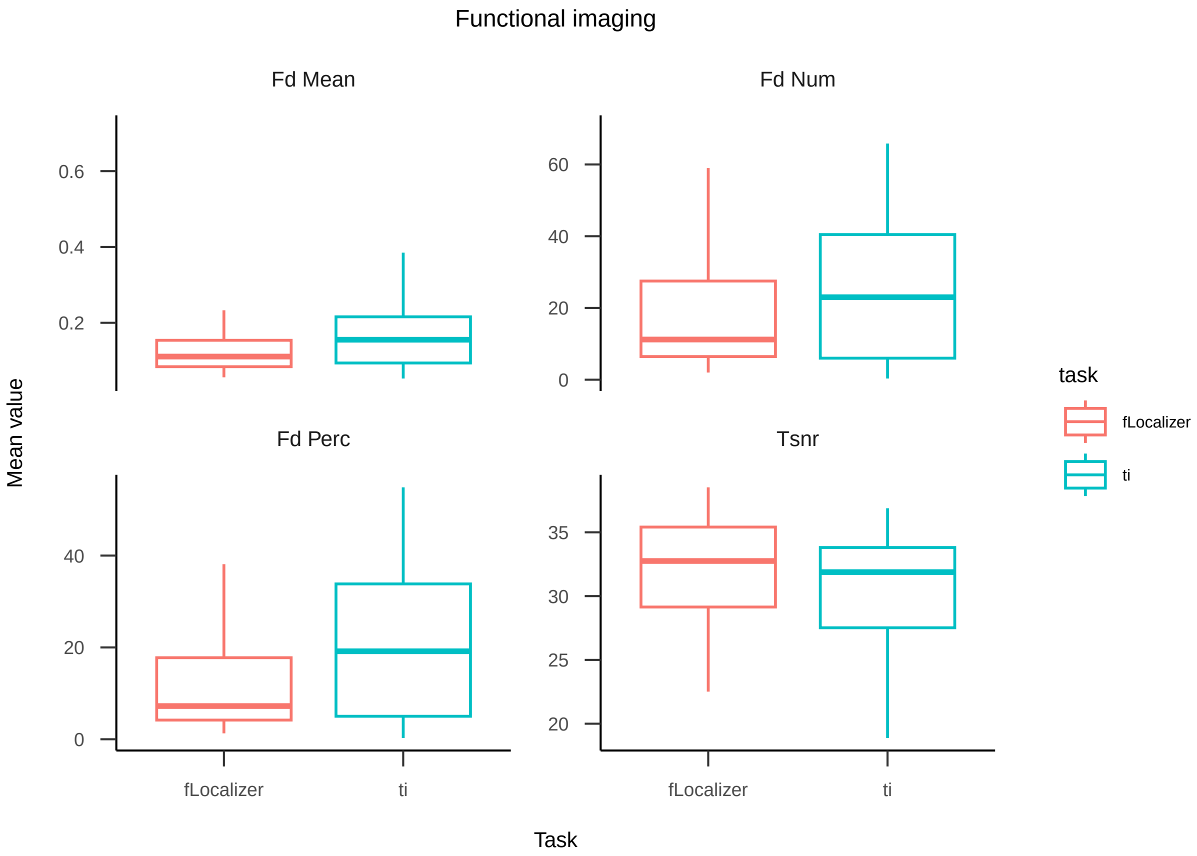


Figure 2: Boxplot of MRIQC metrics across tasks fLocalizer and ti. Four MRIQC metrics are presented: mean framewise displacement $fdmean$, number of framewise displacements over a threshold $fdnum$, percentage of framewise displacements over a threshold $fdperc$, and temporal signal to noise ratio $tsnr$. Each box represents the interquartile range IQR of the metric for a task, with the line inside the box marking the median. Whiskers denote the range of values within 1.5 times the IQR from the first and third quartiles. Outliers are represented by individual points outside the whiskers. The plot provides a comparative view of the distribution of these quality metrics between the two tasks, allowing for evaluation of potential task-related differences in data quality. FD threshold set to 0.5 mm

Estimation of item-level activation patterns

All responses (task-fLocalizer, task-ti) were estimated using the GLMsingle toolbox (Prince et al., 2022), featuring optimized denoising and regularization procedures, to accurately measure changes in brain activity in response to each of our 12 experimental stimuli. To improve the signal-to-noise ratio in estimating beta weights, the toolbox employs a three-stage procedure. First it attempts to identify a hemodynamic response function (HRF) from a library of 20 candidate HRFs for each voxel. Next it finds an ideal number of noise regressors using the previously identified optimal HRFs, principal noise components and cross-validation in order to maximize variance explained across voxels. Lastly, fractional ridge regression is used to regularize ("decorrelate") the beta weight estimation on a per voxel basis. Finally, beta weights for each voxel were averaged over trials on which the same image was shown (three trials per image per run for the Localizer task, two trials per image pair per run for the inference task). For the purpose of noise reduction and enhancement of signal-to-noise ratio, the functional images were subjected to a spatial smoothing process, before running GLMSingle. This was achieved using a Gaussian kernel with a Full Width at Half Maximum (FWHM) of 4mm using 'nilearn.image.smooth_img' (Abraham et al., 2014).

Anatomical ROIs

Drawing on previous research that explored the neural encoding and retrieval of transitive inference, our primary, a priori Regions of Interest (ROIs) - the hippocampus and medial orbitofrontal cortex (mOFC) - due to numerous studies of novel inference identifying their involvement (Backus, Schoffelen, Szebényi, Hanslmayr, & Doeller, 2016; Park et al., 2020a; Margaret L. Schlichting & Preston, 2016a; Zeithamova et al., 2012). The hippocampal and mOFC ROI was delineated for each participant using the automatic segmentation facilitated by FreeSurfer, which was based on both T1-weighted and T2-weighted anatomical scans. Individual participant ROIs were transformed to template space and averaged to create right and left probabilistic masks. The masks were combined and thresholded at 0.1. We followed a similar procedure to generate a whole-brain gray matter (GM) mask that included all cortical and subcortical regions of GM, thresholded at 0.1.

In addition we also explored parahippocampal cortex (PHC) and the inferior frontal gyrus (IFG). PHC has been associated with contextual representation of objects, integrated event structure, formation of cognitive maps as part of the posterior medial network (Aly, Chen, Turk-Browne, & Hasson, 2018; Baldassano, Hasson, & Norman, 2018; Cooper, Kurkela, Davis, & Ritchey, 2021; Morton et al., 2020; Spiers & Barry, 2015). IFG has been associated with relational reasoning, associative inference and querying conceptual cognitive maps (Alfred et al., 2020, 2018; Wendelken, Nakhabenko, Donohue, Carter, & Bunge, 2008; X. Zhang et al., 2022).

PHC ROI was delineated for each participant using the automatic segmentation facilitated by FreeSurfer and then group averaged and thresholded at 0.1. The anatomical ROI for the Inferior Frontal Gyrus (IFG) was delineated based on the Harvard-Oxford Atlas, specifically utilizing the cortical probability map with a threshold of 50%.

We leveraged the 'Highlight Results, Don't Hide Them' approach (P. A. Taylor et al., 2023), employing AFNI's `chauffeur_afni` function, to overlay and visualize Z coefficients from group-level data. By utilizing a combination of voxelwise thresholding and clusterization with translucency, one can enhance the interpretability of their findings—ensuring that voxels within significant clusters appeared with maximum opacity and distinctive boxing, while delineating those above the threshold but outside of sizeable clusters with subtle translucency, thus maintaining methodological transparency and minimizing bias.

5.3.6 Analysis

We used searchlight analyses to examine whether rank-order structure is represented in the brain at delayed test separately for the localizer task and the transitive inference task. After smoothing (4mm) and item level activation estimation via GLMsingle toolbox we separately calculated similarity to an a priori model encoding pairwise distances (Pearson's correlation) per participant per hierarchy (Remote/Recent) via Spearman correlation. We will refer to this measure as "distance similarity" hereafter. This approach allowed us to determine whether the representations of individual items had been affected by time-dependent reorganization aligned with mental model that encodes rank-order ($A > B > C > D > E > F$). Prior to model-based representational similarity analysis, trial-level data from the transitive inference task had to be decomposed into item level representations via leastsquares decomposition (similar to what is used in Berens & Bird (2022b)).

The searchlight analysis was implemented in `rsatoolbox` 0.1.3 (*Representational Similarity Analysis 3.0*, 2023) to test for neural representations of rank-order structure. Separate searchlights were carried out for GM, HPC, PHC, IFG and mOFC. In each searchlight sphere (3 voxel radius), we calculated pairwise correlations between Remote and Recent object patterns separately, estimated based on all runs. This distance similarity statistic was compared with a baseline determined using a permutation test. Objects were randomly permuted the same way for each run, such that object identity was preserved but distance structure was broken. This process was repeated 720 times (6!) to obtain a null distribution for the distance similarity statistic. The same permutation order was used across all searchlight spheres to preserve distance structure in the permuted data. For each sphere, the actual statistic was compared with the permutation distribution to obtain a Z-score (see Fig. 3). To perform cluster correction of the searchlight results within each ROI, we used AFNI 21.1.09 (R. W. Cox, Chen, Glen, Reynolds, & Taylor, 2017). We used the `3dttest++` command with the Clustsim flag initiaties nonparametric one-sample and paired t-test. One-sample t-tests were used for detecting rank-order structure above chance and paired t-tests were used detecting difference in our distance statistic between Remote and Recent. In our analysis, we aimed to explore the differences in model-based representational similarity (i.e.: distance similarity), controlling for potential confounding effects of encoding strength. We used each participant's hierarchy specific (Remote/Recent) mean immediate premise pair performance (i.e., encoding strength) as a covariate of no interest in the model to account for potential confounding effects and improve the precision of our inferential statistics (~ANCOVA). We chose this covariate based on our own prior work showing higher premise pair performance for the Recent hierarchy at delayed recall and findings suggesting that encoding strength may influence time-dependent inference performance (Foldes, Santamaria, & Lewis, 2023). Encoding strength was mean centered and treated as a continuous variable.

The Clustsim flag initiates a nonparametric approach to significance testing. Following the primary t-tests, the residuals from are leveraged in conjunction with sign-flipping to generate 10,000 simulated null 3D datasets. The primary objective of utilizing the Clustsim flag is to compute cluster-size thresholds, eschewing reliance on a fixed model for the spatial autocorrelation function (ACF) of the noise, thereby offering a more robust method to control for false positives. We estimated cluster extent thresholds to control the false-positive rate at 0.05, using three nearest neighbor clustering, one-sided thresholding for the cortical ROIs, two-sided for the HPC and a voxel-wise threshold of 0.005.

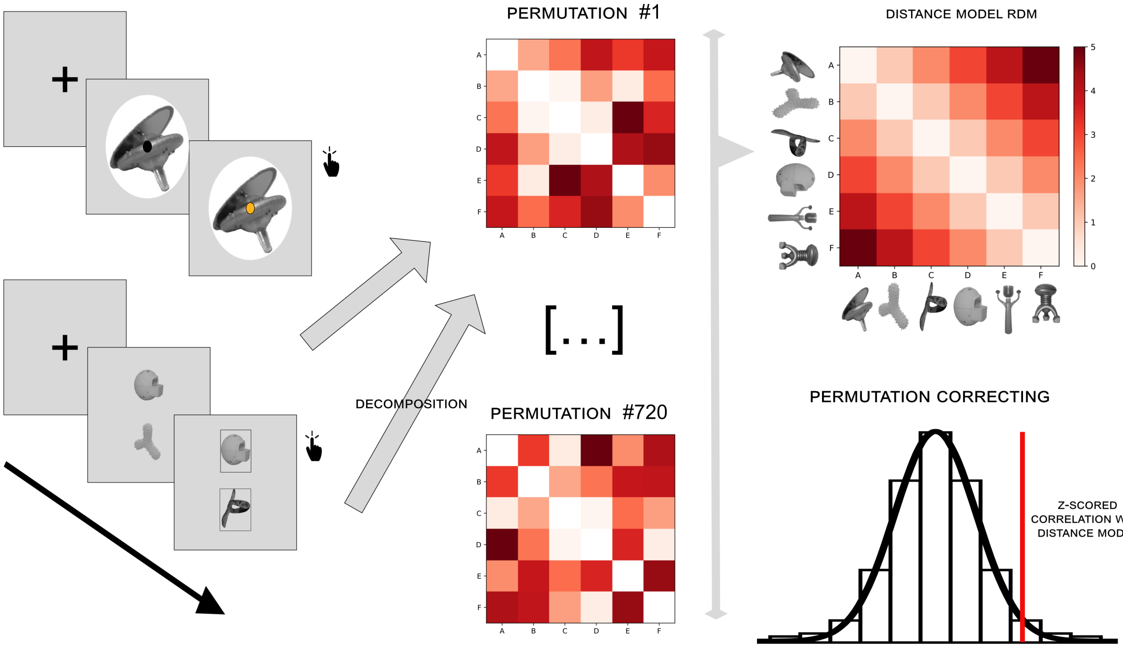


Figure 3: Experiment 4: RSA analysis pipeline: For both the Localizer and Transitive Inference tasks, trial-level data during the response window were subjected to pairwise correlation analyses within each searchlight sphere, conducted separately for the Remote and Recent conditions.The derived representational similarity matrix was permuted 720 times, ensuring the preservation of object identity and correlated with a prior distance model. The neural similarity statistic for each sphere was compared to the permutation distribution, resulting in a Z score (i.e.: distance similarity). Decomposition of TI data was done in line with Berens & Bird (2022).

5.4 Results

5.4.1 Behavioural Results

Accuracy

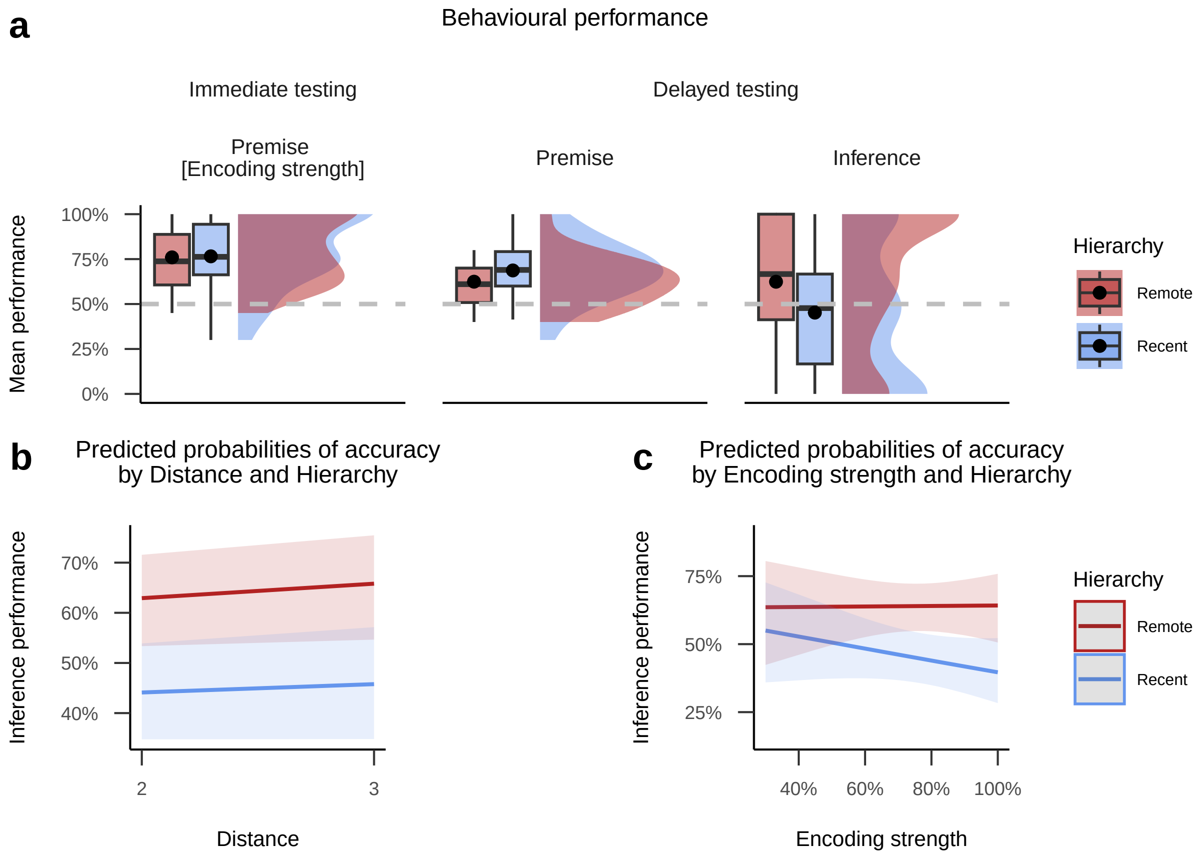


Figure 4: Experiment 4: Behavioural performance and factors predicting inference accuracy at delayed test. (a) Raincloud plot with mean and median performance across Remote and Recent hierarchy conditions. Dashed line represents chance performance at 50% and dot represents mean values (b) Predicted probabilities of accuracy by Distance and Hierarchy, with distance levels of 2 and 3. (c) Predicted probabilities of accuracy by Encoding strength and Hierarchy. In panels (b) and (c), shadowed areas represent 95% confidence intervals.

As in Experiment 3, mean inference performance was significantly higher than chance for the Remote hierarchy $M = 0.62$, 95% CI $[0.55, \infty]$, $t(51) = 2.69$, $p = .005$, but not for the Recent hierarchy ($M = 0.45$, 95% CI $[0.38, \infty]$, $t(51) = -1.06$, $p = .853$), which suggest time-dependent consolidation of relational memory (Fig. 4A). For detailed mean performance see Table S4.

To test for between Hierarchy baseline differences in premise pair learning, we performed a two-way ANOVA on mean premise pair performance with within-subject factors Hierarchy and Session. This revealed an effect of Session ($F(1, 25) = 19.43$,

$p < .001$, $\hat{\eta}^2_G = .098$, 90% CI $[.000, .306]$) with higher performance at Immediate testing compared to Delayed testing ($\Delta M = 0.11$, 95% CI $[0.06, 0.16]$, $t(25) = 4.41$, $p < .001$, Fig. S5). There was no main effect of Hierarchy nor a Session X Hierarchy interaction (smallest p-value=0.16, for details see Table S5). In addition we tested condition specific relationship between mean premise pair and mean inference performance, but found no significant correlations (smallest p-value=0.78, Fig. S6)

These findings suggest that for premise pairs there is a gradual overall decline in performance, whereby participants had higher accuracy for premises at Immediate test (online) than at Delayed test (in the scanner).

We fitted a logistic mixed model (estimated using ML and BOBYQA optimizer) to predict inference accuracy with Hierarchy (formula: inference accuracy ~ Hierarchy). The model included participant as random effect (formula: ~1 | participant). The model's total explanatory power is moderate (conditional R2 = 0.20) and the part related to the fixed effects alone (marginal R2) is of 0.04. The model's intercept, corresponding to Hierarchy = Remote, is at 0.57 (95% CI [0.20, 0.94], p = 0.003). Within this model:

- The effect of Hierarchy [Recent] is statistically significant and negative (beta = -0.78, 95% CI [-1.06, -0.50], p < .001; Std. beta = -0.78, 95% CI [-1.06, -0.50])

These findings suggest improved accuracy for the remotely learned hierarchy compared to the recently learned one, suggesting time-dependent consolidation of relational memory.

Distance and Encoding strength We fitted a logistic mixed model (estimated using ML and BOBYQA optimizer) to predict inference accuracy with Hierarchy and Distance (formula: inference accuracy ~ Hierarchy * Distance). The model included participant as random effect (formula: ~1 | participant). The model's total explanatory power is moderate (conditional R2 = 0.20) and the part related to the fixed effects alone (marginal R2) is of 0.04. The model's intercept, corresponding to Hierarchy = Remote and Distance = 0, is at 0.57 (95% CI [0.20, 0.94], p = 0.002). Within this model:

- The effect of Hierarchy [Recent] is statistically significant and negative (beta = -0.78, 95% CI [-1.07, -0.50], p < .001; Std. beta = -0.78, 95% CI [-1.07, -0.50])
- The effect of Distance is statistically non-significant and positive (beta = 0.13, 95% CI [-0.30, 0.55], p = 0.561; Std. beta = 0.06, 95% CI [-0.14, 0.26])

- The effect of Hierarchy [Recent] × Distance is statistically non-significant and negative (beta = -0.06, 95% CI [-0.65, 0.53], p = 0.845; Std. beta = -0.03, 95% CI [-0.31, 0.25]).

To further investigate the Hierarchy[Recent] × Distance interaction, we calculated the coefficients of simple effects at close (Remote: $\text{logit}(p) = 0.53$, 95% CI $[0.14, 0.92]$, Recent: $\text{logit}(p) = -0.24$, 95% CI $[-0.63, 0.16]$, $\text{OR}_{\text{Remote/Recent}} = 2.15$, SE = 1.19, $p < 0.001$) and distant (Remote: $\text{logit}(p) = 0.65$, 95% CI $[0.19, 1.12]$, Recent: $\text{logit}(p) = -0.17$, 95% CI $[-0.63, 0.29]$, $\text{OR}_{\text{Remote/Recent}} = 2.28$, SE = 1.28, $p < 0.001$) pair values of Distance, suggesting greater relative accuracy benefit for distant, compared to close pairs (Fig. 4B).

Additionally we fitted a logistic mixed model (estimated using ML and BOBYQA optimizer) to predict inference accuracy with Hierarchy and Encoding strength (formula: inference accuracy ~ Hierarchy * Encoding strength). The model included participant as random effect (formula: ~1 | participant). The model's total explanatory power is moderate (conditional R2 = 0.21) and the part related to the fixed effects alone (marginal R2) is of 0.04. The model's intercept, corresponding to Hierarchy = Remote and Encoding strength = 0, is at 0.54 (95% CI [-0.80, 1.89], p = 0.426). Within this model:

- The effect of Hierarchy [Recent] is statistically non-significant and negative (beta = -0.08, 95% CI [-1.56, 1.41], p = 0.918; Std. beta = -0.78, 95% CI [-1.06, -0.50])
- The effect of Encoding strength is statistically non-significant and positive (beta = 0.04, 95% CI [-1.66, 1.74], p = 0.964; Std. beta = 6.96e-03, 95% CI [-0.29, 0.31])
- The effect of Hierarchy [Recent] × Encoding strength is statistically non-significant and negative (beta = -0.93, 95% CI [-2.85, 1.00], p = 0.346; Std. beta = -0.16, 95% CI [-0.50, 0.18])

Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using a Wald z-distribution approximation. These findings are illustrated in Fig. 4C.

Response time

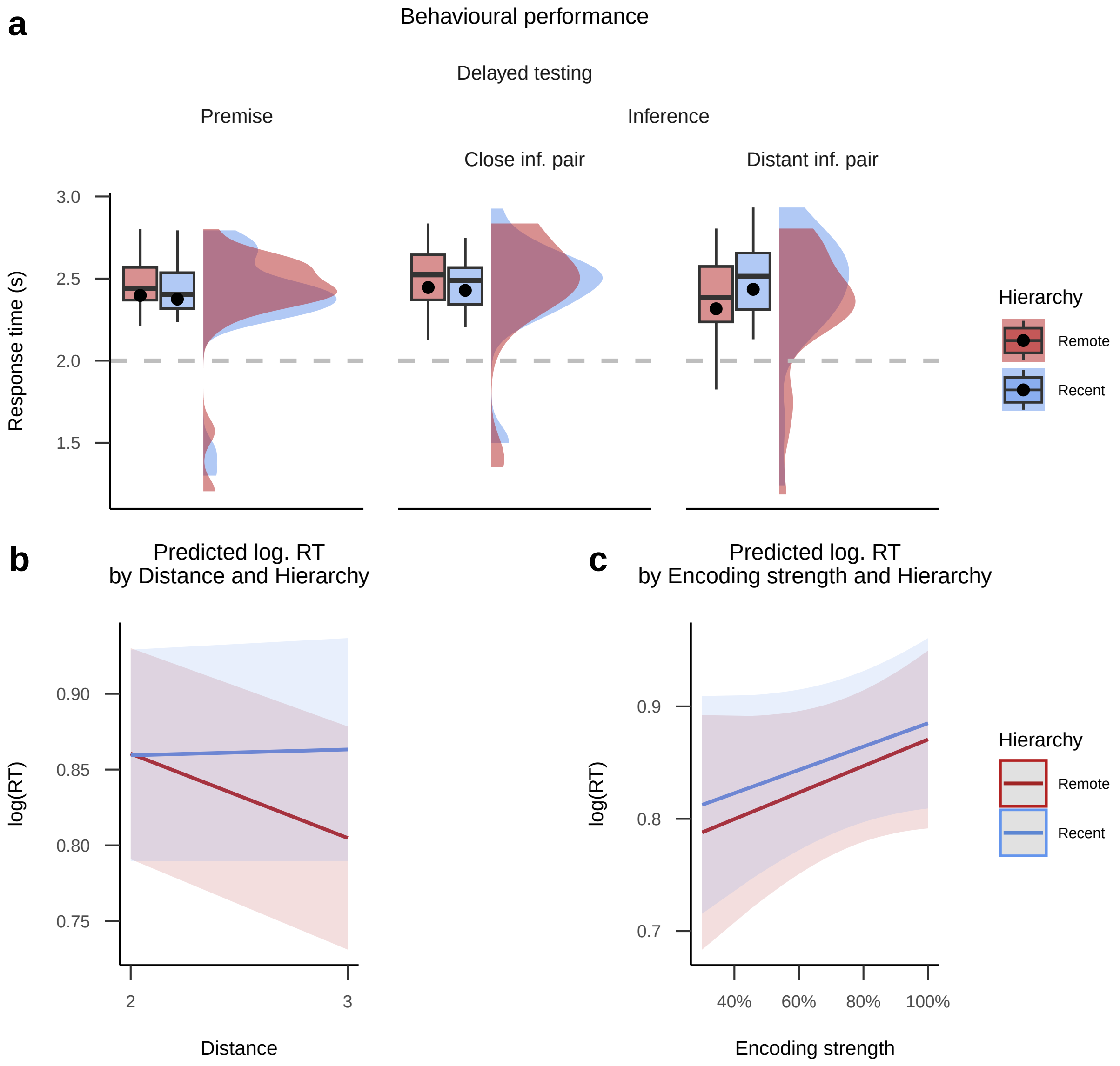


Figure 5: Experiment 4: Behavioural performance and factors predicting RT at delayed test. (a) Raincloud plot with mean and median RT across Remote and Recent hierarchy conditions. Dashed line represents start for response window and dot represents mean values (b) Predicted logRT by Distance and Hierarchy, with distance levels of 2 and 3. (c) Predicted logRT by Encoding strength and Hierarchy. In panels (b) and (c), shadowed areas represent 95% confidence intervals.

To further examine the time-dependent symbolic distance effect, we performed a log transformation on the reaction time data. This transformation was employed to normalize the distribution, reduce the influence of extreme values, and stabilize the variance across observations, which is a common procedure when dealing with reaction time data characterized by positive skewness (for untransformed data see Fig. 5A).

Distance and Encoding strength We fitted a linear mixed model (estimated using REML and nloptwrap optimizer) to predict logRT with Hierarchy and Distance (formula: logRT ~ Hierarchy * Distance). The model included participant as random effect (formula: ~1 | participant). The model's total explanatory power is substantial (conditional R2 = 0.40) and the part related to the fixed effects alone (marginal R2) is of 5.84e-03. The model's intercept, corresponding to Hierarchy = Remote and Distance = 0, is at 0.84 (95% CI [0.77, 0.91], t(912) = 24.20, p < .001). Within this model:

- The effect of Hierarchy [Recent] is statistically non-significant and positive (beta = 0.02, 95% CI [-8.60e-03, 0.05], t(912) = 1.35, p = 0.179; Std. beta = 0.07, 95% CI [-0.03, 0.17])
- The effect of Distance is statistically significant and negative (beta = -0.06, 95% CI [-0.10, -0.01], t(912) = -2.66, p = 0.008; Std. beta = -0.10, 95% CI [-0.17, -0.03])
- The effect of Hierarchy [Recent] × Distance is statistically significant and positive (beta = 0.06, 95% CI [1.56e-03, 0.12], t(912) = 2.02, p = 0.044; Std. beta = 0.10, 95% CI [2.72e-03, 0.21]).

To further investigate the Hierarchy[Recent] × Distance interaction, we calculated the coefficients of partial effects at close (Remote: $M = 0.86$, 95% CI $[0.79, 0.93]$, Recent: $M = 0.86$, 95% CI $[0.79, 0.93]$, $t(889.04) = 0.05$, $p = .958$) and distant (Remote: $M = 0.81$, 95% CI $[0.73, 0.88]$, Recent: $M = 0.86$, 95% CI $[0.79, 0.94]$, $t(889.02) = -2.42$, $p = .016$) pair values of Distance, suggesting greater relative accuracy benefit for distant, compared to close pairs. This is depicted in Fig. 5B.

Additionally we fitted a linear mixed model (estimated using REML and nloptwrap optimizer) to predict logRT with Hierarchy and Encoding strength (formula: logRT ~ Hierarchy * Encoding strength). The model included participant as random effect (formula: ~1 | participant). The model's total explanatory power is substantial (conditional R2 = 0.39) and the part related to the fixed effects alone (marginal R2) is of 6.32e-03. The model's intercept, corresponding to Hierarchy = Remote and Encoding strength = 0, is at 0.75 (95% CI [0.60, 0.90], t(912) = 9.95, p < .001). Within this model:

- The effect of Hierarchy [Recent] is statistically non-significant and positive (beta = 0.03, 95% CI [-0.12, 0.18], t(912) = 0.38, p = 0.701; Std. beta = 0.07, 95% CI [-0.04, 0.17])

- The effect of Encoding strength is statistically non-significant and positive (beta = 0.01, 95% CI [-5.63e-03, 0.03], t(912) = 1.33, p = 0.184; Std. beta = 0.08, 95% CI [-0.04, 0.19])
- The effect of Hierarchy [Recent] × Encoding strength is statistically non-significant and negative (beta = -1.44e-03, 95% CI [-0.02, 0.02], t(912) = -0.15, p = 0.882; Std. beta = -9.40e-03, 95% CI [-0.13, 0.11])

Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using a Wald t-distribution approximation. This is depicted in Fig. 5C.

5.4.2 Imaging Results

RSA Searchlight: Distance similarity

Localizer In our searchlight analysis, our primary aim was to identify regions with above-chance rank-order structure for the Remote hierarchy. This was to determine whether this structure could support the observed time-dependent inference benefit. We identified significant activation in the right PPC (peak MNI coordinates: x = 46.18, y = -53.43, z = 21.80; cluster size: 56 voxels). This activation exhibited a peak Z-value of 4.16. Importantly, this was significant at a whole gray matter mask with a cluster-corrected p-value of $p < 0.05$. Furthermore, we detected significant activations in the left mOFC (peak MNI coordinates: x = -1.96, y = 54.87, z = -15.60; cluster size: 35 voxels; peak Z-value = 4.22), left PHC (peak MNI coordinates: x = -22.58, y = -41.39, z = -10.50; cluster size: 12 voxels; peak Z-value = 3.80), and right IFG (peak MNI coordinates: x = 53.05, y = 23.93, z = 4.80; cluster size: 21 voxels; peak Z-value = 3.51). These findings are illustrated in Fig. 6. Complementing this, we also sought evidence of structural encoding for the Recent hierarchy within our predefined regions of interest. A notable cluster was observed solely in the mOFC (peak MNI coordinates: x = -3.67, y = 13.61, z = -17.30; cluster size: 21 voxels; peak Z-value = 3.65). All reported results are consistent with a cluster-corrected $p_{\text{FWE}} < 0.05$.

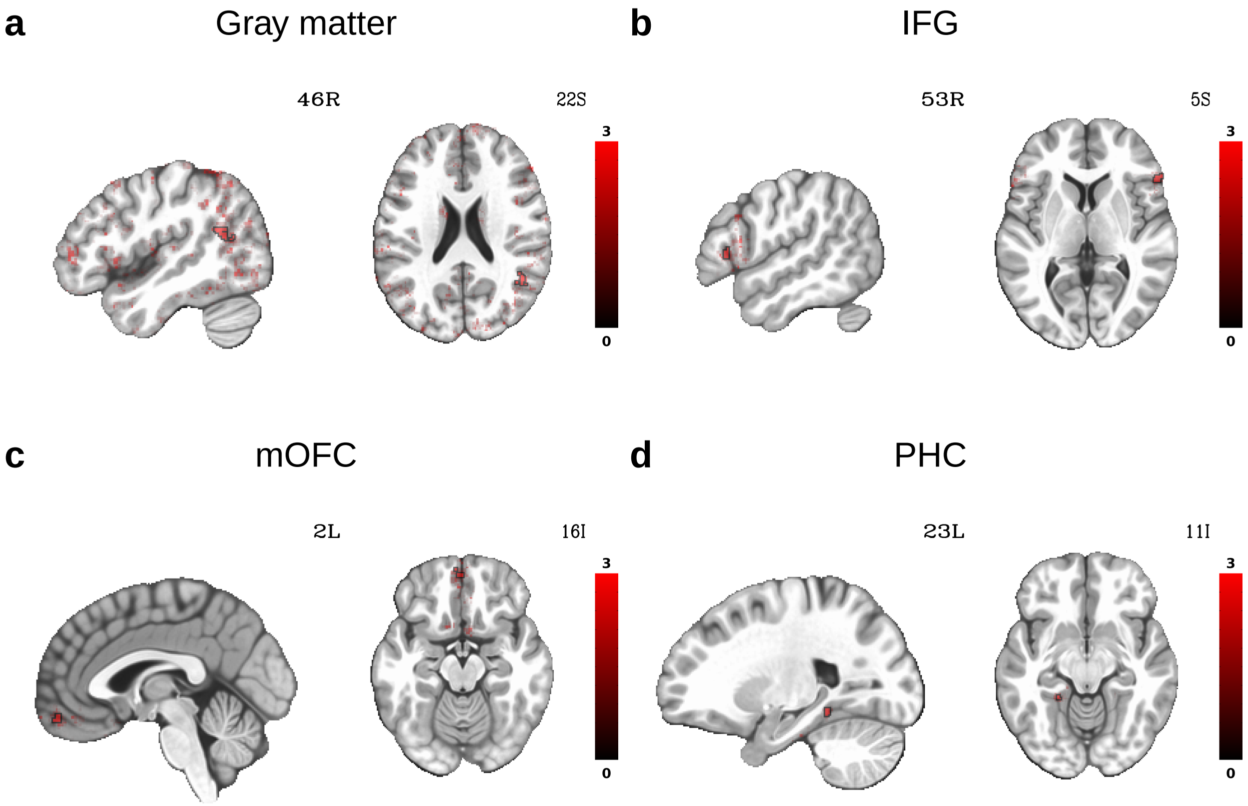


Figure 6: Experiment 4. Searchlight results: Remote > chance (task-Localizer). Searchlight results for the contrast Remote > chance are depicted, indicating distance similarity statistics greater than 0. Subplots (A) through (D) present cluster-corrected effects within the general gray matter, Inferior Frontal Gyrus (IFG), medial Orbitofrontal Cortex (mOFC), and Parahippocampal Cortex (PHC) masks, respectively. All images display Z values as highlighted by the colorbar, with results being cluster-corrected ($p_{\mathrm{FWE}}<0.05$, voxelwise $p<0.005$). Opacity of Z coefficients is modulated by the Z-statistic, as recommended by Taylor and colleagues (2023). Significant clusters are outlined in black, quadratic opacity fading is applied to below cluster threshold voxels.

Next, to determine whether there is increased evidence of rank-order structure in cortical areas consistent with system consolidation theories, we conducted a paired t-test contrasting the distance similarity measures observed for the Remote and Recent hierarchies. We once again identified significant activation in the right PPC (peak MNI coordinates: x = 44.46, y = -55.15, z = 20.10; cluster size: 78 voxels), with a peak Z-value of 3.91. Additional significant activations were found in the left mOFC (peak MNI coordinates: x = -8.83, y = 53.15, z = -12.20; cluster size: 18 voxels; peak Z-value = 3.85) and right IFG (peak MNI coordinates: x = 54.77, y = 20.49, z = 16.70; cluster size: 30 voxels; peak Z-value = 3.44). For our HPC ROI, a two-tailed paired t-test revealed a negative cluster, suggesting stronger encoding for the more recent hierarchy compared to the remote hierarchy (peak MNI coordinates: x = -19.15, y = -15.61, z = -17.30; cluster size: 21 voxels; peak Z-value = -4.21). These results are depicted in Fig. 7.

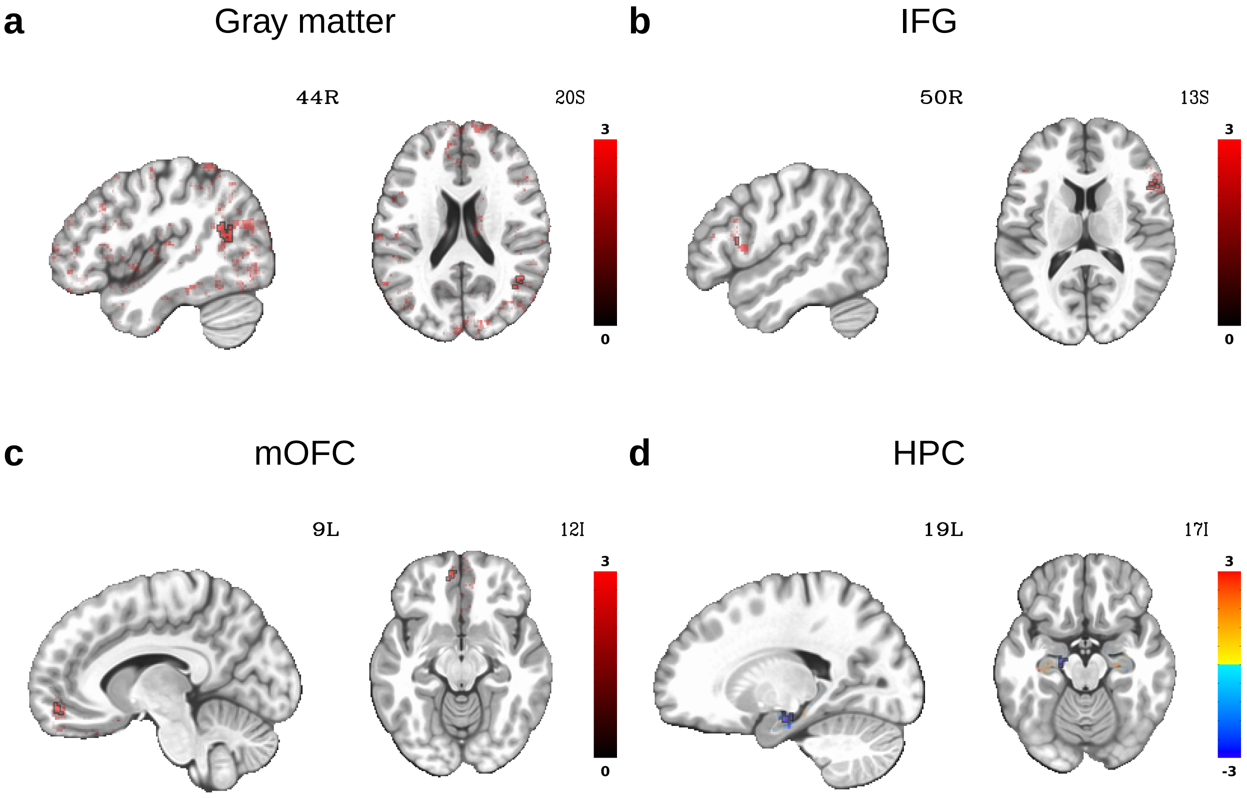


Figure 7: Experiment 4. Searchlight results: Remote > Recent (task-Localizer). Searchlight results for the contrast Remote > Recent are depicted, indicating a relative difference in distance similarity statistics. Subplots (A) through (D) present cluster-corrected effects within the general gray matter, Inferior Frontal Gyrus (IFG), medial Orbitofrontal Cortex (mOFC), and Hippocampus (HPC) masks, respectively. All images display Z values as highlighted by the colorbar, with results being cluster-corrected ($p_{\mathrm{FWE}}<0.05$, voxelwise $p<0.005$). Opacity of Z coefficients is modulated by the Z-statistic, as recommended by Taylor and colleagues (2023). Significant clusters are outlined in black, quadratic opacity fading is applied to below cluster threshold voxels.

TI While our in-scanner experiment was primarily tailored to ensure optimal power for the localizer task, we sought to explore whether our pattern similarity measure would exhibit analogous results during the Transitive Inference task within our principal regions of interest, namely the HPC and mOFC. We found significant activation in the mOFC for the Remote > chance contrast (peak MNI coordinates: x = 6.64, y = 53.15, z = -7.10; cluster size: 76 voxels; peak Z-value = 4.22) as well as the Remote > Recent contrast (peak MNI coordinates: x = 8.36, y = 53.15, z = -8.80; cluster size: 110 voxels; peak Z-value = 3.97). These results are depicted in Fig. 8. For a list of searchlight results see Table 1 and Table 2.

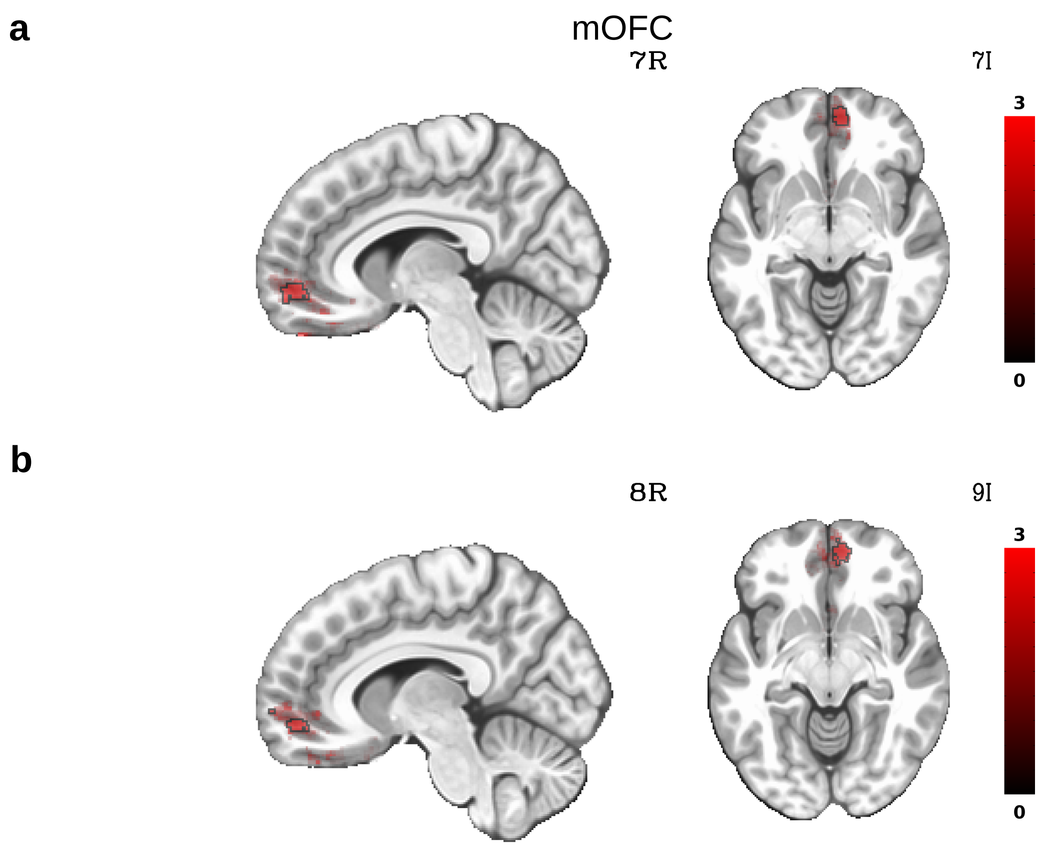


Figure 8: Experiment 4. Searchlight results: task-TI. A) Searchlight results for the contrast Remote > 0 are depicted, indicating distance similarity statistics greater than 0 B) Searchlight results for the contrast Remote > Recent are depicted, indicating a relative difference in distance similarity statistics. All images display Z values as highlighted by the colorbar, with results being cluster-corrected ($p_{\text{FWE}}<0.05$, voxelwise $p<0.005$). Opacity of Z coefficients is modulated by the Z-statistic, as recommended by Taylor and colleagues (2023). Significant clusters are outlined in black, quadratic opacity fading is applied to below cluster threshold voxels.

Brain-behavior correlations

For our brain-behavior correlation analysis, we used lateral masks aligned with the regions where significant clusters emerged in our primary analysis. These lateral masks were based on the anatomical ROIs and serve to focus the analysis on the specific subregions of interest. This decision aims to hone in on the areas most likely engaged by our task while still avoiding potential circularity issues and to align our findings more closely with established neuroanatomical regions, enhancing the interpretability and reproducibility of our results. Notably, we found a negative correlation between mean value across the left HPC mask and Remote inference performance ($r = -.41$, 95% CI $[-.69, -.03]$, $t(24) = -2.23$, $p = .035$) and a positive correlation in the left PHC ($r = .39$, 95% CI $[.00, .68]$, $t(24) = 2.08$, $p = .048$). This supports the hypothesis of hippocampal disengagement and is indicative of relational memory consolidation and that time-dependent inference benefits are supported by cortical structures.

To further evaluate the specificity of our results, we performed a parallel analysis, correlating within-subject inference performance with distance similarity difference scores, using the entire ROI masks for both the HPC and PHC. We found similar result in both HPC ($r = -.41$, 95% CI $[-.69, -.03]$, $t(24) = -2.23$, $p = .036$) and PHC ($r = .34$, 95% CI $[-.06, .64]$, $t(24) = 1.75$, $p = .094$), however latter finding was not significant (Fig. 9).

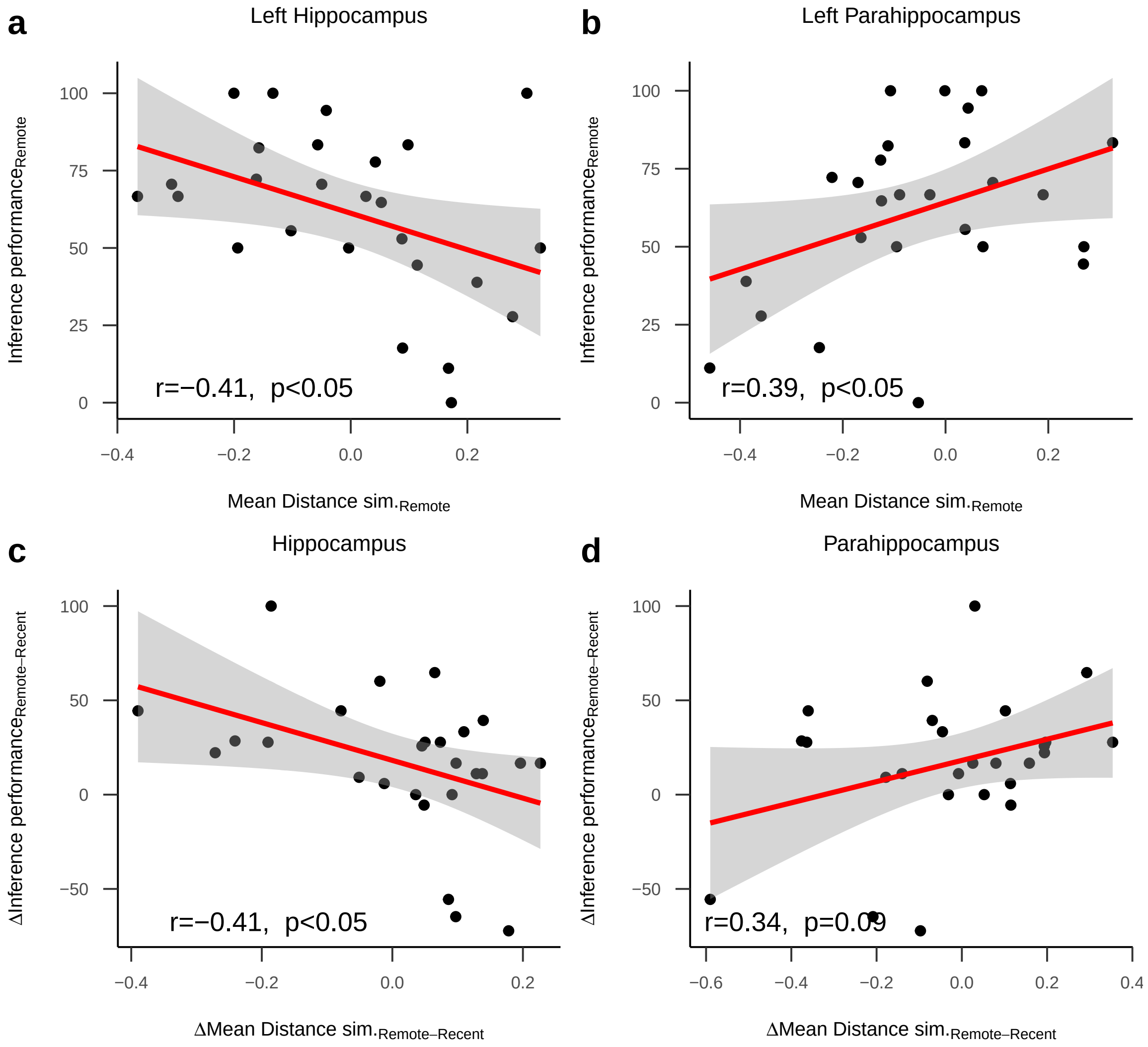


Figure 9: Experiment 4. Brain-behavior correlation: task-Localizer (A) Correlation between Mean Remote distance similarity within the Left Hippocampal mask and Remote inference performance. (B) Similar correlation as in (A) but for the Left Parahippocampal Cortex. (C) Correlation between Mean $\Delta_{\text{Remote-Recent}}$ distance similarity and $\Delta_{\text{Remote-Recent}}$ inference performance within the bilateral Hippocampus. (D) Identical correlation as in (C) but observed within the bilateral Parahippocampal Cortex. Shadowed areas represent 95% confidence intervals.

We next investigated the correlations between distance similarity and behavior in our transitive inference (TI) task. We utilized our PPC (Right Angular Gyrus) cluster, which notably survived rigorous whole-brain gray mask correction from our earlier Localizer task, as our functional region of interest (ROI). We specifically examined if mean Remote distance similarity in this ROI correlated with inference performance for the Remote condition. Additionally, we conducted a similar difference score analysis as we had done previously.

While we did not find significant positive relationship with Remote inference behavior ($r = .09$, 95% CI $[-.31, .46]$, $t(24) = 0.45$, $p = .659$), we did find a positive difference effect ($r = .43$, 95% CI $[.05, .70]$, $t(24) = 2.32$, $p = .029$, Fig. 10). This suggests that the relative differential in pattern similarity within this functional ROI is indicative of relative differential in inference performance at delayed test. Specifically, participants who showed a bigger increase in Remote distance similarity relative to Recent distance similarity also showed a bigger time-dependent benefit in inference performance. This highlights the importance of identifying multivariate neural patterns in tracking the efficiency of cognitive processes such as inference.

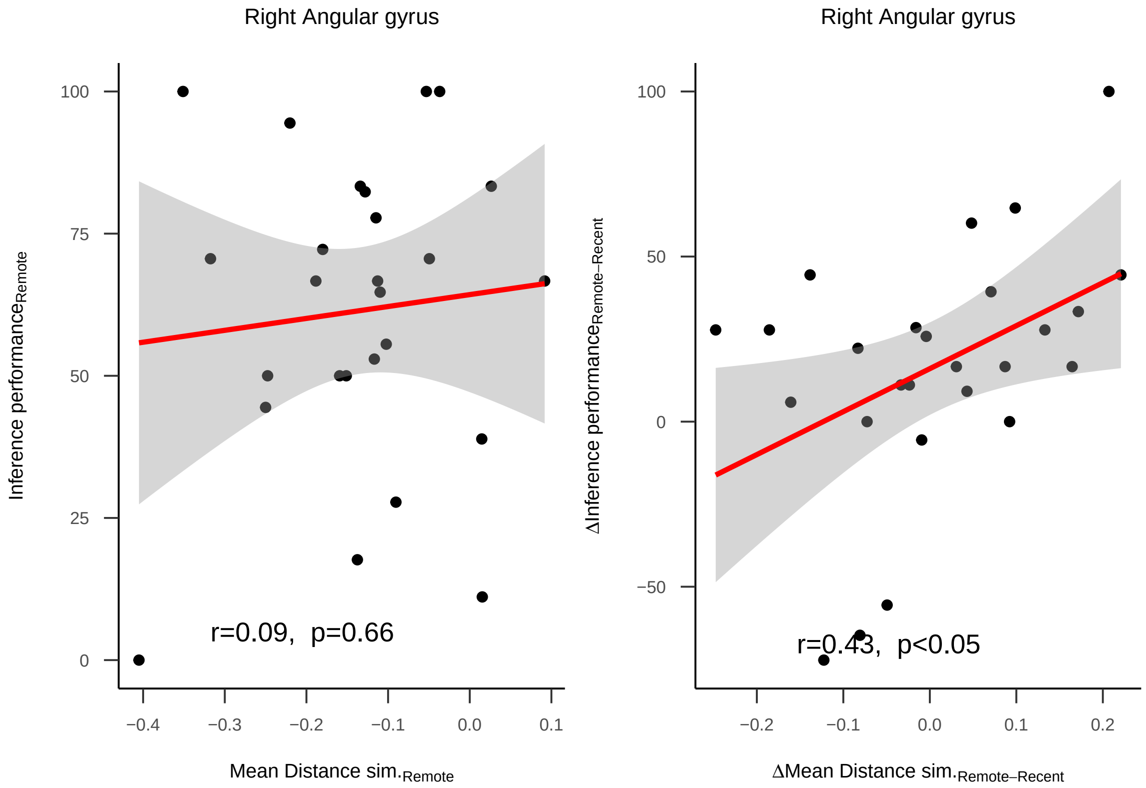


Figure 10: Experiment 4. Brain-behavior correlation: task-Localizer. (A) Correlation between Mean Remote distance similarity within the Right Entorhinal Cortex and Remote inference performance. (B) Similar correlation as in (A) but for the Left Entorhinal Cortex. (C) Correlation between Mean $\Delta_{\text{Remote-Recent}}$ distance similarity and $\Delta_{\text{Remote-Recent}}$ inference performance within the Right Entorhinal Cortex. (D) Identical correlation as in (C) but observed within the Left Entorhinal Cortex. Shadowed areas represent 95% confidence intervals.

5.5 Discussion

Over the course of our two-day experimental design, participants engaged in encoding two distinct transitive hierarchies that offered differing windows for sleep-dependent consolidation. Specifically, they studied adjacent pairs either 27 hours or a mere 3 hours prior to their recall, which we then monitored within a functional magnetic resonance imaging (fMRI) scan session.

Our data yielded several compelling insights. Firstly, there was a clear distinction in memory performance based on the elapsed time since encoding. Participants showcased a notably superior inference performance for the hierarchy learned 27 hours prior compared to the one learned just 3 hours before. This suggests that the period of consolidation plays a pivotal role in the retrieval efficacy of encoded hierarchy in transitive inference. This is in line with Experiment 3 (called Experiment 2 in Foldes et al. (2023)). Building on this, we observed an interesting dynamic in reaction times. This was manifested in faster reaction times for distant inference pairs as opposed to close ones, alluding to a time-dependent symbolic distance effect indicative of heightened memory integration. This provides evidence of the cognitive evolution that relational memories undergo over time, becoming more abstract and integrated with elapsed consolidation.

fMRI data from our Localizer task further enriched these findings, revealing distinct neural similarities consistent with a model of symbolic distance, or the so-called distance similarity We found that structured, generalized memory traces were present in specific regions: the Posterior Parietal Cortex (PPC/Angular gyrus), medial Orbitofrontal Cortex (mOFC), Inferior Frontal Gyrus (IFG), and Parahippocampal Cortex (PHC). It's noteworthy that such patterns were exclusive to the remotely learned hierarchy with the exception of mOFC which showed signs of rank-order representations for the recent hierarchy as well. Further, when contrasting the distance pattern similarity scores of the two hierarchies, we discovered heightened rank-order representations in the majority of the cortical areas, namely in the PPC, mOFC, and IFG. However a contrasting reverse effect emerged in the Hippocampus (HPC). A key highlight was the discernible manifestation in the mOFC during recall in the transitive inference task, echoing our findings from the Localizer task.

Lastly, the interplay between neural activity and performance was evident in the correlations we identified. A significant negative correlation existed between the distance similarity and the inference performance in the HPC, when considering the difference between the remote and recent hierarchies. On the other hand, the PPC demonstrated a positive correlation.

Interpreting our findings necessitates a comparative analysis with the work of Berens & Bird (2022b), who executed a closely aligned 2-day study on transitive inference. Notably, their "interleaved" training and retention interval conditions bear resemblance to the training scheme we adopted, underscoring the value of comparing results. In a pattern that mirrors their findings, our study also showcased an escalation in reaction time determined by the time of encoding and symbolic distance. This convergence, combined with the distance similarity we discerned during the localizer task, nudges us towards positing an encoding-based model as the primary mechanism underlying transitive inference, given that we observe rank-order structure in the memory traces, absent of inference task requirements and that the RT during inference trials is indicative of SDE. By encoding-based mechanism we refer to an overlap-based integration taking place over time, following encoding, although we cannot rule out encoding-based mechanisms at encoding accounting for the superior Remote performance due to e.g., reduced interference.

It's pivotal to note that while Berens & Bird employed a distinct method to compute "distance similarity", for the sake of our analysis, we'll consider both measures equivalent. Their observations revealed that the HPC predominantly encodes structural task representations when behavioral performance doesn't typically reflect encoding-based generalization. In contrast, the medial Prefrontal Cortex (mPFC) displayed a converse pattern, with the anticipated distance effects primarily manifesting in the remote condition. This was particularly evident when participants' behavioral data strongly suggested encoding-based generalizations. This trend parallels our findings concerning the differences in correlations within the HPC and mOFC.

Nevertheless, a salient divergence between our study and theirs emerges concerning the time-dependent improvement in inference accuracy. Our results align more with prior transitive inference research on time and sleep-dependent consolidation (Ellenbogen et al., 2007; Werchan & Gómez, 2013c). The root cause of this discrepancy, as well as a more in-depth discussion, is explored comprehensively in Chapter 2 and 3.

System consolidation theory (SMC) posits that memories initially anchored in the hippocampus eventually migrate to the cortex through a process that involves repeated reactivation, especially during rest and sleep. Our observation of enhanced memory performance for the hierarchy encoded 27 hours prior (as opposed to the 3-hour counterpart) complements this idea. The time-window potentially allowed for sleep-dependent consolidation, where the hippocampus might have actively replayed the memories, facilitating their eventual strengthening and stabilization in cortical structures. The pronounced generalized memory traces in cortical regions, particularly in the PPC, mOFC, and IFG for the 27-hour hierarchy, seem to underscore SMC's assertion. The memories seem to have transitioned from a hippocampal-centric locus (evident from the contrasting reverse effect in the hippocampus) to a more distributed cortical representation. The significant correlations between the $\Delta_{\text{distance similarity}}$ and $\Delta_{\text{inference performance}}$, especially the negative correlation in the HPC, might not only hint at the decreasing reliance on the hippocampus, but that for the recently learned hierarchy, there's an ongoing process where related memories are becoming integrated, which helps in abstracting common elements as predicted by the Trace Transformation Theory (Sekeres, Winocur, & Moscovitch, 2018b). Follow-up analyses will have to clarify negative correlation between the remote distance similarity and inference performance. Orthogonalized memory traces have been reported in the posterior HPC (Berens, Horst, & Bird, 2018; Kwon, Lee, & Lee, 2023; Margaret L. Schlichting, Mumford, & Preston, 2015b) and could produce the observed negative distance similarity scores during our Localizer task.

Emerging evidence suggests that the HPC, OFC, and EC play pivotal roles in the learning, abstraction, and application of explicit structural representations. These representations are instrumental in guiding novel inferences and generalizations after extensive training (Erie D. Boorman et al., 2021a; Morton & Preston, 2021). However, our findings highlight that following a prolonged retention interval, significant neural and behaviorally predictive changes manifest predominantly in the Posterior Medial Episodic (PME) network. This is particularly evident in regions like the posterior hippocampus, PPC, and parahippocampal cortex. The PME network is intrinsically linked with placing events in spatial contexts and recalling the contextual nuances of episodic memories (Davachi, 2006; Diana, Yonelinas, & Ranganath, 2007; Ranganath & Ritchey, 2012; Ritchey & Cooper, 2020). This is in contrast to the Anterior Temporal network, which is more geared towards retrieving distinct item-specific details from episodic memories and processing semantic memory. The observed temporal shifts in the PME further bolster the encoding-based perspective on transitive inference. This stands in contrast to retrieval-based models which posit that the time-dependent advantages in transitive inference originate from strengthened connection between item and conjunctive representations (Kumaran & McClelland, 2012a).

A noteworthy limitation of the current transitive inference task is its inability to distinctly ascertain whether the regions of interest represent all pairwise symbolic distances during retrieval, or if they merely depict the absolute ranks of each item. This uncertainty mirrors the longstanding debate regarding the role of the OFC in either predicting and monitoring the value of options when deciding under uncertainty (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Padoa-Schioppa & Assad, 2006), or in the encoding of cognitive maps (Behrens et al., 2018; Morton & Preston, 2021). To truly disentangle these possibilities, variations of the current paradigm are required. Specifically, those that emphasize the flexible manipulation of a learned rank-order structure, such as the transfer of transitive knowledge between sets (S. Chen, Swartz, & Terrace, 1997; Gazes, Lazareva, Bergene, & Hampton, 2014; Kao, Jensen, Michaelcheck, Ferrera, & Terrace, 2018), the successful merger of two hierarchies into a combined 2D cognitive map (Erie D. Boorman et al., 2021a; Park et al., 2021, 2020a), the utilization of a reversed reward gradient (Jensen, Alkan, et al., 2019) or combined categorization and serial learning (Jensen et al., 2020; Luo & Zhao, 2018).

Moreover, our reliance on fMRI data solely during the retrieval phase limits the depth of our insights into relational encoding. To more definitively pinpoint the role of learned knowledge representations, it would be prudent for future research to incorporate inference testing at immediate test and the recording of neural activity throughout the learning process and during rest and/or sleep. Elements of such an approach has been effectively employed in associative inference (Margaret L. Schlichting & Preston, 2016b) and sequential integration paradigms (Tompary & Davachi, 2022). Furthermore, when considering knowledge structures during retrieval, the mere detection of these structures doesn't inherently validate their necessity or relevance for task performance. To further elucidate the direct implications of these structures, it would be beneficial to explore neural correlates that reflect behaviorally predictive trial-by-trial variations in task-set representations. Additionally, empirical efforts that assess the effects of perturbing these representations on learning and/or retrieval would offer invaluable insights (Wikenheiser, Marrero-Garcia, & Schoenbaum, 2017).

Lastly, a logical and potentially enlightening approach would involve comparing neural representations across different species while utilizing suggested extensions of this task. Such cross-species comparisons could provide deeper insights into the universality and evolutionary aspects of these cognitive processes (Barron et al., 2020; Jensen et al., 2017b).

In summary, our study provides compelling evidence that time-dependent memory consolidation plays a pivotal role in transitive inference. In two-day transitive inference experiment, we observed superior inference performance and faster reaction times on distant inference trials for hierarchies encoded further in the past compared to recently learned ones. fMRI analyses revealed pronounced distance-based representations in cortical areas like the PPC, mOFC and PHC for remote hierarchies, indicative of systems consolidation and model-based trace transformation. The negative correlation between hippocampal distance similarity and inference performance substantiates this perspective.

Overall, our work aligns with and expands on prior findings underscoring the benefits of consolidation on relational reasoning. Our paradigm and analytic approach could aid future efforts to track the temporal trajectory of relational memories. This can offer insights into how we extract commonalities from experiences to fuel generalization and transfer of knowledge.

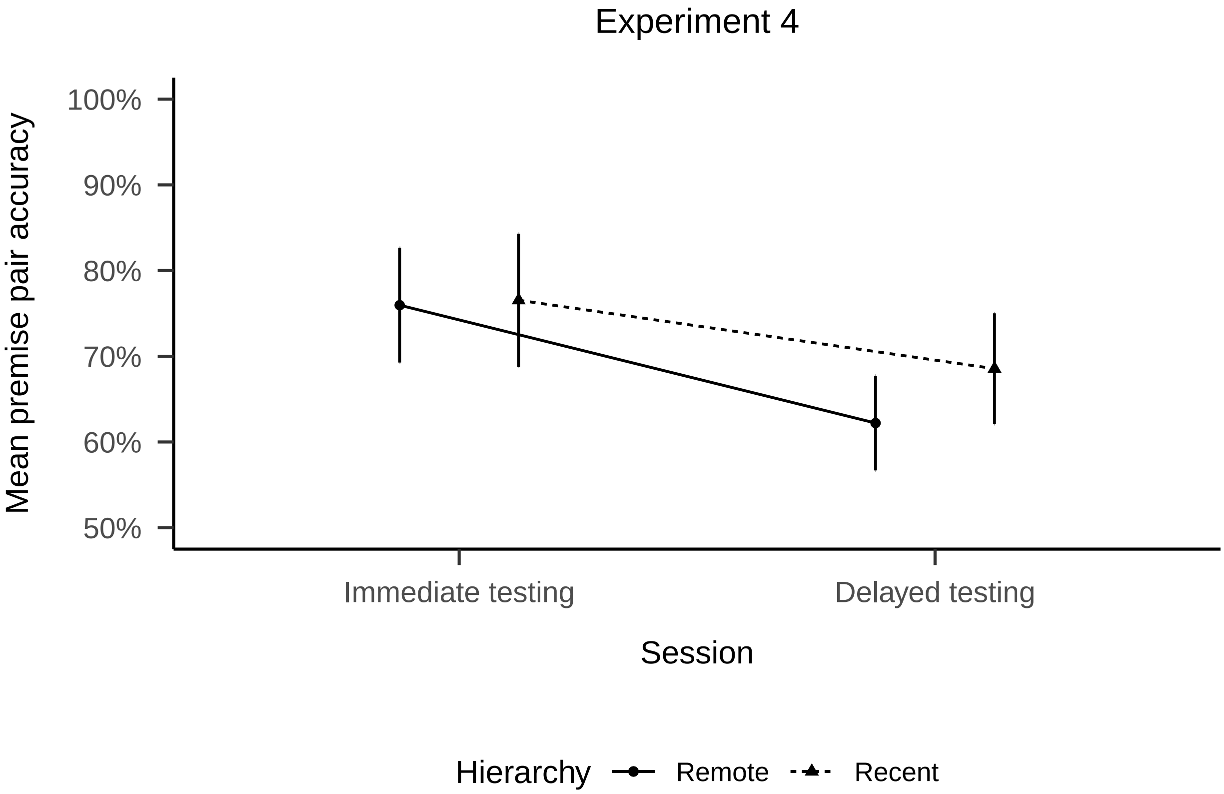
5.6 Supplemental materials

Supplemental Table S4: Experiment 4: Descriptive Statistics for Delayed testing performance by Remote and Recent conditions. The table presents the median, mean, and standard deviation (SD) for performance during delayed testing, grouped by Remote and Recent conditions. Data is further broken down by Pair values.

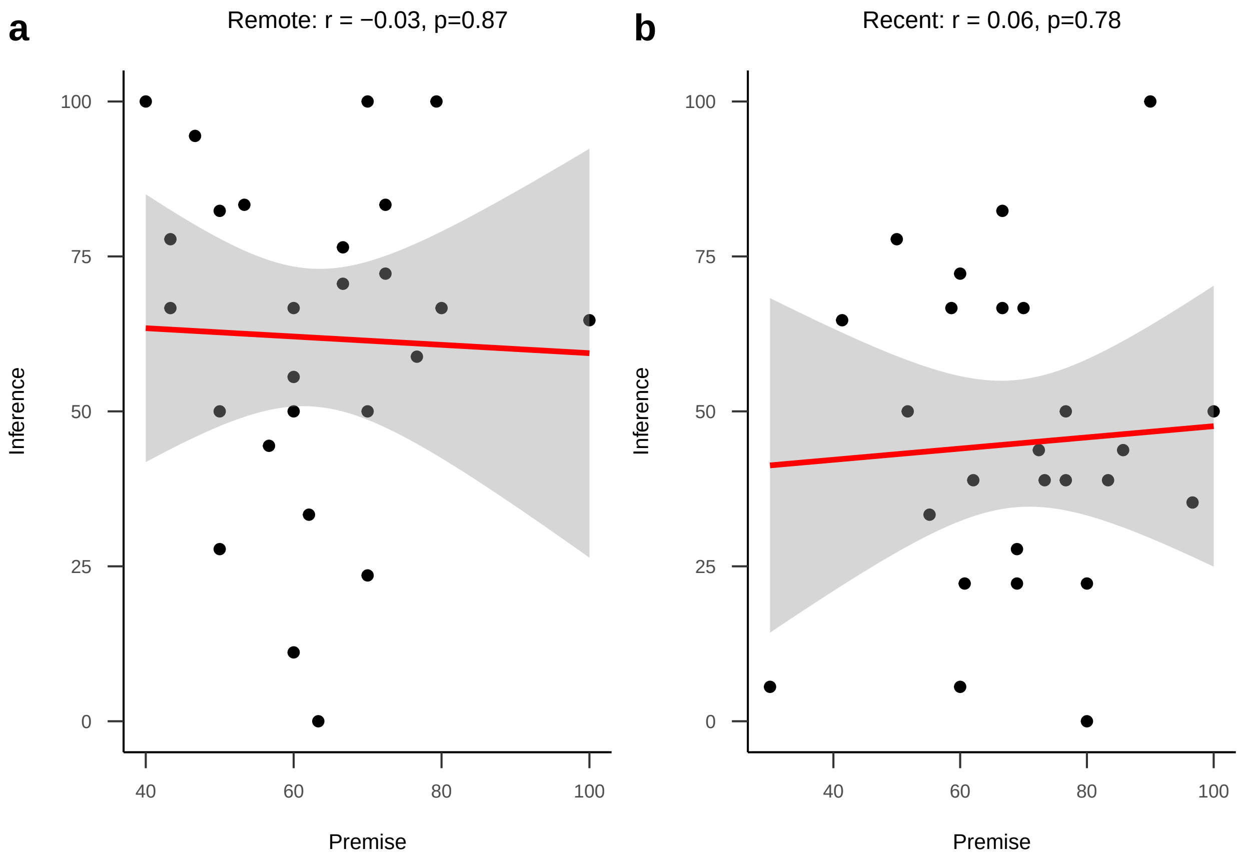
		Hierarchy					
		Remote			**Recent**		
Session	Pair	Median	Mean	SD	Median	Mean	SD
Delayed testing	AB	83.33	69.87	40.01	100.00	83.97	27.28
Delayed testing	BC	83.33	64.74	39.53	66.67	60.77	37.62
Delayed testing	CD	45.00	51.41	40.79	66.67	59.55	40.18
Delayed testing	DE	66.67	59.62	39.21	83.33	59.74	42.34
Delayed testing	EF	83.33	65.38	42.14	100.00	78.72	29.52
Delayed testing	BD	66.67	64.74	36.61	45.00	46.73	40.77
Delayed testing	CE	63.33	57.44	30.71	18.33	42.05	43.71
Delayed testing	BE	83.33	63.72	40.93	41.67	45.90	39.69
Delayed testing	AF	83.33	72.31	30.12	100.00	83.85	26.88
Immediate testing	Overall	73.75	75.96	16.57	76.25	76.54	19.21

Supplemental Table S5: Experiment 4: ANOVA table of baseline differences

Effect	$\hat{\eta}^2_G$	90% CI	F	df^{GG}	$df^{\mathrm{GG}}_{\mathrm{res}}$	p
Hierarchy	.011	[.000, .155]	1.04	1	25	.317
Session	.102	[.000, .311]	20.73	1	25	$< .001$
Hierarchy × Session	.008	[.000, .141]	2.07	1	25	.163



Supplemental Figure S5: Experiment 4. Effect of delay. Mean premise pair accuracy for both Remote and Recent conditions across sessions. The error bars represent the standard error of the mean.



Supplemental Figure S6: Experiment 4: Relationship between mean premise pair performance and inference performance in a transitive inference task. Panel (a) shows the correlation for the Remote condition, while panel (b) presents the correlation for the Recent condition. Each data point represents a participant's mean performance. Linear regression lines (in red) depict the trend of the data, shadowed areas represent 95% confidence intervals.

Table 1: Experiment 4. Descriptive statistics for fMRI Searchlight peaks by Contrast and Mask for the Localizer task

Contrast	Mask	Peak x	Peak y	Peak z	Peak value	Volume mm	Nvox	Aal
Remote > 0	Gray	46.18	-53.43	21.8	4.16	281.31	56	Angular R
	Ifg	53.05	23.93	4.8	3.51	105.49	21	Frontal Inf Tri R
	Mofc	-1.96	54.87	-15.6	4.22	175.82	35	Rectus L
	Parahipp	-22.58	-41.39	-10.5	3.80	60.28	12	Lingual L
Recent > 0	Mofc	-3.67	13.61	-17.3	3.65	105.49	21	Olfactory L
Remote > Recent	Gray	44.46	-55.15	20.1	3.91	391.83	78	Temporal Mid R
	Hipp	-19.15	-15.61	-17.3	-4.21	105.49	21	Hippocampus L
	Ifg	54.77	20.49	16.7	3.44	150.70	30	Frontal Inf Tri R
	Mofc	-8.83	53.15	-12.2	3.85	90.42	18	Frontal Med Orb L

Table 2: Experiment 4. Descriptive statistics for fMRI Searchlight peaks by Contrast and Mask for the TI task

Contrast	Mask	Peak x	Peak y	Peak z	Peak value	Volume mm	Nvox	Aal
Remote > 0	Mofc	6.64	53.15	-7.1	4.22	381.78	76	Frontal Med Orb R
Remote > Recent	Mofc	8.36	53.15	-8.8	3.97	552.58	110	Frontal Med Orb R

Supplemental Table S6: Experiment 4. Table of brain-behavior correlations: Correlation between Mean Remote distance similarity and Remote inference performance (task-Localizer).

Mask	Correlation	P value	P fdr
Right IFG	-0.103	0.615	0.769
Left Hippocampus	-0.414	0.035	0.120
Left mOFC	-0.172	0.402	0.670
Right Angular Gyrus	-0.037	0.857	0.857
Left Parahippocampal cortex	0.391	0.048	0.120

General conclusion

This dissertation aimed to investigate the mechanisms of offline reactivation-dependent consolidation and generalization following episodic learning, with a particular emphasis on the transitive inference paradigm. Through a combination of computational modeling, behavioral experiments, meta-analysis, and neuroimaging, several key findings emerged.

In *Chapter 2* a comparative analysis conducted between REMERGE and MINERVA2 underscores the significance of model-based approaches in deciphering similarity-based generalization as a direct consequence of reactivation dynamics. Both models exhibit similar core mechanics, despite their unique characteristics. REMERGE is a recurrent neural network model that calculates similarity between feature and conjunctive codes, while MINERVA2 is a global matching model that represents memories as vectors and calculates similarity between an external (or internally sampled) cue and all stored traces in parallel. However, it is argued that there are compelling reasons to favor MINERVA2 over REMERGE, particularly in the realm of time and sleep-dependent generalization research. Firstly, REMERGE does not propose a learning algorithm that delineates how the conjunctive layer is formed or what the feature-conjunctive weights should be post-learning. This lack of specification can lead to ambiguity in model interpretation and application. Secondly, the application of REMERGE becomes less clear when dealing with tasks where conjunctive nodes are not readily discernible from the task structure, such as in statistical learning tasks. Thirdly, due to the recurrent nature of REMERGE, the model is highly sensitive to initial conditions. This includes the exact tuning of the "external cueing strength", weights, and temperature parameter. This sensitivity could make the model less robust and more challenging to implement across diverse scenarios. Finally, REMERGE does not propose a mechanism for the storage or integration of newly formed generalizations (aka "stored generalizations" or "integrated representations"). This is a crucial component for any model that aims to contribute to our understanding of how representations form or decay over time. The final point bears significant relevance to this thesis due to its implications for delayed transitive inference effects. REMERGE postulates that enhancements in inferential performance, derived from sleep, can be elucidated by a consolidation mechanism that strengthens conjunctive representations, such as premise pairs like AB and BC. Conversely, when configured to store "intermediate" generalizations, MINERVA2 can model sleep and time-dependent benefits in inferential performance flexibly, without the assumption of bolstered premise pair performance at delayed recall. This represents a crucial, empirically testable distinction between the two models' predictions. Furthermore, through a modest number of simulations, this chapter illustrates that MINERVA2 can replicate findings related to time and sleep-dependency from paired associate, transitive inference, and serial reaction time tasks. This replication is achieved by modeling offline reactivation as internally generated cued recall. Notable phenomena such as resistance to retroactive interference and step wise associative chaining were realized through the selective manipulation of activation functions, referred to as MINERVA2 variants. However, not all targets were successfully modeled, as MINERVA2 encountered difficulties replicating the symbolic distance effect in transitive inference tasks, a challenge that REMERGE also faces. Yet, despite these limitations, the comparative analysis underscores the effectiveness of model-based approaches and positions MINERVA2 as a valuable "minimal" exemplar baseline model for time and sleep-dependent consolidation research.

Chapter 3 explored the relationship between the extent of information processing and storage in memory during the encoding phase, referred to as encoding strength, and the performance on the transitive inference task. This task is a popular measure of relational reasoning and involves making inferences about the relationships between items based on learning a hierarchy of neighboring items commonly called premise pairs. Performance is gauged by the accuracy of novel item pairs like B?D , C?E and B?E , tested after a delay. This investigation takes into account retention intervals that either include or exclude sleep periods. Initial studies showed that inference performance and the symbolic distance effect (SDE) - whereby participants respond more quickly and/or accurately as a function of distance between the items along the rank (or "dominance") dimension (e.g.: $\overline{B?D_{accuracy} + C?E_{accuracy}} < \overline{B?E_{accuracy}}$) - increased after retention intervals with sleep compared to wakefulness. However, subsequent studies with modified designs did not reliably observe these effects. An additional measure to SDE is studied to better understand the representations forming over time, termed joint rank effect (JRE). Here instead of focusing on the relative distance between the presented items as probed by SDE, one can ask whether the absolute summed rank of the items affects inference performance ($\overline{B?D_{accuracy}} \lessgtr \overline{B?E_{accuracy}} \lessgtr \overline{C?E_{accuracy}}$). This provides a systematic approach to studying whether participants are biased towards representing one end of the hierarchy more accurately than the other end (e.g.: high ranking items). The chapter discusses three experiments that investigate the effects of lowered learning criterion, broader stimulus sets, between- and within-subject designs on sleep-related TI benefits. The first two experiments employed AM-PM designs: one session is conducted in the morning (AM), and the other is conducted in the evening (PM). Experiment 1 (lab-based, between-subjects design) found increased TI performance, SDE an after 12 hour interval with sleep compared to wake. Experiment 2 (web-based, between-subjects design) did not find sleep benefits in TI or SDE, partially due to above chance inference performance of the "wake" group. Experiment 3 (web-based, within-subject design) controlled for time of day effect and found increased TI performance after 27 hour interval including sleep compared to 3 hour interval wake. Again no sleep-dependent SDE was found. Sleep-dependent JRE effects were observed in Experiment 1 and 3, whereby following periods containing sleep, participants' inference performance was increased for the dominant inference pair (B?D) compared to low dominance inference pair (C?E). Notably, when combining all three datasets (N=167) in a mixed-effects logistic regression a small, but significant three-way interaction was observed between type of retention interval (includes sleep/does not), encoding strength and joint rank . This interaction suggests that the impact of a period containing sleep on inference performance is moderated by the absolute summed rank of the inference items and varies across different encoding strength scores. When unpacking this interaction, it was shown that for inference involving the high dominance pair (B?D), a period of sleep had a significant and positive effect on accuracy, and this effect increased as pre-delay encoding strength increased. In contrast, for inference involving the low dominance item (C?E), sleep did not lead to higher post-delay performance, regardless of encoding strength. Additionally an inverse joint rank effect was found after sleep, suggesting a rank-based reorganization of relational knowledge. Overall, the studies provide strong support for sleep-related improvements in TI based on pre-sleep encoding strength. The fact that two out of three studies did not find the symbolic distance effect could suggest that the sleep-dependent SDE may not replicate well across studies and depend on specific designs (e.g.: lab-based, between-subject designs). The results also suggest that the joint rank effect may be a more robust measure of sleep-related reorganization. The chapter concludes with a discussion of the implications of these findings for the role of sleep in relational memory consolidation.

Chapter 4 presents re-analyses of two already published datasets examining time and sleep-dependent consolidation effects on transitive inference performance as well as a mini meta-analysis involving 15 studies by six different labs in an effort to further explore the possible presence and impacts of encoding strength, symbolic distance, and joint rank effects. The two datasets were initially collected and analyzed for different research questions, and as such, the potential effects of encoding strength, and joint rank were not directly investigated in the original studies, whilst symbolic distance was only examined in one. The rationale for this reanalysis was twofold. First, given the existing datasets' richness and comprehensiveness, they provided an opportunity to examine these effects in a context where they were not the primary focus, thus potentially offering novel insights. Second, by reanalyzing these datasets, we aimed to utilize existing resources efficiently and ethically, reducing the need for additional data collection and participant involvement. This reanalysis should be seen as exploratory, aimed at uncovering potential effects that may have been overlooked in the original analyses. The study conducted by Berens & Bird (2022) using a within-subject design, showed an inverse time-dependent consolidation effect, with better inference performance in the Recent condition than the Remote condition. This is contrary to the results from Experiment 3 that employed a similar design and showed the opposite pattern. We were unable to test for an interaction between the retention interval and encoding strength because participants were trained to ceiling. There was no evidence for sleep-dependent symbolic distance effect, but we did find a joint rank effect that is time-dependent - although opposite to what we expected - suggesting that participants did better on the Recent hierarchy as a function of joint rank. Reanalyzing data by Matorina and Poppenk (2021) we did not find a main effect of retention interval in a multi-session within-subject design, suggesting that for their study, the length of the retention interval did not impact the results. Additionally we did not find a sleep-dependent symbolic distance effect. We did find however an interaction between the retention interval and encoding strength, which corroborated our previous findings of stronger pre-sleep encoding strength resulting in better post-delay inference performance in combined dataset analysis (composed of Experiment 1,2 and 3). Further, we found a joint rank effect that was time-dependent, indicating that the greatest time-dependent benefit was focused on the dominant B?D inference pair, which is again in line with what was observed in the combined dataset analysis. In order to mitigate potential Type I and Type II errors due to variable sample sizes across our five experiments, we conducted a mini meta-analysis. This analysis integrated findings from studies that examined the role of time and sleep in transitive inference consolidation. This approach enhances our interpretation by providing a more robust statistical overview. The meta-analysis showed that inference performance was significantly above chance overall at delayed test. Furthermore, a meta-regression was conducted to predict inference performance based on the Retention interval (Sleep/no sleep). Although the results were not statistically significant, they did suggest a positive relationship between sleep and improved inference performance at the delayed test. Additionally, our analysis showed descriptively that increased encoding strength seemed to further enhance the benefits of sleep, in line with findings in Experiment 1,2, 3 and the Matorina & Poppenk (2021) study. We also observed small descriptive effect of superior performance on distant pairs as compared to close pairs when the retention interval included sleep. However, these findings were again descriptive in nature. In a similar vein, the results regarding sleep-dependent joint rank were not definitive.

Chapter 5 presents an fMRI study examining time-dependent changes in representational similarity (RSA) of two 6-item hierarchies encoded on two consecutive days, following a similar design to that of Experiment 3. Behavioral results showed superior inference for remote versus recent hierarchy, suggesting consolidation benefits. While accuracy did not, reaction times did exhibit symbolic distance effects only for the remote hierarchy indicative of distance-based reorganization. A searchlight RSA revealed stronger distance-based representations for the remote hierarchy in PPC, mOFC, IFG and PHC during stimulus viewing ("localizer task"). The hippocampus however showed an unexpected inverse pattern, a negative correlation with the *a priori* distance model. In addition, there was a significant negative correlation between ΔNeural distance similarity$_{\text{Remote-Recent}}$ and ΔInference performance$_{\text{Remote-Recent}}$ inference performance in the left hippocampus and positive correlation between the two measures in the right angular gyrus. The shift towards integrated cortical representations is argued to support time-dependent emergence of model-based inference, aligning with systems consolidation theory, whilst the hippocampal pattern can either signal ongoing consolidation of the recent hierarchy or a type of distance-based pattern separation of the remote hierarchy, where the closer items are in latent rank the more separated they become over time. Limitations around assessing representations during encoding and probing flexible use of learned structure are noted. In summary, results demonstrate reorganization of latent relational structure over time towards more integrated mnemonic representations. The study highlights the benefit of model-based multivariate fMRI analyses in tracking memory transformation processes.

General discussion

Together, these findings make important contributions to our comprehension of offline reactivation-dependent consolidation. The adoption of diverse techniques paves the way for interdisciplinary collaborations and rigorous experiments situated at the intersection of competing theories of generalization.

In Marrian terms the overarching computational level question can be summarized as “Why are periods of quiet rest/sleep important for enabling generalization across different experiences or learned information?”. Both wakeful rest and sleep are thought to play a crucial role in reorganizing memories of wakeful experiences. This process involves the consolidation and assimilation of experiences, as well as the generation of novel inferences from learned relationships. Given its evolutionary conservation, sleep in particular is believed to be essential for learning in a way that cannot be accomplished during wakefulness alone (Yoshida & Toyoizumi, 2023). The primary reasons put forward so far largely argue that low sensory states afford brains time and space to produce neural activity driven by an internally built generative model, instead being largely driven by external stimulation. Relatedly observed activation patterns based on awake experiences are replayed during idle and sleep states (Wittkuhn, Chien, Hall-McMaster, et al., 2021). While the exact conditions of when replay occurs based on a task-related, internal generative model is debated, developments in both machine learning and neuroscience research suggest that it contributes to inference and generalization (for reviews see (Kumaran & McClelland, 2012a; Lewis et al., 2018b; Momennejad et al., 2018; Wittkuhn, Chien, Hall-McMaster, et al., 2021)). The algorithmic/representational level question involves understanding the specific processes and mechanisms by which offline periods aid in memory consolidation: “How does the passage of time restructure the representations of memories related to the transitive inference task?”. Lastly, the implementation level would delve into the neural and physiological mechanisms underlying the process. Questions like “Which brain networks or regions are involved in the observed consolidation following periods containing sleep or quiet rest?” or “What are the neural patterns observed after sleep that correlate with improved performance in the transitive inference task?" are relevant here.

This dissertation primarily concentrates on the final two questions, operating under the assumption that periods of rest, particularly sleep, are crucial to the process of generalization. This refers to the creation of abstract representations of learned data, which can then be utilized in novel situations. In essence, the hypothesis proposes that incorporating sleep into the retention interval enhances memory consolidation and integration, thereby bolstering the ability to discern patterns and principles from experiences. The focal point of this dissertation is to explore the correlation between a post-learning retention interval, which includes sleep, and the generalization of memory evident in transitive inference. Despite recent debates surrounding the perceived "criticality" of sleep in memory consolidation and generalization by various authors (Cordi & Rasch, 2021; Dastgheib, Kulanayagam, & Dringenberg, 2022; Dringenberg, 2019a; Lerner & Gluck, 2019a; Mantua, 2018; Nemeth et al., 2023; Newbury et al., 2021; Newbury & Monaghan, 2019a; Siegel, 2021; Talamini et al., 2022b; Ujma, 2022; Vaseghi, Arjmandi-Rad, Kholghi, & Nasehi, 2021; Vertes & Linley, 2021; Vertes & Siegel, 2005; Voderholzer et al., 2011), this dissertation saught to revisit seminal findings, evaluate potential representational mechanisms, and determine the boundary conditions for detectable and pertinent effects. Such a comprehensive approach would shed light on the complex relationship between 'time/sleep-dependent consolidation and generalization'.

5.7 Simulations

The primary candidate mechanism proposed and discussed in the thesis is reactivation and replay, most directly addressed in the simulations of *Chapter 2*. Despite a vast array of observations (Foster, 2017; Hayes et al., 2021b; Wittkuhn, Chien, Hall-McMaster, et al., 2021), it remains unclear how to characterize the nature and purpose of replay. Whilst the last decade appears has been largely dominated by deep neural network (Hayes et al., 2021b) and reinforcement learning (Roscow et al., 2021) based modelling work in terms of 'offline' memory research, there are several reasons why vector-symbolic approaches might be complementary in covering the existing literature of sleep and memory, key among them is their adaptability to various tasks and added interpretability afforded by the limited number of parameters and transparent information processing mechanics.

In regards to such cognitive models, discussion primarily centers around retrieval and encoding-based models (Kumaran, 2012) . While these argue for different stages of the learning trajectory being crucial for generalization, most formal instantiations lack detailed prediction on the time-course and the requisite conditions of successful inference formation. A notable exception includes REMERGE, that posits that strong source memory at recall, strengthened during sleep for example, can underlie successful inference when done in an on-the-fly manner. An alternative formulation presented in this thesis using MINERVA2 argues that the abstraction can happen by way of stored generalizations that can form via spontanous offline replay. Crucially this process while benefitting from strong source memory, does not rely on it at delayed recall. Given the small yet noteworthy forgetting effect observed in the combined analysis of Experiment 1, 2, and 3, in the performance of premise pairs from immediate to delayed testing, one could make a case for supporting a "stored generalization" model of generalization in line with encoding-based models, rather than a retrieval-based model when it comes to time and sleep-dependent models of inference. It is important to note that the way MINERVA2 was utilized makes it a hybrid model when considering retrieval and encoding-based models as two ends of the spectrum. It's recurrent similarity computation is distinctly retrieval-esque, but the way it represents premise memories as overlapping representations is in line with encoding-based models (M. W. Howard et al., 2005a).

Researchers have already highlighted key differences between REMERGE and context based models (e.g.: TCM) for transitive inference (Kumaran & McClelland, 2012a). REMERGE for example predicts that changes in associative rules in one end of the hierarchy (for example learning trials where F beats A, leading to circular value structure) will affect response accuracy in neighboring items (e.g. in B?E), where context-based models that store item-context associations predict a localized, non-spreading effect. Another notable prediction in terms of the content of replay activity in these models is that both classes of models propose that during offline periods, the replay activity in the hippocampal system should exhibit a generalized nature (e.g., ABC), rather than solely relating to a single experience (e.g., AB). However, REMERGE goes a step further by suggesting that periods of generalized replay may be interspersed with episodes of conventional single-episode replay (e.g., AB episode). This is due to the fact that network output in REMERGE is influenced by the varying levels of inhibition that are likely to fluctuate throughout the duration of an offline delay. In light of this future research should investigate replay quality and content as it relates to generalization performance, by quantifying the degree to which reactivation probabilities systematically follows a particular probability distribution or pattern (Y. Liu, Dolan, et al., 2021; Y. Liu et al., 2019; N. W. Schuck & Niv, 2019; Schwartenbeck et al., 2023; Wittkuhn, Krippner, & Schuck, n.d.).

Empirically contrasting MINERVA2 from other context-based models in terms of reactivation-related generalization behavior will be hard. While in TCM, associations between individual items (e.g., A and B) are mediated by their shared temporal context, rather than through direct item–item associations as in the current implementation of MINERVA2, both can produce a "symbolic distance" model of A..F over time, albeit through different means. Broadly defined, MINERVA2 relies on feature-based similarity, TCM on context-based similarity. When considering the generalization and integrated representation formation resulting from replay, the Temporal Context Model (TCM) would likely propose that the spatiotemporal order during learning and replay significantly impacts this formation process. TCM also suggests that the memory context alters systematically over time and experience. As a result, evidence demonstrating abrupt changes in memory context, in contrast to the TCM's presumptive gradual drift, would challenge this model's validity. Finally, the Temporal Context Model suggests that the likelihood of recalling something does not fundamentally rely on the quantity of items to be learned. Adding further items to the set does not inherently increase the difficulty of recalling each item, as they each bear an association with, and can be evoked by their unique spatiotemporal context. However, diversity in generalization effects based on set size could introduce additional considerations into the model's applicability, since the model does not inherently predict effects like the SDE.

While MINERVA2 has the capability to integrate contextual information, it was not utilized in the current simulations. However, it can be assumed that this would be crucial for an exact model-based analysis of replays. A recently proposed model, CMR-replay, which builds upon the Temporal Context Model (TCM), puts forth the idea that during periods of quiescence, awake rest and sleep, replay occurs when spontaneous memory reactivation triggers a series of subsequent memory reactivations (Zhenglong Zhou, Kahana, & Schapiro, n.d.) . These are driven by the reinstatement of memories' associated contexts, which code for the salience of each experience. Contrary to reinforcement learning models, which suggest replay unfolds based on the usefulness of memories for predicting learning value, CMR-replay theorizes that replay erupts from a rudimentary memory mechanism that continually operates on pre-established associations between contexts and experiences. This model successfully explains multiple replay-related phenomena observed in both human and animal studies. These include the enhancement of sequence memory linked to replay, the preferential strengthening of rewarded memories and most significantly for this thesis, improved inference effects (Gupta et al., 2010; Y. Liu et al., 2019). This renders the CMR-replay model the most comprehensive cognitive interpretation of rest and sleep-related reactivation to date. The existence of multiple cognitive models of reactivation such as MINERVA2, CMR-replay, and REMERGE is undoubtedly beneficial. However, to distinguish between these regarding time and sleep-dependent consolidation in generalization tasks, a combination of behavioral and neuroimaging observations is likely required. These should encompass the stages of encoding, rest, and delayed recall. Unfortunately, at the time of writing, such comprehensive observations are currently absent.

considerations into the model's applicability, since the model does not inherently predict effects like the SDE.

While MINERVA2 has the capability to integrate contextual information, it was not utilized in the current simulations. However, it can be assumed that this would be crucial for an exact model-based analysis of replays. A recently proposed model, CMR-replay, which builds upon the Temporal Context Model (TCM), puts forth the idea that during periods of quiescence, awake rest and sleep, replay occurs when spontaneous memory reactivation triggers a series of subsequent me

5.8 Behavioral experiments

In terms of time and sleep-dependent effects, the strength at which requisite paired associations are encoded appears predictive of delayed inference performance. This strength seems to interact with whether the retention interval included sleep or not, suggesting that sleep provides an additional advantage when encoding strength is high. Interestingly, this benefit does not seem to be driven by strengthened premise pair representations at the moment of delayed recall, as would be expected by retrieval-based models. Instead, we observed a small but significant decrease in premise pair performance. Additionally, the robustness of sleep-dependent symbolic distance effects remains ambiguous, which limits the extent to which the transitive inference paradigm can be employed to examine the relative changes in rank-based integrated representations. However, we do present provisional evidence for a joint rank effect, which assesses absolute rather than relative ordinal reorganization occurring over a rest period, including sleep. This evidence is tentative, as while it was consistent in our behavioral experiments, it did not emerge as expected in the study conducted by Berens & Bird (2022b), and only surfaced seven days after encoding in the Matorina & Poppenk (2021b) study.

As previously noted, future research interested in exploring sleep-related impacts on integrated representations and relational memory at large may find it advantageous to employ more intricate latent structure learning tasks. While transitive inference paradigm has the advantage of being historically studied in multiple species (Gazes & Lazareva, 2021a)and has numerous mathematical models to its name (Ciranka et al., 2021a; Jensen, Terrace, et al., 2019a; Lippl, Kay, Jensen, Ferrera, & Abbott, 2023) future tasks should encompass structures that are more intricate than a one-dimensional linear hierarchy, such as a two-dimensional rank-order topology or other advanced graphical structures as these can better disentangle relative versus absolute ordinal representations as they involve larger number of diverse inference trials while also being representative of everyday structures humans arguably use (G. B. Feld et al., 2022a; Javadi, Tolat, & Spiers, 2015; Park et al., 2021). Observing sleep-mediated transfer effects in these structures would be particularly exciting as its been long considered a key aspect of cognitive flexibility and intelligence, allowing us to adapt our knowledge to new circumstances . Transfer effects in generalization refer to the phenomenon where learning information in one context or situation is applied, or “transferred”, to another context or situation that was not directly encountered during learning. Essentially, it involves the process of generalizing learned knowledge or skills across different scenarios or tasks, not just within. If offline reactivation is indeed serving essential learning that cannot be achieved in wakefulness alone, more difficult generalization paradigms should produce stronger differential effects, whilst also reducing confounds related to rapid generalization effects emerging before offline periods begin.

This thesis primarily relied on online methodologies for two out of the three behavioral experiments conducted, with the final experiment being partially conducted online. While online studies significantly benefit sleep and memory research for two major reasons, they have certain limitations. Firstly, online studies aid in establishing sleep-dependent benefits in low complexity tasks, (such as associative inference tasks), which can be learned reliably without the active presence of an experimenter. Secondly, they allow participants to sleep in the comfort of their own beds, effectively eliminating the “First-Night Effect,” characterized by a decrease in sleep quality on the first night in an unfamiliar environment (Agnew, Webb, & Williams, 1966). However, these advantages are accompanied by a loss in control over the time-gap between training and bedtime and individual variations in “sleep hygiene” and “pre-sleep” rituals. These factors can influence offline consolidation in terms of behavior and may also result in diminished EEG signal quality for at-home recording designs. Future studies could address these limitations by incorporating actigraph measures to determine the time-gap between encoding and bedtime accurately, as well as self-reported measures related to sleep hygiene during the task period.

In summary the results from these experiments suggest that a retention interval containing sleep improves inference performance. However, this effect should be confirmed with high-powered pre-registered studies to strengthen this claim significantly.

5.9 Neuroimaging experiment

Through a two-day transitive inference experiment, we provide evidence that time-dependent memory consolidation plays a pivotal role in supporting inference. Behaviorally, we observed superior inference performance and faster reaction times for distant pairs in a hierarchy encoded 27 hours prior to recall compared to one encoded just 3 hours before recall. This aligns with predictions of sleep-dependent consolidation theories and indicates the emergence of integrated, model-based representations over time. While the earlier point regarding the ambiguity of the sleep-dependent symbolic distance effect remains valid, a decision was nonetheless made to utilize a model-based RSA to probe the distance-based representations. This analysis still provides an interpretable measure of relative change between the Remote and Recent conditions and also enhances comparability with the small number of published studies that have employed this technique in analyzing this and other similar tasks (Alfred et al., 2020; Berens & Bird, 2022b; Nelli et al., 2023; Park et al., 2021).

It is promising to observe that the multivariate pattern analysis (MVPA) appears sensitive to consolidation effects at the item level, as probed by our ‘localizer task’ - the key optimization target in our study - and that potential correlations with inference performance also begin to surface. However, the recommended sample size for reliable brain-behavior correlations, as suggested by Spisak, Bingel, & Wager (2023) and Tervo-Clemmens et al. (2023), adds a note of caution. Additional limitation was in the data derived from our in-scanner ‘inference task’, where the format posed a significant challenge in gleaning clear item-level effect estimates. We hypothesize that the reduced cluster-corrected findings in the inference task could have been due to an increase in motion during the second half of the experiment and signal loss stemming from our simplistic approach in decomposing paired presentation trials into items. This approach was adopted for the sake of comparability with the localizer data. For future studies interested in utilizing MVPA to examine inference, we recommend a trial presentation sequence that features only one item at a given time, similar to the approach used in Park et al. (2021) . The wider presence of comparable activation patterns in both the localizer data and the inference task as was observed in the mOFC would have produced stronger evidence for the formation of integrated representations that are causally relevant for successful inference. Relatedly, neuroimaging at encoding and comparing RDMs at recall with those observed immediately after encoding would have also strengthened the argument that our ROI are relevant for inference.

Further analyses should be directed towards better understanding the time-dependent role of the medial temporal lobe in hippocampal-dependent inference tasks, which could be better understood with complementary computational modelling and task-based functional connectivity analyses targeting the interaction of cortical and MTL regions. While the computational models used in this thesis have not been formulated as trial-level models of inference, considering recent advancements in likelihood-free estimations, a simulation models such as Minerva2 and REMERGE can be fitted to behavioral data to generate trial-level regressors (Palestro, Sederberg, Osth, Van Zandt, & Turner, 2018).

5.10 Closing remarks

Generalization is of vital importance. Its enhancement is a significant practical consideration in numerous fields such as education (Mata-Pereira & Ponte, 2017), and psychotherapy (Stokes & Baer, 1977). Another area where it plays a pivotal role is artificial intelligence. Ensuring that machine learning systems can adequately generalize is a growing concern, particularly with the mounting reliance on Large Language Models in question and answer settings (Gendron, Bao, Witbrock, & Dobbie, n.d.). Furthermore, the value of understanding generalization extends into the realm of mental health. Relational memory deficits are increasingly being identified as a reliable cognitive indicator of schizophrenia. Similarly, abnormalities in relational function related neuronal activity may emerge as a potential biomarker (Avery et al., 2021; Lepage et al., 2015). Given these various implications, it's no surprise that numerous disciplines and subfields are investigating generalization (R. W. Howard, 2000; J. E. Taylor et al., 2021). Despite this, there remains considerable scope for progress, especially in terms of integrating insights from all these interconnected fields.

By elucidating the conditions promoting generalization, this research may have implications for optimizing learning strategies and facilitating memory consolidation in healthy populations. In conclusion, this dissertation advanced our understanding of offline reactivation-dependent consolidation and identified fruitful directions for future research. The synthesized findings underscore the significance of model-based, multifaceted approaches in propelling the field forward.

List of Tables

List of Figures

List of Supplemental Tables

List of Supplemental Figures

References

Abel, M., Nickl, A. T., Reßle, A., Unger, C., & Bäuml, K.-H. T. (2023). The role of sleep for memory consolidation: does sleep protect memories from retroactive interference? *Psychonomic Bulletin & Review*. `https://doi.org/10.3758/s13423-023-02264-8`

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., . . . Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*. `https://doi.org/10.3389/fninf.2014.00014`

Acuna, B. D., Eliassen, J. C., Donoghue, J. P., & Sanes, J. N. (2002). Frontal and parietal lobe activation during transitive inference in humans. *Cerebral Cortex*, *12*(12), 1312–1321. `https://doi.org/cwz5nt`

Agnew, H. W., Webb, W. B., & Williams, R. L. (1966). THE FIRST NIGHT EFFECT: AN EEG STUDYOF SLEEP. *Psychophysiology*, *2*(3), 263–266. `https://doi.org/10.1111/j.1469-8986.1966.tb02650.x`

Alfred, K. L., Connolly, A. C., Cetron, J. S., & Kraemer, D. J. M. (2020). Mental models use common neural spatial structure for spatial and abstract content. *Communications Biology*, *3*(1), 1–11. `https://doi.org/gr8nf6`

Alfred, K. L., Connolly, A. C., & Kraemer, D. J. M. (2018). Putting the pieces together: Generating a novel representational space through deductive reasoning. *NeuroImage*, *183*(August), 99–111. `https://doi.org/gfjpqn`

Alger, S. E., & Payne, J. D. (2016). The differential effects of emotional salience on direct associative and relational memory during a nap. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(6), 1150–1163. `https://doi.org/10.3758/s13415-016-0460-1`

Alvarez, P., & Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: a simple network model. *Proceedings of the National Academy of Sciences*, *91*(15), 7041–7045. `https://doi.org/10.1073/pnas.91.15.7041`

Aly, M., Chen, J., Turk-Browne, N. B., & Hasson, U. (2018). Learning Naturalistic Temporal Structure in the Posterior Medial Network. *Journal of Cognitive*

Neuroscience, *30*(9), 1345–1365. `https://doi.org/10.1162/jocn_a_01308`

Andersen, P. (2007). *The hippocampus book*. Oxford university press.

Antony, J. W., & Schapiro, A. C. (2019). Active and effective replay: systems consolidation reconsidered again. *Nature Reviews Neuroscience*, *20*(8), 506–507. `https://doi.org/10.1038/s41583-019-0191-8`

Araujo Sanchez, M. A. de, & Zeithamova, D. (2023). Generalization and false memory in acquired equivalence. *Cognition*, *234*, 105385. `https://doi.org/10.1016/j.cognition.2023.105385`

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Sciences*, *9*(2), 83–89. `https://doi.org/10.1016/j.tics.2004.12.003`

Ashby, F. G., & Rosedahl, L. (2017). A neural interpretation of exemplar theory. *Psychological Review*, *124*(4), 472–482. `https://doi.org/10.1037/rev0000064`

Ashton, J. E., Staresina, B. P., & Cairney, S. A. (2022). Sleep bolsters schematically incongruent memories. *PLOS ONE*, *17*(6), e0269439. `https://doi.org/10.1371/journal.pone.0269439`

Ashton, J., & Cairney, S. A. (2021). Future-relevant memories are not selectively strengthened during sleep. *Plos One*, *16*(11). `https://doi.org/10.1371/journal.pone.0258110`

Audrain, S., & McAndrews, M. P. (2022). Schemas provide a scaffold for neocortical integration of new memories over time. *Nature Communications*, *13*(1), 5795. `https://doi.org/gq37vv`

August, D. A., & Levy, W. B. (1999). Temporal Sequence Compression by an Integrate-and-Fire Model of Hippocampal Area CA3. *Journal of Computational Neuroscience*, *6*(1), 71–90. `https://doi.org/10.1023/A:1008861001091`

Augustinack, J. C., Kouwe, A. J. W. van der, Salat, D. H., Benner, T., Stevens, A. A., Annese, J., … Corkin, S. (2014). H.M.'s contributions to neuroscience: A review and autopsy studies. *Hippocampus*, *24*(11), 1267–1286. `https://doi.org/10.1002/hipo.22354`

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. `https://doi.org/10.1016/j.media.2007.06.004`

Avery, S. N., Armstrong, K., McHugo, M., Vandekar, S., Blackford, J. U., Woodward, N. D., & Heckers, S. (2021). Relational memory in the early stage of psychosis: A 2-year follow-up study. *Schizophrenia Bulletin*, *47*(1), 7586. `https://doi.org/`

10.1093/schbul/sbaa081

Backus, A. R., Schoffelen, J.-M., Szebényi, S., Hanslmayr, S., & Doeller, C. F. (2016). Hippocampal-prefrontal theta oscillations support memory integration. *Current Biology*, *26*(4), 450–457. https://doi.org/10.1016/j.cub.2015.12.048

Baddeley, A. D., & Scott, D. (1971). Short Term Forgetting in the Absence of Proactive Interference. *Quarterly Journal of Experimental Psychology*, *23*(3), 275–283. https://doi.org/10.1080/14640746908401822

Baldassano, C., Hasson, U., & Norman, K. A. (2018). *Representation of real-world event schemas during narrative perception*. Retrieved from http://dx.doi.org/10.1101/252718

Banino, A., Koster, R., Hassabis, D., & Kumaran, D. (2016). Retrieval-Based Model Accounts for Striking Profile of Episodic Memory and Generalization. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep31330

Baram, Alon Boaz, Muller, T. H., Nili, H., Garvert, M. M., & Behrens, T. E. J. (2021). Entorhinal and ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning problems. *Neuron*, *109*(4), 713–723.e7. https://doi.org/ghqmjw

Baram, Alon B., Muller, T. H., Nili, H., Garvert, M., & Behrens, T. E. J. (2019). *Entorhinal and ventromedial prefrontal cortices abstract and generalise the structure of reinforcement learning problems*. Retrieved from http://dx.doi.org/10.1101/827253

Barron, H. C., Reeve, H. M., Koolschijn, R. S., Perestenko, P. V., Shpektor, A., Nili, H., … Dupret, D. (2020). Neuronal computation underlying inferential reasoning in humans and mice. *Cell*, *183*(1), 228–243.e21. https://doi.org/ghd7jp

Barry, D. N., & Maguire, E. A. (2019). Remote memory and the hippocampus: A constructive critique. *Trends in Cognitive Sciences*, *23*(2), 128–142. https://doi.org/10.1016/j.tics.2018.11.005

Basile, B. M., Templer, V. L., Gazes, R. P., & Hampton, R. R. (2020). Preserved visual memory and relational cognition performance in monkeys with selective hippocampal lesions. *Science Advances*, *6*(29). https://doi.org/10.1126/sciadv.aaz0484

Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). *Fitting linear mixed-effects models using {lme4}*. *67*. https://doi.org/10.18637/jss.v067.i01

Bayley, P. J., Gold, J. J., Hopkins, R. O., & Squire, L. R. (2005). The Neuroanatomy of Remote Memory. *Neuron*, *46*(5), 799–810. https://doi.org/10.1016/j.neuron.2005.04.034

Becker, S. (2005). A computational principle for hippocampal learning and neurogenesis. *Hippocampus*, *15*(6), 722–738. `https://doi.org/10.1002/hipo.20095`

Behrens, T. E. J., Muller, T. H., Whittington, J., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. *Neuron*, *100*(2), 490–509. `https://doi.org/10.1016/j.neuron.2018.10.002`

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, *37*(1), 90–101. `https://doi.org/10.1016/j.neuroimage.2007.04.042`

Bendor, D., & Wilson, M. A. (2012). Biasing the content of hippocampal replay during sleep. *Nature Neuroscience*, *15*(10), 1439–1444. `https://doi.org/10.1038/nn.3203`

Berens, S. C., & Bird, C. M. (2022b). Hippocampal and medial prefrontal cortices encode structural task representations following progressive and interleaved training schedules. *PLOS Computational Biology*, *18*(10), e1010566. `https://doi.org/10.1371/journal.pcbi.1010566`

Berens, S. C., & Bird, C. M. (2022a). Hippocampal and medial prefrontal cortices encode structural task representations following progressive and interleaved training schedules. *PLOS Computational Biology*, *18*. `https://doi.org/gr8ngr`

Berens, S. C., Horst, J. S., & Bird, C. M. (2018). Cross-Situational Learning Is Supported by Propose-but-Verify Hypothesis Testing. *Current Biology*, *28*(7), 1132–1136.e5. `https://doi.org/10.1016/j.cub.2018.02.042`

Berres, S., & Erdfelder, E. (2021). The sleep benefit in episodic memory: An integrative review and a meta-analysis. *Psychological Bulletin*, *147*(12), 1309–1353. `https://doi.org/10.1037/bul0000350`

Boccia, M., Teghil, A., & Guariglia, C. (2019). Looking into recent and remote past: Meta-analytic evidence for cortical re-organization of episodic autobiographical memories. *Neuroscience & Biobehavioral Reviews*, *107*, 84–95. `https://doi.org/10.1016/j.neubiorev.2019.09.003`

Bonnici, H. M., Chadwick, M. J., Lutti, A., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2012). Detecting representations of recent and remote autobiographical memories in vmPFC and hippocampus. *Journal of Neuroscience*, *32*(47), 1698216991. `https://doi.org/10.1523/JNEUROSCI.2475-12.2012`

Bonnici, H. M., & Maguire, E. A. (2018). Two years later–revisiting autobiographical memory representations in vmPFC and hippocampus. *Neuropsychologia*, *110*, 159169. `https://doi.org/10.1016/j.neuropsychologia.2017.05.014`

Boorman, Erie D., Sweigart, S. C., & Park, S. A. (2021a). Cognitive maps and novel inferences: a flexibility hierarchy. *Current Opinion in Behavioral Sciences*, *38*, 141–149. `https://doi.org/10.1016/j.cobeha.2021.02.017`

Boorman, Erie D., Sweigart, S. C., & Park, S. A. (2021b). Cognitive maps and novel inferences: A flexibility hierarchy. *Current Opinion in Behavioral Sciences*, *38*, 141–149. `https://doi.org/gr8ngh`

Boorman, Erie D., Sweigart, S. C., & Park, S. A. (2021c). Cognitive maps and novel inferences: A flexibility hierarchy. *Current Opinion in Behavioral Sciences*, *38*, 141–149. `https://doi.org/gr8ngh`

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research*, *76*(2), 192–203. `https://doi.org/dwndpt`

Bosshardt, S., Degonda, N., Schmidt, C. F., Boesiger, P., Nitsch, R. M., Hock, C., & Henke, K. (2005). One month of human memory consolidation enhances retrieval-related hippocampal activity. *Hippocampus*, *15*(8), 1026–1040. `https://doi.org/10.1002/hipo.20105`

Bower, G. H. (1970). Organizational factors in memory. *Cognitive Psychology*, *1*(1), 18–46. `https://doi.org/10.1016/0010-0285(70)90003-4`

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`

Bragin, A., Engel, J., Wilson, C. L., Fried, I., & Buzsáki, G. (1999). High-frequency oscillations in human brain. *Hippocampus*, *9*(2), 137–142. `https://doi.org/10.1002/(sici)1098-1063(1999)9:2%3C137::aid-hipo5%3E3.0.co;2-0`

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional Imaging of Neural Responses to Expectancy and Experience of Monetary Gains and Losses. *Neuron*, *30*(2), 619–639. `https://doi.org/10.1016/s0896-6273(01)00303-8`

Brodt, S., Gais, S., Beck, J., Erb, M., Scheffler, K., & Schönauer, M. (2018). Fast track to the neocortex: A memory engram in the posterior parietal cortex. *Science*, *362*(6418), 1045–1048. `https://doi.org/10.1126/science.aau2528`

Brodt, Svenja, Inostroza, M., Niethard, N., & Born, J. (2023). Sleep—A brain-state serving systems memory consolidation. *Neuron*, *111*(7), 1050–1075. `https://doi.org/10.1016/j.neuron.2023.03.005`

Brodt, Svenja, Pöhlchen, D., Täumer, E., Gais, S., & Schönauer, M. (2018). Incubation, not sleep, aids problem-solving. *Sleep*, *41*(10). `https://doi.org/10.1093/`

`sleep/zsy155`

Brown, R. M., & Robertson, E. M. (2007). Inducing motor skill improvements with a declarative task. *Nature Neuroscience*, *10*(2), 148–149. `https://doi.org/10.1038/nn1836`

Buckmaster, C. A., Eichenbaum, H., Amaral, D. G., Suzuki, W. A., & Rapp, P. R. (2004). Entorhinal Cortex Lesions Disrupt the Relational Organization of Memory in Monkeys. *The Journal of Neuroscience*, *24*(44), 9811–9825. `https://doi.org/10.1523/jneurosci.1532-04.2004`

Burghardt, J., & Bodansky, A. N. (2021). Why psychology needs to stop striving for novelty and how to move towards theory-driven research. *Frontiers in Psychology*, *12*. `https://doi.org/gr8ngn`

Buzsáki, G. (1989). Two-stage model of memory trace formation: A role for "noisy" brain states. *Neuroscience*, *31*(3), 551–570. `https://doi.org/10.1016/0306-4522(89)90423-5`

Buzsáki, G. (2015). Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus*, *25*(10), 1073–1188. `https://doi.org/10.1002/hipo.22488`

Buzsáki, G., & Fernández-Ruiz, A. (2019). Utility of the Idling Brain: Abstraction of New Knowledge. *Cell*, *178*(3), 513–515. `https://doi.org/10.1016/j.cell.2019.07.004`

Cai, D. J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., … Silva, A. J. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. *Nature*, *534*(7605), 115–118. `https://doi.org/10.1038/nature17955`

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`

Cazé, R., Khamassi, M., Aubin, L., & Girard, B. (2018). Hippocampal replays under the scrutiny of reinforcement learning models. *Journal of Neurophysiology*, *120*(6), 2877–2896. `https://doi.org/10.1152/jn.00145.2018`

Cellini, N., Mercurio, M., Vanzetti, V., Bergamo, D., & Sarlo, M. (2020). Comparing the effect of daytime sleep and wakefulness on mnemonic discrimination. *Physiology and Behavior*, *224*, 113078. `https://doi.org/gr8ngp`

Chan, S. C. Y., Applegate, M. C., Morton, N. W., Polyn, S. M., & Norman, K. A. (2016). *Lingering representations of stimulus history at encoding influence recall organization.* Retrieved from `http://dx.doi.org/10.1101/068353`

Chanales, A. J. H., Oza, A., Favila, S. E., & Kuhl, B. A. (2017). Overlap among Spatial Memories Triggers Repulsion of Hippocampal Representations. *Current Biology*, *27*(15), 2307–2317.e5. `https://doi.org/10.1016/j.cub.2017.06.057`

Chatburn, A., Lushington, K., & Kohler, M. J. (2014). Complex associative memory processing and sleep: A systematic review and meta-analysis of behavioural evidence and underlying EEG mechanisms. *Neuroscience and Biobehavioral Reviews*, *47*, 646–655. `https://doi.org/f6vm4j`

Chen, P.-C., Niknazar, H., Alaynick, W. A., Whitehurst, L. N., & Mednick, S. C. (2021). The sleeping brain switches between working memory and long-term memory processing. *bioRxiv*, 2021.03.31.437952. `https://doi.org/gr8ngq`

Chen, S., Swartz, K. B., & Terrace, H. S. (1997). Knowledge of the Ordinal Position of List Items in Rhesus Monkeys. *Psychological Science*, *8*(2), 80–86. `https://doi.org/10.1111/j.1467-9280.1997.tb00687.x`

Ciranka, S., Linde-Domingo, J., Padezhki, I., Wicharz, C., Wu, C. M., & Spitzer, B. (2021a). Asymmetric learning facilitates human inference of transitive relations. *Nature Human Behaviour*, *6*(4), 1–33. `https://doi.org/gpgprv`

Ciranka, S., Linde-Domingo, J., Padezhki, I., Wicharz, C., Wu, C. M., & Spitzer, B. (2021b). Asymmetric learning facilitates human inference of transitive relations. *Nature Human Behaviour*, *6*(4), 1–33. `https://doi.org/gpgprv`

Ciranka, S., Linde-Domingo, J., Padezhki, I., Wicharz, C., Wu, C. M., & Spitzer, B. (2022). Asymmetric reinforcement learning facilitates human inference of transitive relations. *Nature Human Behaviour*, *6*(4), 555–564. `https://doi.org/10.1038/s41562-021-01263-w`

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, *3*(1), 37–60. `https://doi.org/10.3758/bf03210740`

Clifford, S., & Jerit, J. (2014). Is There a Cost to Convenience? An Experimental Comparison of Data Quality in Laboratory and Online Studies. *Journal of Experimental Political Science*, *1*(2), 120–131. `https://doi.org/10.1017/xps.2014.5`

Cohen, Neal J., & Eichenbaum, H. (1993b). *Memory, amnesia, and the hippocampal system*. MIT press.

Cohen, Neal J., & Eichenbaum, H. (1993a). *Memory, Amnesia, and the Hippocampal System*. MIT Press.

Cohen, Neal J., Poldrack, R. A., Eichenbaum, H., Cohen, N. J., Poldrack, R. A., & Eichenbaum, H. (1997). Memory for items and memory for relations in the procedural / declarative memory framework. *Memory*, *5*(1-2), 131–178. `https:`

//doi.org/bcgjjq

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

Comrie, A. E., Frank, L. M., & Kay, K. (2022). Imagination as a fundamental function of the hippocampus. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1866). https://doi.org/10.1098/rstb.2021.0336

Cooper, R. A., Kurkela, K. A., Davis, S. W., & Ritchey, M. (2021). Mapping the organization and dynamics of the posterior medial network during movie watching. *NeuroImage*, *236*, 118075. https://doi.org/10.1016/j.neuroimage.2021.118075

Cordi, M. J., & Rasch, B. (2021). How robust are sleep-mediated memory benefits? *Current Opinion in Neurobiology*, *67*, 1–7. https://doi.org/10.1016/j.conb.2020.06.002

Cordi, M. J., Schreiner, T., & Rasch, B. (2023). Is prior knowledge essential? Additional training opportunities restore sleep-associated memory benefits under conditions of low prior knowledge. *Journal of Sleep Research*, e13834. https://doi.org/10.1111/jsr.13834

Cousins, J. N., El-Deredy, W., Parkes, L. M., Hennies, N., & Lewis, P. A. (2016). Cued reactivation of motor learning during sleep leads to overnight changes in functional brain activity and connectivity. *PLoS Biology*, *14*(5), e1002451. https://doi.org/f8pp3d

Covington, N. V., Brown-Schmidt, S., & Duff, M. C. (2018). The Necessity of the Hippocampus for Statistical Learning. *Journal of Cognitive Neuroscience*, *30*(5), 680–697. https://doi.org/10.1162/jocn_a_01228

Cowan, E. T., Liu, A. A., Henin, S., Kothare, S., Devinsky, O., & Davachi, L. (2021). Time-dependent transformations of memory representations differ along the long axis of the hippocampus. *Learning and Memory*, *28*(9), 329–340. https://doi.org/gpdd9p

Cox, G., & Shiffrin, R. M. (2017). A dynamic approach to recognition memory. *Psychological Review*, *124*(6), 795–860. https://doi.org/gcvj5s

Cox, R. W., Chen, G., Glen, D. R., Reynolds, R. C., & Taylor, P. A. (2017). FMRI Clustering in AFNI: False-Positive Rates Redux. *Brain Connectivity*, *7*(3), 152–171. https://doi.org/10.1089/brain.2016.0475

Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine*, *10*(4-5), 171–178. https://doi.org/10.1002/

`(SICI)1099-1492(199706/08)10:4/5%3C171::AID-NBM453%3E3.0.CO;2-L`

Cox, R., Bronkhorst, M. L. V. van, Bayda, M., Gomillion, H., Cho, E., Parr, M. E., . . . Stickgold, R. (2018). Sleep Selectively Stabilizes Contextual Aspects of Negative Memories. *Scientific Reports*, *8*(1), 17861. `https://doi.org/10.1038/s41598-018-35999-9`

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

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. *PLoS ONE*, *8*(3), e57410. `https://doi.org/10.1371/journal.pone.0057410`

Curtis, E. T., & Jamieson, R. K. (2018). Computational and empirical simulations of selective memory impairments: Converging evidence for a single-system account of memory dissociations. *Quarterly Journal of Experimental Psychology*, *72*(4), 798–817. `https://doi.org/10.1177/1747021818768502`

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage*, *9*(2), 179–194. `https://doi.org/10.1006/nimg.1998.0395`

Dandolo, L. C., & Schwabe, L. (2018). Time-dependent memory transformation along the hippocampal anterior–posterior axis. *Nature Communications*, *9*(1). `https://doi.org/10.1038/s41467-018-03661-7`

Dastgheib, M., Kulanayagam, A., & Dringenberg, H. C. (2022). Is the role of sleep in memory consolidation overrated? *Neuroscience & Biobehavioral Reviews*, *140*, 104799. `https://doi.org/10.1016/j.neubiorev.2022.104799`

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*, *16*(6), 693–700. `https://doi.org/10.1016/j.conb.2006.10.012`

Dayan, P., Hinton, G. E., Neal, R. M., & Zemel, R. S. (1995). The Helmholtz Machine. *Neural Computation*, *7*(5), 889–904. `https://doi.org/10.1162/neco.1995.7.5.889`

De Brigard, F., & Sinnott-Armstrong, W. (Eds.). (2022). *Neuroscience and philosophy*. The MIT Press. `https://doi.org/10.7551/mitpress/12611.001.0001`

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From Creatures of Habit to Goal-Directed Learners. *Psychological Science*, *27*(6), 848–858. `https://doi.org/10.1177/0956797616639301`

Denis, D., Bottary, R., Cunningham, T. J., Tcheukado, M.-C., & Payne, J. D. (2023). The influence of encoding strategy on associative memory consolidation across wake and sleep. *Learn. Mem.*, *30*(9), 185–191.

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

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

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

Deperrois, N., Petrovici, M. A., Senn, W., & Jordan, J. (2022). Learning cortical representations through perturbed and adversarial dreaming. *eLife*, *11*. `https://doi.org/10.7554/elife.76384`

Deuker, L., Bellmund, J. L., Navarro Schröder, T., & Doeller, C. F. (2016). An event map of memory space in the hippocampus. *eLife*, *5*. `https://doi.org/10.7554/elife.16534`

DeVito, L. M., Kanter, B. R., & Eichenbaum, H. (2009). The hippocampus contributes to memory expression during transitive inference in mice. *Hippocampus*, NA–NA. `https://doi.org/10.1002/hipo.20610`

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

Diba, K., & Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience*, *10*(10), 1241–1242. `https://doi.org/10.1038/nn1961`

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

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, *208*(2), 425–429. `https://doi.org/10.1016/j.bbr.2009.12.021`

Diekelmann, S., Wilhelm, I., & Born, J. (2009). The whats and whens of sleep-dependent memory consolidation. *Sleep Medicine Reviews*, *13*(5), 309–321. `https://doi.org/10.1016/j.smrv.2008.08.002`

Dimsdale-Zucker, H. R., Montchal, M. E., Reagh, Z. M., Wang, S.-F., Libby, L. A., & Ranganath, C. (2022). Representations of complex contexts: A role for hippocampus. *Journal of Cognitive Neuroscience*, *35*(1), 90110. `https://doi.org/10.1162/jocn_a_01919`

Dimsdale-Zucker, H. R., Ritchey, M., Ekstrom, A. D., Yonelinas, A. P., & Ranganath, C. (2018). CA1 and CA3 differentially support spontaneous retrieval of episodic contexts within human hippocampal subfields. *Nature Communications*, *9*(1). `https://doi.org/gcv5js`

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language*, *59*(4), 447–456. `https://doi.org/10.1016/j.jml.2007.11.004`

Doxey, C. R., Hodges, C. B., Bodily, T. A., Muncy, N. M., & Kirwan, C. B. (2018). The effects of sleep on the neural correlates of pattern separation. *Hippocampus*, *28*(2), 108–120. `https://doi.org/ggrvbq`

Dringenberg, H. C. (2019a). *Chapter 32 - sleep and memory consolidation: Conceptual and methodological challenges* (H. C. Dringenberg, Ed.). In (pp. 489–501). Elsevier. `https://doi.org/10.1016/B978-0-12-813743-7.00032-3`

Dringenberg, H. C. (2019b). *Sleep and memory consolidation: Conceptual and methodological challenges* (Vol. 30). Elsevier B.V. Retrieved from `http://dx.doi.org/10.1016/B978-0-12-813743-7.00032-3`

Drosopoulos, S., Schulze, C., Fischer, S., & Born, J. (2007a). Sleep's function in the spontaneous recovery and consolidation of memories. *Journal of Experimental Psychology: General*, *136*(2), 169. `https://doi.org/10.1037/0096-3445.136.2.169`

Drosopoulos, S., Schulze, C., Fischer, S., & Born, J. (2007b). Sleep's function in the spontaneous recovery and consolidation of memories. *Journal of Experimental Psychology: General*, *136*(2), 169–183. `https://doi.org/10.1037/0096-3445.136.2.169`

Du, X., Zhan, L., Chen, G., Guo, D., Li, C., Moscovitch, M., & Yang, J. (2019). Differential activation of the medial temporal lobe during item and associative memory across time. *Neuropsychologia*, *135*, 107252. `https://doi.org/10.1016/j.neuropsychologia.2019.107252`

DuBrow, S., Rouhani, N., Niv, Y., & Norman, K. A. (2017). Does mental context drift or shift? *Current Opinion in Behavioral Sciences*, *17*, 141146. `https://doi.org/10.1016/j.cobeha.2017.08.003`

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annu. Rev. Psychol.*, *55*, 51–86. `https://doi.org/dkv4fk`

Dudai, Y., Karni, A., & Born, J. (2015). The Consolidation and Transformation of Memory. *Neuron*, *88*(1), 20–32. `https://doi.org/10.1016/j.neuron.2015.09.004`

Dumas, D. (2016). Relational Reasoning in Science, Medicine, and Engineering. *Educational Psychology Review*, *29*(1), 73–95. `https://doi.org/10.1007/s10648-016-9370-6`

Dumas, D., Alexander, P. A., Baker, L. M., Jablansky, S., & Dunbar, K. N. (2014). Relational reasoning in medical education: Patterns in discourse and diagnosis. *Journal of Educational Psychology*, *106*(4), 1021–1035. `https://doi.org/10.1037/a0036777`

Durrant, S. J., Cairney, S. A., McDermott, C., & Lewis, P. A. (2015). Schema-conformant memories are preferentially consolidated during REM sleep. *Neurobiology of Learning and Memory*, *122*, 41–50. `https://doi.org/ggrwk2`

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/cpp8sv`

Dusek, J. A., & Eichenbaum, H. (1997a). The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(13), 7109–7114. `https://doi.org/c8ntn5`

Dusek, J. A., & Eichenbaum, H. (1997b). The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences*, *94*(13), 7109–7114. `https://doi.org/10.1073/pnas.94.13.7109`

Duszkiewicz, A. J., Rossato, J. I., Moreno, A., Takeuchi, T., Yamasaki, M., Genzel, L., … Morris, R. G. M. (2023). Execution of new trajectories toward a stable goal without a functional hippocampus. *Hippocampus*, *33*(6), 769–786. `https://doi.org/10.1002/hipo.23497`

Eichenbaum, H. (1999). The hippocampus and mechanisms of declarative memory. *Behavioural Brain Research*, *103*(2), 123–133. `https://doi.org/10.1016/s0166-4328(99)00044-3`

Eichenbaum, H. (2003). How does the hippocampus contribute to memory? *Trends in Cognitive Sciences*, *7*(10), 427–429. `https://doi.org/10.1016/j.tics.2003.`

08.008

Eichenbaum, H. (2017). Memory: Organization and control. *Annual Review of Psychology*, *68*(1), 19–45. `https://doi.org/10.1146/annurev-psych-010416-044131`

Ellenbogen, J., 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`

Ellenbogen, J., Hulbert, J. C., Jiang, Y., & Stickgold, R. (2009). The sleeping brain's influence on verbal memory: Boosting resistance to interference. *PLoS ONE*, *4*(1), 2–5. `https://doi.org/bx4f5x`

Ellenbogen, J., Payne, J. D., & Stickgold, R. (2006). The Role of Sleep in Declarative Memory Consolidation: Passive, Permissive, Active or None? *Current Opinion in Neurobiology*, *16*(6), 716–722. `https://doi.org/10.1016/j.conb.2006.10.006`

Epstein, R. A., Patai, E. Z., Julian, J. B., & Spiers, H. J. (2017). The cognitive map in humans: spatial navigation and beyond. *Nature Neuroscience*, *20*(11), 1504–1513. `https://doi.org/10.1038/nn.4656`

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLOS ONE*, *12*(9), e0184661. `https://doi.org/10.1371/journal.pone.0184661`

Esteban, O., Blair, R., Markiewicz, C. J., Berleant, S. L., Moodie, C., Ma, F., … Gorgolewski, K. J. (2018). fMRIPrep. *Software*. `https://doi.org/10.5281/zenodo.852659`

Esteban, O., Markiewicz, C., Blair, R. W., Moodie, C., Isik, A. I., Erramuzpe Aliaga, A., … Gorgolewski, K. J. (2018). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*. `https://doi.org/10.1038/s41592-018-0235-4`

Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, *57*(2), 94–107. `https://doi.org/10.1037/h0058559`

Favila, S. E., Chanales, A. J., & Kuhl, B. A. (2016). Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. *Nature Communications*, *7*(1), 11066. Retrieved from `https://www.nature.com/articles/ncomms11066`

Feeley, T. H. (2020). Assessing Study Quality in Meta-Analysis. *Human Communication Research*, *46*(2-3), 334–342. `https://doi.org/10.1093/hcr/hqaa001`

Feeney, A., & Heit, E. (Eds.). (2010). *Inductive reasoning*. Cambridge, England: Cambridge University Press.

Feld, G. B., Bernard, M., Rawson, A. B., & Spiers, H. J. (2022a). Sleep targets highly connected global and local nodes to aid consolidation of learned graph networks. *Scientific Reports*, *12*(1). `https://doi.org/10.1038/s41598-022-17747-2`

Feld, G. B., Bernard, M., Rawson, A. B., & Spiers, H. J. (2022b). Sleep targets highly connected global and local nodes to aid consolidation of learned graph networks. *Scientific Reports*, *12*(1), 1–15. `https://doi.org/gr8nhg`

Feld, Gordon B., Weis, P. P., & Born, J. (2016a). The limited capacity of sleep-dependent memory consolidation. *Frontiers in Psychology*, *7*, 1368. `https://doi.org/10.3389/fpsyg.2016.01368`

Feld, Gordon B., Weis, P. P., & Born, J. (2016b). The limited capacity of sleep-dependent memory consolidation. *Frontiers in Psychology*, *7*(SEP), 1368. `https://doi.org/f84cq3`

Fenn, K. M., & Hambrick, D. Z. (2010). *Individual differences in working memory capacity predict sleep-dependent memory consolidation*. American Psychological Association (APA). `https://doi.org/10.1037/e520592012-094`

Fenn, K. M., & Hambrick, D. Z. (2012). Individual differences in working memory capacity predict sleep-dependent memory consolidation. *Journal of Experimental Psychology: General*, *141*(3), 404–410. `https://doi.org/10.1037/a0025268`

Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2013). Sleep restores loss of generalized but not rote learning of synthetic speech. *Cognition*, *128*(3), 280–286. `https://doi.org/f46knf`

Ficca, G., Lombardo, P., Rossi, L., & Salzarulo, P. (2000). Morning recall of verbal material depends on prior sleep organization. *Behavioural Brain Research*, *112*(1-2), 159–163. `https://doi.org/bx5r85`

Findlay, G., Tononi, G., & Cirelli, C. (2020). The evolving view of replay and its functions in wake and sleep. *SLEEP Advances*, *1*(1). `https://doi.org/10.1093/sleepadvances/zpab002`

Fleiss, J. (1993). Review papers : The statistical basis of meta-analysis. *Statistical Methods in Medical Research*, *2*(2), 121–145. `https://doi.org/10.1177/096228029300200202`

Foldes, T., Santamaria, L., & Lewis, P. A. (2023). Sleep-related benefits to transitive inference are modulated by encoding strength and joint rank. *Learning & Memory*, *30*(9), 201–211. `https://doi.org/10.1101/lm.053787.123`

Folkerts, S., Rutishauser, U., & Howard, M. W. (2018). Human Episodic Memory Retrieval Is Accompanied by a Neural Contiguity Effect. *The Journal of Neuroscience*, *38*(17), 4200–4211. `https://doi.org/10.1523/jneurosci.2312-17.2018`

Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, *47, Supplement 1*, S102. `https://doi.org/10.1016/S1053-8119(09)70884-5`

Fosse, M. J., Fosse, R., Hobson, J. A., & Stickgold, R. J. (2003). Dreaming and Episodic Memory: A Functional Dissociation? *Journal of Cognitive Neuroscience*, *15*(1), 1–9. `https://doi.org/10.1162/089892903321107774`

Foster, D. J. (2017). Replay comes of age. *Annual Review of Neuroscience*, *40*(1), 581–602. `https://doi.org/ggb5gg`

Fox, J., Dennis, S., & Osth, A. F. (2020). Accounting for the build-up of proactive interference across lists in a list length paradigm reveals a dominance of item-noise in recognition memory. *Journal of Memory and Language*, *110*, 104065. `https://doi.org/10.1016/j.jml.2019.104065`

Frank, M. J., O'Reilly, R. C., & Curran, T. (2006). When Memory Fails, Intuition Reigns. *Psychological Science*, *17*(8), 700–707. `https://doi.org/10.1111/j.1467-9280.2006.01769.x`

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`

Fried, E. I. (2020). Theories and models: What they are, what they are for, and what they are about. *Psychological Inquiry*, *31*(4), 336–344. `https://doi.org/gjzbzp`

Furman, O., Mendelsohn, A., & Dudai, Y. (2012). The episodic engram transformed: Time reduces retrieval-related brain activity but correlates it with memory accuracy. *Learning & Memory*, *19*(12), 575–587. `https://doi.org/10.1101/lm.025965.112`

Gais, S., Albouy, G., Boly, M., Dang-Vu, T. T., Darsaud, A., Desseilles, M., ... Peigneux, P. (2007). Sleep transforms the cerebral trace of declarative memories. *Proceedings of the National Academy of Sciences*, *104*(47), 18778–18783. `https://doi.org/10.1073/pnas.0705454104`

Gais, S., & Born, J. (2004). Declarative memory consolidation : Mechanisms acting during human sleep. *Learning & Memory*, *11*(6), 679–685. `https://doi.org/10.1101/lm.80504.An`

Gardner, M. P., & Schoenbaum, G. (2021). The orbitofrontal cartographer. *Behavioral Neuroscience*, *135*(2), 267. `https://doi.org/10.1037/bne0000463`

Garvert, M. M., Dolan, R. J., & Behrens, T. E. J. (2017). A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *eLife*, *6*, e17086. `https://doi.org/f95g5z`

Gazes, R. P., & Lazareva, O. F. (2021a). Does cognition differ across species, and how do we know? Lessons from research in transitive inference. *Journal of Experimental Psychology: Animal Learning and Cognition*, *47*(3), 223–233. `https://doi.org/gr8nhp`

Gazes, R. P., & Lazareva, O. F. (2021b). Does cognition differ across species, and how do we know? Lessons from research in transitive inference. *Journal of Experimental Psychology: Animal Learning and Cognition*, *47*(3), 223–233. `https://doi.org/10.1037/xan0000301`

Gazes, R. P., Lazareva, O. F., Bergene, C. N., & Hampton, R. R. (2014). Effects of spatial training on transitive inference performance in humans and rhesus monkeys. *Journal of Experimental Psychology: Animal Learning and Cognition*, *40*(4), 477–489. `https://doi.org/10.1037/xan0000038`

Gazes, R. P., Templer, V. L., & Lazareva, O. F. (2022). Thinking about order: a review of common processing of magnitude and learned orders in animals. *Animal Cognition*, *26*(1), 299–317. `https://doi.org/10.1007/s10071-022-01713-6`

Gendron, G., Bao, Q., Witbrock, M., & Dobbie, G. (n.d.). *Large language models are not strong abstract reasoners.*

Gentner, D., & Namy, L. L. (2006). Analogical Processes in Language Learning. *Current Directions in Psychological Science*, *15*(6), 297–301. `https://doi.org/10.1111/j.1467-8721.2006.00456.x`

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, *19*(5), 847–857. `https://doi.org/10.3758/s13423-012-0296-9`

Gershman, S. J. (2018). The Successor Representation: Its Computational Logic and Neural Substrates. *The Journal of Neuroscience*, *38*(33), 7193–7200. `https://doi.org/10.1523/jneurosci.0151-18.2018`

Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, *68*(1), 101–128. `https://doi.org/10.1146/annurev-psych-122414-033625`

Gershman, S. J., & Niv, Y. (2010). Learning latent structure: carving nature at its joints. *Current Opinion in Neurobiology*, *20*(2), 251–256. `https://doi.org/10.1016/j.conb.2010.02.008`

Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour*, *66*(1), 1536. `https://doi.org/10.1006/anbe.2003.2174`

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

Gilboa, A., & Moscovitch, M. (2021). No consolidation without representation: Correspondence between neural and psychological representations in recent and remote memory. *Neuron*, *109*(14), 2239–2255. `https://doi.org/gmcwfz`

Gilboa, A., Winocur, G., Grady, C. L., Hevenor, S. J., & Moscovitch, M. (2004). Remembering our past: Functional neuroanatomy of recollection of recent and very remote personal events. *Cerebral Cortex*, *14*(11), 1214–1225. `https://doi.org/10.1093/cercor/bhh082`

Gillespie, A. K., Astudillo Maya, D. A., Denovellis, E. L., Liu, D. F., Kastner, D. B., Coulter, M. E., … Frank, L. M. (2021). Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron*, *109*(19), 3149–3163.e6. `https://doi.org/10.1016/j.neuron.2021.07.029`

Giuditta, A., Ambrosini, M. V., Montagnese, P., Mandile, P., Cotugno, M., Zucconi, G. G., & Vescia, S. (1995). The sequential hypothesis of the function of sleep. *Behavioural Brain Research*, *69*(1-2), 157–166. `https://doi.org/d3dq5m`

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., … Jenkinson, M. (2013). The minimal preprocessing pipelines for the human connectome project. *NeuroImage*, *80*, 105–124. `https://doi.org/10.1016/j.neuroimage.2013.04.127`

Gluck, M. A., Meeter, M., & Myers, C. E. (2003). Computational models of the hippocampal region: linking incremental learning and episodic memory. *Trends in Cognitive Sciences*, *7*(6), 269–276. `https://doi.org/10.1016/s1364-6613(03)00105-0`

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini Meta-Analysis of Your Own Studies: Some Arguments on Why and a Primer on How. *Social and Personality Psychology Compass*, *10*(10), 535–549. `https://doi.org/10.1111/spc3.12267`

Golkashani, H. A., Ghorbani, S., Leong, R. L. F., Ong, J. L., & Chee, M. W. L. (2023a). Advantage conferred by overnight sleep on schema-related memory may last only a day. *SLEEP Advances*, 1–43. `https://doi.org/gr8nhs`

Golkashani, H. A., Ghorbani, S., Leong, R. L. F., Ong, J. L., & Chee, M. W. L. (2023b). Advantage conferred by overnight sleep on schema-related memory may last only a day. *SLEEP Advances*, 1–43. `https://doi.org/gr8nhs`

Golkashani, H., Leong, R. L. F., Ghorbani, S., Ong, J. L., Fernández, G., & Chee, M. W. L. (2022a). A sleep schedule incorporating naps benefits the transformation of hierarchical knowledge. *Sleep*, *45*(4), zsac025. `https://doi.org/gr8nht`

Golkashani, H., Leong, R. L. F., Ghorbani, S., Ong, J. L., Fernández, G., & Chee, M. W. L. (2022b). A sleep schedule incorporating naps benefits the transformation of hierarchical knowledge. *Sleep*, *45*(4), zsac025. `https://doi.org/gr8nht`

Gorgolewski, K. J., Esteban, O., Markiewicz, C. J., Ziegler, E., Ellis, D. G., Notter, M. P., … Ghosh, S. (2018). Nipype. *Software*. `https://doi.org/10.5281/zenodo.596855`

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. (2011). Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, *5*, 13. `https://doi.org/10.3389/fninf.2011.00013`

Greenwald, A. G. (1976). Within-subjects designs: To use or not to use? *Psychological Bulletin*, *83*(2), 314–320. `https://doi.org/10.1037/0033-2909.83.2.314`

Gregory, S. E. A., Winstone, N. E., Ridout, N., & Nash, R. A. (2019). Weak memory for future-oriented feedback: investigating the roles of attention and improvement focus. *Memory*, *28*(2), 216–236. `https://doi.org/10.1080/09658211.2019.1709507`

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72. `https://doi.org/10.1016/j.neuroimage.2009.06.060`

Groch, S., Schreiner, T., Rasch, B., Huber, R., & Wilhelm, I. (2017). Prior knowledge is essential for the beneficial effect of targeted memory reactivation during sleep. *Scientific Reports*, *7*(1). `https://doi.org/10.1038/srep39763`

Groch, S., Wilhelm, I., Diekelmann, S., Sayk, F., Gais, S., & Born, J. (2011). Contribution of norepinephrine to emotional memory consolidation during sleep. *Psychoneuroendocrinology*, *36*(9), 1342–1350. `https://doi.org/10.1016/j.psyneuen.2011.03.006`

Grossberg, S. (1986). *The adaptive self-organization of serial order in behavior: Speech, language, and motor control*. ACADEMIC PRESS, INC. `https://doi.org/10.1016/b978-0-12-631403-8.50011-4`

Guest, O., & Martin, A. E. (2021). How computational modeling can force theory building in psychological science. *Perspectives on Psychological Science*. `https://doi.org/ghvcwr`

Gupta, A. S., Meer, M. A. A. van der, Touretzky, D. S., & Redish, A. D. (2010). Hippocampal Replay Is Not a Simple Function of Experience. *Neuron*, *65*(5), 695–705. `https://doi.org/10.1016/j.neuron.2010.01.034`

Haas, M., Scheibe, S., El Khawli, E., Künzi, M., Ihle, A., Ballhausen, N., … Zuber, S. (2021). Online assessment of cognitive functioning across the adult lifespan using the eCOGTEL: a reliable alternative to laboratory testing. *European Journal of Ageing*, *19*(3), 609–619. `https://doi.org/10.1007/s10433-021-00667-x`

Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: the foundation of higher cognition. *Trends in Cognitive Sciences*, *14*(11), 497–505. `https://doi.org/10.1016/j.tics.2010.08.005`

Hanert, A., Weber, F. D., Pedersen, A., Born, J., & Bartsch, T. (2017). Sleep in humans stabilizes pattern separation performance. *Journal of Neuroscience*, *37*(50), 12238–12246. `https://doi.org/gcq22d`

Harvey, B., Fracasso, A., Petridou, N., & Dumoulin, S. (2015). Overlapping topographic representations of numerosity and object size in human parietal cortex. *Journal of Vision*, *15*(12), 1283. `https://doi.org/10.1167/15.12.1283`

Hasselmo, Michael E. (2008). Temporally structured replay of neural activity in a model of entorhinal cortex , hippocampus and postsubiculum. *European Journal of Neuroscience*, *28*(August), 1301–1315. `https://doi.org/dphqnc`

Hasselmo, Michael E., & Hinman, J. R. (2017). Marr's influence on the standard model of hippocampus, and the need for more theoretical advances. In *The title of book here* (pp. 133–158). Oxford University Press. Retrieved from `https://experts.illinois.edu/en/publications/marrs-influence-on-the-standard-model-of-hippocampus-and-the-need`

Havas, V., Taylor, J., Vaquero, L., Diego-Balaguer, R. de, Rodríguez-Fornells, A., & Davis, M. H. (2018). Semantic and phonological schema influence spoken word learning and overnight consolidation. *Quarterly Journal of Experimental Psychology*, *71*(6), 1469–1481. `https://doi.org/10.1080/17470218.2017.1329325`

Hayes, T. L., Krishnan, G. P., Bazhenov, M. A., Siegelmann, H. T., Sejnowski, T. J., & Kanan, C. (2021a). *Replay in deep learning : Current approaches and missing biological elements*. 1–53. Retrieved from `http://arxiv.org/abs/2104.04132`

Hayes, T. L., Krishnan, G. P., Bazhenov, M., Siegelmann, H. T., Sejnowski, T. J., & Kanan, C. (2021b). Replay in deep learning: Current approaches and missing biological elements. *Neural Computation*, *33*(11), 29082950.

Hebb, D. O. (1949). *The organisation of behaviour: A neuropsychological theory.* Science Editions New York.

Hebscher, M., Wing, E., Ryan, J., & Gilboa, A. (2019). Rapid Cortical Plasticity Supports Long-Term Memory Formation. *Trends in Cognitive Sciences*, *23*(12), 989–1002. `https://doi.org/10.1016/j.tics.2019.09.009`

Hecker, U. von, Klauer, K. C., & Aßfalg, A. (2019). A robust anchoring effect in linear ordering. *Quarterly Journal of Experimental Psychology*, *72*(11), 2680–2689. `https://doi.org/10.1177/1747021819855234`

Hecker, U. von, Müller, E., Kirian Dill, S., & Christoph Klauer, K. (2023). Mental representation of equivalence and order. *Quarterly Journal of Experimental Psychology*, 174702182311539. `https://doi.org/10.1177/17470218231153974`

Heckers, S., Zalesak, M., Weiss, A. P., Ditman, T., & Titone, D. (2004). Hippocampal activation during transitive inference in humans. *Hippocampus*, *14*(2), 153–162. `https://doi.org/10.1002/hipo.10189`

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

Helm, E. van der, Gujar, N., Nishida, M., & Walker, M. P. (2011). Sleep-Dependent Facilitation of Episodic Memory Details. *PLOS ONE*, *6*(11), e27421. `https://doi.org/10.1371/journal.pone.0027421`

Henin, S., Turk-Browne, N. B., Friedman, D., Liu, A., Dugan, P., Flinker, A., . . . Melloni, L. (2019). *Learning hierarchical sequence representations across human cortex and hippocampus.* Retrieved from `http://dx.doi.org/10.1101/583856`

Hennies, N., Lambon Ralph, M. A., Kempkes, M., Cousins, J. N., & Lewis, P. A. (2016). Sleep Spindle Density Predicts the Effect of Prior Knowledge on Memory Consolidation. *The Journal of Neuroscience*, *36*(13), 3799–3810. `https://doi.org/10.1523/jneurosci.3162-15.2016`

Himmer, L., Müller, E., Gais, S., & Schönauer, M. (2017). Sleep-mediated memory consolidation depends on the level of integration at encoding. *Neurobiology of Learning and Memory*, *137*, 101–106. `https://doi.org/10.1016/j.nlm.2016.11.019`

Hinton, E. C., Dymond, S., Hecker, U. von, & Evans, C. J. (2010a). Neural correlates of relational reasoning and the symbolic distance effect: involvement of parietal cortex. *Neuroscience*, *168*(1), 138–148. `https://doi.org/10.1016/j.`

`neuroscience.2010.03.052`

Hinton, E. C., Dymond, S., Hecker, U. von, & Evans, C. J. (2010b). Neural correlates of relational reasoning and the symbolic distance effect: Involvement of parietal cortex. *Neuroscience*, *168*(1), 138–148. `https://doi.org/dprcqh`

Hinton, G. E., Dayan, P., Frey, B. J., & Neal, R. M. (1995). The "Wake-Sleep" Algorithm for Unsupervised Neural Networks. *Science*, *268*(5214), 1158–1161. `https://doi.org/10.1126/science.7761831`

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, *16*(2), 96–101. `https://doi.org/fx78p6`

Hintzman, D. L. (1986a). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*(4), 411–428. `https://doi.org/br5xp9`

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

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*(4), 528–551. `https://doi.org/10.1037/0033-295x.95.4.528`

Hintzman, D. L., & Ludlam, G. (1980). Differential forgetting of prototypes and old instances: Simulation by an exemplar-based classification model. *Memory & Cognition*, *8*(4), 378–382. `https://doi.org/10.3758/bf03198278`

Hoel, E. (2021). The overfitted brain: Dreams evolved to assist generalization. *Patterns*, *2*(5), 100244. `https://doi.org/10.1016/j.patter.2021.100244`

Hok, V., Save, E., Lenck-Santini, P. P., & Poucet, B. (2005). Coding for spatial goals in the prelimbic/infralimbic area of the rat frontal cortex. *Proceedings of the National Academy of Sciences*, *102*(12), 4602–4607. `https://doi.org/10.1073/pnas.0407332102`

Hołda, M., Głodek, A., Dankiewicz-Berger, M., Skrzypińska, D., & Szmigielska, B. (2020). Ill-defined problem solving does not benefit from daytime napping. *Frontiers in Psychology*, *11*. `https://doi.org/10.3389/fpsyg.2020.00559`

Holyoak, K. J., & Lu, H. (2021). Emergence of relational reasoning. *Current Opinion in Behavioral Sciences*, *37*, 118–124. `https://doi.org/10.1016/j.cobeha.2020.11.012`

Horst, J. S., & Hout, M. C. (2016). The novel object and unusual name (NOUN) database: A collection of novel images for use in experimental research. *Behavior Research Methods*, *48*(4), 1393–1409. `https://doi.org/10.3758/S13428-015-`

`0647-3/FIGURES/9`

Howard, M. D., Skorheim, S. W., & Pilly, P. K. (2019). A model of bi-directional interactions between complementary learning systems for memory consolidation of sequential experiences. *Frontiers in Systems Neuroscience*, *16*, 1–23. `https://doi.org/10.3389/fnsys.2022.972235`

Howard, M. W., Fotedar, M. S., Datey, A. V., & Hasselmo, M. E. (2005a). The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychological Review*, *112*(1), 75. Retrieved from `https://psycnet.apa.org/record/2004-22409-004`

Howard, M. W., Fotedar, M. S., Datey, A. V., & Hasselmo, M. E. (2005b). The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychological Review*, *112*(1), 75. Retrieved from `https://psycnet.apa.org/record/2004-22409-004`

Howard, M. W., Fotedar, M. S., Datey, A. V., & Hasselmo, M. E. (2005c). The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychological Review*, *112*(1), 75. Retrieved from `https://psycnet.apa.org/record/2004-22409-004`

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269–299. `https://doi.org/c6x596`

Howard, R. W. (2000). Generalization and Transfer: An Interrelation of Paradigms and a Taxonomy of Knowledge Extension Processes. *Review of General Psychology*, *4*(3), 211–237. `https://doi.org/10.1037/1089-2680.4.3.211`

Hu, X., Cheng, L. Y., Chiu, M. H., & Paller, K. A. (2020). Promoting memory consolidation during sleep: A meta-analysis of targeted memory reactivation. *Psychological Bulletin*, *146*(3), 218. Retrieved from `https://psycnet.apa.org/record/2020-08052-001`

Hughes, G. I., & Thomas, A. K. (2022). When Memory and Metamemory Align: How Processes at Encoding Influence Delayed Judgment-of-Learning Accuracy. *Journal of Intelligence*, *10*(4), 101. `https://doi.org/10.3390/jintelligence10040101`

Humiston, G. B., Tucker, M. A., Summer, T., & Wamsley, E. J. (2019). Resting States and Memory Consolidation: A Preregistered Replication and Meta-Analysis. *Scientific Reports*, *9*(1). `https://doi.org/10.1038/s41598-019-56033-6`

Igarashi, K. M., Lee, J. Y., & Jun, H. (2022). Reconciling neuronal representations of schema, abstract task structure, and categorization under cognitive maps in the entorhinal-hippocampal-frontal circuits. *Current Opinion in Neurobiology*, *77*, 102641. `https://doi.org/10.1016/j.conb.2022.102641`

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`

Jackson, D., & Turner, R. (2017). Power analysis for random-effects meta-analysis. *Research Synthesis Methods*, *8*(3), 290–302. `https://doi.org/10.1002/jrsm.1240`

Jacobs, J., Weidemann, C. T., Miller, J. F., Solway, A., Burke, J. F., Wei, X.-X., … Kahana, M. J. (2013). Direct recordings of grid-like neuronal activity in human spatial navigation. *Nature Neuroscience*, *16*(9), 1188–1190. `https://doi.org/10.1038/nn.3466`

Jadhav, Shantanu P., Rothschild, G., Roumis, Demetris K., & Frank, Loren M. (2016). Coordinated Excitation and Inhibition of Prefrontal Ensembles during Awake Hippocampal Sharp-Wave Ripple Events. *Neuron*, *90*(1), 113–127. `https://doi.org/10.1016/j.neuron.2016.02.010`

Jamieson, R. K., Johns, B. T., Vokey, J. R., & Jones, M. N. (2022). Instance theory as a domain-general framework for cognitive psychology. *Nature Reviews Psychology*, *1*(3), 174–183. `https://doi.org/10.1038/s44159-022-00025-3`

Jamieson, R. K., & Mewhort, D. J. K. (2009). Applying an exemplar model to the serial reaction-time task: Anticipating from experience. *Quarterly Journal of Experimental Psychology*, *62*(9), 1757–1783. `https://doi.org/cds843`

Javadi, A.-H., Tolat, A., & Spiers, H. J. (2015). Sleep enhances a spatially mediated generalization of learned values. *Learning & Memory*, *22*, 532–536. `https://doi.org/ggrv4v`

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

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825–841. `https://doi.org/10.1006/nimg.2002.1132`

Jensen, G., Alkan, Y., Ferrera, V. P., & Terrace, H. S. (2019). Reward associations do not explain transitive inference performance in monkeys. *Science Advances*, *5*(7). `https://doi.org/10.1126/sciadv.aaw2089`

Jensen, G., Alkan, Y., Muñoz, F., Ferrera, V. P., & Terrace, H. S. (2017a). Transitive inference in humans (homo sapiens) and rhesus macaques (macaca mulatta) after massed training of the last two list items. *Journal of Comparative Psychology*, *131*(3), 231–245. `https://doi.org/gbsdvf`

Jensen, G., Alkan, Y., Muñoz, F., Ferrera, V. P., & Terrace, H. S. (2017b). Transitive inference in humans (Homo sapiens) and rhesus macaques (Macaca mulatta) after massed training of the last two list items. *Journal of Comparative Psychology*, *131*(3), 231–245. `https://doi.org/10.1037/com0000065`

Jensen, G., Altschul, D., Danly, E., & Terrace, H. (2013). Transfer of a serial representation between two distinct tasks by rhesus macaques. *PLoS ONE*, *8*(7). `https://doi.org/f5c5rb`

Jensen, G., Kao, T., Michaelcheck, C., Borge, S. S., Ferrera, V. P., & Terrace, H. S. (2020). *Category learning in a transitive inference paradigm.* Retrieved from `http://dx.doi.org/10.31234/osf.io/5re9w`

Jensen, G., Munoz, F., Meaney, A., Terrace, H. S., & Ferrera, V. P. (2021). Transitive inference after minimal training in rhesus macaques (macaca mulatta). *Journal of Experimental Psychology: Animal Learning and Cognition*, *47*(4), 464–475. `https://doi.org/gr8nh6`

Jensen, G., Muñoz, F., Alkan, Y., Ferrera, V. P., & Terrace, H. S. (2015). Implicit value updating explains transitive inference performance: The betasort model. *PLoS Computational Biology*, *11*(9), 1–27. `https://doi.org/gr8nh5`

Jensen, G., Terrace, H. S., & Ferrera, V. P. (2019a). Discovering implied serial order through model-free and model-based learning. *Frontiers in Neuroscience*, *13*(August), 1–24. `https://doi.org/gf7p7r`

Jensen, G., Terrace, H. S., & Ferrera, V. P. (2019b). Discovering implied serial order through model-free and model-based learning. *Frontiers in Neuroscience*, *13*(August), 1–24. `https://doi.org/gf7p7r`

Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, *10*(1), 100–107. `https://doi.org/c2gxfx`

Johns, B. T., Jamieson, R. K., & Jones, M. N. (2023). Scalable cognitive modelling: Putting simon's (1969) ant back on the beach. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale.* `https://doi.org/10.1037/cep0000306`

Joo, H. R., & Frank, L. M. (2018b). The hippocampal sharp wave–ripple in memory retrieval for immediate use and consolidation. *Nature Reviews Neuroscience*, *19*(12), 744–757. `https://doi.org/10.1038/s41583-018-0077-1`

Joo, H. R., & Frank, L. M. (2018a). The hippocampal sharp wave–ripple in memory retrieval for immediate use and consolidation. *Nature Reviews Neuroscience*, *19*(12), 744–757. `https://doi.org/10.1038/s41583-018-0077-1`

Jurewicz, K., Cordi, M. J., Staudigl, T., & Rasch, B. (2016). No evidence for memory decontextualization across one night of sleep. *Frontiers in Human Neuroscience*, *10*. `https://doi.org/10.3389/fnhum.2016.00007`

Kahana, M. J. (2012). *Foundations of human memory*. New York, NY, US: Oxford University Press.

Káli, S., & Dayan, P. (2004). Off-line replay maintains declarative memories in a model of hippocampal-neocortical interactions. *Nature Neuroscience*, *7*(3), 286–294. `https://doi.org/cg3fc2`

Kalm, K., & Norris, D. (2018). Sequence learning creates new cortical representations instead of strengthening initial ones. *bioRxiv*. `https://doi.org/gfwf4j`

Kandel, E. R., Dudai, Y., & Mayford, M. R. (2014). The molecular and systems biology of memory. *Cell*, *157*(1), 163186. `https://doi.org/10.1016/j.cell.2014.03.001`

Kao, T., Jensen, G., Michaelcheck, C., Ferrera, V. P., & Terrace, H. S. (2018). *Absolute and relative knowledge of ordinal position*. (1913), 1–18.

Kao, T., Jensen, G., Michaelcheck, C., Ferrera, V. P., & Terrace, H. S. (2019). Absolute and relative knowledge of ordinal position on implied lists. *Journal of Experimental Psychology. Learning, Memory, and Cognition*.

Kao, T., Jensen, G., Michaelcheck, C., Ferrera, V. P., & Terrace, H. S. (2020). Absolute and relative knowledge of ordinal position on implied lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(12), 2227–2243. `https://doi.org/10.1037/xlm0000783`

Karlsson, M. P., & Frank, L. M. (2009). Awake replay of remote experiences in the hippocampus. *Nature Neuroscience*, *12*(7), 913–918. `https://doi.org/10.1038/nn.2344`

Kelly, M. A., Mewhort, D. J. K., & West, R. L. (2017). The memory tesseract: Mathematical equivalence between composite and separate storage memory models. *Journal of Mathematical Psychology*, *77*, 142–155. `https://doi.org/f9426m`

Kemp, C., & Jern, A. (2013). A taxonomy of inductive problems. *Psychonomic Bulletin & Review*, *21*(1), 23–46. `https://doi.org/10.3758/s13423-013-0467-3`

Kern, S., Nagel, J., Gerchen, M. F., Guersoy, C., Meyer-Lindenberg, A., Kirsch, P., … Feld, G. B. (2023). *Reactivation strength during cued recall is modulated by graph distance within cognitive maps.* Retrieved from `http://dx.doi.org/10.1101/2023.07.31.551234`

Kesteren, M. T. R. van, 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). `https://doi.org/10.1073/pnas.0914892107`

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

Kim, J. J., & Fanselow, M. S. (1992). Modality-Specific Retrograde Amnesia of Fear. *Science*, *256*(5057), 675–677. `https://doi.org/10.1126/science.1585183`

Kim, J. S., & Lee, S. A. (2023). Hippocampal orchestration of associative and sequential memory networks for episodic retrieval. *Cell Reports*, *42*(8), 112989. `https://doi.org/10.1016/j.celrep.2023.112989`

Kinder, A., & Shanks, D. R. (2003). Neuropsychological dissociations between priming and recognition: A single-system connectionist account. *Psychological Review*, *110*(4), 728–744. `https://doi.org/10.1037/0033-295x.110.4.728`

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., & Rusu, A. A. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, *114*(13), 3521–3526. `https://doi.org/gfvjcp`

Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., … Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLOS Computational Biology*, *13*(2), e1005350. `https://doi.org/10.1371/journal.pcbi.1005350`

Kleiner, M., Brainard, D., & Pelli, D. (2007). *What's new in psychtoolbox-3?*

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

Knowlton, B. J., & Squire, L. R. (1993). The Learning of Categories: Parallel Brain Systems for Item Memory and Category Knowledge. *Science*, *262*(5140), 1747–1749. `https://doi.org/10.1126/science.8259522`

Kolibius, L. D., Born, J., & Feld, G. B. (2020). *Vast amounts of encoded items nullify but do not reverse the effect of sleep on declarative memory.* Retrieved from `http://dx.doi.org/10.1101/2020.07.23.218735`

Kolibius, L. D., Born, J., & Feld, G. B. (2021a). Vast amounts of encoded items nullify but do not reverse the effect of sleep on declarative memory. *Frontiers in Psychology*, *11*. `https://doi.org/10.3389/fpsyg.2020.607070`

Kolibius, L. D., Born, J., & Feld, G. B. (2021b). Vast amounts of encoded items nullify but do not reverse the effect of sleep on declarative memory. *Frontiers in Psychology*, *11*. `https://doi.org/10.3389/fpsyg.2020.607070`

Korsakoff. (1889). Etude médico-psychologique sur une forme des maladies de la mémoire. *Revue Philosophique de La France Et de l'étranger*, *28*, 501530. Retrieved from `https://www.jstor.org/stable/41075158?casa_token=Uz_gRPCWCZIAAAAA:0M1lS9yrEzT_qyJc7jxZNb-kyJW7SlIH57BwF8WIGz6IszmBa4HtMAOGH4BCHSPxB`

Koster, R., Chadwick, M. J., Chen, Y., Berron, D., Banino, A., Düzel, E., … Kumaran, D. (2018a). Big-loop recurrence within the hippocampal system supports integration of information across episodes. *Neuron*, *99*(6), 13421354. `https://doi.org/10.1016/j.neuron.2018.08.009`

Koster, R., Chadwick, M. J., Chen, Y., Berron, D., Banino, A., Düzel, E., … Kumaran, D. (2018b). Big-loop recurrence within the hippocampal system supports integration of information across episodes. *Neuron*, *99*(6), 13421354. `https://doi.org/10.1016/j.neuron.2018.08.009`

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*(NOV). `https://doi.org/fh5dft`

Kuhl, B. A., Shah, A. T., DuBrow, S., & Wagner, A. D. (2010). Resistance to forgetting associated with hippocampus-mediated reactivation during new learning. *Nature Neuroscience*, *13*(4), 501–506. `https://doi.org/10.1038/nn.2498`

Kumaran, D. (2012). What representations and computations underpin the contribution of the hippocampus to generalization and inference? *Frontiers in Human Neuroscience*, *6*. `https://doi.org/10.3389/fnhum.2012.00157`

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016a). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences*, *20*(7), 512–534. `https://doi.org/f8sj7h`

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016b). 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., & Ludwig, H. (2013). Transitivity performance, relational hierarchy knowledge and awareness: Results of an instructional framing manipulation. *Hip-*

pocampus, *23*(12), 1259–1268. `https://doi.org/10.1002/hipo.22163`

Kumaran, D., & McClelland, J. L. (2012a). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, *119*(3), 573–616. `https://doi.org/f3365p`

Kumaran, D., & McClelland, J. L. (2012b). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, *119*(3), 573–616. `https://doi.org/f3365p`

Kumaran, D., & McClelland, J. L. (2012c). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, *119*(3), 573–616. `https://doi.org/f3365p`

Kumaran, D., & McClelland, J. L. (2012d). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, *119*(3), 573–616. `https://doi.org/10.1037/a0028681`

Kurinec, C. A., Whitney, P., Hinson, J. M., Hansen, D. A., & Van Dongen, H. P. (2021). Sleep deprivation impairs binding of information with its context. *Sleep*, *44*(8), zsab113. `https://doi.org/10.1093/sleep/zsab113`

Kuriyama, K., Stickgold, R., & Walker, M. P. (2004). Sleep-dependent learning and motor-skill complexity. *Learning & Memory*, *11*(6), 705–713. `https://doi.org/10.1101/lm.76304`

Kwon, M., Lee, S. W., & Lee, S.-H. (2023). Hippocampal integration and separation processes with different temporal and spatial dynamics during learning for associative memory. *Human Brain Mapping*, *44*(9), 3873–3884. `https://doi.org/10.1002/hbm.26319`

Lanczos, C. (1964). Evaluation of noisy data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, *1*(1), 76–85. `https://doi.org/10.1137/0701007`

Lau, H., Tucker, M. A., & Fishbein, W. (2010). Daytime napping: Effects on human direct associative and relational memory. *Neurobiology of Learning and Memory*, *93*(4), 554–560. `https://doi.org/10.1016/j.nlm.2010.02.003`

Lazareva, O. F., Kandray, K., & Acerbo, M. J. (2014). Hippocampal lesion and transitive inference: Dissociation of inference-based and reinforcement-based strategies in pigeons. *Hippocampus*, *25*(2), 219–226. `https://doi.org/10.1002/hipo.22366`

Lazareva, O. F., Paxton Gazes, R., Elkins, Z., & Hampton, R. (2020). Associative models fail to characterize transitive inference performance in rhesus monkeys (macaca mulatta). *Learning and Behavior*, *48*(1), 135–148. `https://doi.org/`

`gr8njh`

Le Van Quyen, M., Bragin, A., Staba, R., Crepon, B., Wilson, C. L., & Engel, J. (2008). Cell Type-Specific Firing during Ripple Oscillations in the Hippocampal Formation of Humans. *Journal of Neuroscience*, *28*(24), 6104–6110. `https://doi.org/10.1523/jneurosci.0437-08.2008`

Lee, J. L. C., Nader, K., & Schiller, D. (2017). An Update on Memory Reconsolidation Updating. *Trends in Cognitive Sciences*, *21*(7), 531–545. `https://doi.org/10.1016/j.tics.2017.04.006`

Lenth, R. V. (2023). *Emmeans: Estimated marginal means, aka least-squares means.* Retrieved from `https://CRAN.R-project.org/package=emmeans`

Lepage, M., Hawco, C., & Bodnar, M. (2015). Relational Memory as a Possible Neurocognitive Marker of Schizophrenia. *JAMA psychiatry*, *72*(9), 946–947. `https://doi.org/10.1001/jamapsychiatry.2015.0488`

Lerner, I. (2017). Sleep is for the brain. *Computational Models of Brain and Behavior*, 245–256. `https://doi.org/10.1002/9781119159193.ch18`

Lerner, I., & Gluck, M. A. (2019a). Sleep and the extraction of hidden regularities: A systematic review and the importance of temporal rules. *Sleep Medicine Reviews*, *47*, 39–50. `https://doi.org/10.1016/j.smrv.2019.05.004`

Lerner, I., & Gluck, M. A. (2019b). Sleep and the extraction of hidden regularities: A systematic review and the importance of temporal rules. *Sleep Medicine Reviews*, *47*, 39–50. `https://doi.org/ggrtwr`

Levy, S. D., & Gayler, R. W. (2009). *Lateral inhibition in a fully distributed connectionist architecture.* Retrieved from `https://simondlevy.academic.wlu.edu/files/publications/levy_gayler_4_16.pdf`

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*, *15*(8), 343–351. `https://doi.org/10.1016/j.tics.2011.06.004`

Lewis, P. A., Knoblich, G., & Poe, G. (2018b). How Memory Replay in Sleep Boosts Creative Problem-Solving. *Trends in Cognitive Sciences*, *22*(6), 491–503. `https://doi.org/10.1016/j.tics.2018.03.009`

Lewis, P. A., Knoblich, G., & Poe, G. (2018a). How memory replay in sleep boosts creative problem-solving. *Trends in Cognitive Sciences*, *22*(6), 491–503. `https://doi.org/gf5wvg`

Lewis, P. A., Manning, L. M., & Critchley, H. D. (2011). Sleep mediates retention and post-retrieval processing of source memory. *Neuropsychologia*, *49*(9), 2619–2629. `https://doi.org/10.1016/j.neuropsychologia.2011.05.009`

Libben, M., & Titone, D. (2008). The role of awareness and working memory in human transitive inference. *Behavioural Processes*, *77*(1), 43–54. `https://doi.org/10.1016/j.beproc.2007.06.006`

Lippl, S., Kay, K., Jensen, G., Ferrera, V. P., & Abbott, L. F. (2023). A mathematical theory of relational generalization in transitive inference. *bioRxiv: The Preprint Server for Biology*, 2023.08.22.554287. `https://doi.org/10.1101/2023.08.22.554287`

Lisman, J., Buzsáki, G., Eichenbaum, H., Nadel, L., Ranganath, C., & Redish, A. D. (2017). Viewpoints: how the hippocampus contributes to memory, navigation and cognition. *Nature Neuroscience*, *20*(11), 1434–1447. `https://doi.org/10.1038/nn.4661`

Liu, A., Henin, S., Abbaspoor, S., Bragin, A., Buffalo, E. A., Farrell, J. S., … Buzsáki, G. (2022). A consensus statement on detection of hippocampal sharp wave ripples and differentiation from other fast oscillations. *Nature Communications*, *13*(1). `https://doi.org/10.1038/s41467-022-33536-x`

Liu, X. L., O'Reilly, R. C., & Ranganath, C. (2021). *Effects of retrieval practice on tested and untested information: Cortico-hippocampal interactions and error-driven learning* (Vol. 75). Elsevier Inc. Retrieved from `http://dx.doi.org/10.1016/bs.plm.2021.07.003`

Liu, X., Ranganath, C., & O'Reilly, R. C. (2022). *A complementary learning systems model of how sleep moderates retrieval practice effects.* `https://doi.org/10.31219/osf.io/5aqwp`

Liu, Y., Dolan, R. J., Higgins, C., Penagos, H., Woolrich, M., Ólafsdóttir, F., … Behrens, T. E. (2021). Temporally delayed linear modelling (TDLM) measures replay in both animals and humans. *eLife*, *10*, e66917.

Liu, Y., Dolan, R. J., Kurth-Nelson, Z., & Behrens, T. E. J. (2019). Human replay spontaneously reorganizes experience. *Cell*, *178*(3), 640–652.e14. `https://doi.org/gf9gxx`

Liu, Y., Mattar, M. G., Behrens, T. E. J., Daw, N. D., & Dolan, R. J. (2021). Experience replay is associated with efficient non-local learning. *Science (New York, N.Y.)*, *372*(6544), eabf1357. `https://doi.org/gj6kcx`

Lüdecke, D. (2023). *sjPlot: Data visualization for statistics in social science.* Retrieved from `https://CRAN.R-project.org/package=sjPlot`

Luo, Y., & Zhao, J. (2018). Statistical Learning Creates Novel Object Associations via Transitive Relations. *Psychological Science*, *29*(8), 1207–1220. `https://doi.org/10.1177/0956797618762400`

Lutz, N. D., Diekelmann, S., Hinse-Stern, P., Born, J., & Rauss, K. (2017). Sleep supports the slow abstraction of GIST from visual perceptual memories. *Scientific Reports*, *7*(1), 1–9. `https://doi.org/f9rvgg`

Mack, L., & Preston, A. R. (2016). Decisions about the past are guided by reinstatement of specific memories in the hippocampus and perirhinal cortex. *NeuroImage*, *127*, 144–157. `https://doi.org/10.1016/j.neuroimage.2015.12.015`

Mack, Michael L., Preston, Alison R., & Love, Bradley 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`

Maguire, E. A., & Hassabis, D. (2011). Role of the hippocampus in imagination and future thinking. *Proceedings of the National Academy of Sciences*, *108*(11). `https://doi.org/10.1073/pnas.1018876108`

Maguire, E. A., & Mullally, S. L. (2013). The hippocampus: A manifesto for change. *Journal of Experimental Psychology: General*, *142*(4), 1180–1189. `https://doi.org/10.1037/a0033650`

Mak, M. H., Curtis, A. J., Rodd, J. M., & Gaskell, M. G. (2022). *Episodic memory and sleep are involved in the maintenance of context-specific lexical information.* Retrieved from `http://dx.doi.org/10.31234/osf.io/4dscu`

Mak, M. H., Curtis, A. J., Rodd, J. M., & Gaskell, M. G. (2023). *Recall and recognition of discourse memory across sleep and wake.* Retrieved from `http://dx.doi.org/10.31234/osf.io/6vqh9`

Mak, M. H., O'Hagan, A., Horner, A. J., & Gaskell, M. G. (2023a). *Sleep, relative to wake, increases both veridical and false memory in the DRM paradigm: A registered report.* Retrieved from `http://dx.doi.org/10.31234/osf.io/83fyx`

Mak, M. H., O'Hagan, A., Horner, A. J., & Gaskell, M. G. (2023b). *Sleep, relative to wake, increases both veridical and false memory in the DRM paradigm: A registered report.* Retrieved from `http://dx.doi.org/10.31234/osf.io/83fyx`

Mantua, J. (2018). Sleep physiology correlations and human memory consolidation: Where do we go from here? *Sleep*, *41*(2), zsx204. `https://doi.org/10.1093/sleep/zsx204`

Martini, M., Marhenke, R., Martini, C., Rossi, S., & Sachse, P. (2020). Individual differences in working memory capacity moderate effects of post-learning activity on memory consolidation over the long term. *Scientific Reports*, *10*(1). `https://doi.org/10.1038/s41598-020-74760-z`

Mata-Pereira, J., & Ponte, J.-P. da. (2017). Enhancing students' mathematical reasoning in the classroom: teacher actions facilitating generalization and justification. *Educational Studies in Mathematics*, *96*(2), 169–186. `https://doi.org/10.1007/s10649-017-9773-4`

Matorina, N., & Poppenk, J. (2021b). Memory decay distinguishes subtypes of gist. *Neurobiology of Learning and Memory*, *185*, 107519. `https://doi.org/10.1016/j.nlm.2021.107519`

Matorina, N., & Poppenk, J. (2021a). Memory decay distinguishes subtypes of gist. *Neurobiology of Learning and Memory*, *185*, 107519. `https://doi.org/gr8njt`

Mattar, M. G., & Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature Neuroscience*, *21*(11), 1609–1617. `https://doi.org/10.1038/s41593-018-0232-z`

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

McClelland, J. L., McNaughton, B. L., & Lampinen, A. K. (2020). Integration of new information in memory: New insights from a complementary learning systems perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1799), 20190637. `https://doi.org/10.1098/rstb.2019.0637`

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/dvtrsx`

Mcclelland, J. L., Mirman, D., Bolger, D. J., & Khaitan, P. (2014). Interactive activation and mutual constraint satisfaction in perception and cognition. *Cognitive Science*, *38*(6), 1139–1189. `https://doi.org/f6dxfz`

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375–407. `https://doi.org/10.1037/0033-295x.88.5.375`

McCloskey, M., & Cohen, N. J. (1989). *Catastrophic interference in connectionist networks: The sequential learning problem*. Elsevier. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0079742108605368`

McDevitt, E. A., Duggan, K. A., & Mednick, S. C. (2015). REM sleep rescues learning from interference. *Neurobiology of Learning and Memory*, *122*, 51–62. `https://doi.org/10.1016/j.nlm.2014.11.015`

McGonigle, B. O., & Chalmers, M. (1977). Are monkeys logical? *Nature*, *267*(5613), 694–696. `https://doi.org/10.1038/267694a0`

McKenzie, S., & Eichenbaum, H. (2011). Consolidation and Reconsolidation: Two Lives of Memories? *Neuron*, *71*(2), 224–233. `https://doi.org/10.1016/j.neuron.2011.06.037`

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207–238. `https://doi.org/b4wm44`

Mermillod, M., Bugaiska, A., & Bonin, P. (2013). *The stability-plasticity dilemma: Investigating the continuum from catastrophic forgetting to age-limited learning effects.* Retrieved from `https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00504/full`

Mewhort, D. J. K., & Johns, E. E. (2005). Sharpening the echo: An iterative-resonance model for short-term recognition memory. *Memory*, *13*(3-4), 300–307. `https://doi.org/10.1080/09658210344000242`

Minda, J., & Smith, J. D. (2001). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(3), 775–799. `https://doi.org/10.1037/0278-7393.27.3.775`

Molitor, R. J., Sherril, K. R., Morton, N. W., Miller, A. A., & Preston, A. R. (2021). Memory reactivation during learning simultaneously promotes dentate gyrus/CA2,3pattern differentiation and CA1memory integration. *Journal of Neuroscience*, *41*(4), 726–738. `https://doi.org/ghmkqh`

Molitor, R. J., Sherrill, K. R., Morton, N. W., Miller, A. A., & Preston, A. R. (2020). Memory Reactivation during Learning Simultaneously Promotes Dentate Gyrus/CA$_{2,3}$Pattern Differentiation and CA$_1$Memory Integration. *The Journal of Neuroscience*, *41*(4), 726–738. `https://doi.org/10.1523/jneurosci.0394-20.2020`

Momennejad, I. (2020). Learning Structures: Predictive Representations, Replay, and Generalization. *Current Opinion in Behavioral Sciences*, *32*, 155–166. `https://doi.org/10.1016/j.cobeha.2020.02.017`

Momennejad, I., Otto, A. R., Daw, N. D., & Norman, K. A. (2018). Offline replay supports planning in human reinforcement learning. *eLife*, *7*, 1–25. `https://doi.org/gfscxx`

Morgan, A., & Stickgold, R. (2017). A nap has no effect on the transitive inference task. *Sleep*, *40*(suppl_1), A82–A82. `https://doi.org/10.1093/sleepj/zsx050.221`

Morton, N. W., & Preston, A. R. (2021). Concept formation as a computational cognitive process. *Current Opinion in Behavioral Sciences*, *38*, 83–89. `https:`

`//doi.org/gkgmmd`

Morton, N. W., Schlichting, M. L., & Preston, A. R. (2020). Representations of common event structure in medial temporal lobe and frontoparietal cortex support efficient inference. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(47), 29338–29345. `https://doi.org/gj7c87`

Moscovitch, M. (2012). *Memory before and after HM: An impressionistic historical perspective.* Oxford University Press Oxford, UK. Retrieved from `https://books.google.com/books?hl=en&lr=&id=P1dVtAFlUd0C&oi=fnd&pg=PA19&dq=Moscovitch,+M.+(2012).+Memory+before+and+after+HM:+An+Impressionistic+Historical+Perspective.+In+A.+Zeman+Jones-Gotman,+M.,+Kapur,+N.+(Ed.),+Epilepsy+and+Memory.+Oxford+University+Press.&ots=USpC8bo5K1&sig=mmc7k6jmpWB7j19q9g9qP7-ej2c`

Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic memory and beyond: The hippocampus and neocortex in transformation. *Annual Review of Psychology*, *67*(1), 105–134. `https://doi.org/gdz7rv`

Moscovitch, M., & Gilboa, A. (2021). *Systems consolidation, transformation and reorganization: Multiple trace theory, trace transformation theory and their competitors.* Retrieved from `http://dx.doi.org/10.31234/osf.io/yxbrs`

Moscovitch, M., Nadel, L., Winocur, G., Gilboa, A., & Rosenbaum, R. S. (2006). The cognitive neuroscience of remote episodic, semantic and spatial memory. *Current Opinion in Neurobiology*, *16*(2), 179–190. `https://doi.org/10.1016/j.conb.2006.03.013`

Moser, E. I., Moser, M.-B., & McNaughton, B. L. (2017). Spatial representation in the hippocampal formation: a history. *Nature Neuroscience*, *20*(11), 1448–1464. `https://doi.org/10.1038/nn.4653`

Moses, S. N., Brown, T. M., Ryan, J. D., & Randal McIntosh, A. (2010). Neural system interactions underlying human transitive inference. *Hippocampus*, NA–NA. `https://doi.org/10.1002/hipo.20735`

Moses, S. N., Villate, C., Binns, M. A., Davidson, P. S. R., & Ryan, J. D. (2008). Cognitive integrity predicts transitive inference performance bias and success. *Neuropsychologia*, *46*(5), 1314–1325. `https://doi.org/10.1016/j.neuropsychologia.2007.12.009`

Moses, S., Villate, C., & Ryan, J. (2006). An Investigation of Learning Strategy Supporting Transitive Inference Performance in Humans Compared to Other Species. *Neuropsychologia*, *44*(8), 1370–1387. `https://doi.org/10.1016/j.neuropsychologia.2006.01.004`

Moyano, M. D., Carbonari, G., Bonilla, M., Pedreira, M. E., Brusco, L. I., Kaczer, L., & Forcato, C. (2022). Non-linear susceptibility to interferences in declarative memory formation. *PLOS ONE*, *17*(6), e0270678. `https://doi.org/10.1371/journal.pone.0270678`

Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. *Cognitive Psychology*, *8*(2), 228–246. `https://doi.org/10.1016/0010-0285(76)90025-6`

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), 1519–1520. `https://doi.org/10.1038/2151519a0`

Müller, G. E., & Pilzecker, A. (1900). *Experimentelle beiträge zur lehre vom gedächtniss* (Vol. 1). JA Barth.

Munoz, F., Jensen, G., Kennedy, B. C., Alkan, Y., Terrace, H. S., & Ferrera, V. P. (2020a). Learned Representation of Implied Serial Order in Posterior Parietal Cortex. *Scientific Reports*, *10*(1). `https://doi.org/10.1038/s41598-020-65838-9`

Munoz, F., Jensen, G., Kennedy, B. C., Alkan, Y., Terrace, H. S., & Ferrera, V. P. (2020b). Learned representation of implied serial order in posterior parietal cortex. *Scientific Reports*, *10*(1), 1–14. `https://doi.org/gr8nj6`

Munoz, F., Jensen, G., Kennedy, B. C., Alkan, Y., Terrace, H. S., & Ferrera, V. P. (2020c). Learned representation of implied serial order in posterior parietal cortex. *Scientific Reports*, *10*(1), 1–14. `https://doi.org/gr8nj6`

Murdock, B. B. (1993). TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review*, *100*(2), 183–203. `https://doi.org/fwc536`

Musa, A., Khan, S., Mujahid, M., & El-Gaby, M. (2022). The shallow cognitive map hypothesis: A hippocampal framework for thought disorder in schizophrenia. *Schizophrenia*, *8*(1). `https://doi.org/10.1038/s41537-022-00247-7`

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, *7*(2), 217227. `https://doi.org/10.1016/S0959-4388(97)80010-4`

Nadel, L., & Moscovitch, M. (1998). Hippocampal contributions to cortical plasticity. *Neuropharmacology*, *37*(4), 431–439. `https://doi.org/10.1016/S0028-3908(98)00057-4`

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%3C352::aid-hipo2%3E3.0.co;2-d`

Nelli, S., Braun, L., Dumbalska, T., Saxe, A., & Summerfield, C. (2023). Neural knowledge assembly in humans and neural networks. *Neuron*, *111*(9), 1504–1516.e9. `https://doi.org/10.1016/j.neuron.2023.02.014`

Nelson, A. B., & Shiffrin, R. M. (2013). The co-evolution of knowledge and event memory. *Psychological Review*, *120*(2), 356–394. `https://doi.org/f4t9md`

Nemeth, D., Gerbier, E., Born, J., Rickard, T., Diekelman, S., Fogel, S., … Dresler, M. (2023). *Optimizing methodology of sleep and memory research in humans.* `https://doi.org/10.20944/preprints201908.0208.v4`

Nemeth, D., Gerbier, E., & Janacsek, K. (2019). *Four pitfalls in sleep and memory research and how to avoid them.* Retrieved from `http://dx.doi.org/10.20944/preprints201908.0208.v1`

Newbury, C. R., Crowley, R., Rastle, K., & Tamminen, J. (2021). Sleep deprivation and memory: Meta-analytic reviews of studies on sleep deprivation before and after learning. *Psychological Bulletin*, *147*(11), 1215. `https://doi.org/10.1037/bul0000348`

Newbury, C. R., & Monaghan, P. (2019a). When does sleep affect veridical and false memory consolidation? A meta-analysis. *Psychonomic Bulletin & Review*, *26*(2), 387–400. `https://doi.org/10.3758/s13423-018-1528-4`

Newbury, C. R., & Monaghan, P. (2019b). When does sleep affect veridical and false memory consolidation? A meta-analysis. *Psychonomic Bulletin & Review*, *26*(2), 387–400. `https://doi.org/10.3758/s13423-018-1528-4`

Nielson, D. M., Smith, T. A., Sreekumar, V., Dennis, S., & Sederberg, P. B. (2015). Human hippocampus represents space and time during retrieval of real-world memories. *Proceedings of the National Academy of Sciences*, *112*(35), 11078–11083. `https://doi.org/10.1073/pnas.1507104112`

Nieuwenhuis, I. L. C., & Takashima, A. (2011). The role of the ventromedial prefrontal cortex in memory consolidation. *Behavioural Brain Research*, *218*(2), 325–334. `https://doi.org/10.1016/j.bbr.2010.12.009`

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1–32. `https://doi.org/10.1016/0010-0285(87)90002-8`

Norman, K. A., Newman, E. L., & Perotte, A. J. (2005). Methods for reducing interference in the Complementary Learning Systems model: Oscillating inhibition and autonomous memory rehearsal. *Neural Networks*, *18*(9), 1212–1228. `https:`

//doi.org/10.1016/j.neunet.2005.08.010

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 104–114. `https://doi.org/c4tjck`

Nosofsky, R. M., & Kruschke, J. K. (1992). *Investigations of an exemplar-based connectionist model of category learning.* Elsevier. `https://doi.org/10.1016/s0079-7421(08)60491-0`

Nosofsky, R. M., Little, D. R., & James, T. W. (2011). Activation in the neural network responsible for categorization and recognition reflects parameter changes. *Proceedings of the National Academy of Sciences*, *109*(1), 333–338. `https://doi.org/10.1073/pnas.1111304109`

Nour, M. M., Liu, Y., Arumuham, A., Kurth-Nelson, Z., & Dolan, R. J. (2021). Impaired neural replay of inferred relationships in schizophrenia. *Cell*, *184*(16), 4315–4328.e17. `https://doi.org/10.1016/j.cell.2021.06.012`

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a Cognitive Map.* Oxford, UK: Oxford University Press. Retrieved from `http://www.cognitivemap.net/`

O'Reilly, R. C., & Rudy, J. W. (2001). Conjunctive representations in learning and memory: Principles of cortical and hippocampal function. *Psychological Review*, *108*(2), 311. `https://doi.org/10.1037/0033-295X.108.2.311`

Ólafsdóttir, F., Barry, C., Saleem, A. B., Hassabis, D., & Spiers, H. J. (2015). Hippocampal place cells construct reward related sequences through unexplored space. *eLife*, *4*. `https://doi.org/10.7554/elife.06063`

Ólafsdóttir, F., Carpenter, F., & Barry, C. (2017). Task Demands Predict a Dynamic Switch in the Content of Awake Hippocampal Replay. *Neuron*, *96*(4), 925–935.e6. `https://doi.org/10.1016/j.neuron.2017.09.035`

Ólafsdóttir, H., Bush, D., & Barry, C. (2018). The Role of Hippocampal Replay in Memory and Planning. *Current Biology*, *28*(1), R37–R50. `https://doi.org/10.1016/j.cub.2017.10.073`

Osth, A., & Dennis, S. (2020b). *Global matching models of recognition memory.* 1–55. `https://doi.org/gr8nj9`

Osth, A., & Dennis, S. (2020a). *Global matching models of recognition memory.* `https://doi.org/10.31234/osf.io/mja6c`

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226. `https://doi.org/10.1038/nature04676`

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. `https://doi.org/10.1016/j.jbef.2017.12.004`

Palestro, J. J., Sederberg, P. B., Osth, A. F., Van Zandt, T., & Turner, B. M. (2018). *Likelihood-free methods for cognitive science*. Cham: Springer International Publishing. `https://doi.org/10.1007/978-3-319-72425-6`

Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. *Neural Networks*, *113*, 54–71. `https://doi.org/gfvx7d`

Park, S. A., Miller, D. S., & Boorman, E. D. (2021). Inferences on a multidimensional social hierarchy use a grid-like code. *Nature Neuroscience*, *24*(9), 1292–1301. `https://doi.org/10.1038/s41593-021-00916-3`

Park, S. A., Miller, D. S., Nili, H., Ranganath, C., & Boorman, E. D. (2020a). Map Making: Constructing, Combining, and Inferring on Abstract Cognitive Maps. *Neuron*, *107*(6), 1226–1238.e8. `https://doi.org/10.1016/j.neuron.2020.06.030`

Park, S. A., Miller, D. S., Nili, H., Ranganath, C., & Boorman, E. D. (2020b). Map making: Constructing, combining, and inferring on abstract cognitive maps. *Neuron*, *107*(6), 1226–1238.e8. `https://doi.org/gg78bp`

Payne, J. D., Tucker, M. A., Ellenbogen, J., Wamsley, E. J., Walker, M. P., Schacter, D. L., & Stickgold, R. (2012). Memory for Semantically Related and Unrelated Declarative Information: The Benefit of Sleep, the Cost of Wake. *PLoS ONE*, *7*(3), e33079. `https://doi.org/10.1371/journal.pone.0033079`

Peer, M., Brunec, I. K., Newcombe, N. S., & Epstein, R. A. (2021a). Structuring Knowledge with Cognitive Maps and Cognitive Graphs. *Trends in Cognitive Sciences*, *25*(1), 37–54. `https://doi.org/10.1016/j.tics.2020.10.004`

Peer, M., Brunec, I. K., Newcombe, N. S., & Epstein, R. A. (2021b). Structuring knowledge with cognitive maps and cognitive graphs. *Trends in Cognitive Sciences*, *25*(1), 37–54. `https://doi.org/ghmdn7`

Peigneux, P., Laureys, C. A. S., Delbeuck, X., & Maquet, P. (2001). Sleeping brain , learning brain . The role of sleep for memory systems. *Neuroreport*, *12*(18), 111–124. `https://doi.org/cj8h9v`

Peigneux, P., Laureys, S., Fuchs, S., Destrebecqz, A., Collette, F., Delbeuck, X., . . . Maquet, P. (2003). Learned material content and acquisition level modulate cerebral reactivation during posttraining rapid-eye-movements sleep. *NeuroImage*, *20*(1), 125–134. `https://doi.org/df5pqz`

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13. `https://doi.org/10.1016/j.jneumeth.2006.11.017`

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*(JAN). `https://doi.org/10.3389/NEURO.11.010.2008/FULL`

Pereira, S. I. R., Santamaria, L., Andrews, R., Schmidt, E., Van Rossum, M., & Lewis, P. A. (2023). Rule abstraction is facilitated by auditory cueing in REM sleep. *The Journal of Neuroscience*. `https://doi.org/10.1523/jneurosci.1966-21.2022`

Petzka, M., Charest, I., Balanos, G. M., & Staresina, B. P. (2021). Does sleep-dependent consolidation favour weak memories? *Cortex*, *134*, 65–75. `https://doi.org/10.1016/j.cortex.2020.10.005`

Pfeiffer, B. E. (2018). The content of hippocampal “replay”. *Hippocampus*, *30*(1), 6–18. `https://doi.org/10.1002/hipo.22824`

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, *497*(7447), 74–79. `https://doi.org/10.1038/nature12112`

Piaget, J. (1921). *Une forme verbale de la comparaison chez l’enfant: Un cas de transition entre le jugement prédicatif et le jugement de relation.*

Plancher, G., Massol, S., Dorel, T., & Chainay, H. (2018). Effect of negative emotional content on attentional maintenance in working memory. *Cognition and Emotion*, *33*(7), 1489–1496. `https://doi.org/10.1080/02699931.2018.1561420`

Plihal, W., & Born, J. (1997). Effects of early and late nocturnal sleep on declarative and procedural memory. *Journal of Cognitive Neuroscience*, *9*(4), 534–547. `https://doi.org/cg525d`

Poh, J. H., & Cousins, J. N. (2018). Is there a role for pattern separation during sleep? *Journal of Neuroscience*, *38*(17), 4062–4064. `https://doi.org/ggrvw4`

Pöhlchen, D., & Schönauer, M. (2020). Sleep-dependent memory consolidation in the light of rapid neocortical plasticity. *Current Opinion in Behavioral Sciences*, *33*, 118–125. `https://doi.org/10.1016/j.cobeha.2020.02.001`

Poldrack, R., Clark, J., Paré-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature*, *414*(6863), 546–550. `https://doi.org/10.1038/35107080`

Poldrack, R., & Packard, M. G. (2003). Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia*, *41*(3), 245–251. `https://doi.org/10.1016/s0028-3932(02)00157-4`

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, *116*(1), 129. `https://doi.org/10.1037/a0014420`

Poulter, S., Austen, J. M., Kosaki, Y., Dachtler, J., Lever, C., & McGregor, A. (2019). En route to delineating hippocampal roles in spatial learning. *Behavioural Brain Research*, *369*, 111936. `https://doi.org/10.1016/j.bbr.2019.111936`

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, *59*(3), 2142–2154. `https://doi.org/10.1016/j.neuroimage.2011.10.018`

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, *84*(Supplement C), 320–341. `https://doi.org/10.1016/j.neuroimage.2013.08.048`

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

Preston, Alison 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`

Prince, J. S., Charest, I., Kurzawski, J. W., Pyles, J. A., Tarr, M. J., & Kay, K. N. (2022). GLMsingle: A toolbox for improving single-trial fMRI response estimates. *bioRxiv*, 2022.01.31.478431. `https://doi.org/gpdqvs`

Prolific. (2021). *Prolific online research platform.* London, UK. Retrieved from `https://www.prolific.com`

Qualtrics. (2020). *Qualtrics survey software.* Retrieved from `https://www.qualtrics.com`

Randomizer. (2022). Retrieved from `https://www.qualtrics.com/support/survey-platform/survey-module/survey-flow/standard-elements/randomizer/`

Ranganath, C. (2010). Binding Items and Contexts: The Cognitive Neuroscience of Episodic Memory. *Current Directions in Psychological Science*, *19*(3), 131–137. `https://doi.org/10.1177/0963721410368805`

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, *13*(10), 713–726. `https://doi.org/10.1038/nrn3338`

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`

Rennó-Costa, C., Silva, A. C. C. da, Blanco, W., & Ribeiro, S. (2019). Computational models of memory consolidation and long-term synaptic plasticity during sleep. *Neurobiology of Learning and Memory*, *160*(September), 32–47. `https://doi.org/ggrvc4`

Representational similarity analysis 3.0. (2023). RSA development group. Retrieved from `https://github.com/rsagroup/rsatoolbox`

Ribot, T. (1887). *Diseases of memory.* Appleton. Retrieved from `https://books.google.com/books?hl=en&lr=&id=7e53B4p8GzIC&oi=fnd&pg=PA9&dq=Ribot,+T.+(1882).+Diseases+of+memory:+Essays+in+the+positive+psychology.+Kegan+Paul,+Trench+%26+Co.+https://doi.org/10.1037/12818-000&ots=2Hspj5rgs0&sig=TOl8LAxHJSeY3qCgmK94ZgOjrgY`

Ritchey, M., & Cooper, R. A. (2020). Deconstructing the Posterior Medial Episodic Network. *Trends in Cognitive Sciences*, *24*(6), 451–465. `https://doi.org/10.1016/j.tics.2020.03.006`

Ritvo, V. J. H., Turk-Browne, N. B., & Norman, K. A. (2019). Nonmonotonic plasticity: How memory retrieval drives learning. *Trends in Cognitive Sciences*, *23*(9), 726–742. `https://doi.org/gg7799`

Robertson, E. M. (2018). Memory instability as a gateway to generalization. *PLOS Biology*, *16*(3), e2004633. `https://doi.org/10.1371/journal.pbio.2004633`

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, III, H. L. (2008). Relativity of Remembering: Why the Laws of Memory Vanished. *Annual Review of Psychology*, *59*(1), 225–254. `https://doi.org/10.1146/annurev.psych.57.102904.190139`

Roscow, E. L., Chua, R., Costa, R. P., Jones, M. W., & Lepora, N. (2021). Learning offline: memory replay in biological and artificial reinforcement learning. *Trends in Neurosciences*, *44*(10), 808–821. `https://doi.org/10.1016/j.tins.2021.07.007`

Roumis, D. K., & Frank, L. M. (2015). Hippocampal sharp-wave ripples in waking and sleeping states. *Current Opinion in Neurobiology*, *35*, 6–12. `https://doi.org/10.1016/j.conb.2015.05.001`

Rudebeck, P. H., & Rich, E. L. (2018). Orbitofrontal cortex. *Current Biology*, *28*(18), R1083–R1088. https://doi.org/10.1016/j.cub.2018.07.018

Ryan, L., Cox, C., Hayes, S. M., & Nadel, L. (2008). Hippocampal activation during episodic and semantic memory retrieval: Comparing category production and category cued recall. *Neuropsychologia*, *46*(8), 2109–2121. https://doi.org/10.1016/j.neuropsychologia.2008.02.030

Sanders, H. I., & Warrington, E. K. (1971). Memory for Remote Events in Amnesic Patients. *Brain*, *94*(4), 661–668. https://doi.org/10.1093/brain/94.4.661

Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., … Wolf, D. H. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage*, *64*(1), 240–256. https://doi.org/10.1016/j.neuroimage.2012.08.052

Saxe, A. M., McClelland, J. L., & Ganguli, S. (2019). A mathematical theory of semantic development in deep neural networks. *Proceedings of the National Academy of Sciences of the United States of America*, *166*(23), 11537–11546. https://doi.org/gf32sx

Schacter, D. L. (2012). Constructive memory: past and future. *Dialogues in Clinical Neuroscience*, *14*(1), 7–18. https://doi.org/10.31887/DCNS.2012.14.1/dschacter

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

Schapiro, A., McDevitt, E. A., Chen, L., Norman, K. A., Mednick, S. C., & Rogers, T. T. (2017). Sleep Benefits Memory for Semantic Category Structure While Preserving Exemplar-Specific Information. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-12884-5

Schapiro, A., McDevitt, E. A., Rogers, T. T., Mednick, S. C., & Norman, K. A. (2018). Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-06213-1

Schapiro, A., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nature Neuroscience*, *16*(4), 486–492. https://doi.org/10.1038/nn.3331

Schapiro, A., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2016). *Complementary learning systems within the hippocampus: A neural network modeling approach to reconciling episodic memory with statistical learning.* Retrieved from `http://dx.doi.org/10.1101/051870`

Schapiro, A., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017a). Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160049. `https://doi.org/10.1098/rstb.2016.0049`

Schapiro, A., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017b). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711). `https://doi.org/gdz7tb`

Schapiro, A., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2015). Statistical learning of temporal community structure in the hippocampus. *Hippocampus*, *26*(1), 3–8. `https://doi.org/10.1002/hipo.22523`

Schechtman, E., Heilberg, J., & Paller, K. A. (2023a). Context matters: changes in memory over a period of sleep are driven by encoding context. *Learning & Memory*, *30*(2), 36–42. `https://doi.org/10.1101/lm.053634.122`

Schechtman, E., Heilberg, J., & Paller, K. A. (2023b). Memory consolidation during sleep involves context reinstatement in humans. *Cell Reports*, *42*(4), 112331. `https://doi.org/10.1016/j.celrep.2023.112331`

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

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

Schlichting, Margaret L., & Preston, A. R. (2015). Memory integration: neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, *1*, 1–8. `https://doi.org/10.1016/j.cobeha.2014.07.005`

Schlichting, Margaret L., & Preston, A. R. (2016a). Hippocampal–medial prefrontal circuit supports memory updating during learning and post-encoding rest. *Neu-*

robiology of Learning and Memory, *134*, 91–106. `https://doi.org/10.1016/j.nlm.2015.11.005`

Schlichting, Margaret L., & Preston, A. R. (2016b). Hippocampal–medial prefrontal circuit supports memory updating during learning and post-encoding rest. *Neurobiology of Learning and Memory*, *134*, 91–106. `https://doi.org/10.1016/j.nlm.2015.11.005`

Schlichting, Margaret L., Zeithamova, D., & Preston, A. R. (2014). CA_1 subfield contributions to memory integration and inference. *Hippocampus*, *24*(10), 1248–1260. `https://doi.org/10.1002/hipo.22310`

Schlichting, M., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal Structure Predicts Statistical Learning and Associative Inference Abilities during Development. *Journal of Cognitive Neuroscience*, *29*(1), 37–51. `https://doi.org/10.1162/jocn_a_01028`

Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. *Science*, *364*(6447), eaaw5181. `https://doi.org/10.1126/science.aaw5181`

Schuck, Nicolas W., Cai, M., Wilson, Robert C., & Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron*, *91*(6), 1402–1412. `https://doi.org/10.1016/j.neuron.2016.08.019`

Schwartenbeck, P., Baram, A., Liu, Y., Mark, S., Muller, T., Dolan, R., … Behrens, T. (2023). Generative replay underlies compositional inference in the hippocampal-prefrontal circuit. *Cell*, *186*(22), 4885–4897.e14. `https://doi.org/10.1016/j.cell.2023.09.004`

Scoville, W. B., & Milner, B. (1957). LOSS OF RECENT MEMORY AFTER BILATERAL HIPPOCAMPAL LESIONS. *Journal of Neurology, Neurosurgery & Psychiatry*, *20*(1), 11–21. `https://doi.org/10.1136/jnnp.20.1.11`

Scullin, M. K., Trotti, L. M., Wilson, A. G., Greer, S. A., & Bliwise, D. L. (2012). Nocturnal sleep enhances working memory training in Parkinson's disease but not Lewy body dementia. *Brain*, *135*(9), 2789–2797. `https://doi.org/10.1093/brain/aws192`

Sekeres, M. J., Winocur, G., & Moscovitch, M. (2018b). The hippocampus and related neocortical structures in memory transformation. *Neuroscience Letters*, *680*(August 2017), 39–53. `https://doi.org/gd5gn6`

Sekeres, M. J., Winocur, G., & Moscovitch, M. (2018a). The hippocampus and related neocortical structures in memory transformation. *Neuroscience Letters*, *680*, 39–53. `https://doi.org/10.1016/j.neulet.2018.05.006`

Sekeres, M. J., Winocur, G., Moscovitch, M., Anderson, J. A. E., Pishdadian, S., Martin Wojtowicz, J., … Grady, C. L. (2018). Changes in patterns of neural activity underlie a time-dependent transformation of memory in rats and humans. *Hippocampus*, *28*(10), 745–764. `https://doi.org/10.1002/hipo.23009`

Semon, R. (1921). *The mneme london*.

Sheldon, S., & Levine, B. (2013). Same as it ever was: Vividness modulates the similarities and differences between the neural networks that support retrieving remote and recent autobiographical memories. *NeuroImage*, *83*, 880–891. `https://doi.org/10.1016/j.neuroimage.2013.06.082`

Sherman, Brynn E., Graves, K. N., & Turk-Browne, N. B. (2020). The prevalence and importance of statistical learning in human cognition and behavior. *Current Opinion in Behavioral Sciences*, *32*, 15–20. `https://doi.org/10.1016/j.cobeha.2020.01.015`

Sherman, Brynn E., Turk-Browne, N. B., & Goldfarb, E. V. (2023). Multiple Memory Subsystems: Reconsidering Memory in the Mind and Brain. *Perspectives on Psychological Science*. `https://doi.org/10.1177/17456916231179146`

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`

Siapas, A. G., & Wilson, M. A. (1998). Coordinated Interactions between Hippocampal Ripples and Cortical Spindles during Slow-Wave Sleep. *Neuron*, *21*(5), 1123–1128. `https://doi.org/10.1016/s0896-6273(00)80629-7`

Siegel, J. M. (2021). Memory consolidation is similar in waking and sleep. *Current Sleep Medicine Reports*, *7*(1), 15–18. `https://doi.org/gr8nkt`

Singh, D., Norman, K. A., & Schapiro, A. C. (2022). A model of autonomous interactions between hippocampus and neocortex driving sleep-dependent memory consolidation. *bioRxiv*, 2022.01.31.478475. Retrieved from `https://www.biorxiv.org/content/10.1101/2022.01.31.478475v1%0Ahttps://www.biorxiv.org/content/10.1101/2022.01.31.478475v1.abstract`

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2023). *Afex: Analysis of factorial experiments*. Retrieved from `https://CRAN.R-project.org/package=afex`

Smith, J. D., & Minda, J. P. (2001). Journey to the center of the category: The dissociation in amnesia between categorization and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(4), 984–1002. `https://doi.org/10.1037/0278-7393.27.4.984`

Smith, J. F., Alexander, G. E., Chen, K., Husain, F. T., Kim, J., Pajor, N., & Horwitz, B. (2010). Imaging systems level consolidation of novel associate memories: A longitudinal neuroimaging study. *NeuroImage*, *50*(2), 826–836. `https://doi.org/10.1016/j.neuroimage.2009.11.053`

Söderlund, H., Moscovitch, M., Kumar, N., Mandic, M., & Levine, B. (2012). As time goes by: Hippocampal connectivity changes with remoteness of autobiographical memory retrieval. *Hippocampus*, *22*(4), 670–679. `https://doi.org/10.1002/hipo.20927`

Spiers, H. J., & Barry, C. (2015). Neural systems supporting navigation. *Current Opinion in Behavioral Sciences*, *1*, 47–55. `https://doi.org/10.1016/j.cobeha.2014.08.005`

Spisak, T., Bingel, U., & Wager, T. D. (2023). Multivariate BWAS can be replicable with moderate sample sizes. *Nature*, *615*(7951), E4–E7. `https://doi.org/10.1038/s41586-023-05745-x`

Squire, L. R. (2009). The legacy of patient HM for neuroscience. *Neuron*, *61*(1), 69. `https://doi.org/10.1016/j.neuron.2008.12.023`

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory Consolidation. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a021766. `https://doi.org/10.1101/cshperspect.a021766`

Squire, L. R., Horst, A. S. van der, McDuff, S. G. R., Frascino, J. C., Hopkins, R. O., & Mauldin, K. N. (2010). Role of the hippocampus in remembering the past and imagining the future. *Proceedings of the National Academy of Sciences*, *107*(44), 19044–19048. `https://doi.org/10.1073/pnas.1014391107`

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

Sterpenich, V., Albouy, G., Darsaud, A., Schmidt, C., Vandewalle, G., Vu, T. T. D., … Maquet, P. (2009). Sleep Promotes the Neural Reorganization of Remote Emotional Memory. *Journal of Neuroscience*, *29*(16), 5143–5152. `https://doi.org/10.1523/JNEUROSCI.0561-09.2009`

Stickgold, R. (2009). How do I remember? Let me count the ways. *Sleep Medicine Reviews*, *13*(5), 305–308. `https://doi.org/10.1016/j.smrv.2009.05.004`

Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: Evolving generalization through selective processing. *Nature Neuroscience*, *16*(2), 139–145. `https://doi.org/f4j6rb`

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization1. *Journal of Applied Behavior Analysis*, *10*(2), 349–367. `https://doi.org/10.1901/jaba.1977.10-349`

Strien, N. M. van, Cappaert, N. L. M., & Witter, M. P. (2009). The anatomy of memory: an interactive overview of the parahippocampal–hippocampal network. *Nature Reviews Neuroscience*, *10*(4), 272–282. `https://doi.org/10.1038/nrn2614`

Sučević, J., & Schapiro, A. C. (2022). *A neural network model of hippocampal contributions to category learning.* Retrieved from `http://dx.doi.org/10.1101/2022.01.12.476051`

Sun, W., Advani, M., Spruston, N., Saxe, A., & Fitzgerald, J. E. (2021). *Organizing memories for generalization in complementary learning systems.* Retrieved from `http://dx.doi.org/10.1101/2021.10.13.463791`

Sun, W., Advani, M., Spruston, N., Saxe, A., & Fitzgerald, J. E. (2023). Organizing memories for generalization in complementary learning systems. *Nature Neuroscience*, *26*(8), 1438–1448. `https://doi.org/10.1038/s41593-023-01382-9`

Surprenant, A. M., & Neath, I. (2013). *Principles of memory.* `https://doi.org/10.4324/9780203848760`

Sutherland, R. J., Lee, J. Q., McDonald, R. J., & Lehmann, H. (2020). Has multiple trace theory been refuted? *Hippocampus*, *30*(8), 842–850. `https://doi.org/10.1002/hipo.23162`

Sutterer, D. W., & Awh, E. (2015). Retrieval practice enhances the accessibility but not the quality of memory. *Psychonomic Bulletin & Review*, *23*(3), 831–841. `https://doi.org/10.3758/s13423-015-0937-x`

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction.* MIT press. Retrieved from `https://books.google.com/books?hl=en&lr=&id=uWV0DwAAQBAJ&oi=fnd&pg=PR7&dq=sutton+and+barto&ots=miwNq5_2p4&sig=NJuPAqLf0-TAD7QvnSL-DmY3LjA`

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia*, *44*(12), 21892208. `https://doi.org/10.1016/j.neuropsychologia.2006.05.023`

Sweegers, C. C. G., Takashima, A., Fernández, G., & Talamini, L. M. (2014). Neural mechanisms supporting the extraction of general knowledge across episodic memories. *NeuroImage*, *87*, 138–146. `https://doi.org/10.1016/j.neuroimage.2013.10.063`

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

Takashima, A., Petersson, K. M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M. J., … Fernández, G. (2006). Declarative memory consolidation in humans: A prospective functional magnetic resonance imaging study. *Proceedings of the National Academy of Sciences*, *103*(3), 756–761. `https://doi.org/10.1073/pnas.0507774103`

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

Takehara-Nishiuchi, K., & McNaughton, B. L. (2008). Spontaneous Changes of Neocortical Code for Associative Memory During Consolidation. *Science*, *322*(5903), 960–963. `https://doi.org/10.1126/science.1161299`

Tal, A., Schechtman, E., Caughran, B., Paller, K. A., & Davachi, L. (2023). *The reach of reactivation: Effects of consciously-triggered versus unconsciously-triggered reactivation of associative memory.* Retrieved from `http://dx.doi.org/10.1101/2023.07.26.546400`

Talamini, L. M., Moorselaar, D. van, Bakker, R., Bulath, M., Szegedi, S., Sinichi, M., & De Boer, M. (2022a). No evidence for a preferential role of sleep in episodic memory abstraction. *Frontiers in Neuroscience*, *16*, 871188. `https://doi.org/10.3389/fnins.2022.871188`

Talamini, L. M., Moorselaar, D. van, Bakker, R., Bulath, M., Szegedi, S., Sinichi, M., & De Boer, M. (2022b). No evidence for a preferential role of sleep in episodic memory abstraction. *Frontiers in Neuroscience*, *16*, 871188. `https://doi.org/10.3389/fnins.2022.871188`

Tallman, C. W., Clark, R. E., & Smith, C. N. (2022). Human brain activity and functional connectivity as memories age from one hour to one month. *Cognitive Neuroscience*, *13*(3-4), 115–133. `https://doi.org/10.1080/17588928.2021.2021164`

Tannen, D., Labov, W., & Fanshel, D. (1981). Therapeutic discourse: Psychotherapy as conversation. *Language*, *57*(2), 481. `https://doi.org/10.2307/413709`

Tavares, R., Mendelsohn, A., Grossman, Y., Williams, C., Shapiro, M., Trope, Y., & Schiller, D. (2015). A Map for Social Navigation in the Human Brain. *Neuron*, *87*(1), 231–243. `https://doi.org/10.1016/j.neuron.2015.06.011`

Taylor, J. E., Cortese, A., Barron, H. C., Pan, X., Sakagami, M., & Zeithamova, D. (2021). How do we generalize? *Neurons, Behavior, Data Analysis, and Theory*, *1*. `https://doi.org/10.51628/001c.27687`

Taylor, P. A., Reynolds, R. C., Calhoun, V., Gonzalez-Castillo, J., Handwerker, D. A., Bandettini, P. A., … Chen, G. (2023). Highlight results, don't hide them: Enhance interpretation, reduce biases and improve reproducibility. *NeuroImage*, *274*, 120138. `https://doi.org/10.1016/j.neuroimage.2023.120138`

Tervo-Clemmens, B., Marek, S., Chauvin, R. J., Van, A. N., Kay, B. P., Laumann, T. O., … Dosenbach, N. U. F. (2023). Reply to: Multivariate BWAS can be replicable with moderate sample sizes. *Nature*, *615*(7951), E8–E12. `https://doi.org/10.1038/s41586-023-05746-w`

Theves, S., Fernandez, G., & Doeller, C. F. (2019). The Hippocampus Encodes Distances in Multidimensional Feature Space. *Current Biology*, *29*(7), 1226–1231.e3. `https://doi.org/10.1016/j.cub.2019.02.035`

Theves, S., Fernández, G., & Doeller, C. F. (2020). The Hippocampus Maps Concept Space, Not Feature Space. *The Journal of Neuroscience*, *40*(38), 7318–7325. `https://doi.org/10.1523/jneurosci.0494-20.2020`

Thompson, R., & McConnell, J. (1955). Classical conditioning in the planarian, Dugesia dorotocephala. *Journal of Comparative and Physiological Psychology*, *48*(1), 65–68. `https://doi.org/10.1037/h0041147`

Tingley, D., & Peyrache, A. (2020). On the methods for reactivation and replay analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1799), 20190231. `https://doi.org/10.1098/rstb.2019.0231`

Tompary, A., & Davachi, L. (2017). Consolidation promotes the emergence of representational overlap in the hippocampus and medial prefrontal cortex. *Neuron*, *96*(1), 228–241.e5. `https://doi.org/gb2x29`

Tompary, A., & Davachi, L. (2022). *Consolidation-dependent behavioral integration of sequences related to mPFC neural overlap and hippocampal-cortical connectivity.* Retrieved from `http://dx.doi.org/10.1101/2022.10.20.513126`

Tononi, G., & Cirelli, C. (2014). Sleep and the price of plasticity: From synaptic and cellular homeostasis to memory consolidation and integration. *Neuron*, *81*(1), 12–34. `https://doi.org/f5pb94`

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

Tucker, M. A., & Fishbein, W. (2008). Enhancement of Declarative Memory Performance Following a Daytime Nap Is Contingent on Strength of Initial Task Acquisition. *Sleep*, *31*(2), 197–203. `https://doi.org/10.1093/sleep/31.2.197`

Tulving, E. (1972). Episodic and semantic memory. *Organization of Memory*, *1*(381-403), 1. Retrieved from `http://alicekim.ca/EMSM72.pdf`

Turner, B. M., Forstmann, B. U., & Steyvers, M. (2019). *Joint models of neural and behavioral data*. Springer International Publishing. `https://doi.org/10.1007/978-3-030-03688-1`

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. `https://doi.org/10.1109/TMI.2010.2046908`

Ujma, P. P. (2022). *Meta-analytic evidence suggests no correlation between sleep spindles and memory*. `https://doi.org/10.31234/osf.io/hjdex`

Umbach, G., Kantak, P., Jacobs, J., Kahana, M., Pfeiffer, B. E., Sperling, M., & Lega, B. (2020). Time cells in the human hippocampus and entorhinal cortex support episodic memory. *Proceedings of the National Academy of Sciences*, *117*(45), 28463–28474. `https://doi.org/10.1073/pnas.2013250117`

Vanasse, T. J., Boly, M., Allen, E. J., Wu, Y., Naselaris, T., Kay, K., … Tononi, G. (2022). Multiple traces and altered signal-to-noise in systems consolidation: Evidence from the 7T fMRI natural scenes dataset. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(44), 1–10. `https://doi.org/gr8nmg`

Vaseghi, S., Arjmandi-Rad, S., Kholghi, G., & Nasehi, M. (2021). Inconsistent effects of sleep deprivation on memory function. *EXCLI Journal*, *20*, 1011–1027. `https://doi.org/10.17179/excli2021-3764`

Vertes, R. P., & Linley, S. B. (2021). No cognitive processing in the unconscious, anesthetic-like, state of sleep. *Journal of Comparative Neurology*, *529*(3), 524–538. `https://doi.org/10.1002/cne.24963`

Vertes, R. P., & Siegel, J. M. (2005). Time for the sleep community to take a critical look at the purported role of sleep in memory processing. *Sleep*, *28*(10), 12281229. `https://doi.org/10.1093/sleep/28.10.1228`

Viganò, S., & Piazza, M. (2020). Distance and Direction Codes Underlie Navigation of a Novel Semantic Space in the Human Brain. *The Journal of Neuroscience*, *40*(13), 2727–2736. `https://doi.org/10.1523/jneurosci.1849-19.2020`

Vlach, H. A. (2014). The Spacing Effect in Children's Generalization of Knowledge: Allowing Children Time to Forget Promotes Their Ability to Learn. *Child Development Perspectives*, *8*(3), 163–168. `https://doi.org/10.1111/cdep.12079`

Voderholzer, U., Piosczyk, H., Holz, J., Landmann, N., Feige, B., Loessl, B., … Nissen, C. (2011). Sleep restriction over several days does not affect long-term recall of declarative and procedural memories in adolescents. *Sleep Medicine*, *12*(2), 170–178. `https://doi.org/10.1016/j.sleep.2010.07.017`

Walker, S., Henderson, L. M., Fletcher, F. E., Knowland, V. C. P., Cairney, S. A., & Gaskell, M. G. (2019). Learning to live with interfering neighbours: the influence of time of learning and level of encoding on word learning. *Royal Society Open Science*, *6*(4), 181842. `https://doi.org/10.1098/rsos.181842`

Wamsley, E. J. (2022). Offline memory consolidation during waking rest. *Nature Reviews Psychology*, *1*(8), 441–453. `https://doi.org/10.1038/s44159-022-00072-w`

Wamsley, E. J., Arora, M., Gibson, H., Powell, P., & Collins, M. (2023). Memory Consolidation during Ultra-short Offline States. *Journal of Cognitive Neuroscience*, 1–18. `https://doi.org/10.1162/jocn_a_02035`

Wamsley, E. J., Hamilton, K., Graveline, Y. M., Manceor, S., & Parr, E. L. (2016). Test Expectation Enhances Memory Consolidation across Both Sleep and Wake. *Plos One*, *11*(10). `https://doi.org/10.1371/journal.pone.0165141`

Wamsley, E. J., & Summer, T. (2020). Spontaneous entry into an “offline” state during wakefulness: A mechanism of memory consolidation? *Journal of Cognitive Neuroscience*, *32*(9), 1714–1734. `https://doi.org/10.1162/jocn_a_01587`

Wang, J.-Y., Weber, F. D., Zinke, K., Noack, H., & Born, J. (2017). Effects of sleep on word pair memory in children–separating item and source memory aspects. *Frontiers in Psychology*, *8*, 1533. `https://doi.org/10.3389/fpsyg.2017.01533`

Weber, F. D., Wang, J.-Y., Born, J., & Inostroza, M. (2014). Sleep benefits in parallel implicit and explicit measures of episodic memory. *Learning & Memory*, *21*(4), 190–198. `https://doi.org/10.1101/lm.033530.113`

Webler, R. D., Berg, H., Fhong, K., Tuominen, L., Holt, D. J., Morey, R. A., … Lissek, S. (2021). The neurobiology of human fear generalization: meta-analysis and working neural model. *Neuroscience & Biobehavioral Reviews*, *128*, 421–436. `https://doi.org/10.1016/j.neubiorev.2021.06.035`

Wei, Y., Krishnan, G. P., & Bazhenov, M. (2016). Synaptic mechanisms of memory consolidation during sleep slow oscillations. *Journal of Neuroscience*, *36*(15), 4231–4247. `https://doi.org/f8jkqq`

Wendelken, C., Nakhabenko, D., Donohue, S. E., Carter, C. S., & Bunge, S. A. (2008). “Brain Is to Thought as Stomach Is to ??”: Investigating the Role of Rostrolateral Prefrontal Cortex in Relational Reasoning. *Journal of Cognitive*

Neuroscience, *20*(4), 682–693. `https://doi.org/10.1162/jocn.2008.20055`

Werchan, D. M., & Gómez, R. L. (2013a). Generalizing memories over time: Sleep and reinforcement facilitate transitive inference. *Neurobiology of Learning and Memory*, *100*, 70–76. `https://doi.org/10.1016/j.nlm.2012.12.006`

Werchan, D. M., & Gómez, R. L. (2013b). Generalizing memories over time: Sleep and reinforcement facilitate transitive inference. *Neurobiology of Learning and Memory*, *100*, 70–76. `https://doi.org/ggrttf`

Werchan, D. M., & Gómez, R. L. (2013c). Generalizing memories over time: Sleep and reinforcement facilitate transitive inference. *Neurobiology of Learning and Memory*, *100*, 70–76. `https://doi.org/ggrttf`

Werchan, D. M., & Gómez, R. L. (2016). Corrigendum to "Generalizing memories over time: Sleep and reinforcement facilitate transitive inference" [Neurobiol. Learn. Memory 100 (2013) 70–76]. *Neurobiology of Learning and Memory*, *130*, 213. `https://doi.org/10.1016/j.nlm.2016.02.005`

Wernette, E. M. D., & Fenn, K. M. (2023a). Consolidation without intention: Sleep strengthens veridical and gist representations of information after incidental encoding. *Psychonomic Bulletin & Review*, *30*(4), 1475–1483. `https://doi.org/10.3758/s13423-023-02247-9`

Wernette, E. M. D., & Fenn, K. M. (2023b). Consolidation without intention: Sleep strengthens veridical and gist representations of information after incidental encoding. *Psychonomic Bulletin and Review*, *1*, 1–9. `https://doi.org/10.3758/S13423-023-02247-9/TABLES/2`

Whittington, J., McCaffary, D., Bakermans, J. J., & Behrens, T. E. (2022). How to build a cognitive map. *Nature Neuroscience*, *25*(10), 12571272. `https://doi.org/10.1038/s41593-022-01153-y`

Whittington, J., Muller, T. H., Mark, S., Chen, G., Barry, C., Burgess, N., & Behrens, T. E. J. (2019a). *The tolman-eichenbaum machine: Unifying space and relational memory through generalisation in the hippocampal formation.* https://doi.org/`https://doi.org/10/gkwg4m`

Whittington, J., Muller, T. H., Mark, S., Chen, G., Barry, C., Burgess, N., & Behrens, T. E. J. (2019b). The tolman-eichenbaum machine: Unifying space and relational memory through generalization in the hippocampal formation. *Cell*, *183*, 1249–1263.e23. `https://doi.org/ghjs77`

Wierzynski, C. M., Lubenov, E. V., Gu, M., & Siapas, A. G. (2009). State-Dependent Spike-Timing Relationships between Hippocampal and Prefrontal Circuits during Sleep. *Neuron*, *61*(4), 587–596. `https://doi.org/10.1016/j.neuron.2009.01.`

`011`

Wikenheiser, A. M., Marrero-Garcia, Y., & Schoenbaum, G. (2017). Suppression of Ventral Hippocampal Output Impairs Integrated Orbitofrontal Encoding of Task Structure. *Neuron*, *95*(5), 1197–1207.e3. `https://doi.org/10.1016/j.neuron.2017.08.003`

Wikenheiser, A. M., & Schoenbaum, G. (2016). Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. *Nature Reviews Neuroscience*, *17*(8), 513–523. `https://doi.org/10.1038/nrn.2016.56`

Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Mölle, M., & Born, J. (2011). Sleep Selectively Enhances Memory Expected to Be of Future Relevance. *The Journal of Neuroscience*, *31*(5), 1563–1569. `https://doi.org/10.1523/jneurosci.3575-10.2011`

Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of Hippocampal Ensemble Memories During Sleep. *Science*, *265*(5172), 676–679. `https://doi.org/10.1126/science.8036517`

Wilson, R. C., & Collins, A. G. E. E. G. (2019). Ten simple rules for the computational modeling of behavioral data. *PsyArXiv*, *8*, 1–35. `https://doi.org/gr7gr8`

Wiltgen, B. J., & Silva, A. J. (2007). Memory for context becomes less specific with time. *Learning & Memory*, *14*(4), 313–317. `https://doi.org/10.1101/lm.430907`

Wimmer, G. E., & Shohamy, D. (2012). Preference by Association: How Memory Mechanisms in the Hippocampus Bias Decisions. *Science*, *338*(6104), 270–273. `https://doi.org/10.1126/science.1223252`

Winocur, G., & Moscovitch, M. (2011). Memory Transformation and Systems Consolidation. *Journal of the International Neuropsychological Society*, *17*(05), 766–780. `https://doi.org/10.1017/s1355617711000683`

Wislowska, M., Heib, D. P. J., Griessenberger, H., Hoedlmoser, K., & Schabus, M. (2017). Individual baseline memory performance and its significance for sleep-dependent memory consolidation. *Sleep Spindles & Cortical Up States*, *1*(1), 2–13. `https://doi.org/10.1556/2053.1.2016.001`

Wittkuhn, L., Chien, S., Hall-mcmaster, S., & Schuck, N. W. (2021a). Replay in minds and machines. *bioRxiv.*

Wittkuhn, L., Chien, S., Hall-McMaster, S., & Schuck, N. W. (2021). Replay in minds and machines. *Neuroscience & Biobehavioral Reviews*, *129*, 367388.

Wittkuhn, L., Chien, S., Hall-mcmaster, S., & Schuck, N. W. (2021b). Replay in minds and machines. *bioRxiv.*

Wittkuhn, L., Krippner, L. M., & Schuck, N. W. (n.d.). *Statistical learning of successor representations is related to on-task replay.* `https://doi.org/10.1101/2022.02.02.478787`

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`

Wright, B. C., & Howells, D. (2008). Getting one step closer to deduction: Introducing an alternative paradigm for transitive inference. *Thinking & Reasoning*, *14*(3), 244–280. `https://doi.org/10.1080/13546780802110131`

Wu, C. M., Schulz, E., & Gershman, S. J. (2019). *Generalization as diffusion: Human function learning on graphs.* Retrieved from `http://dx.doi.org/10.1101/538934`

Wu, C.-H., Ramos, R., Katz, D. B., & Turrigiano, G. G. (2021). Homeostatic synaptic scaling establishes the specificity of an associative memory. *Current Biology*, *31*(11), 2274–2285.e5. `https://doi.org/10.1016/j.cub.2021.03.024`

Wynne, C. D. L. (1995). Reinforcement accounts for transitive inference performance. *Animal Learning & Behavior*, *23*(2), 207–217. `https://doi.org/10.3758/bf03199936`

Yassa, M. A., & Reagh, Z. M. (2013). Competitive trace theory: A role for the hippocampus in contextual interference during retrieval. *Frontiers in Behavioral Neuroscience*, *7*(JUL), 1–13. `https://doi.org/gdz7vc`

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

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. `https://doi.org/fgw736`

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

Yonelinas, A. P., Ranganath, C., Ekstrom, A. D., & Wiltgen, B. J. (2019b). Reply to ‘Active and effective replay: systems consolidation reconsidered again’. *Nature Reviews Neuroscience*, *20*(8), 507–508. `https://doi.org/10.1038/s41583-019-0192-7`

Yoshida, K., & Toyoizumi, T. (2023). Computational role of sleep in memory reorganization. *Current Opinion in Neurobiology*, *83*, 102799. `https://doi.org/10.1016/j.conb.2023.102799`

Yu, L. Q., Park, S. A., Sweigart, S. C., Boorman, E. D., & Nassar, M. R. (2021). Do grid codes afford generalization and flexible decision-making? *arXiv Preprint arXiv:2106.16219*.

Zaki, S. R., & Nosofsky, R. M. (2001a). A single-system interpretation of dissociations between recognition and categorization in a task involving object-like stimuli. *Cognitive, Affective, & Behavioral Neuroscience*, *1*(4), 344–359. `https://doi.org/10.3758/cabn.1.4.344`

Zaki, S. R., & Nosofsky, R. M. (2001b). *Single-system accounts of categorization and recognition dissociations.* American Psychological Association (APA). `https://doi.org/10.1037/e537102012-731`

Zaki, S. R., Nosofsky, R. M., Jessup, N. M., & Unverzagt, F. W. (2003). Categorization and Recognition Performance of a Memory-Impaired Group: Evidence for Single-System Models. *Journal of the International Neuropsychological Society*, *9*(3), 394–406. `https://doi.org/10.1017/S1355617703930050`

Zalesak, M., & Heckers, S. (2009). The role of the hippocampus in transitive inference. *Psychiatry Research: Neuroimaging*, *172*(1), 24–30. `https://doi.org/10.1016/j.pscychresns.2008.09.008`

Zeidman, P., & Maguire, E. A. (2016). Anterior hippocampus: the anatomy of perception, imagination and episodic memory. *Nature Reviews Neuroscience*, *17*(3), 173–182. `https://doi.org/10.1038/nrn.2015.24`

Zeithamova, D., & Bowman, C. R. (2020a). Generalization and the hippocampus: More than one story? *Neurobiology of Learning and Memory*, *175*, 107317. `https://doi.org/10.1016/j.nlm.2020.107317`

Zeithamova, D., & Bowman, C. R. (2020c). Generalization and the hippocampus: More than one story? *Neurobiology of Learning and Memory*, *175*, 107317. `https://doi.org/10.1016/j.nlm.2020.107317`

Zeithamova, D., & Bowman, C. R. (2020b). Generalization and the hippocampus: More than one story? *Neurobiology of Learning and Memory*, *175*. `https://doi.org/ghdrtb`

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

Zeithamova, D., & Preston, A. R. (2017). Temporal Proximity Promotes Integration of Overlapping Events. *Journal of Cognitive Neuroscience*, *29*(8), 1311–1323. `https://doi.org/10.1162/jocn_a_01116`

Zhang, X., Qiu, Y., Li, J., Jia, C., Liao, J., Chen, K., … Huang, R. (2022). Neural correlates of transitive inference: An SDM meta-analysis on 32 fMRI studies. *NeuroImage*, *258*, 119354. `https://doi.org/10.1016/j.neuroimage.2022.119354`

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, *20*(1), 45–57. `https://doi.org/10.1109/42.906424`

Zhou, J., Montesinos-Cartagena, M., Wikenheiser, A. M., Gardner, M. P. H., Niv, Y., & Schoenbaum, G. (2019). Complementary Task Structure Representations in Hippocampus and Orbitofrontal Cortex during an Odor Sequence Task. *Current Biology*, *29*(20), 3402–3409.e3. `https://doi.org/10.1016/j.cub.2019.08.040`

Zhou, Zhenglong, Kahana, M. J., & Schapiro, A. C. (n.d.). *Replay as context-driven memory reactivation.* `https://doi.org/10.1101/2023.03.22.533833`

Zhou, Z., Kahana, M. J., & Schapiro, A. C. (2023). Replay as context-driven memory reactivation. *bioRxiv*.

Zion, D. B., Nevat, M., Prior, A., & Bitan, T. (2019). Prior knowledge predicts early consolidation in second language learning. *Frontiers in Psychology*, *10*. `https://doi.org/10.3389/fpsyg.2019.02312`