Universitätsklinikum Hamburg-Eppendorf

Institut für Systemische Neurowissenschaften
Prof. Dr. med. Christian Büchel

Constructing Memory in Context: Environmental, Neural, and Temporal Dynamics of Multisensory Associations

Dissertation
zur Erlangung des Doktorgrades Dr. rer. biol. hum.
an der Medizinischen Fakultät der Universität Hamburg.

vorgelegt von:
Marike Christiane Maack
aus Heide (Holstein)

Hamburg, 2025

Betreuer:in / Gutachter:in der Dissertation: Prof. Dr. Michael Rose
Gutachter:in der Dissertation: PD Dr. Patrick Bruns

Vorsitz der Prüfungskommission: Prof. Dr. Michael Rose
Mitglied der Prüfungskommission: PD Dr. Patrick Bruns
Mitglied der Prüfungskommission: PD Dr. Gregor Leicht

Datum der mündlichen Prüfung: 02.10.2025

Table of Contents

List of abbreviations

ABC	Arousal-Biased Competition
ABE	Attentional Boost Effect
BIC	Binding of Item and Context
CMR	Context Maintenance and Retrieval
DMN	Default Mode Network
EEG	Electroencephalography
fMRI	functional Magnetic Resonance Imaging
IPS	Intraparietal Sulcus
MVPA	Multivariate Pattern Analysis
NE	No Entrainment
PPA	Parahippocampal Place Area
SEM	Standard Error of the Mean
SME	Subsequent Memory Effects
STS	Superior Temporal Sulcus
TCM	Temporal Context Model
TMR	Targeted Memory Reactivation
RSC	Retrosplenial Cortex
vmPFC	ventromedial Prefrontal Cortex
VR	Virtual Reality

List of Figures

1. Why Context Matters: Episodic Memory Beyond the Laboratory

Imagine navigating a new city to find a café. You rely on visual landmarks, auditory cues like the sound of honking traffic, and perhaps a remembered recommendation. These inputs arrive sequentially and across modalities, however your brain integrates them into a coherent and memorable experience. How does the brain bind such fragmented multisensory information into enduring episodic memories, and how are these memories later retrieved?

Episodic memory allows us to retrieve and relive past events by binding together not only the event content but also the spatial, temporal, and emotional context in which an event occurred (Levine et al., 1998; Tulving, 2002; Wheeler et al., 1997). In real-world settings, our experiences are inherently multisensory, with visual, auditory, tactile, and spatial modalities co-occurring dynamically over time. The brain integrates these modalities across sensory channels and time, a process referred to as multisensory associative memory formation (Clouter et al., 2017; Kim & Lee, 2023; Michelmann et al., 2018). This type of memory supports essential functions such as object recognition, spatial navigation, and social communication (Lee et al., 2017; Okray et al., 2023; Ursino et al., 2014). The process by which the brain combines individual, cross-modal sensory information into a cohesive and enduring memory is fundamental to understanding episodic memory. However, the mechanisms by which the brain selects, binds, and retrieves such multisensory associations remain unclear, particularly in real-world environments. One promising factor increasingly recognised as central to memory formation, as well as retrieval, is context.

When we experience an event with multiple sensory inputs, such as sights, sounds, and smells, our brain integrates these diverse streams into a unified, detailed representation (Cao et al., 2019; French & DeAngelis, 2020; Macaluso, 2006). This process constructs a context-rich depiction that encompasses the environment, sensory experiences, and contextual details (Hyman et al., 2012; Tovar et al., 2020; van Atteveldt et al., 2014). It is not merely about storing isolated items but about encoding the associations between what happened, where, when, and in what sensory and attentional state it took place (Baldassano et al., 2017; Horner et al., 2015; Horner & Burgess, 2013). Context encompasses not only the external environment (e.g., physical surroundings, spatial cues, sensory richness) but also internal neural states (e.g., attentional preparation, oscillatory phase) and temporal structure (e.g., the sequence in which modalities are experienced). For instance, the order in which stimuli are presented, such as seeing a face before hearing a voice or vice versa, influences the strength of memory formation and retrieval accuracy (Heusser et al., 2016; Peteranderl & Oberauer, 2018; X. Xu et al., 2024).

To this end, context is not a singular backdrop for memory, but a multidimensional construct comprising temporal, spatial, sensory, cognitive, and internal state features, all of which shape how experiences are encoded, stored, and retrieved (Z. Liu et al., 2024; Marks et al., 2022; Staudigl & Hanslmayr, 2013). For example, when walking through a new city to find a café, we do not simply register isolated landmarks; instead, we integrate them into a larger context. We encode a sequence of visual, auditory, and spatial cues that unfold over time, often occurring both simultaneously and in sequence, all embedded within a unique internal state such as curiosity or stress (Otten et al., 2002; Richter et al., 2016). These dimensions of context are not processed independently; instead, they interact to form an integrated, experience-specific memory trace (Marks et al., 2022; Yadav et al., 2022). Notably, the way context is bound into memory may differ depending on whether events are unisensory, crossmodal, or multisensory (Lehmann & Murray, 2005; Li & Deng, 2023; Thelen et al., 2015; Thelen & Murray, 2013). While unisensory learning involves processing information through a single modality (e.g., vision), crossmodal and multisensory learning integrate input across multiple senses, often yielding enhanced encoding and retrieval through mechanisms such as temporal synchrony and spatial congruence (Bruns & Röder, 2023; Okray et al., 2023; Senkowski et al., 2008). These richer sensory experiences are more likely to engage context-binding mechanisms in the brain and may lead to more robust or explicit memory formation. However, the interaction between different types of sensory processing and various dimensions of context, such as temporal order, internal states, or environmental enrichment, remains underexplored. Addressing this gap is essential for understanding how episodic memories are formed under naturalistic conditions and for clarifying the neural and cognitive mechanisms that support memory in real-world settings.

While existing literature provides insights into the brain regions and dynamic interactions involved in multisensory associative memory processes, this thesis advances our understanding by dissociating the roles of distinct contextual dimensions - external, internal, and temporal - in the formation and retrieval of multisensory associations. Specifically, it investigates how these contextual factors interact to support the encoding, consolidation, and retrieval of multisensory memory traces. To truly capture the complexity and dynamism of memory in natural settings, episodic memory research must explicitly incorporate multisensory, context-rich, and temporally structured environments. Such an approach is essential for understanding how episodic memories are formed, retrieved, and applied in real-world scenarios (Moscovitch et al., 2016; Xue, 2018). Moreover, it highlights the adaptive functions of episodic memory, including its contributions to flexible decision-making, the integration of perception with higher-order cognition, and support for complex, temporally extended learning (Xue, 2022; Zhou et al., 2025). Ultimately, this thesis aims to refine theoretical models of memory and offer practical implications for education, neurorehabilitation, and learning technologies.

1.1. Contextual Integration Drives Multisensory Memory Formation

Episodic memory emerges from the integration of diverse sensory inputs embedded within rich contextual frameworks that span external environments, internal neural states, and temporal dynamics. These contextual dimensions not only shape the encoding of information but also critically influence how memories are stored and later retrieved. Understanding how these layers of context contribute to memory formation is crucial for developing comprehensive models of memory that more accurately reflect the complexity of real-world experiences.

In daily life, experiences are inherently multisensory as we see, hear, touch, and move through space in dynamic and temporally unfolding environments. The brain combines these co-occurring sensory events into cohesive memory representations, a process that goes beyond sensory integration (Noel et al., 2018; Senkowski & Engel, 2024; van Atteveldt et al., 2014). Theoretical models such as the Binding of Item and Context (BIC) framework suggest that episodic memory forms through hippocampal-mediated binding of item-specific and contextual information (Diana et al., 2007; Ranganath, 2010), while the Context Maintenance and Retrieval (CMR) model highlights the role of ongoing contextual states in guiding memory search and retrieval (Polyn et al., 2009). Both models emphasise that context, whether external, internal, or temporal, is not secondary to memory but is integral to its core mechanisms.

A key distinction in understanding sensory memory processes concerns unisensory versus multisensory encoding and retrieval. Unisensory processing involves information from a single modality and typically engages modality-specific cortical areas (e.g., visual or auditory cortex; (Schroeder & Foxe, 2005; Tan & Hsieh, 2016). While often considered isolated, recent evidence shows that even primary sensory cortices can be modulated by input from other modalities (Schroeder & Foxe, 2005; Thunell et al., 2025). Prior multisensory experience can alter unisensory perception and memory, influencing sensory learning and retrieval dynamics (Shams et al., 2011; Thelen & Murray, 2013). Unisensory memory traces are often less vivid and detailed, especially under cognitive load (Junker et al., 2021; Shams & Seitz, 2008), and retrieval typically involves modality-specific reactivation shaped by previous multisensory associations (Butler & James, 2011; Han et al., 2022).

In contrast, multisensory processing involves the presentation of stimuli across multiple modalities, such as pairing an image with sound or tactile information with visual cues. This process recruits distributed associative networks, including multisensory integration hubs like the posterior superior temporal sulcus and parietal cortices, which promote the formation of richer, more redundant memory representations (Duarte et al., 2025; Pecher & Zeelenberg, 2022; Thelen et al., 2015). Empirical findings demonstrate that multisensory encoding generally enhances memory accuracy, vividness, and resistance to interference, facilitated by semantic congruency, shared meaning, and cross-modal binding (Lehmann & Murray, 2005; Yu et al., 2021). These experiences also induce oscillatory activity in theta (~4–7 Hz) and alpha (~8–12 Hz) bands, coordinating neural communication and timing across sensory channels (Jensen, 2002; Khader & Rösler, 2011). Importantly, neural signatures of multisensory encoding are often reactivated during retrieval, even when only one modality is presented, which highlights the durable cross-modal dependencies established during initial learning (Keil & Senkowski, 2018; H. Park & Kayser, 2019).

These distinctions, however, are not merely categorical. Unisensory and multisensory processing exist along a continuum, with the degree of sensory integration influenced by the richness of external cues, internal neural states, and temporal structure. External context, such as environmental features, spatial location, and sensory richness, modulates the strength of binding and the vividness of memories, particularly in naturalistic or immersive environments, where engagement of perceptual and spatial networks enhances encoding and retrieval (S. S. Cohen & Parra, 2016; Hendriks et al., 2024). Internal neural states, including oscillatory rhythms (e.g., theta and alpha), dynamically regulate encoding efficacy by gating sensory information and facilitating the binding process (Clouter et al., 2017; Hanslmayr et al., 2012). Temporal context, describing how stimuli are sequenced, synchronised, and arranged, provides a structural scaffold that supports episodic memories within a coherent timeline, further strengthening recall and fine-grained reconstruction (Davachi & DuBrow, 2015; Michelmann et al., 2018). Importantly, these layers of context interact synergistically to create comprehensive and durable memory traces. Rich multisensory experiences that are embedded within meaningful external environments, supported by optimal internal neural states, and structured with temporal regularities tend to produce more vivid and resilient memories than simple unisensory episodes (Gershman & Daw, 2017). Conversely, unisensory experiences can benefit from prior multisensory integration, implying that the brain retains cross-modal traces even in contexts where sensory input is limited at retrieval (Murray et al., 2005; Yu et al., 2021).

In short, multisensory learning not only enhances memory formation but fundamentally alters how memories are organised and stored. It promotes the creation of contextually richer representations that span sensory modalities and semantic categories. Retrieval, in turn, is supported by the reactivation of both item-specific and context-specific features, including the original modality, order of presentation, and associated sensory cues. As such, episodic memory is shaped not just by content, but especially by the given context, the structure and state of the system in which the initial content is embedded in. Open questions, relate to (i) the exact mechanisms of Contextual Binding in Multisensory Episodic Memory and (ii) the role of contextual binding in light of different learning demands (implicit/explicit), stimulus contents (naturalistic/artificial), and temporal scales (e.g. sequences).

1.1.1 Contextual Binding in Multisensory Episodic Memory: Models and Mechanisms

Unlike semantic memory, which encodes abstract knowledge, episodic memory preserves the details of experiences, including the *where*, *when*, and *how* (Ekstrom & Yonelinas, 2020; M. W. Howard, 2017; Tulving, 2002). This precision arises from contextual binding, a core mechanism that links items to the perceptual, emotional, and temporal contexts in which they occur (Ranganath, 2010; Yonelinas et al., 2019). Successful retrieval depends on the reinstatement of these bindings, allowing memory systems to reconstruct past experiences (Diana et al., 2007; M. W. Howard, 2017; Ranganath, 2010).

The BIC model provides a neural framework for understanding how discrete elements of experience are integrated into coherent memory traces (Diana et al., 2007; Eichenbaum, 2017; Hunsaker et al., 2013). According to this model, item identity is processed by the perirhinal cortex, spatial and environmental context by the parahippocampal cortex, and their integration by the hippocampus. This conjunctive coding enables flexible, relational memory representations. Complementing this, the CMR model posits that memory relies on a continually evolving internal context shaped by recent perceptual and cognitive experiences (M. W. Howard & Kahana, 2002; Polyn et al., 2009). Items are associated with this internal state during encoding, and reinstating it during retrieval facilitates access to the original memory trace.

Traditional models emphasised spatial and semantic cues as primary scaffolds for encoding and recall (Chang et al., 2024; Packard et al., 2017). However, recent evidence highlights the importance of temporal structure, including the order and timing of events, as a core component of contextual binding (Diamond & Levine, 2020; Heuer & Rolfs, 2021; Pu et al., 2022). Transitions in sensory input can serve as event boundaries, shaping how experiences are segmented and stored (Clewett et al., 2019; DuBrow & Davachi, 2013; Horner et al., 2016; Van De Ven et al., 2021). The hippocampus plays a central role in encoding these transitions through temporal binding and compression mechanisms (Ranganath & Hsieh, 2016).

However, it remains unclear whether modality sequences act as contextual characteristics that are reinstated during retrieval or whether they primarily scaffold associative links at encoding (Bramão et al., 2022; DuBrow & Davachi, 2016; Gerver et al., 2020; Ritchey et al., 2013). Growing neuroimaging evidence suggests that retrieval involves the reactivation of neural patterns that mirror those active during encoding, a process termed neural reinstatement or pattern completion (Bainbridge et al., 2021; Horner et al., 2015; Kuhl & Chun, 2014; Wing et al., 2015). EEG and fMRI studies show that this reinstatement is supported by oscillatory dynamics, particularly in the theta (4–7 Hz) and alpha (8–12 Hz) bands, and involves phase synchronisation and cross-frequency coupling across hippocampal and frontotemporal circuits (Friese et al., 2013; Fuentemilla, 2018; Kerrén et al., 2018).

Multisensory integration amplifies these dynamics. It facilitates the formation of complex, associative networks that interlink sensory, spatial, temporal, and emotional elements into cohesive episodic structures (Senkowski & Engel, 2024; van Atteveldt et al., 2014). This interconnectedness enhances memory richness and accessibility, making them more resistant to interference (Alwashmi et al., 2024; Barutchu et al., 2019; X. Tang et al., 2016). Evidence from immersive contexts such as Virtual Reality (VR) supports this. Compared to 2D paradigms, VR enhances the binding and integration of multisensory information through increased engagement with spatial and contextual features (N. Cooper et al., 2021; Jeong et al., 2024; Johnsdorf et al., 2023). Such environments recruit hippocampal networks more strongly, resulting in richer, more vivid episodic memories (Kerrén et al., 2025; Nikolaev et al., 2023; Ventura et al., 2019).

Importantly, contextual binding operates across multiple dimensions: external context (e.g., environmental and sensory richness), internal context (e.g., preparatory neural oscillations), and temporal context (e.g., stimulus sequence and timing). These dimensions interact during encoding and are reinstated during retrieval, determining not only whether an event is remembered, but how, whether as an isolated item, a bound association, or a coherent episode. Crucially, the strength of contextual binding may influence whether a memory remains implicit or becomes available to conscious awareness. Importantly, contextual binding not only shapes the content and structure of memory traces but may also influence their accessibility.

Whether an experience is remembered implicitly or explicitly may depend on how strongly contextual features are encoded and reinstated. This leads to a key question: Under what conditions does passive exposure to multisensory input yield implicit knowledge, and when does it lead to explicit retrieval and awareness?

1.1.2. Learning Type: From Implicit Exposure to Explicit Awareness Through Contextual Binding

Episodic memory formation does not always rely on conscious intent (Cleeremans et al., 1998; Reber & Squire, 1994; Williams, 2005). Instead, learning can emerge along a continuum, from passive, unconscious acquisition to active, deliberate encoding, commonly referred to as implicit and explicit learning. Though these learning types are often studied separately, recent research shows that they may lie on a continuum, and that contextual factors can shape how and when implicit knowledge becomes explicitly accessible (Esser et al., 2022; Goujon et al., 2014; Rose et al., 2010; Wessel et al., 2012). While traditionally treated as distinct (DeKeyser, 2003; Seger, 1994), accumulating evidence suggests that these learning modes engage overlapping but functionally differentiated neural systems and can dynamically interact, depending on contextual factors (Sun et al., 2005; Willingham & Goedert-Eschmann, 1999; J. Yang & Li, 2012).

Implicit learning refers to the incidental acquisition of environmental regularities without conscious awareness (Forkstam & Petersson, 2005; Frensch & Rünger, 2003). In contrast, explicit learning is typically associated with focused attention and deliberate encoding strategies, often resulting in declarative memory traces (Eichenbaum, 1997; Kirkhart, 2001). However, explicit knowledge can also emerge during incidental learning, particularly when learners become aware of patterns or regularities during the task (Clos et al., 2018; Gabay et al., 2023; Weinberger & Green, 2022). Neuroimaging and electrophysiological studies demonstrate that implicit and explicit learning are supported by distinct but partially overlapping brain networks. These networks differ in their connectivity and regional involvement, but also share some core structures, reflecting the complex interplay between conscious and unconscious learning processes (Destrebecqz et al., 2005; Loonis et al., 2017; J. Yang & Li, 2012). Implicit learning engages a frontal–striatal network, particularly the striatum, and is associated with increased theta-band synchrony during early learning phases (Batterink et al., 2019; M. X. Cohen, 2011; Destrebecqz et al., 2005; J. Yang & Li, 2012). In contrast, explicit learning recruits a broader network including the insula, medial prefrontal cortex, and ventrolateral prefrontal cortex, with greater involvement of alpha/beta-band synchrony during encoding (Loonis et al., 2017; Rameson et al., 2010; J. Yang & Li, 2012). These patterns suggest that implicit learning reflects automatic, low-resource acquisition, while explicit learning requires attention, working memory, and cognitive control. This distinction is also evident in oscillatory dynamics: while theta activity typically dominates during statistical learning and early-stage pattern acquisition (Colgin, 2013; Feng et al., 2015; Herweg et al., 2020), alpha and beta synchrony increase as learning becomes explicit and rule-based (Brincat & Miller, 2015; Buschman et al., 2012).

Notably, the transition from implicit to explicit learning appears to be modulated by contextual features, including stimulus complexity, environmental richness, task structure, and cognitive load (Ayala & Henriques, 2021; Bond & Taylor, 2015; Goujon et al., 2015). Environmental richness increases attentional engagement and emotional salience, enhancing the chances that implicitly learned regularities reach awareness (Chun & Jiang, 1999; Duncan et al., 2025; Zhao et al., 2013). Temporal structure and sensory modality order can act as contextual scaffolds, helping to organise information into episodic representations that support explicit retrieval (Diamond & Levine, 2020; Sadeh & Moscovitch, 2024; Stern et al., 2020). Even neural context, such as the pre-stimulus alpha state, can influence whether upcoming information is encoded implicitly or explicitly by modulating attentional preparation and sensory gating (Fell et al., 2011; Ostrowski & Rose, 2024; Zazio et al., 2022). Furthermore, cognitive load can modulate the interplay between implicit and explicit learning. Under high load, individuals may default to implicit strategies (Schnotz & Kürschner, 2007; Wierzchoń & Derda, 2019). However, particular task demands, such as semantic conflict or divided attention, can trigger explicit shifts, reflected in increased alpha activity and engagement of prefrontal control systems (Ji et al., 2017; Qu et al., 2021). So indeed, the transformation from implicit exposure to explicit memory is particularly sensitive to contextual features.

Importantly, these systems do not function in isolation. Dual-process models propose that implicit and explicit systems operate in parallel and can influence one another depending on task demands, attentional state, and contextual cues (Evans, 2008; Sun et al., 2005; Willingham & Goedert-Eschmann, 1999). Explicit awareness may facilitate or disrupt implicit learning, depending on the timing and nature of task demands (Kerz et al., 2017; Song et al., 2007). Conversely, implicit learning may lay the foundation for later explicit access, particularly when the learning context supports prediction, structure, and attentional engagement (Weinberger & Green, 2022). A key mechanism linking these systems is contextual binding, the process by which perceptual, temporal, and emotional features are encoded alongside item information to form a unified memory trace (Jiménez et al., 2006; Weinberger & Green, 2022). Whether an implicitly learned association becomes accessible for explicit recall may depend on how richly it is embedded within a contextual framework (Goujon et al., 2015).

In the context of multisensory learning, this dynamic relationship becomes especially important. Many multisensory associations, such as audiovisual sequences, are first learned implicitly through passive exposure. However, when contextual features such as environmental richness or modality sequence are present, these associations are more likely to become explicitly accessible at retrieval (D. M. Smith et al., 2022; Stern et al., 2020; Zhang et al., 2018). For example, semantically congruent audiovisual stimuli can accelerate unisensory retrieval, suggesting that multisensory context at encoding supports explicit memory performance, even when only one modality is tested later (Murray et al., 2004; Yu et al., 2021). Thus, the transition from implicit exposure to explicit memory is not fixed, but context-dependent, shaped by sensory richness, temporal structure, neural preparation, and task demands (Mulligan, 2011; Parker et al., 2007). These findings underscore that contextual binding is not only a mechanism of memory organisation but also a determinant of memory accessibility. Whether an experience remains unconscious or gets available for conscious recall may depend on how strongly it is embedded within, and reinstated by, its contextual features.

1.1.3 Stimulus Type: Naturalistic vs. Artificial Inputs in Contextual Binding

The nature of the stimuli used during learning, whether naturalistic (e.g., faces, scenes, narratives) or artificial (e.g., abstract shapes, meaningless sounds, or unrelated word lists), profoundly influences how memory traces are formed and organised (Jääskeläinen et al., 2021; Virk et al., 2024). This distinction reflects not only the perceptual and semantic richness of the input but also its ability to engage in contextual binding mechanisms across sensory, spatial, temporal, and emotional dimensions (Pooja et al., 2024; Robertson, 2003; Yonelinas et al., 2019). Naturalistic stimuli more closely resemble everyday experiences and tend to evoke stronger emotional responses, activate pre-existing semantic schemas, and facilitate the formation of coherent event representations embedded in spatiotemporal contexts (Jääskeläinen et al., 2021; Nanni-Zepeda et al., 2024; Saarimäki, 2021). Conversely, artificial stimuli provide experimental precision by minimising confounding variables but generally lack the multimodal complexity and semantic depth characteristic of real-world memory encoding (Mudrik et al., 2024; Parsons, 2015; Waskom et al., 2019).

These differences in stimulus type are reflected in the neural dynamics of encoding and retrieval. Neuroimaging and electrophysiological studies have demonstrated that naturalistic stimuli provoke stronger and more widespread oscillatory responses, including enhanced theta–

gamma coupling and alpha suppression, especially in hippocampal–prefrontal–parietal circuits involved in associative binding and contextual integration (Griffiths et al., 2021; Karakaş, 2020; Kota et al., 2020; Zhu et al., 2020). These oscillatory patterns are believed to underpin mechanisms that support segmenting continuous experience and organising complex multimodal inputs into cohesive memory traces. In fMRI studies, the encoding of naturalistic content correlates with increased activity in a distributed network, including the hippocampus, parahippocampal place area (PPA), retrosplenial cortex (RSC), and ventromedial prefrontal cortex (vmPFC), regions involved in spatial navigation, schema-based encoding, and episodic simulation (Barnett et al., 2024; Hebscher et al., 2021; Khosla et al., 2021). During retrieval, these regions often show content-specific reinstatement, providing a neural basis for the vividness and structural coherence of naturalistic memories (Nyberg et al., 2000; Staudigl & Hanslmayr, 2019; Wing et al., 2015; Xiao et al., 2017). Conversely, artificial stimuli tend to engage domain-general cognitive control processes, such as those mediated by the lateral prefrontal cortex, and rely on effortful, strategy-based encoding (Hodgson et al., 2024; Panichello & Buschman, 2021). Although these designs allow precise manipulation of task parameters, they may underrepresent the distributed and integrative processes characterising episodic memory in natural settings. Importantly, electrophysiological evidence indicates that mechanisms like hippocampal theta synchronisation, alpha suppression, and theta–gamma coupling are more robust and better predict later memory success when participants encode naturalistic content (Chanaz et al., 2023; Kragel et al., 2020; Lega et al., 2016; Murray et al., 2016).

The use of naturalistic versus artificial stimuli is therefore not merely a methodological choice but a factor that fundamentally alters the dynamics of memory formation and retrieval. Stimulus type determines not only which brain regions and oscillatory mechanisms are recruited but also how richly and flexibly information is embedded within a broader contextual framework. Consequently, the ecological validity of memory studies and their relevance to everyday cognitive functioning depend critically on whether the employed stimuli support the multidimensional binding processes that underlie episodic memory.

1.1.4 Sequences: Temporal Structure as Contextual Framework

Temporal structure is a fundamental dimension of episodic memory that supports the integration of what, where, and when into coherent representations of past experiences (Eichenbaum et al., 2012; C. Liu et al., 2022; Torres-Morales & Cansino, 2024). This scaffold becomes particularly intricate when events unfold across multiple sensory modalities, as is frequently the case in real-world environments. In such multisensory contexts, the temporal order of incoming information, whether auditory, visual, or crossmodal, serves not only as a passive timeline but as an active contextual framework that enhances encoding, chunking, and later retrieval.

Sequential learning mechanisms have long been recognised for their role in organising information into temporally structured memory traces (Farrell, 2012; M. D. Howard et al., 2022). Within episodic memory, these mechanisms enable the association of temporally adjacent elements, facilitating the formation of coherent sequences that reflect the temporal dynamics of experience (DuBrow & Davachi, 2016; Heusser et al., 2018; Sols et al., 2017). When sequences span across modalities, such as hearing a sound followed by seeing a related image, temporal structure becomes a binding feature that links otherwise discrete sensory inputs (Kayser & Logothetis, 2007). Behavioural studies suggest that such crossmodal sequences enhance memory formation, particularly when sensory features are semantically congruent and temporally predictable, as they promote chunking and facilitate the construction of meaningful associative units (Akyürek et al., 2017; Fonollosa et al., 2015; Wu et al., 2023). From a contextual binding perspective, the order of modality presentation itself may constitute a salient feature of the episodic trace. For instance, experiencing an auditory-visual (AV) sequence versus a visual-auditory (VA) sequence, even with identical content, may lead to distinct temporal and contextual imprints. While traditional memory research has focused extensively on temporal order effects within a single sensory modality, the question of how modality order is encoded, and whether it is reinstated during retrieval, remains largely unexplored.

Neurophysiological studies provide converging evidence that oscillatory dynamics play a key role in the encoding and retrieval of sequential information. Theta oscillations (4–7 Hz), particularly in hippocampal and frontotemporal circuits, are associated with the temporal organization of events and are enhanced during sequential learning in both unimodal and multisensory contexts (Benchenane et al., 2010; Gruber et al., 2018; Siapas et al., 2005; Su et al., 2024). These rhythms facilitate the temporal binding of information across time and modality. Moreover, theta–gamma coupling is proposed to provide a temporal coding scheme that supports the segmentation of continuous input into discrete memory units (Heusser et al., 2016; Ursino et al., 2023; Ursino & Pirazzini, 2024). Simultaneously, alpha (8–12 Hz) and beta (13–30 Hz) desynchronization during sequential encoding reflect anticipatory attention and the processing of temporal regularities, consistent with the principles of predictive coding (Bastos et al., 2020; Bauer et al., 2014; Capotosto et al., 2017; Strube et al., 2021).

Functional neuroimaging complements this electrophysiological perspective by demonstrating that the encoding of temporally structured, multisensory information engages a distributed network of brain regions. The medial temporal lobe, and particularly the hippocampus, supports the integration of content with its temporal context, while prefrontal regions contribute to the strategic organization and maintenance of sequences in working memory (DuBrow & Davachi, 2014; Eichenbaum et al., 2012; Lehn et al., 2009; Libby et al., 2014). Multisensory integration areas, such as the superior temporal sulcus (STS) and intraparietal sulcus (IPS), are involved in aligning temporally congruent features across modalities (Beauchamp et al., 2010; Calvert et al., 2001; Straube et al., 2018).

Together, these results suggest that temporal structure in multisensory learning is not a passive timeline, but an active representation of context, one that supports encoding, retrieval, and the reactivation of episodic details. These findings align with the view that contextual binding is a multi-dimensional process, incorporating temporal, sensory, and neural states into the memory trace. While unimodal research has explored temporal context reinstatement (DuBrow & Davachi, 2016), the role of modality order reinstatement in multisensory memory remains elusive: Is the order of sensory modalities encoded as part of the contextual representation and later reinstated during retrieval?

1.2. Reactivating the Past: Neural Signature of Context Reinstatement

A central mechanism by which episodic memories are retrieved is neural reinstatement, the reactivation of brain activity patterns that resemble those present during the original encoding episode (Kragel et al., 2021; Rau et al., 2025; Staresina et al., 2012). This reinstatement process enables the brain to reconstruct not only the content of a memory but also its associated contextual features, such as spatial location, sensory modality, emotional tone, and temporal structure (Hennings et al., 2020; Manning et al., 2011; Schechtman et al., 2023). Reinstatement has thus been widely regarded as a neurobiological basis for pattern completion, wherein partial cues at retrieval trigger the reconstruction of a more complete memory representation (Danker et al., 2011; Norman & O'Reilly, 2003; Ritchey et al., 2013). This notion is consistent with theories of hippocampal indexing and content-addressable memory systems, which propose that episodic retrieval relies on the reactivation of patterns originally formed during encoding (Heinbockel et al., 2024; Horner et al., 2015; Staresina et al., 2012, 2016; Tompary et al., 2016).

Oscillatory dynamics have emerged as a key mechanism supporting reinstatement. Low-frequency rhythms, particularly in the theta (4–7 Hz) and alpha (8–12 Hz) bands, play a critical role in coordinating the retrieval of temporally structured and contextually rich information across widespread neural networks (Menesse & Torres, 2024). Theta oscillations have been implicated in the temporal organisation of memory and the reactivation of sequential elements, reflecting interactions between the hippocampus and prefrontal cortex that support retrieval structure (Heinbockel et al., 2022; Kerrén et al., 2018; Schreiner et al., 2018; Shahbaba et al., 2022). Alpha desynchronization, by contrast, has been linked to attentional orienting during memory retrieval, facilitating access to relevant memory traces while inhibiting interference from irrelevant information (Erickson et al., 2019; Gould et al., 2011; Poch et al., 2014). These oscillatory mechanisms enable dynamic routing of information during recall and are thought to enhance the precision and efficiency of the reinstatement process.

Multivariate pattern analysis (MVPA) of EEG and fMRI data has further shown that reinstated neural patterns can be used to decode specific features of the original memory (Johnson et al., 2009; Rissman & Wagner, 2012). Such patterns can reflect specific perceptual features, such as visual stimulus category or location, as well as higher-order contextual dimensions, including the temporal order of events or the sensory modality involved (Ashton et al., 2022; Bone et al., 2020; Lifanov-Carr et al., 2024; Peelen & Downing, 2023). This suggests that reinstatement is not a unitary process but operates at multiple representational levels, engaging both early sensory regions and higher-order associative networks (Pacheco Estefan et al., 2019; Rau et al., 2025). For instance, content-specific reinstatement has been observed in early visual areas, while broader contextual features are reactivated in parietal and medial temporal structures, including the angular gyrus, precuneus, and hippocampus (Baldassano et al., 2017; Clarke et al., 2022; Hebscher et al., 2019; Sabo & Schneider, 2022). Importantly, reinstatement does not always reflect a perfect replay of encoding activity. Instead, memory retrieval is increasingly understood as a reconstructive process shaped by task demands, attentional focus, and prior knowledge (Linde-Domingo et al., 2019; Xiao et al., 2017). Reinstated activity patterns may be transformed relative to their original state, supporting the flexible use of memory for current goals (Brainerd et al., 2002; Lohnas et al., 2018; Xue, 2022). This constructive view of memory aligns with predictive coding frameworks, in which top-down signals generated during retrieval interact with stored memory representations to generate a most likely reconstruction of past events (Barron et al., 2020; M. Tang et al., 2023). Thus, the process of reinstating neural patterns involves not only reactivating what has been encoded, but dynamically reassembling contextual features to support episodic remembering.

1.3. Enhancing Multisensory Memory: Enrichment, Entrainment and Sequence Structure

While cognitive neuroscience has made significant progress in identifying neural correlates of memory encoding and retrieval, much of this work has remained correlational, often relying on occurring variations in neural or behavioural data. To truly understand how memory is shaped by context and how it can be optimised for learning, experimental approaches that systematically manipulate contextual features are essential. Such methods enable the direct investigation of mechanisms underlying contextual binding and establish causal relationships between specific factors and memory performance.

Research in cognitive neuroscience and education has long emphasised the role of contextual reinstatement in memory performance (Bramão et al., 2017; Manning et al., 2011; Staudigl et al., 2015). Environments rich in spatial, perceptual, and affective features can serve as effective contextual cues, becoming integrated into the memory trace and later facilitating retrieval (Johansson & Johansson, 2014; Lifanov-Carr et al., 2024). Multisensory richness has been shown to enhance attentional engagement, emotional involvement, and memory encoding depth, particularly when environments provide realistic, spatially coherent, and perceptually immersive experiences (Krokos et al., 2019; Parsons, 2015; Repetto et al., 2016).

VR provides a valuable experimental tool for creating such environments, enabling researchers to present controllable, naturalistic learning contexts that simulate real-world experiences while maintaining experimental precision (Bohil et al., 2011; Parsons, 2015). Crucially, VR enables embodied interaction, spatial navigation, and crossmodal sensory convergence (Bhowmik, 2024; Biocca et al., 2001). These are all factors known to support deeper encoding through increased engagement of perceptual and mnemonic systems (Krokos et al., 2019). These behavioural effects are consistent with evidence from research on environmental enrichment, which shows that enriched settings can enhance neuroplasticity and hippocampal neurogenesis, increase dendritic complexity, and promote the expression of neurotrophic factors related to learning and memory (Kempermann, 2019; Young et al., 1999). Moreover, enriched environments influence multiple systems of behavioural control, including sensory processing, motivational regulation, and motor activity, and can reduce conditioned fear and facilitate adaptive learning (Grigoryan, 2023). This empirical foundation motivates the *enriched environment hypothesis* developed later in this thesis, which proposes that sensory and contextual richness systematically improves memory performance by enhancing contextual binding mechanisms.

As contextual binding is a core mechanism of episodic memory, it could be modulated, enhanced or impaired by changes in environmental stimulation (Petrucci et al., 2025; Vedder et al., 2015), preparatory brain states (Hebscher & Voss, 2020; Medvedeva et al., 2019), and the temporal structure of experience (Waldhauser et al., 2016). Each of these domains corresponds to a different layer of context: external, internal, and temporal. Understanding how these layers can be optimised offers powerful leverage for enhancing real-world learning and memory performance. The brain's oscillatory dynamics reflect and regulate attentional, perceptual, and mnemonic functions. Theta (3–7 Hz) and alpha (8–12 Hz) rhythms are especially important in episodic memory, associated with temporal integration and attentional gating, respectively (Benchenane et al., 2011; Cruz et al., 2025; van Ede, 2018). Pre-stimulus oscillatory states, the neural rhythms present before an event begins, can bias encoding success (Guderian et al., 2009; Ostrowski & Rose, 2024; Scholz et al., 2017; Sweeney-Reed et al., 2016). Higher alpha power is linked to enhanced top-down filtering (Klimesch et al., 2010; Magosso & Borra, 2024; Scholz et al., 2021), while theta supports episodic binding during stimulus processing (Griffiths et al., 2021; Herweg et al., 2020; Köster et al., 2018; Nyhus & Curran, 2010).

Temporal structure is a defining feature of episodic memory. The order in which stimuli appear, especially across sensory modalities (e.g., AV vs. VA), may itself become encoded as a contextual feature (Baldassano et al., 2017; Davachi & DuBrow, 2015). Although much research has focused on event order within a single modality, the temporal order of multisensory input remains underexplored as a context cue. The process of contextual binding, which involves linking items with their associated features such as location, time, or modality, is an essential mechanism in associative and episodic memory (Ranganath, 2010; Yonelinas et al., 2019). This binding process is thought to rely on interactions between the hippocampus, which supports integration across temporal and spatial gaps, and the prefrontal cortex, which contributes strategic and organisational processes (Eichenbaum, 2017; Place et al., 2016; Preston & Eichenbaum, 2013; Staresina & Davachi, 2009). Recognising that contextual binding is modifiable has important implications. In educational contexts, enriched sensory environments (Clemenson & Stark, 2015; H. Wang et al., 2020) or oscillatory stimulation could improve learning and retention (ten Oever et al., 2020). In clinical populations such as ageing adults or patients with memory disorders, targeted interventions may support associative binding by enhancing specific contextual features during learning (Hwang et al., 2024; H.-L. Yang et al., 2018). More broadly, this research highlights the importance of designing memory studies and memory-enhancing interventions that move beyond unimodal, static paradigms and instead reflect the complex, dynamic, and multisensory nature of real-world experiences.

Together, these three empirical aspects - enriched sensory environments, rhythmic entrainment, and modality sequence structure - highlight how contextual binding is not fixed but dynamic. They demonstrate that memory can be enhanced through the synchronisation of brain state, environmental richness, and temporal structure, each contributing to more robust encoding and accessible retrieval. This framework provides a new view for understanding how and when episodic memories form, emphasizing the importance of studying contextual binding across levels, not only within items, but across neural states and perceptual timelines. These findings lay the groundwork for the central research questions of this thesis: How do external, internal, and temporal context features interact to shape multisensory episodic memory?

1.4. Research scope and aim

Despite decades of research on episodic memory, fundamental questions remain about how memories are formed and retrieved in the multisensory, temporally structured, and context-rich environments that characterise real-world experiences. So far, most experimental paradigms in memory research have relied on static, unimodal stimuli presented under tightly controlled laboratory conditions, offering critical but limited insights into underlying mechanisms. These simplified designs often overlook how context, both external and internal, as well as temporal factors, actively shape the structure of memory traces and how memory operates in the dynamic environments of everyday life. This thesis addresses three gaps in the literature. First, context is often treated as an extraneous background variable rather than a mechanistic component of memory encoding and retrieval. Although theories such as encoding specificity and context-dependent memory acknowledge contextual effects, they typically frame context as a static cue for recall rather than as an integral, dynamic feature of the memory trace itself. This underestimates the roles that environmental richness, internal brain states, and temporal structure play in shaping memory formation during encoding and memory access at retrieval. Second, while multisensory memory research has shown that learning involving multiple sensory modalities enhances memory performance, few studies have integrated these behavioural effects with their neural underpinnings. There is a lack of frameworks that explicitly link external sensory context, internal oscillatory dynamics, and temporal stimulus structure as mutually reinforcing dimensions of contextual binding. Without this integration, our understanding of how memory emerges from the interaction of brain, body, and environment remains incomplete. Third, memory research continues to be primarily conducted in artificial, decontextualized settings, although real-world memories are multisensory, spatially situated, and embedded in time. The potential of enriched, immersive environments, such as VR, to enhance memory via increased contextual engagement is underexplored, especially in combination with neural or temporal manipulations.

The primary aim of this thesis is to uncover mechanisms of contextual binding in multisensory associative memory. Specifically, it examines how three core dimensions of context — external, internal, and temporal — impact the processes of memory formation and retrieval. External context pertains to the sensory and environmental richness of the learning environment, such as immersive virtual reality versus conventional 2D screen-based settings. Internal context refers to the preparatory neural oscillatory state prior to encoding, exemplified by alpha and theta entrainment. Temporal context refers to the structural sequence of modality presentation, such as the auditory–visual versus visual–auditory order. Investigating these dimensions across three empirical studies, the research employs immersive technologies, EEG time-frequency analyses, and multivariate pattern classification techniques.

The first study focused on whether immersive virtual reality environments enhance explicit memory for incidental multisensory associations compared to traditional screen-based settings. This approach tests the hypothesis that a richly immersive environment facilitates the transformation of implicit regularities into explicit memory by strengthening contextual binding mechanisms. The second study examines how pre-stimulus sensory entrainment at alpha, theta, or arrhythmic frequencies modulates the internal neural state prior to encoding. By manipulating oscillatory activity, the study aims to determine whether alpha- or theta-band oscillations support crossmodal associative encoding, possibly through attentional gating or temporal structure. The third study investigates whether the temporal order of crossmodal stimuli, such as auditory-visual versus visual-auditory, serves as an implicit contextual cue that is encoded and reinstated during recognition. Across these investigations, neural reinstatement serves as a key marker of contextual binding, reflecting how the brain retrieves not only the content of an event but also the *where*, *when*, and sensory or neural contexts in which it was encoded. Together, these studies aim to advance our understanding of the dynamic interplay between external environment, internal neural states, and structural temporal features in shaping episodic memories within naturalistic settings.

2. Experimental studies

2.1. Study I: VR environments enhance the shift from implicit to explicit learning

2.1.1. Background

The transition from implicit to explicit memory is a fundamental but underexplored aspect of episodic learning, especially in multisensory, naturalistic environments. Most memory studies employ unimodal stimuli and artificial laboratory settings, which may underrepresent the mechanisms that govern memory formation in everyday life. The environmental context, including its multisensory richness, spatial structure, and affective properties, may play a crucial role in promoting explicit awareness of associative regularities during incidental learning (Goujon et al., 2014; S. M. Smith & Vela, 2001). Drawing on theories of contextual binding, predictive coding, and the enriched environment hypothesis, this study tested whether encoding in an immersive, multisensory virtual reality (VR) environment facilitates the emergence of explicit memory compared to a traditional 2D screen setting. Prior literature suggests that environmental richness enhances attention, motivation, and memory performance, particularly for episodic and associative memory (Clemenson & Stark, 2015; Krokos et al., 2019; Makowski et al., 2017). However, it remains unclear whether such enrichment affects the transition from implicit to explicit knowledge, especially in tasks where learning is not goal-directed.

2.1.2. Methods

A total of 102 participants were recruited and assigned to either an immersive VR group or one of two 2D screen-based groups (non-enriched; PC-short and PC-long). The study employed a between-subjects design with the learning environment (VR vs. a 2D screen) served as the main manipulation. Participants engaged in an incidental sequential association learning task, during which four sound–image pairs were presented in a fixed, sequential order across multiple repeated blocks. Importantly, participants were not informed about the associative nature of the task; their only instruction was to attend to the stimuli, allowing for the investigation of naturalistic encoding processes. Memory performance was assessed after the learning phase through several measures, including a completion task in which participants identified the correct sound–image pairs, a free recall task to evaluate spontaneous retrieval, and confidence ratings to distinguish explicit from implicit responses. To further capture participants' subjective experiences, stimulus ratings, such as pleasantness and realism, were collected, providing insights into perceived environmental richness and immersion.

2.1.3. Results

The results demonstrated that although the total number of learned sequence pairs did not differ significantly between groups, the quality of memory, specifically, the degree of explicit access, was significantly enhanced in the VR condition. Participants in the VR environment recalled a greater proportion of pairs with high confidence in both the completion task and the free recall, particularly when compared to the 2D PC-long condition. This selective enhancement of explicit memory suggests that the immersive, perceptually enriched VR context facilitated deeper encoding and stronger conscious access to learned associations. Importantly, these effects could not be attributed to differences in stimulus exposure or subjective stimulus salience, as there were no significant differences in overall stimulus ratings between conditions. Furthermore, ANCOVA analyses confirmed that the observed interaction between environment and memory type remained robust when controlling for stimulus ratings. These findings highlight that the environmental richness of VR can improve the accessibility and awareness of incidentally learned multisensory associations, an effect with important implications for understanding how contextual embedding supports explicit memory formation.

2.1.4. Conclusion

The study demonstrates that environmental enrichment, operationalised via immersive virtual reality, enhances the emergence of explicit memory in an incidental multisensory sequential-association learning task. Importantly, these effects occurred without explicit memorisation instructions, suggesting that enriched environments may facilitate spontaneous awareness of regularities through increased attentional engagement, emotional salience, and context binding. The findings support the view that external context is not a passive backdrop, but an active component of the memory trace. They also challenge accounts that attribute learning improvements solely to stimulus fluency or exposure repetition. Instead, enrichment may promote the transition from implicit to explicit learning by enhancing the salience of prediction errors or increasing cognitive resources available for binding across modalities. By anchoring memory formation in realistic, multisensory environments, this study contributes to a more ecologically valid understanding of how episodic memories emerge in everyday life. It also provides the first layer of support for the broader thesis claim: that external, internal, and temporal contexts each play a critical role in contextual binding within multisensory episodic memory.

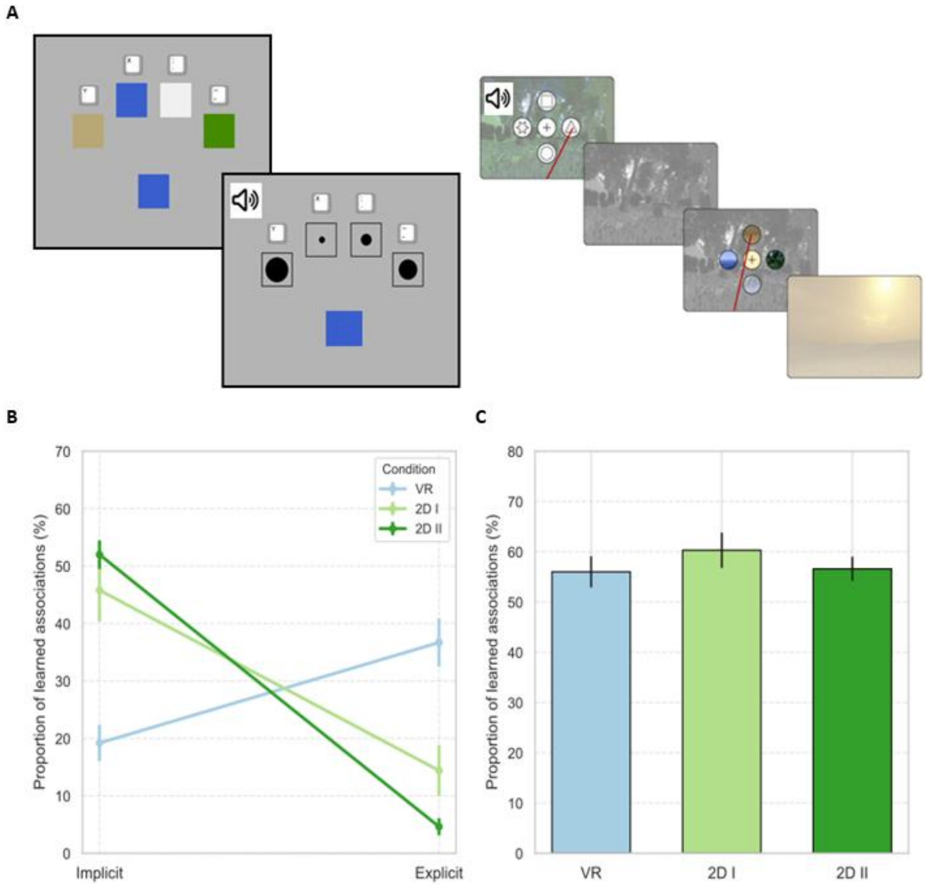


Figure 1. Task and memory performance Study I. **(A)** *Left* shows the task for the 2D with alternating visual squares and auditory tones in short (260 trials) and long (520 trials) versions. *Right* shows the VR task with immersive landscapes and tones. (B) Completion task performance by Learning-Type (implicit, explicit) and Condition (VR, 2D-short, 2D-long); lines indicate group means with SEM. (C) Proportion of learned associations (%) across all applications, with SEM error bars.

2.2. Study II: Pre-Stimulus Entrainment Dissociates the Roles of Oscillations in Crossmodal Associative Memory Formation

2.2.1. Background

While the external environment shapes memory encoding, the brain’s internal neural state, particularly oscillatory activity, plays a crucial role. Alpha (8–12 Hz) and theta (3–7 Hz) rhythms have been consistently linked to episodic memory, with theta supporting associative binding and temporal sequencing, and alpha indexing attentional filtering and preparation to encode (Herweg et al., 2020; Klimesch et al., 2010; Waldhauser et al., 2012). Recent evidence suggests that oscillatory dynamics prior to stimulus onset can bias encoding outcomes, especially in crossmodal tasks (Fell et al., 2011; Guderian et al., 2009; Otten et al., 2010). However, most findings are correlational. To test causality, sensory entrainment provides a non-invasive method for modulating endogenous rhythms through rhythmic visual stimulation. We hypothesised that visual flicker would successfully entrain neural oscillations and that both theta and alpha entrainment would improve memory performance compared to arrhythmic controls. We further expected differential effects between theta and alpha conditions, and overall enhancement relative to a no-entrainment baseline.

2.2.2. Methods

A total of 150 participants were randomly assigned to one of four between-subjects entrainment conditions: Theta (5 Hz), Alpha (9 Hz), arrhythmic control (phase-scrambled stimulation), or a no-entrainment (NE). Participants performed a Subsequent Memory Effects (SME) task, in which simultaneously presented auditory–visual stimulus pairs were encoded. Each trial was preceded by a 2-second visual flicker stimulus designed to entrain brain oscillations at the assigned frequency. Entrainment was presented via centrally flickering naturalistic images. EEG was recorded throughout the session, including both the encoding and recognition phases, to assess entrainment success and to analyse pre-stimulus spectral power differences across conditions. After encoding, participants completed a brief intermission task to prevent recency effects and to reorient attention. Memory performance was then assessed using an old/new recognition task with previously studied and novel audiovisual pairs. Behavioural outcomes were analysed using recognition accuracy and signal detection measures (d′). Bayesian statistical methods were employed to examine differences in memory performance and pre-stimulus brain activity across entrainment conditions.

2.2.3. Results

Time–frequency analyses confirmed successful frequency-specific entrainment during the pre-stimulus interval. Participants in the theta condition exhibited significantly increased power in the 3–7 Hz range throughout the late entrainment period, while the alpha condition showed elevated power in the 6–10 Hz range across the entire pre-stimulus window. These effects were absent post-stimulus and were spatially centered over occipital and parieto-occipital electrodes, indicating frequency-specific modulation of neural activity by rhythmic visual stimulation. Comparisons with both the control and NE groups confirmed that entrainment selectively enhanced power at the targeted frequencies. Behaviourally, the alpha entrainment group demonstrated significantly better memory performance, as indexed by higher d′ scores and hit rates, compared to the control group. This effect was not accompanied by an increase in false alarms, supporting the interpretation that alpha entrainment enhanced associative memory encoding. In contrast, theta entrainment did not differ significantly from the control or alpha conditions, with Bayes factors indicating weak support for the null hypothesis. Ratings of attention and fatigue changed across the experiment but did not differ between conditions, suggesting that differences in subjective task engagement did not drive the memory effects.

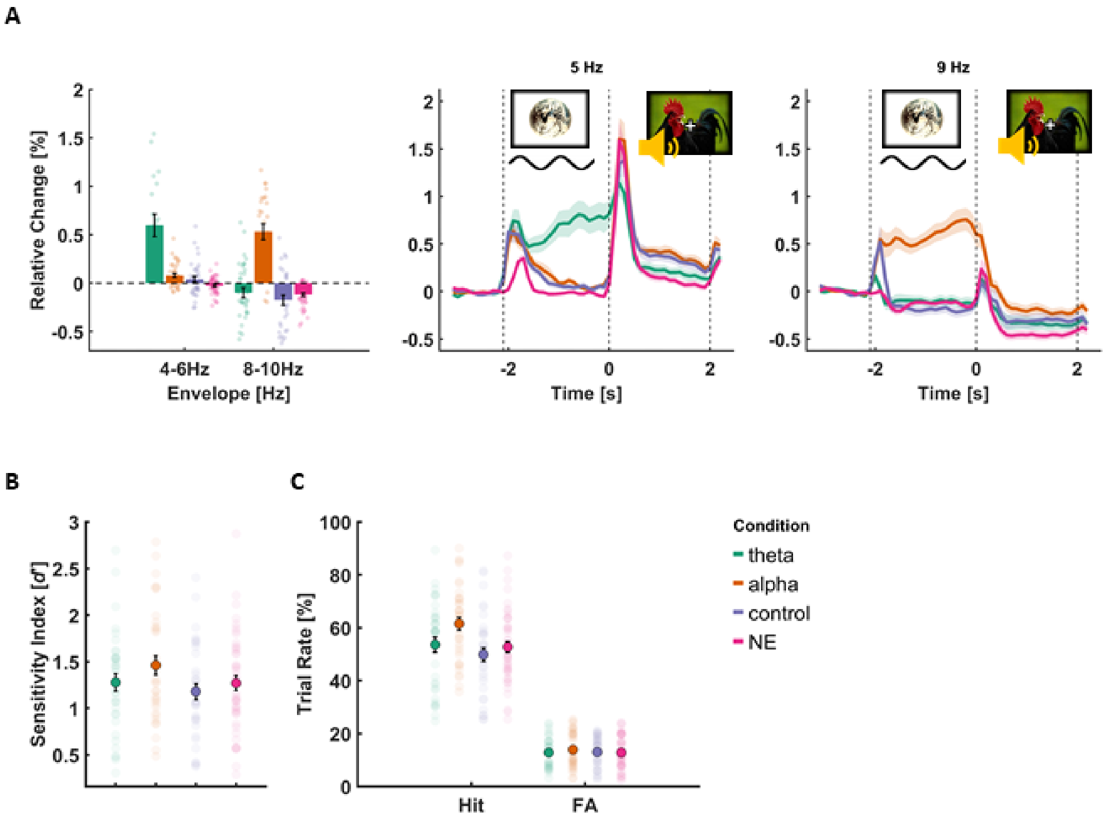


Figure 2. Effects of neural entrainment on recognition memory. **(A)** Left shows mean oscillatory power changes (±1 Hz, 5 Hz and 9 Hz bands) relative to baseline for each participant; black error bars indicate SEM. Right displays average power time courses across the trial period for each group, with SEM shading. All data are from occipital electrodes (O1, O2, Oz). Vertical lines mark stimulus onset and presentation windows. **(B)** Mean sensitivity index (d′) with SEM across the three entrainment conditions. **(C)** Trial rates of hits and false alarms across individuals for each group, with SEM error bars.

2.2.4. Conclusion

This study demonstrates that pre-stimulus alpha entrainment enhances crossmodal associative memory, likely by modulating attentional preparation and sensory gating mechanisms prior to stimulus presentation. In contrast, theta entrainment did not improve performance, suggesting that theta-related memory benefits may rely more on online, stimulus-bound dynamics than on preparatory brain states. These findings support the growing view that pre-stimulus oscillatory dynamics shape encoding quality (Ostrowski & Rose, 2024; Zazio et al., 2022) and can be modulated causally through rhythmic sensory input (Köster et al., 2019; D. Wang, 2024). They also emphasise the importance of timing specificity in entrainment studies: while alpha may act as a preparatory gatekeeper, theta may support binding during or after stimulus presentation. The results provide converging neural and behavioural evidence for the role of internal brain states as a layer of contextual binding in episodic memory. They suggest that preparatory alpha activity facilitates the binding of crossmodal associations, particularly under conditions of low semantic congruency. By using entrainment to manipulate internal context at encoding, this study complements the findings of Study 1 on external context. It sets the stage for Study 3, which investigates how temporal structure, in particular, the order of sensory modalities, is encoded and reinstated as a context feature during retrieval.

2.3. Study III: The order of multisensory associative sequences is reinstated as context-feature during recognition

2.3.1. Background

Temporal structure is a core component of episodic memory, supporting the binding of events into coherent sequences and enabling the brain to reconstruct not only what happened, but also in what order (Clewett & Davachi, 2017). While extensive research has examined the temporal context in unimodal memory tasks, the role of modality sequence order, i.e. the specific order in which multisensory stimuli are presented (e.g., auditory-visual vs. visual-auditory), remains underexplored. Multisensory episodic memory typically involves asynchronous and sequential input. Modality order may thus act as a temporal context feature, shaping how the brain encodes and later reconstructs multisensory experiences (Bramão et al., 2017; DuBrow & Davachi, 2014; Kim & Lee, 2023). However, it is unclear whether this order is stored as part of the memory trace and whether it is reinstated during retrieval, a key mechanism of pattern completion in contextual binding models. To address this gap, this study investigated whether the sequence of sensory modalities during encoding is neuronally reinstated at recognition, even when perceptual cues about the order are absent. By using EEG and multivariate pattern analysis (MVPA), this study tested whether modality sequence functions as a retrievable contextual tag, supporting the broader thesis aim that temporal structure contributes to contextual binding in multisensory memory.

2.3.2. Methods

Thirty-two participants were tested in a subsequent memory paradigm, in which they encoded audiovisual stimulus pairs presented sequentially in one of two fixed modality orders: auditory–visual (AV) or visual–auditory (VA). This temporal sequence served as the critical experimental manipulation. During the recognition phase, all stimulus pairs were presented simultaneously to eliminate perceptual cues about their original order, and participants completed an old/new recognition task. EEG data were recorded continuously during both encoding and recognition. Time–frequency decomposition was applied to examine oscillatory dynamics across the frequency spectrum of 1 to 40 Hz. MVPA classifiers were trained to distinguish between AV and VA encoding trials based on EEG patterns and subsequently tested on EEG data during recognition to assess neural reinstatement of the original modality order.

2.3.3. Results

Behavioural analysis showed that recognition accuracy did not significantly differ between the AV and VA sequences, ruling out potential performance confounds related to sequence order. MVPA analysis of EEG data revealed that the encoding order (AV vs. VA) could be successfully decoded from neural activity during recognition, particularly within the theta (3–7 Hz) and low beta (13–21 Hz) bands. Decoding accuracy exceeded chance levels in a post-stimulus window from approximately 500 to 1000 ms, suggesting reinstatement of the original modality sequence during retrieval. Spatially, decoding effects were most pronounced over fronto-temporal and parietal electrodes, consistent with regions implicated in episodic retrieval and crossmodal integration. These findings demonstrate that the temporal structure of multisensory encoding episodes is preserved in the neural signal and can be reinstated during recognition, even when the sensory input lacks sequential information.

2.3.4. Conclusion

This study provides the first direct evidence that the order of sensory modalities during encoding is neurally reinstated during recognition, even in the absence of perceptual cues to that order. The findings support the claim that modality sequence acts as a contextual feature, embedded within the memory trace and accessible via neural reactivation mechanisms. From a theoretical perspective, the results align with models of temporal context coding (e.g., CMR) and support the idea that contextual reinstatement includes not only spatial or semantic features but also perceptual-temporal structure (Heald et al., 2023). The involvement of theta and beta oscillations suggests that these rhythms play a role in reconstructing temporal and structural aspects of memory, extending prior work on theta–gamma coupling in episodic binding. Critically, the reinstatement of modality order occurred in a task where sequence was not relevant for performance, highlighting the automaticity of contextual encoding and retrieval. This supports the broader thesis claim that episodic memory is richly contextual, even for features that are not explicitly task-relevant. By demonstrating that temporal structure in multisensory encoding leaves retrievable neural signatures, Study 3 adds a third temporal dimension to the contextual binding framework developed throughout this thesis. Together with the findings from enriched environments (external context) and neural entrainment (internal context), this study strengthens the view that episodic memory depends on multi-layered, dynamic context integration.

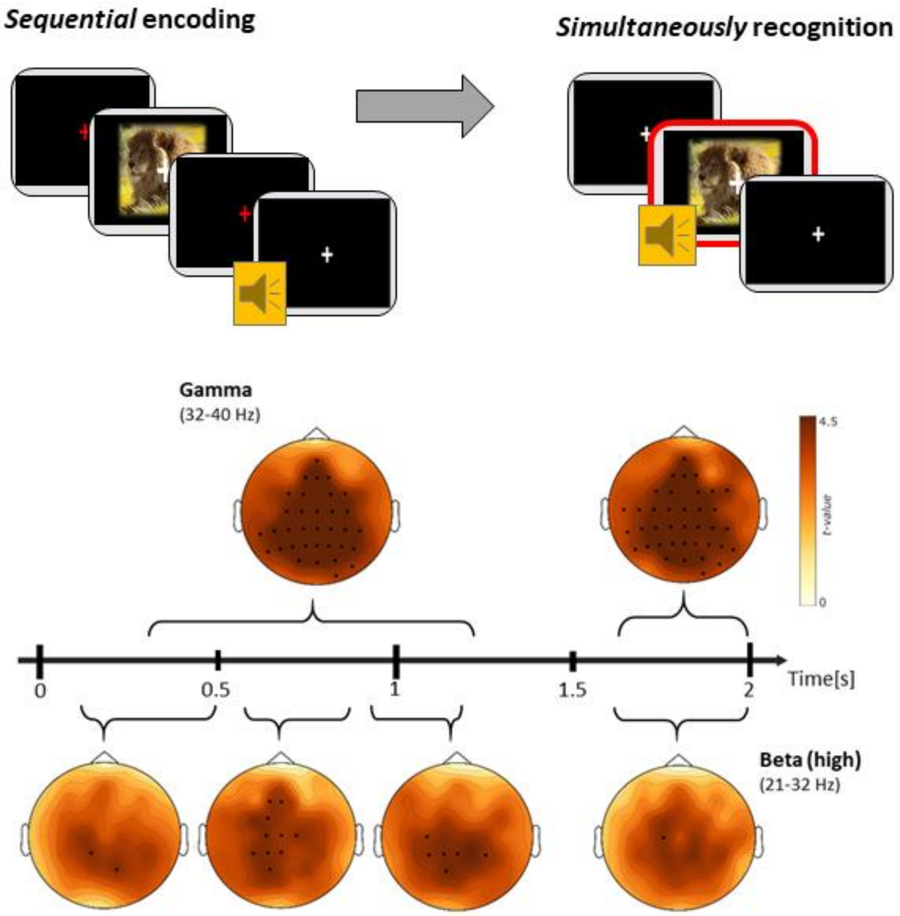


Figure 3. Decoding modality order from neural activity during recognition. During encoding, AV and VA pairs were presented *sequentially*, but during recognition, pairs were presented *simultaneously* (red frame). MVPA successfully decoded the original modality order (AV vs. VA) from neural activity, indicating reinstatement of temporal context. Bottom panels show time-resolved topographies of decoding accuracy (t-values) in gamma (32–40 Hz) and high beta (21–32 Hz) bands. Significant effects occurred between 400–2000 ms, mainly in centro-posterior and centro-parietal regions.

3. General Discussion

Episodic memory enables the reconstruction of past experiences, integrating not only the “what,” but also the “where” and “when” of events (Ngo et al., 2019; Tulving, 2002; Yonelinas et al., 2019). This reconstruction critically depends on the binding of contextual features during encoding and their reinstatement during retrieval (Gilmore et al., 2021; Heinbockel et al., 2024; Staudigl & Hanslmayr, 2013; Uncapher et al., 2006). However, memory research has often relied on unimodal and abstract stimuli, limiting our understanding of how contextual binding operates in complex, multisensory environments (S. S. Cohen & Parra, 2016; Hendriks et al., 2024; Pecher & Zeelenberg, 2022). The present studies addressed this limitation by examining how external, internal, and temporal contextual factors interact with memory formation and retrieval across multiple levels of analysis.

Study 1 investigated the external context by comparing memory encoding in immersive VR to traditional desktop environments. Despite similar overall learning outcomes, immersive VR led to more explicit and confident memory retrieval, suggesting that enriched spatial and sensory contexts act as external scaffolds that support strategic retrieval, even under incidental learning conditions (Kisker et al., 2021b; Krokos et al., 2019). This finding is supported by theories that propose contextual richness enhances encoding through elevated emotional engagement, attentional depth, and self-relevance, potentially via the stronger recruitment of the default mode network and memory-related regions (Bréchet et al., 2019; Cadet & Chainay, 2020).

Study 2 focused on internal context by manipulating pre-encoding brain states via rhythmic visual stimulation. Entrainment in the alpha band (9 Hz) enhanced associative memory performance, aligning with the view that alpha oscillations facilitate attentional filtering and preparatory inhibition (Foxe & Snyder, 2011; Klimesch et al., 2011; Poch et al., 2018; Waldhauser et al., 2012). The effect was frequency-specific, as theta-band entrainment, despite modulating neural power, did not result in behavioural improvements. These findings suggest that oscillatory power enhancement prior to stimulus onset can enhance the brain's preparation for contextual binding, consistent with predictive coding and selective attention frameworks (Addante et al., 2011; Arnal & Giraud, 2012; Schneider & Rose, 2016; Zareian et al., 2020).

Study 3 explored temporal structure by demonstrating that sequences across modalities, although not explicitly relevant to the task, were spontaneously reinstated during retrieval. Multivariate decoding of EEG patterns showed that participants reinstated modality order shortly after recognition cues, supporting the idea that temporal context is embedded within the memory trace and reactivated even without conscious effort. This finding supports temporal context models (M. W. Howard & Kahana, 2002; Lohnas et al., 2023; Polyn et al., 2009) and extends reinstatement theories by showing that abstract structural information can guide memory retrieval in the absence of explicit cues.

Together, these studies converge on a multidimensional framework of contextual binding in multisensory episodic memory. Rather than operating in isolation, external environments, internal neural states, and temporal regularities interact to shape the dynamics of encoding and retrieval. This framework extends traditional memory models by integrating ecological, neural, and temporal layers into a unified account. It also emphasises that episodic memory is not a passive record of events, but a constructive process shaped by attentional, oscillatory, and structural scaffolds.

3.1. Integrating Contextual Binding Across Layers

Episodic memory is fundamentally a context-dependent construct, relying on the integration of diverse sensory, temporal, and internal cues into a coherent, retrievable trace (Staudigl & Hanslmayr, 2013; Yonelinas et al., 2019). Contextual binding enables the segmentation, encoding, and retrieval of experiences in a way that supports flexible cognition and adaptive behaviour (Agafonov et al., 2023; Bramão et al., 2017; Yonelinas et al., 2019). Extending classical models such as the BIC theory (Diana et al., 2007; Ranganath, 2010) and the CMR model (Polyn et al., 2009), this framework proposes that contextual features operate along partially dissociable but interacting axes. These axes encompass external environments, internal neural states, and temporal structure, jointly defining the conditions under which episodic memories are formed and later reconstructed.

The role of environmental context in episodic memory has been established, with enriched perceptual environments functioning as more than passive backgrounds (Z. Liu et al., 2024; J. L. Park & Donaldson, 2019). Immersive VR settings, for instance, have been shown to enhance explicit memory expression compared to conventional 2D displays, despite equivalent exposure and cognitive demands (Kisker et al., 2021a; Schöne et al., 2023). The mnemonic advantage of VR is attributed to a suite of factors: increased presence, attentional engagement, motivational and affective relevance, and enhanced multisensory integration (Beitner et al., 2023; Hurter et al., 2024; Monaro et al., 2024; Schöne et al., 2023). These immersive qualities facilitate the integration of sensorimotor and spatial cues, resulting in holistic scene representations that enhance both object and relational memory (Bréchet et al., 2019; Kisker et al., 2021b; Ventura et al., 2019). Noteably, the context itself becomes encoded as an element of the episodic trace, facilitating richer associative representations, consistent with models such as CMR (Polyn et al., 2009). Environmental reinstatement at retrieval provides a direct benefit by reinvoking the internal context present at encoding, demonstrating that memory performance depends on the integration of contextual features into the representation of the episode.

Beyond environmental factors, internal neural states, particularly those indexed by pre-stimulus oscillations, emerge as determinants of encoding efficacy and memory organisation. Oscillatory dynamics in the theta (3–7 Hz) and alpha (8–12 Hz) frequency bands have repeatedly been linked to successful encoding and associative binding (Cruzat et al., 2021; Ostrowski & Rose, 2024). Increased pre-stimulus theta power correlates with enhanced contextual binding and source memory, presumably reflecting a preparatory neural state (Addante et al., 2011). Pre-stimulus alpha activity, in turn, has been implicated in attentional filtering and sensory gating, facilitating the establishment of conditions favourable to encoding complex, multisensory information (Foxe & Snyder, 2011; Jensen & Mazaheri, 2010; Strunk & Duarte, 2019). Modulation of these oscillatory states, such as via alpha entrainment, can causally enhance associative encoding, underscoring their functional contribution (Schneider & Rose, 2016; Michael et al., 2023; Hanslmayr et al., 2019). Notably, while alpha oscillations play a preparatory role before stimulus onset, theta rhythms appear most crucial during the active integration of incoming information (Buzsáki & Moser, 2013; Herweg et al., 2020). Memory enhancements achieved through theta band entrainment typically arise when stimulation coincides with stimulus presentation, likely reflecting an increasing capacity for network binding at those critical moments (Hanslmayr et al., 2019; Köster et al., 2018). Collectively, these observations establish brain state oscillations as dynamically modulatable features that shape the probability and richness of contextual binding.

While external and internal contexts influence how well memory traces are formed, temporal structure determines how events are organised and recalled. Temporal sequencing of sensory events represents a still underappreciated, but fundamental dimension of episodic memory. While classical models emphasise modality-independent order codes (Depoorter & Vandierendonck, 2009; Vandierendonck, 2016), converging evidence demonstrates that the specific order of sensory modalities (temporal context) can be encoded and later reinstated as an integral component of the memory trace. Recognition of audiovisual sequences distinguished by modality order (auditory–visual versus visual–auditory), even in the absence of perceptual cues at retrieval, highlights the embedding of modality sequence into episodic representation. This form of temporal binding aligns with theories of temporal context reinstatement, whereby dynamic states present at encoding are recapitulated at retrieval to facilitate memory access (DuBrow & Davachi, 2016; M. W. Howard & Kahana, 2002). Further, the neural reinstatement of modality order, particularly within the theta and low beta frequency ranges, implicates oscillatory network coordination in the retrieval of sequence information (Michelmann et al., 2016; Sols et al., 2017; Yaffe et al., 2014, 2017). Neuroimaging findings robustly demonstrate hippocampal and cortical replay of temporal sequences, directly supporting the subjective phenomenon of “jumping back in time” during episodic memory retrieval (Thavabalasingam et al., 2019; Vaz et al., 2020). Importantly, these replay mechanisms preserve modality-specific temporal structures, which function as integral contextual characteristics (Michelmann et al., 2016). Such modality-dependent temporal tags appear essential for discriminating and reconstructing complex, overlapping memories, challenging models that conceptualise temporal order as abstracted solely from sensory modality (Folkerts et al., 2018; Foudil et al., 2021). Instead, these findings underscore that, in the domain of episodic memory, the hippocampus and cortex collaborate to reinstate both the content and the temporal context of experiences. This highlights the importance of a modality-specific temporal scaffold for accurate memory reflection and pattern completion in dynamic environments.

These three contextual dimensions (external, internal, and temporal) do not act in isolation. Instead, they dynamically interact to define the quality and durability of multisensory episodic memories. A rich external setting may heighten internal preparation. Rhythmic entrainment may sharpen attention to temporal sequences, and structured temporal input may enhance the coherence of multisensory binding. The convergence of these factors reflects a systems-level model of episodic memory that emphasises context as an active and integrative component of the memory trace. Episodic memory is not merely the result of encoding isolated items but emerges from the structured interplay of contextual features that shape how experiences are organised and accessed.

3.2. Neural Mechanisms of Reinstatement and Implications for Episodic Memory

Current perspectives on episodic memory emphasise its reconstructive nature, wherein retrieval reflects the dynamic reinstatement of neural states that were present during encoding (Heinbockel et al., 2024; Horner et al., 2015; Staresina et al., 2012; Yaffe et al., 2014). This reinstatement process is central to the binding of contextual information, encompassing environmental, temporal, and neural states, into flexible and retrievable episodic representations. Importantly, converging evidence suggests that reinstatement acts as a core mechanism of episodic memory, operating across multiple phases and at varying levels of abstraction.

Traditionally, research has focused on the phenomenon of reinstatement during retrieval (Pacheco Estefan et al., 2019; Wing et al., 2015; Xiao et al., 2017; Yaffe et al., 2014). However, emerging neurocognitive models advance a more nuanced perspective, proposing that reinstatement processes may be initiated prior to the conscious act of recall and may extend dynamically across distinct memory phases (Gordon et al., 2014; Tompary et al., 2016; Yaffe et al., 2014). Crucially, temporal context information, such as the sequential order of sensory modalities, is decodable from neural activity early in the recognition process, sometimes preceding deliberate retrieval efforts. This temporal specificity suggests that contextual reinstatement is not solely a strategic, top-down process, but can also reflect anticipatory or automatic mechanisms, consistent with models of predictive coding and the gradual reactivation of memory traces (DuBrow & Davachi, 2016; Folkerts et al., 2018). As such, it becomes essential to distinguish between various forms of reinstatement, pre-stimulus (preparatory reactivation of context before a cue), early online reactivation (matching sensory input with stored context), and classic retrieval-phase reinstatement (deliberate, strategic recall). Notably, early decoding of modality order during recognition suggests that the reinstatement of context may occur during the initial sensory analysis, independent of explicit retrieval intentions. This challenges the view that reinstatement is limited to strategic recollection, highlighting partial automatic contextual reactivation. A critical theoretical consideration involves the content of reinstatement.

Whereas much of the literature has focused on item-specific reactivation (such as objects or words; Pacheco Estefan et al., 2019; Rau et al., 2025), recent evidence demonstrates the reinstatement of complex, structural contextual features (Clarke et al., 2022; Manning et al., 2011), including the order of modality presentation. These findings are essential because item and context reinstatement implicate overlapping but functionally dissociable brain networks and fulfil distinct roles in episodic memory. Item-based reinstatement predominantly engages posterior neocortical areas, such as the lateral occipital cortex, processing object identity and semantic content (Bencze et al., 2024; Pacheco Estefan et al., 2019). By contrast, the reinstatement of context, encompassing scene features, modality, or temporal order, relies more heavily on parahippocampal and hippocampal regions (Hayes et al., 2007; L. R. Howard et al., 2011; Staresina et al., 2012). The decoding of modality order from neural dynamics during recognition underscores the hippocampus's role in integrating relational and structural features within episodic traces (R. A. Cooper & Ritchey, 2020; Ranganath, 2010). The retrieval of such contextual features, even when not explicitly relevant to the task, highlights the automaticity of contextual binding and suggests that contextual information may be encoded and reinstated by default.

The reactivation of modality order closely aligns with models of temporal context drift and "jump-back-in-time" processes in episodic memory (Folkerts et al., 2018; M. W. Howard & Kahana, 2002; Lohnas & Healey, 2021). These frameworks propose that the brain continuously encodes a slowly evolving temporal context, with retrieval cues capable of reinstating both the content and the contextual state in which the information was embedded (M. W. Howard, 2017; Polyn & Kahana, 2008). Evidence for the reinstatement of structural temporal features, such as modality sequence, extends these theories, suggesting a broader neural capacity for reactivating abstract contextual information. Notably, such reinstatement of non-task-relevant structural features underscores the automaticity and potential need of contextual binding. This observation challenges classic dual-process models, which restrict contextual reinstatement to strategic, recollection-dependent processes, in favour of a graded model in which both familiarity and recollection can involve variable degrees of contextual retrieval (Addante et al., 2024; Kuhl & Chun, 2014).

Reinstatement dynamics further differ in temporal profile depending on the nature of the contextual feature. The earlier reinstatement of modality order in Study 3 contrasts with item-based reactivation, which typically emerges later in the trial and may rely more heavily on conscious retrieval strategies (Rau et al., 2025; Xiao et al., 2017). This timing distinction highlights the need to characterise when, during the trial, different features are reactivated. Pre-stimulus reinstatement, online encoding reactivation, and post-cue retrieval dynamics may each serve complementary roles in memory reconstruction, with early context reinstatement providing a scaffold for later item-specific search or decision processes (Bramão et al., 2017; Manning et al., 2011; Xue, 2022). The temporal structure of reinstatement observed in Study 3 was also supported by specific frequency bands, particularly theta and low beta oscillations. These frequencies have been implicated in temporal sequence replay, cross-modal integration, and context reinstatement (Sols et al., 2017; Vaz et al., 2020; Yaffe et al., 2014). The presence of such oscillatory reinstatement supports the idea that memory is not merely a static reactivation of stored traces but a temporally patterned reconstruction, dynamically aligned with the unfolding of contextual cues.

Collectively, these findings refine the understanding of episodic memory as a temporally distributed, context-sensitive, and dynamically constructive process. Reinstatement emerges as a multifaceted mechanism, extending from preparatory neural states to explicit retrieval, capable to flexibly reactivate both, items, and their contextual structure highlighting the importance of cross-level (behavioural, oscillatory, neural) investigations to unravel the systems-level architecture supporting episodic memory.

3.3. The Role of Oscillatory Entrainment in Memory Encoding

Oscillatory brain dynamics play a fundamental role in shaping how information is encoded into episodic memory (Hanslmayr et al., 2009, 2016; Minarik et al., 2018; Nyhus & Curran, 2010; Staudigl & Hanslmayr, 2013). Study 2 examined the causal impact of pre-stimulus oscillatory states on memory formation by using rhythmic visual stimulation to entrain neural activity at alpha and theta frequencies. While prior work has relied mainly on correlational evidence linking oscillatory power to subsequent memory performance, this approach allowed for the direct manipulation of endogenous brain rhythms prior to encoding, offering critical insights into the temporal dynamics and functional role of internal context. Rhythmic sensory stimulation, also known as neural entrainment, has emerged as a powerful method for modulating brain oscillations and assessing their functional contributions to cognition (Haegens & Zion Golumbic, 2018; Pomper et al., 2023). By aligning neural excitability phases to predictable external rhythms, entrainment can enhance perception, attentional filtering, and memory (Hanslmayr et al., 2019; Michael et al., 2023; Roberts et al., 2018). Stimulation at alpha (8–12 Hz) and theta (4-7 Hz) frequencies has been shown to modulate large-scale network activity, particularly in parietal and sensory cortices, facilitating cognitive preparation for incoming stimuli (Alagapan et al., 2019; Ma et al., 2025).

The results revealed that only alpha-band entrainment significantly enhanced associative memory. Participants exposed to alpha rhythmic stimulation showed increased recognition sensitivity (d′), primarily driven by elevated hit rates. This effect was not accompanied by changes in false alarm rates, indicating an enhancement in memory strength rather than altered response bias (Criss, 2010; Higham et al., 2009). These findings align with the alpha inhibition hypothesis, which posits that alpha oscillations reflect active functional inhibition, supporting attentional gating and the suppression of irrelevant input (Hakim et al., 2020; Poch et al., 2018). Pre-stimulus alpha power has been repeatedly linked to successful memory performance, particularly in contexts requiring crossmodal integration or attentional anticipation (Burke et al., 2015; Fell et al., 2011; Ostrowski & Rose, 2024). By synchronizing cortical excitability phases, entrainment may facilitate optimal engagement of encoding networks, including parietal and hippocampal structures (Dave et al., 2022; Raud et al., 2023). These findings suggest that alpha entrainment enhances preparatory states, thereby increasing preparation for associative binding at the moment of stimulus arrival. In contrast, theta-band entrainment did not yield behavioural benefits despite successful modulation of oscillatory power. This null effect highlights the frequency specificity and temporal sensitivity of oscillatory contributions to memory. While theta oscillations are critical for associative encoding and flexible memory updating (Herweg et al., 2020; Kota et al., 2020), their functional role may be more pronounced during stimulus presentation rather than in the pre-stimulus interval. Theta effects often depend on phase synchrony across hippocampal-cortical circuits and tight temporal alignment with stimulus onset conditions, which cannot be fully met in the pre-stimulus entrainment design (Hanslmayr et al., 2019; Köster et al., 2018).

Neural entrainment in this context extends beyond modulating spectral power as it also impacts the temporal architecture of neural excitability, influencing the brain's ability to prepare for and encode incoming stimuli. These results align with predictive coding frameworks, which propose that preparatory oscillatory states support sensory selection and attentional filtering (Arnal & Giraud, 2012; Calderone et al., 2014; Jensen, 2002; Lakatos et al., 2019). Whereas alpha and beta oscillations facilitate sustained attentional focus and top-down inhibition (Clayton et al., 2015; C. Wang et al., 2016), theta oscillations support sequential binding and temporal integration (Griffiths et al., 2021; Heusser et al., 2016; Nyhus & Curran, 2010)). The dissociation observed in Study 2 reinforces that distinct frequencies contribute differently to encoding. Alpha may facilitate preparatory attentional processes. At the same time, theta supports the dynamic integration of information during stimulus processing. This suggests that oscillatory influences on memory are not uniform but instead operate through frequency-specific and temporally defined mechanisms.

Beyond their theoretical significance, these findings have applied relevance. Sensory entrainment offers a non-invasive method for enhancing cognitive preparation and memory performance. The ability of alpha entrainment to boost associative encoding may be leveraged in educational or clinical settings, particularly in populations with deficits in attentional control or memory function. Future research should explore synergistic effects of entrainment applied both before and during stimulus presentation, as well as investigate individual differences in entrainment responsiveness. In sum, these findings establish a causal link between internal neural states and memory formation, demonstrating that pre-stimulus alpha entrainment enhances associative encoding by shaping the brain's preparatory landscape. The results contribute to a systems-level understanding of contextual binding, where internal oscillatory context interacts with external and temporal features to determine the quality of episodic memory representations.

3.4. External Context and Motivation: Lessons from Enriched Environments

Perceptual rich environments do more than supply sensory detail. They shape the cognitive and emotional scaffolding of memory encoding (Makowski et al., 2017; Ritchey et al., 2013; Tambini et al., 2017). Encoding in immersive VR environments leads not to an increase in the pure amount of learned information. However, there is a qualitative shift in memory expression as participants exhibit more explicit knowledge, greater confidence, and increased strategic accessibility of memory content, even under incidental learning conditions. This divergence in memory quality underscores the transformative role of external context in shaping memory format, facilitating a transition from implicit familiarity to explicit recollection. Several mechanisms may underlie this effect. Immersive VR reliably induces stronger emotional responses than traditional 2D desktop settings, and emotional arousal is known to enhance the vividness and accessibility of episodic memory (Beitner et al., 2023; Kisker et al., 2021b). This observation aligns with the Arousal-Biased Competition (ABC) theory, which proposes that emotional arousal amplifies the processing of salient stimuli while suppressing the processing of irrelevant input (Mather & Sutherland, 2011; Sakaki et al., 2014; Warren et al., 2016). Elevated arousal in VR likely increases the salience of contextual cues, enhancing their encoding and later reactivation. Moreover, the subjective sense of presence and bodily self-location within the virtual scene deepens the self-relevance of encoded events, a factor known to facilitate autobiographical memory (Kilteni et al., 2012; S. A. Smith & Mulligan, 2021).

Beyond emotion, immersive environments also boost intrinsic motivation and attentional engagement, which are catalysts of memory consolidation (Lüddecke & Felnhofer, 2022). VR has been shown to enhance attentional depth, cognitive involvement, and satisfaction during learning tasks (Allcoat & Mühlenen, 2018; Portuguez-Castro & Santos Garduño, 2024). These features likely promote deeper semantic and associative processing, consistent with the Attentional Boost Effect (ABE), whereby task-relevant events transiently enhance encoding of concurrent stimuli (Mulligan et al., 2023; Swallow et al., 2009, 2022). Remarkably, even in incidental learning paradigms, immersion in a contextually rich environment leads to greater explicit retrieval. This suggests that enriched environments lower the threshold for strategic retrieval by supporting automatic contextual binding and deep-level encoding. Recent work has shown that distinct immersive contexts can spontaneously trigger explicit memory without deliberate effort, thereby challenging dual-process models that strictly separate automatic and strategic retrieval (Helbing et al., 2020; Shin et al., 2021). These behavioural findings are complemented by neuroimaging evidence, as reinstating the encoding context reactivates the associated memory traces, as shown by increased hippocampal and parietal activation during retrieval (Bramão et al., 2017; Essoe et al., 2022; Tanaka et al., 2014; Wing et al., 2015). Immersive VR, in particular, can strongly engages the default mode network (DMN), including the hippocampus, posterior parietal cortex, and ventromedial prefrontal cortex (Seinfeld et al., 2021; Wong et al., 2014). These regions are implicated in scene construction, self-referential processing, and schema-based encoding (Hassabis & Maguire, 2007; Smallwood et al., 2021; Wen et al., 2020). The presence of bodily cues further strengthens these activations by linking memory encoding to self-location and self-motion, anchoring experiences in first-person perspective (Bayramova et al., 2021; Iriye et al., 2024; Penaud et al., 2023). The engagement of spatial, sensorimotor, and emotional systems supports the creation of multimodal scene representations, consistent with enriched environment theories in both animal and human learning (Bréchet et al., 2019; Peney & Skarratt, 2024; Ventura et al., 2019).

Taken together, immersive VR environments serve not merely as sensory background but as powerful external scaffolds for contextual binding. They offer dense arrays of perceptual, emotional, and motivational cues that embed experiences into richly detailed and retrievable memory traces. These findings support the theoretical view that the external context operates as an active component in episodic memory construction, shaping what is remembered, how vividly, and under what conditions it can be retrieved. From an applied perspective, these results carry implications for education, rehabilitation, and clinical interventions. In educational settings, immersive environments may foster durable learning by enhancing attentional engagement and emotional salience. In clinical populations, such as individuals with amnesia or mild cognitive impairment, enriched contexts may support memory reactivation via contextual reinstatement, offering new paths for therapeutic design. In conclusion, contextual richness, like internal oscillatory states and temporal structure, plays a central role in organising and retrieving episodic experiences.

3.5. Methodological Contributions

The present work makes significant methodological contributions to the study of episodic memory by integrating innovative approaches that enhance ecological validity, experimental control, and mechanistic understanding. A key strength is the use of immersive VR environments in Study 1, which facilitates the creation of rich, naturalistic contexts that better approximate real-world memory experiences. This approach advances beyond traditional laboratory paradigms by enabling the investigation of memory within dynamic, multisensory, and naturalistic settings, allowing for more applicable insights into everyday memory processes. The use of VR also highlights the potential for multisensory, immersive learning environments, such as VR-based education, to enhance memory encoding and retrieval, with significant implications for real-world learning and training. Furthermore, the application of neural entrainment design prior to stimulus presentation represents a novel method for causally testing the role of oscillatory neural rhythms in memory formation. By selectively modulating brain states before encoding, this approach sheds light on the mechanistic influence of neural oscillations on attention and memory, addressing questions about the causal dynamics underlying successful binding and subsequent retrieval. The consistent paradigm utilised across two out of three studies, combining behavioural tasks, EEG recordings, and pattern classification, enables direct comparison of effects across different levels of context (neural and temporal), thereby strengthening the interpretability and integration of findings. The use of MVPA to decode modality-specific reinstatement during retrieval is an insightful methodological approach. This technique enables the precise tracking of neural signatures associated with the reactivation of sensory and contextual features, providing compelling evidence for the neural mechanisms underlying the reconstruction of episodic memory. Such analyses enhance the field's capacity to link neural oscillations with specific memory processes, thereby bridging the gap between neural dynamics and experiential retrieval. As research advances, it will be essential to enhance ecological validity by integrating mobile EEG with augmented or mixed reality to capture memory processes in semi-naturalistic real-world settings. These include areas such as education, clinical interventions, and technology-enhanced learning.

3.6. Limitations

While the findings of this research provide valuable insights into the mechanisms of contextual binding in multisensory episodic memory, potential limitations should be acknowledged. The use of between-subject designs in all three studies, although necessary for controlling environmental variables and stimulation conditions, inherently reduces sensitivity to individual differences and diminishes statistical power compared to within-subject approaches (Montoya, 2023; Mullet & Chasseigne, 2018). This may limit the ability to detect subtle, but meaningful, variations in responsiveness to contextual cues or neural entrainment processes. Future studies could incorporate within-subject designs, especially in neural modulation paradigms, to better capture individual variability and identify responder profiles. Although EEG offered excellent temporal resolution to assess oscillatory dynamics and reinstatement, it has limited spatial localisation capabilities, particularly concerning deep brain structures such as the hippocampus and parahippocampal cortex, which are known to play critical roles in contextual binding and episodic memory retrieval (Copara et al., 2014; Hayes et al., 2007; Pacheco Estefan et al., 2019; Staudigl & Hanslmayr, 2013; Yonelinas et al., 2019; Yoo et al., 2021). Without direct measures from these regions, interpretations regarding their involvement remain indirect. Integrating multimodal imaging techniques, such as fMRI or MEG, could provide a more comprehensive understanding of the neural circuits engaged in these processes. Additionally, the fixed stimulation frequencies of 5 Hz for theta and 9 Hz for alpha may not optimally align with each participant’s intrinsic oscillatory frequencies, which are known to vary substantially across individuals (Duecker et al., 2024; Gulbinaite et al., 2017; Zaehle et al., 2010). This variability could have impacted the efficacy of entrainment, particularly for theta stimulation, potentially resulting in null effects in some cases. Future research should consider employing frequency-tuning protocols based on each individual's resting-state EEG to maximise resonance and entrainment success. Moreover, all three studies relied on recognition tasks to assess memory, which, although methodologically advantageous in terms of control and reduced demands, may underestimate the reconstructive and effortful aspects of episodic retrieval (Cabeza et al., 1997; Chen et al., 2017; Malmberg, 2008). Recognition tasks tend to capture familiarity and recognition processes but may overlook the richness, vividness, and phenomenological qualities characteristic of accurate episodic recall (Cabeza et al., 2003; Ozubko et al., 2021). Employing free or cued recall paradigms could provide a deeper understanding of how context influences the vividness and detail of memories, offering a more ecologically valid perspective on episodic retrieval (Kahana, 1996; Lohnas & Kahana, 2014; Wilson & Criss, 2017).

Overall, while the methodological strategies employed facilitated precise investigation and yielded important results, these limitations highlight ways for refinement. Future research integrating more personalized neurostimulation protocols, multimodal neuroimaging, diverse memory assessments, and within-subject designs can enhance sensitivity, neural specificity, and ecological validity, thereby advancing our understanding of the neural and cognitive mechanisms underlying contextual memory in real-world settings. Addressing these limitations will not only increase experimental sensitivity but also bridge the gap between mechanistic neuroscience and the complexity of real-world memory, enabling more robust models of contextual binding in dynamic environments.

3.7. Future Directions and Applications

Building on our findings, future investigations should explore how the combination of immersive VR environments and neural entrainment protocols can elucidate their additive or interactive effects on memory processes. While the current studies examined these factors separately, employing a factorial design would allow researchers to assess whether immersive, multisensory contexts amplify the effects of preparatory neural states or whether bottom-up salience in enriched environments can override or modulate the impact of oscillatory modulations. Such work could clarify whether these mechanisms operate synergistically or competitively during encoding, shedding light on how external and internal levels of context interact to influence the formation of episodic memory. A key research direction is to extend this line of inquiry into the long-term consolidation of memory. Considering the critical role of oscillations during sleep, combining VR-based encoding with sleep studies or multi-day designs could reveal how enriched environments and neural entrainment influence offline reactivation and the stabilisation of episodic memories. Techniques such as targeted memory reactivation (TMR) can be employed to cue enriched or entrained episodes during sleep, testing whether preserved or transformed contextual features support durable long-term storage (Abdellahi et al., 2023; Hu et al., 2020; Lewis & Bendor, 2019).

Another promising direction involves individualising oscillatory stimulation protocols by tailoring frequencies to each participant's intrinsic alpha or theta peaks, derived from resting-state EEG. Such personalised approaches could enhance entrainment efficacy, especially for populations with altered neural rhythms, such as older adults or individuals with neurocognitive impairments, potentially leading to more robust and consistent cognitive enhancements. Extending these methods to more ecologically valid, semantically meaningful stimuli, such as speech or music, could further illuminate how high-level integration and semantic schemas bolster contextual binding. When stimuli reflect real-world event structures, effects related to modality sequence and sequence reinstatement might be more pronounced, providing insights applicable to everyday memory.

The transferability potential of these approaches in educational and clinical settings is substantial. Combining immersive VR with neuroadaptive entrainment could improve learning and memory for complex or abstract material, offering tailored interventions for populations with attention deficits or memory impairments. In clinical contexts, such strategies could support neural plasticity and compensatory encoding mechanisms in conditions like mild cognitive impairment, attention-deficit/hyperactivity disorder, or age-related decline (Boller et al., 2021; Corrigan et al., 2023; Wais et al., 2021). Longitudinal studies examining how modulating context through environmental manipulation, stimulation, or task design influences memory trajectories would be valuable. Additionally, integrating portable VR and EEG devices allow for testing memory processes in naturalistic environments such as classrooms, training simulations, or everyday situations (Mathewson et al., 2024; J. L. Park & Donaldson, 2019; J. Xu & Zhong, 2018). This approach promises increased naturalistic and real-time monitoring of neural–environment interactions, with the potential for closed-loop systems that adapt learning conditions dynamically based on neural markers of engagement, attention, or fatigue.

3.8. Conclusion

Episodic memory extends beyond merely encoding sensory information. It is influenced by the contextual structures that support experience. This body of work demonstrates that three interconnected layers of context dynamically influence episodic memory formation and retrieval: external (environmental richness), internal (oscillatory brain states), and temporal (structure of stimulus sequences). Across these dimensions, contextual binding not only determines what is remembered but also how vividly memories are retrieved.

Rich, immersive environments enhance episodic memory not just by increasing the amount of material learned but by transforming representations to be more explicit and strategically accessible. External supports, such as emotional arousal, attentional engagement, and embodied spatial cues, help embed experiences within more accessible mnemonic structures, thereby broadening ecological accounts of memory and contextual binding theories, like the CMR model, to naturalistic, multisensory settings. Neural oscillations also influence encoding success. For example, the entrainment of pre-stimulus alpha oscillations encourages preparatory neural states that facilitate attentional gating and input selection, resulting in stronger associative binding, which aligns with the role of alpha rhythms in managing cortical excitability and information prioritisation. Conversely, theta entrainment mainly affects oscillatory power without improving memory performance, emphasising the importance of accurate phase alignment between oscillatory activity and cognitive processes during key encoding periods. Moreover, temporal structure is crucial in episodic memory, shown by the spontaneous reinstatement of modality-specific sequence patterns during retrieval, regardless of task relevance, highlighting the automatic integration of temporal regularities into the memory trace. Decoding these patterns from EEG signals reveals that the brain reconstructs not only content but also its temporal context, supporting the idea that episodic memory is inherently sequential and dynamically reassembled based on features encoded during learning.

Collectively, these findings propose a multidimensional framework of contextual binding, where episodic memories are constructed and reactivated through the coordinated interplay of environmental richness, neural state preparation, and temporal organisation. This perspective emphasises that memory retrieval is inherently constructive, shaped by internal neural states, such as oscillatory dynamics and preparatory activity, and the structural scaffolds established during encoding, including sensory, spatial, and temporal cues. Rather than peripheral, contextual information forms the architecture that admit structure, salience, and the ability to retrieve sensory content. By combining ecologically valid virtual environments, causal neural modulation, and temporally precise decoding, the present results provide a framework for studying memory in real-world settings. Future research can build on this foundation to develop context-sensitive interventions for education, cognitive training, and memory rehabilitation, ultimately advancing strategies to support adaptive remembering across the lifespan.

References

Abdellahi, M. E. A., Koopman, A. C. M., Treder, M. S., & Lewis, P. A. (2023). Targeting targeted memory reactivation: Characteristics of cued reactivation in sleep. *NeuroImage*, *266*, 119820. https://doi.org/10.1016/j.neuroimage.2022.119820

Addante, R. J., Clise, E., Waechter, R., Bengson, J., Drane, D. L., & Perez-Caban, J. (2024). Context familiarity is a third kind of episodic memory distinct from item familiarity and recollection. *iScience*, *27*(12), 111439. https://doi.org/10.1016/j.isci.2024.111439

Addante, R. J., Watrous, A. J., Yonelinas, A. P., Ekstrom, A. D., & Ranganath, C. (2011). Prestimulus theta activity predicts correct source memory retrieval. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(26), 10702–10707. https://doi.org/10.1073/pnas.1014528108

Agafonov, A. Y., Zolotukhina, A. A., Kryukova, A. P., & Burmistrov, S. N. (2023). Contextual Mediation of Cognitive Activity. *Experimental Psychology (Russia)*, *16*(3), 98–120. https://doi.org/10.17759/exppsy.2023160307

Akyürek, E. G., Kappelmann, N., Volkert, M., & van Rijn, H. (2017). What You See Is What You Remember: Visual Chunking by Temporal Integration Enhances Working Memory. *Journal of Cognitive Neuroscience*, *29*(12), 2025–2036. https://doi.org/10.1162/jocn_a_01175

Alagapan, S., Riddle, J., Huang, W. A., Hadar, E., Shin, H. W., & Fröhlich, F. (2019). Network-Targeted, Multi-site Direct Cortical Stimulation Enhances Working Memory by Modulating Phase Lag of Low-Frequency Oscillations. *Cell Reports*, *29*(9), 2590-2598.e4. https://doi.org/10.1016/j.celrep.2019.10.072

Allcoat, D., & Mühlenen, A. von. (2018). Learning in virtual reality: Effects on performance, emotion and engagement. *Research in Learning Technology*, *26*. https://doi.org/10.25304/rlt.v26.2140

Alwashmi, K., Meyer, G., Rowe, F., & Ward, R. (2024). Enhancing learning outcomes through multisensory integration: A fMRI study of audio-visual training in virtual reality. *NeuroImage*, *285*, 120483. https://doi.org/10.1016/j.neuroimage.2023.120483

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398. https://doi.org/10.1016/j.tics.2012.05.003

Ashton, K., Zinszer, B. D., Cichy, R. M., Nelson, C. A., Aslin, R. N., & Bayet, L. (2022). Time-resolved multivariate pattern analysis of infant EEG data: A practical tutorial. *Developmental Cognitive Neuroscience*, *54*, 101094. https://doi.org/10.1016/j.dcn.2022.101094

Ayala, M. N., & Henriques, D. Y. P. (2021). Differential contributions of implicit and explicit learning mechanisms to various contextual cues in dual adaptation. *PLOS ONE*, *16*(7), e0253948. https://doi.org/10.1371/journal.pone.0253948

Bainbridge, W. A., Hall, E. H., & Baker, C. I. (2021). Distinct Representational Structure and Localization for Visual Encoding and Recall during Visual Imagery. *Cerebral Cortex*, *31*(4), 1898–1913. https://doi.org/10.1093/cercor/bhaa329

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, *95*(3), 709-721.e5. https://doi.org/10.1016/j.neuron.2017.06.041

Barnett, A. J., Nguyen, M., Spargo, J., Yadav, R., Cohn-Sheehy, B. I., & Ranganath, C. (2024). Hippocampal-cortical interactions during event boundaries support retention of complex narrative events. *Neuron*, *112*(2), 319-330.e7. https://doi.org/10.1016/j.neuron.2023.10.010

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

Barutchu, A., Sahu, A., Humphreys, G. W., & Spence, C. (2019). Multisensory processing in event-based prospective memory. *Acta Psychologica*, *192*, 23–30. https://doi.org/10.1016/j.actpsy.2018.10.015

Bastos, A. M., Lundqvist, M., Waite, A. S., Kopell, N., & Miller, E. K. (2020). Layer and rhythm specificity for predictive routing. *Proceedings of the National Academy of Sciences*, *117*(49), 31459–31469. https://doi.org/10.1073/pnas.2014868117

Batterink, L. J., Paller, K. A., & Reber, P. J. (2019). Understanding the Neural Bases of Implicit and Statistical Learning. *Topics in Cognitive Science*, *11*(3), 482–503. https://doi.org/10.1111/tops.12420

Bauer, M., Stenner, M.-P., Friston, K. J., & Dolan, R. J. (2014). Attentional Modulation of Alpha/Beta and Gamma Oscillations Reflect Functionally Distinct Processes. *Journal of Neuroscience*, *34*(48), 16117–16125. https://doi.org/10.1523/JNEUROSCI.3474-13.2014

Bayramova, R., Valori, I., McKenna-Plumley, P. E., Callegher, C. Z., & Farroni, T. (2021). The role of vision and proprioception in self-motion encoding: An immersive virtual reality study. *Attention, Perception, & Psychophysics*, *83*(7), 2865–2878. https://doi.org/10.3758/s13414-021-02344-8

Beauchamp, M. S., Nath, A. R., & Pasalar, S. (2010). fMRI-Guided Transcranial Magnetic Stimulation Reveals That the Superior Temporal Sulcus Is a Cortical Locus of the McGurk Effect. *Journal of Neuroscience*, *30*(7), 2414–2417. https://doi.org/10.1523/JNEUROSCI.4865-09.2010

Beitner, J., Helbing, J., David, E. J., & Vo, M. L.-H. (2023). Investigating the effects of a virtual reality vs. Screen-based testing setup on incidental memory after visual search through scenes. *Journal of Vision*, *23*(9), 5126. https://doi.org/10.1167/jov.23.9.5126

Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P. L., Gioanni, Y., Battaglia, F. P., & Wiener, S. I. (2010). Coherent Theta Oscillations and Reorganization of Spike Timing in the Hippocampal- Prefrontal Network upon Learning. *Neuron*, *66*(6), 921–936. https://doi.org/10.1016/j.neuron.2010.05.013

Benchenane, K., Tiesinga, P. H., & Battaglia, F. P. (2011). Oscillations in the prefrontal cortex: A gateway to memory and attention. *Current Opinion in Neurobiology*, *21*(3), 475–485. https://doi.org/10.1016/j.conb.2011.01.004

Bencze, D., Marián, M., Szőllősi, Á., Pajkossy, P., Nemecz, Z., Keresztes, A., Hermann, P., Vidnyánszky, Z., & Racsmány, M. (2024). Contribution of the lateral occipital and parahippocampal cortices to pattern separation of objects and contexts. *Cerebral Cortex*, *34*(7), bhae295. https://doi.org/10.1093/cercor/bhae295

Bhowmik, A. K. (2024). Virtual and augmented reality: Human sensory-perceptual requirements and trends for immersive spatial computing experiences. *Journal of the Society for Information Display*, *32*(8), 605–646. https://doi.org/10.1002/jsid.2001

Biocca, F., Kim, J., & Choi, Y. (2001). Visual Touch in Virtual Environments: An Exploratory Study of Presence, Multimodal Interfaces, and Cross-Modal Sensory Illusions. *Presence: Teleoperators and Virtual Environments*, *10*(3), 247–265. https://doi.org/10.1162/105474601300343595

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience*, *12*(12), 752–762. https://doi.org/10.1038/nrn3122

Boller, B., Ouellet, É., & Belleville, S. (2021). Using Virtual Reality to Assess and Promote Transfer of Memory Training in Older Adults With Memory Complaints: A Randomized Controlled Trial. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.627242

Bond, K. M., & Taylor, J. A. (2015). Flexible explicit but rigid implicit learning in a visuomotor adaptation task. *Journal of Neurophysiology*, *113*(10), 3836–3849. https://doi.org/10.1152/jn.00009.2015

Bone, M. B., Ahmad, F., & Buchsbaum, B. R. (2020). Feature-specific neural reactivation during episodic memory. *Nature Communications*, *11*(1), 1945. https://doi.org/10.1038/s41467-020-15763-2

Brainerd, C. J., Wright, R., Reyna, V. F., & Payne, D. G. (2002). Dual-Retrieval Processes in Free and Associative Recall. *Journal of Memory and Language*, *46*(1), 120–152. https://doi.org/10.1006/jmla.2001.2796

Bramão, I., Jiang, J., Wagner, A. D., & Johansson, M. (2022). Encoding contexts are incidentally reinstated during competitive retrieval and track the temporal dynamics of memory interference. *Cerebral Cortex*, *32*(22), 5020–5035. https://doi.org/10.1093/cercor/bhab529

Bramão, I., Karlsson, A., & Johansson, M. (2017). Mental reinstatement of encoding context improves episodic remembering. *Cortex*, *94*, 15–26. https://doi.org/10.1016/j.cortex.2017.06.007

Bréchet, L., Mange, R., Herbelin, B., Theillaud, Q., Gauthier, B., Serino, A., & Blanke, O. (2019). First-person view of one's body in immersive virtual reality: Influence on episodic memory. *PLOS ONE*, *14*(3), e0197763. https://doi.org/10.1371/journal.pone.0197763

Brincat, S. L., & Miller, E. K. (2015). Frequency-specific hippocampal-prefrontal interactions during associative learning. *Nature Neuroscience*, *18*(4), 576–581. https://doi.org/10.1038/nn.3954

Bruns, P., & Röder, B. (2023). Development and experience-dependence of multisensory spatial processing. *Trends in Cognitive Sciences*, *27*(10), 961–973. https://doi.org/10.1016/j.tics.2023.04.012

Burke, J. F., Merkow, M. B., Jacobs, J., Kahana, M. J., & Zaghloul, K. A. (2015). Brain computer interface to enhance episodic memory in human participants. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.01055

Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D., & Miller, E. K. (2012). Synchronous Oscillatory Neural Ensembles for Rules in the Prefrontal Cortex. *Neuron*, *76*(4), 838–846. https://doi.org/10.1016/j.neuron.2012.09.029

Butler, A. J., & James, K. H. (2011). Cross-modal versus within-modal recall: Differences in behavioral and brain responses. *Behavioural Brain Research*, S0166432811004803. https://doi.org/10.1016/j.bbr.2011.06.017

Buzsáki, G., & Moser, E. I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature Neuroscience*, *16*(2), 130–138. https://doi.org/10.1038/nn.3304

Cabeza, R., Kapur, S., Craik, F. I. M., McIntosh, A. R., Houle, S., & Tulving, E. (1997). Functional Neuroanatomy of Recall and Recognition: A PET Study of Episodic Memory. *Journal of Cognitive Neuroscience*, *9*(2), 254–265. https://doi.org/10.1162/jocn.1997.9.2.254

Cabeza, R., Locantore, J. K., & Anderson, N. D. (2003). Lateralization of Prefrontal Activity during Episodic Memory Retrieval: Evidence for the Production-Monitoring Hypothesis. *Journal of Cognitive Neuroscience*, *15*(2), 249–259. https://doi.org/10.1162/089892903321208187

Cadet, L. B., & Chainay, H. (2020). Memory of virtual experiences: Role of immersion, emotion and sense of presence. *International Journal of Human-Computer Studies*, *144*. https://doi.org/10.1016/j.ijhcs.2020.102506

Calderone, D. J., Lakatos, P., Butler, P. D., & Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends in Cognitive Sciences*, *18*(6), 300–309. https://doi.org/10.1016/j.tics.2014.02.005

Calvert, G. A., Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of Audio-Visual Integration Sites in Humans by Application of Electrophysiological Criteria to the BOLD Effect. *NeuroImage*, *14*(2), 427–438. https://doi.org/10.1006/nimg.2001.0812

Cao, Y., Summerfield, C., Park, H., Giordano, B. L., & Kayser, C. (2019). Causal Inference in the Multisensory Brain. *Neuron*, *102*(5), 1076-1087.e8. https://doi.org/10.1016/j.neuron.2019.03.043

Capotosto, P., Baldassarre, A., Sestieri, C., Spadone, S., Romani, G. L., & Corbetta, M. (2017). Task and Regions Specific Top-Down Modulation of Alpha Rhythms in Parietal Cortex. *Cerebral Cortex*, *27*(10), 4815–4822. https://doi.org/10.1093/cercor/bhw278

Chanaz, L. S. A. di, Pérez-Bellido, A., Wu, X., Lozano-Soldevilla, D., Pacheco-Estefan, D., Lehongre, K., Conde-Blanco, E., Roldan, P., Adam, C., Lambrecq, V., Frazzini, V., Donaire, A., Carreño, M., Navarro, V., Valero-Cabré, A., & Fuentemilla, L. (2023). Gamma amplitude is coupled to opposed hippocampal theta-phase states during the encoding and retrieval of episodic memories in humans. *Current Biology*, *33*(9), 1836-1843.e6. https://doi.org/10.1016/j.cub.2023.03.073

Chang, M., Hong, B., Savel, K., Du, J., Meade, M. E., Martin, C. B., & Barense, M. D. (2024). Spatial context scaffolds long-term episodic richness of weaker real-world autobiographical memories in both older and younger adults. *Memory*, *32*(4), 431–448. https://doi.org/10.1080/09658211.2024.2334008

Chen, H.-Y., Gilmore, A. W., Nelson, S. M., & McDermott, K. B. (2017). Are There Multiple Kinds of Episodic Memory? An fMRI Investigation Comparing Autobiographical and Recognition Memory Tasks. *Journal of Neuroscience*, *37*(10), 2764–2775. https://doi.org/10.1523/JNEUROSCI.1534-16.2017

Chun, M. M., & Jiang, Y. (1999). Top-Down Attentional Guidance Based on Implicit Learning of Visual Covariation. *Psychological Science*, *10*(4), 360–365. https://doi.org/10.1111/1467-9280.00168

Clarke, A., Crivelli-Decker, J., & Ranganath, C. (2022). Contextual Expectations Shape Cortical Reinstatement of Sensory Representations. *Journal of Neuroscience*, *42*(30), 5956–5965. https://doi.org/10.1523/JNEUROSCI.2045-21.2022

Clayton, M. S., Yeung, N., & Cohen Kadosh, R. (2015). The roles of cortical oscillations in sustained attention. *Trends in Cognitive Sciences*, *19*(4), 188–195. https://doi.org/10.1016/j.tics.2015.02.004

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*(10), 406–416. https://doi.org/10.1016/S1364-6613(98)01232-7

Clemenson, G. D., & Stark, C. E. L. (2015). Virtual Environmental Enrichment through Video Games Improves Hippocampal-Associated Memory. *Journal of Neuroscience*, *35*(49), 16116–16125. https://doi.org/10.1523/JNEUROSCI.2580-15.2015

Clewett, D., & Davachi, L. (2017). The ebb and flow of experience determines the temporal structure of memory. *Current Opinion in Behavioral Sciences*, *17*, 186–193. https://doi.org/10.1016/j.cobeha.2017.08.013

Clewett, D., DuBrow, S., & Davachi, L. (2019). Transcending time in the brain: How event memories are constructed from experience. *Hippocampus*, *29*(3), 162–183. https://doi.org/10.1002/hipo.23074

Clos, M., Sommer, T., Schneider, S. L., & Rose, M. (2018). Enhanced transformation of incidentally learned knowledge into explicit memory by dopaminergic modulation. *PLOS ONE*, *13*(6), e0199013. https://doi.org/10.1371/journal.pone.0199013

Clouter, A., Shapiro, K. L., & Hanslmayr, S. (2017). Theta Phase Synchronization Is the Glue that Binds Human Associative Memory. *Current Biology*, *27*(20), 3143-3148.e6. https://doi.org/10.1016/j.cub.2017.09.001

Cohen, M. X. (2011). Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *NeuroImage*, *55*(3), 1373–1383. https://doi.org/10.1016/j.neuroimage.2010.12.072

Cohen, S. S., & Parra, L. C. (2016). Memorable Audiovisual Narratives Synchronize Sensory and Supramodal Neural Responses. *eNeuro*, *3*(6). https://doi.org/10.1523/ENEURO.0203-16.2016

Colgin, L. L. (2013). Mechanisms and Functions of Theta Rhythms. *Annual Review of Neuroscience*, *36*(1), 295–312. https://doi.org/10.1146/annurev-neuro-062012-170330

Cooper, N., Millela, F., Cant, I., White, M. D., & Meyer, G. (2021). Transfer of training—Virtual reality training with augmented multisensory cues improves user experience during training and task performance in the real world. *PLOS ONE*, *16*(3), e0248225. https://doi.org/10.1371/journal.pone.0248225

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

Copara, M. S., Hassan, A. S., Kyle, C. T., Libby, L. A., Ranganath, C., & Ekstrom, A. D. (2014). Complementary Roles of Human Hippocampal Subregions during Retrieval of Spatiotemporal Context. *Journal of Neuroscience*, *34*(20), 6834–6842. https://doi.org/10.1523/JNEUROSCI.5341-13.2014

Corrigan, N., Păsărelu, C.-R., & Voinescu, A. (2023). Immersive virtual reality for improving cognitive deficits in children with ADHD: A systematic review and meta-analysis. *Virtual Reality*, *27*(4), 3545–3564. https://doi.org/10.1007/s10055-023-00768-1

Criss, A. H. (2010). Differentiation and response bias in episodic memory: Evidence from reaction time distributions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 484–499. https://doi.org/10.1037/a0018435

Cruz, G., Melcón, M., Sutandi, L., Palva, J. M., Palva, S., & Thut, G. (2025). Oscillatory Brain Activity in the Canonical Alpha-Band Conceals Distinct Mechanisms in Attention. *Journal of Neuroscience*, *45*(1). https://doi.org/10.1523/JNEUROSCI.0918-24.2024

Cruzat, J., Torralba, M., Ruzzoli, M., Fernández, A., Deco, G., & Soto-Faraco, S. (2021). The phase of Theta oscillations modulates successful memory formation at encoding. *Neuropsychologia*, *154*, 107775. https://doi.org/10.1016/j.neuropsychologia.2021.107775

Danker, J. F., Fincham, J. M., & Anderson, J. R. (2011). The neural correlates of competition during memory retrieval are modulated by attention to the cues. *Neuropsychologia*, *49*(9), 2427–2438. https://doi.org/10.1016/j.neuropsychologia.2011.04.020

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

Dave, S., VanHaerents, S., Bonakdarpour, B., Mesulam, M.-M., & Voss, J. L. (2022). Stimulation of distinct parietal locations differentiates frontal versus hippocampal network involvement in memory formation. *Current Research in Neurobiology*, *3*, 100030. https://doi.org/10.1016/j.crneur.2022.100030

DeKeyser, R. (2003). Implicit and Explicit Learning. In *The Handbook of Second Language Acquisition* (S. 312–348). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470756492.ch11

Depoorter, A., & Vandierendonck, A. (2009). Evidence for Modality-Independent Order Coding in Working Memory. *Quarterly Journal of Experimental Psychology*, *62*(3), 531–549. https://doi.org/10.1080/17470210801995002

Destrebecqz, A., Peigneux, P., Laureys, S., Degueldre, C., Del Fiore, G., Aerts, J., Luxen, A., Van Der Linden, M., Cleeremans, A., & Maquet, P. (2005). The neural correlates of implicit and explicit sequence learning: Interacting networks revealed by the process dissociation procedure. *Learning & Memory*, *12*(5), 480–490. https://doi.org/10.1101/lm.95605

Diamond, N. B., & Levine, B. (2020). Linking Detail to Temporal Structure in Naturalistic-Event Recall. *Psychological Science*, *31*(12), 1557–1572. https://doi.org/10.1177/0956797620958651

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

Duarte, S. E., Yonelinas, A. P., Ghetti, S., & Geng, J. J. (2025). Multisensory processing impacts memory for objects and their sources. *Memory & Cognition*, *53*(2), 646–665. https://doi.org/10.3758/s13421-024-01592-x

DuBrow, S., & Davachi, L. (2013). The influence of context boundaries on memory for the sequential order of events. *Journal of Experimental Psychology: General*, *142*(4), 1277–1286. https://doi.org/10.1037/a0034024

DuBrow, S., & Davachi, L. (2014). Temporal Memory Is Shaped by Encoding Stability and Intervening Item Reactivation. *The Journal of Neuroscience*, *34*(42), 13998–14005. https://doi.org/10.1523/JNEUROSCI.2535-14.2014

DuBrow, S., & Davachi, L. (2016). Temporal binding within and across events. *Neurobiology of Learning and Memory*, *134*, 107–114. https://doi.org/10.1016/j.nlm.2016.07.011

Duecker, K., Doelling, K. B., Breska, A., Coffey, E. B. J., Sivarao, D. V., & Zoefel, B. (2024). Challenges and Approaches in the Study of Neural Entrainment. *The Journal of Neuroscience*, *44*(40), e1234242024. https://doi.org/10.1523/JNEUROSCI.1234-24.2024

Duncan, D. H., van Moorselaar, D., & Theeuwes, J. (2025). Visual statistical learning requires attention. *Psychonomic Bulletin & Review*, *32*(3), 1240–1253. https://doi.org/10.3758/s13423-024-02605-1

Eichenbaum, H. (1997). DECLARATIVE MEMORY: Insights from Cognitive Neurobiology. *Annual Review of Psychology*, *48*(1), 547–572. https://doi.org/10.1146/annurev.psych.48.1.547

Eichenbaum, H. (2017). Prefrontal–hippocampal interactions in episodic memory. *Nature Reviews Neuroscience*, *18*(9), 547–558. https://doi.org/10.1038/nrn.2017.74

Eichenbaum, H., Sauvage, M., Fortin, N., Komorowski, R., & Lipton, P. (2012). Towards a functional organization of episodic memory in the medial temporal lobe. *Neuroscience & Biobehavioral Reviews*, *36*(7), 1597–1608. https://doi.org/10.1016/j.neubiorev.2011.07.006

Ekstrom, A. D., & Yonelinas, A. P. (2020). Precision, binding, and the hippocampus: Precisely what are we talking about? *Neuropsychologia*, *138*, 107341. https://doi.org/10.1016/j.neuropsychologia.2020.107341

Erickson, Smith, D., Albrecht, M. A., & Silverstein, S. (2019). Alpha-band desynchronization reflects memory-specific processes during visual change detection. *Psychophysiology*, *56*(11), e13442. https://doi.org/10.1111/psyp.13442

Esser, S., Lustig, C., & Haider, H. (2022). What triggers explicit awareness in implicit sequence learning? Implications from theories of consciousness. *Psychological Research*, *86*(5), 1442–1457. https://doi.org/10.1007/s00426-021-01594-3

Essoe, J. K.-Y., Reggente, N., Ohno, A. A., Baek, Y. H., Dell'Italia, J., & Rissman, J. (2022). Enhancing learning and retention with distinctive virtual reality environments and mental context reinstatement. *Npj Science of Learning*, *7*(1), 31. https://doi.org/10.1038/s41539-022-00147-6

Evans, J. S. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology*, *59*(Volume 59, 2008), 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review*, *119*(2), 223–271. https://doi.org/10.1037/a0027371

Fell, J., Ludowig, E., Staresina, B. P., Wagner, T., Kranz, T., Elger, C. E., & Axmacher, N. (2011). Medial Temporal Theta/Alpha Power Enhancement Precedes Successful Memory Encoding: Evidence Based on Intracranial EEG. *The Journal of Neuroscience*, *31*(14), 5392–5397. https://doi.org/10.1523/JNEUROSCI.3668-10.2011

Feng, T., Silva, D., & Foster, D. J. (2015). Dissociation between the Experience-Dependent Development of Hippocampal Theta Sequences and Single-Trial Phase Precession. *The Journal of Neuroscience*, *35*(12), 4890–4902. https://doi.org/10.1523/JNEUROSCI.2614-14.2015

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

Fonollosa, J., Neftci, E., & Rabinovich, M. (2015). Learning of Chunking Sequences in Cognition and Behavior. *PLOS Computational Biology*, *11*(11), e1004592. https://doi.org/10.1371/journal.pcbi.1004592

Forkstam, C., & Petersson, K. M. (2005). Towards an explicit account of implicit learning: *Current Opinion in Neurology*, *18*(4), 435–441. https://doi.org/10.1097/01.wco.0000171951.82995.c4

Foudil, S.-A., Pleche, C., & Macaluso, E. (2021). Memory for spatio-temporal contextual details during the retrieval of naturalistic episodes. *Scientific Reports*, *11*(1), 14577. https://doi.org/10.1038/s41598-021-93960-9

Foxe, J. J., & Snyder, A. C. (2011). The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00154

French, R. L., & DeAngelis, G. C. (2020). Multisensory neural processing: From cue integration to causal inference. *Current Opinion in Physiology*, *16*, 8–13. https://doi.org/10.1016/j.cophys.2020.04.004

Frensch, P. A., & Rünger, D. (2003). Implicit Learning. *CURRENT DIRECTIONS IN PSYCHOLOGICAL SCIENCE*, *12*(1).

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

Fuentemilla, L. (2018). Memory: Theta Rhythm Couples Periodic Reactivation during Memory Retrieval. *Current Biology*, *28*(21), R1243–R1245. https://doi.org/10.1016/j.cub.2018.09.016

Gabay, Y., Karni, A., & Holt, L. L. (2023). Memory for incidentally learned categories evolves in the post-learning interval. *eLife*, *12*, e81855. https://doi.org/10.7554/eLife.81855

Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, *68*(Volume 68, 2017), 101–128. https://doi.org/10.1146/annurev-psych-122414-033625

Gerver, C. R., Overman, A. A., Babu, H. J., Hultman, C. E., & Dennis, N. A. (2020). Examining the Neural Basis of Congruent and Incongruent Configural Contexts during Associative Retrieval. *Journal of Cognitive Neuroscience*, *32*(9), 1796–1812. https://doi.org/10.1162/jocn_a_01593

Gilmore, A. W., Quach, A., Kalinowski, S. E., Gotts, S. J., Schacter, D. L., & Martin, A. (2021). Dynamic Content Reactivation Supports Naturalistic Autobiographical Recall in Humans. *Journal of Neuroscience*, *41*(1), 153–166. https://doi.org/10.1523/JNEUROSCI.1490-20.2020

Gordon, A. M., Rissman, J., Kiani, R., & Wagner, A. D. (2014). Cortical Reinstatement Mediates the Relationship Between Content-Specific Encoding Activity and Subsequent Recollection Decisions. *Cerebral Cortex*, *24*(12), 3350–3364. https://doi.org/10.1093/cercor/bht194

Goujon, A., Didierjean, A., & Poulet, S. (2014). The emergence of explicit knowledge from implicit learning. *Memory & Cognition*, *42*(2), 225–236. https://doi.org/10.3758/s13421-013-0355-0

Goujon, A., Didierjean, A., & Thorpe, S. (2015). Investigating implicit statistical learning mechanisms through contextual cueing. *Trends in Cognitive Sciences*, *19*(9), 524–533. https://doi.org/10.1016/j.tics.2015.07.009

Gould, I. C., Rushworth, M. F., & Nobre, A. C. (2011). Indexing the graded allocation of visuospatial attention using anticipatory alpha oscillations. *Journal of Neurophysiology*, *105*(3), 1318–1326. https://doi.org/10.1152/jn.00653.2010

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

Grigoryan, G. A. (2023). The systemic effects of the enriched environment on the conditioned fear reaction. *Frontiers in Behavioral Neuroscience*, *17*. https://doi.org/10.3389/fnbeh.2023.1227575

Gruber, M. J., Hsieh, L.-T., Staresina, B. P., Elger, C. E., Fell, J., Axmacher, N., & Ranganath, C. (2018). Theta Phase Synchronization between the Human Hippocampus and Prefrontal Cortex Increases during Encoding of Unexpected Information: A Case Study. *Journal of Cognitive Neuroscience*, *30*(11), 1646–1656. https://doi.org/10.1162/jocn_a_01302

Guderian, S., Schott, B. H., Richardson-Klavehn, A., & Düzel, E. (2009). Medial temporal theta state before an event predicts episodic encoding success in humans. *Proceedings of the National Academy of Sciences*, *106*(13), 5365–5370. https://doi.org/10.1073/pnas.0900289106

Gulbinaite, R., Van Viegen, T., Wieling, M., Cohen, M. X., & VanRullen, R. (2017). Individual Alpha Peak Frequency Predicts 10 Hz Flicker Effects on Selective Attention. *The Journal of Neuroscience*, *37*(42), 10173–10184. https://doi.org/10.1523/JNEUROSCI.1163-17.2017

Haegens, S., & Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience & Biobehavioral Reviews*, *86*, 150–165. https://doi.org/10.1016/j.neubiorev.2017.12.002

Hakim, N., Feldmann-Wüstefeld, T., Awh, E., & Vogel, E. K. (2020). Perturbing Neural Representations of Working Memory with Task-irrelevant Interruption. *Journal of Cognitive Neuroscience*, *32*(3), 558–569. https://doi.org/10.1162/jocn_a_01481

Han, X., Xu, J., Chang, S., Keniston, L., & Yu, L. (2022). Multisensory-Guided Associative Learning Enhances Multisensory Representation in Primary Auditory Cortex. *Cerebral Cortex*, *32*(5), 1040–1054. https://doi.org/10.1093/cercor/bhab264

Hanslmayr, S., Axmacher, N., & Inman, C. S. (2019). Modulating Human Memory via Entrainment of Brain Oscillations. *Trends in Neurosciences*, *42*(7), 485–499. https://doi.org/10.1016/j.tins.2019.04.004

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

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

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

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences*, *11*(7), 299–306. https://doi.org/10.1016/j.tics.2007.05.001

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

Heald, J. B., Lengyel, M., & Wolpert, D. M. (2023). Contextual inference in learning and memory. *Trends in Cognitive Sciences*, *27*(1), 43–64. https://doi.org/10.1016/j.tics.2022.10.004

Hebscher, M., Kragel, J. E., Kahnt, T., & Voss, J. L. (2021). Enhanced reinstatement of naturalistic event memories due to hippocampal-network-targeted stimulation. *Current Biology: CB*, *31*(7), 1428-1437.e5. https://doi.org/10.1016/j.cub.2021.01.027

Hebscher, M., Meltzer, J. A., & Gilboa, A. (2019). A causal role for the precuneus in network-wide theta and gamma oscillatory activity during complex memory retrieval. *eLife*, *8*, e43114. https://doi.org/10.7554/eLife.43114

Hebscher, M., & Voss, J. L. (2020). Testing network properties of episodic memory using non-invasive brain stimulation. *Current Opinion in Behavioral Sciences*, *32*, 35–42. https://doi.org/10.1016/j.cobeha.2020.01.012

Heinbockel, H., Wagner, A. D., & Schwabe, L. (2024). Post-retrieval stress impairs subsequent memory depending on hippocampal memory trace reinstatement during reactivation. *Science Advances*, *10*(18), eadm7504. https://doi.org/10.1126/sciadv.adm7504

Heinbockel, H., W.E.M. Quaedflieg, C., Wacker, J., & Schwabe, L. (2022). Spatio-temporal theta pattern dissimilarity in the right centro-parietal area during memory generalization. *Brain and Cognition*, *164*, 105926. https://doi.org/10.1016/j.bandc.2022.105926

Helbing, J., Draschkow, D., & Võ, M. L.-H. (2020). Search superiority: Goal-directed attentional allocation creates more reliable incidental identity and location memory than explicit encoding in naturalistic virtual environments. *Cognition*, *196*, 104147. https://doi.org/10.1016/j.cognition.2019.104147

Hendriks, D., Verkoeijen, P., & Pecher, D. (2024). Encoding variability explains the multisensory benefit in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *50*(6), 920–929. https://doi.org/10.1037/xlm0001305

Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience*, *11*(7), 523–532. https://doi.org/10.1038/nrn2850

Hennings, A. C., McClay, M., Lewis-Peacock, J. A., & Dunsmoor, J. E. (2020). Contextual reinstatement promotes extinction generalization in healthy adults but not PTSD. *Neuropsychologia*, *147*, 107573. https://doi.org/10.1016/j.neuropsychologia.2020.107573

Herweg, N. A., Solomon, E. A., & Kahana, M. J. (2020). Theta Oscillations in Human Memory. *Trends in Cognitive Sciences*, *24*(3), 208–227. https://doi.org/10.1016/j.tics.2019.12.006

Heuer, A., & Rolfs, M. (2021). Incidental encoding of visual information in temporal reference frames in working memory. *Cognition*, *207*, 104526. https://doi.org/10.1016/j.cognition.2020.104526

Heusser, A. C., Ezzyat, Y., Shiff, I., & Davachi, L. (2018). Perceptual boundaries cause mnemonic trade-offs between local boundary processing and across-trial associative binding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(7), 1075–1090. https://doi.org/10.1037/xlm0000503

Heusser, A. C., Poeppel, D., Ezzyat, Y., & Davachi, L. (2016). Episodic sequence memory is supported by a theta–gamma phase code. *Nature Neuroscience*, *19*(10), 1374–1380. https://doi.org/10.1038/nn.4374

Higham, P. A., Perfect, T. J., & Bruno, D. (2009). Investigating strength and frequency effects in recognition memory using type-2 signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 57–80. https://doi.org/10.1037/a0013865

Hodgson, V. J., Ralph, M. A. L., & Jackson, R. L. (2024). Disentangling the neural correlates of semantic and domain-general control: The roles of stimulus domain and task process. *Imaging Neuroscience*, *2*, 1–21. https://doi.org/10.1162/imag_a_00092

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

Horner, A. J., Bisby, J. A., Wang, A., Bogus, K., & Burgess, N. (2016). The role of spatial boundaries in shaping long-term event representations. *Cognition*, *154*, 151–164. https://doi.org/10.1016/j.cognition.2016.05.013

Horner, A. J., & Burgess, N. (2013). The associative structure of memory for multi-element events. *Journal of Experimental Psychology: General*, *142*(4), 1370–1383. https://doi.org/10.1037/a0033626

Howard, L. R., Kumaran, D., Ólafsdóttir, H. F., & Spiers, H. J. (2011). Double Dissociation between Hippocampal and Parahippocampal Responses to Object–Background Context and Scene Novelty. *Journal of Neuroscience*, *31*(14), 5253–5261. https://doi.org/10.1523/JNEUROSCI.6055-10.2011

Howard, M. D., Skorheim, S. W., & Pilly, P. K. (2022). A model of bi-directional interactions between complementary learning systems for memory consolidation of sequential experiences. *Frontiers in Systems Neuroscience*, *16*, 972235. https://doi.org/10.3389/fnsys.2022.972235

Howard, M. W. (2017). Temporal and spatial context in the mind and brain. *Current Opinion in Behavioral Sciences*, *17*, 14–19. https://doi.org/10.1016/j.cobeha.2017.05.022

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

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), 218–244. https://doi.org/10.1037/bul0000223

Hunsaker, M. R., Chen, V., Tran, G. T., & Kesner, R. P. (2013). The medial and lateral entorhinal cortex both contribute to contextual and item recognition memory: A test of the binding ofitems and context model. *Hippocampus*, *23*(5), 380–391. https://doi.org/10.1002/hipo.22097

Hurter, C., Rogowitz, B., Truong, G., Andry, T., Romat, H., Gardy, L., Amini, F., & Riche, N. H. (2024). Memory Recall for Data Visualizations in Mixed Reality, Virtual Reality, 3D and 2D. *IEEE Transactions on Visualization and Computer Graphics*, *30*(10), 6691–6706. https://doi.org/10.1109/TVCG.2023.3336588

Hwang, J., Whitman, K., & Umanath, S. (2024). Associative memory in older adults: Making sense of associative memory deficits and hyperbinding effects. *Psychology and Aging*, *39*(8), 871–883. https://doi.org/10.1037/pag0000851

Hyman, J. M., Ma, L., Balaguer-Ballester, E., Durstewitz, D., & Seamans, J. K. (2012). Contextual encoding by ensembles of medial prefrontal cortex neurons. *Proceedings of the National Academy of Sciences*, *109*(13), 5086–5091. https://doi.org/10.1073/pnas.1114415109

Iriye, H., Chancel, M., & Ehrsson, H. H. (2024). Sense of own body shapes neural processes of memory encoding and reinstatement. *Cerebral Cortex*, *34*(1), bhad443. https://doi.org/10.1093/cercor/bhad443

Jääskeläinen, I. P., Sams, M., Glerean, E., & Ahveninen, J. (2021). Movies and narratives as naturalistic stimuli in neuroimaging. *NeuroImage*, *224*, 117445. https://doi.org/10.1016/j.neuroimage.2020.117445

Jensen, O. (2002). Oscillations in the Alpha Band (9-12 Hz) Increase with Memory Load during Retention in a Short-term Memory Task. *Cerebral Cortex*, *12*(8), 877–882. https://doi.org/10.1093/cercor/12.8.877

Jensen, O., & Mazaheri, A. (2010). Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition. *Frontiers in Human Neuroscience*, *4*. https://doi.org/10.3389/fnhum.2010.00186

Jeong, S., Kim, J., & Lee, J. (2024). The Differential Effects of Multisensory Attentional Cues on Task Performance in VR Depending on the Level of Cognitive Load and Cognitive Capacity. *IEEE Transactions on Visualization and Computer Graphics*, *30*(5), 2703–2712. https://doi.org/10.1109/TVCG.2024.3372126

Ji, E., Lee, K. M., & Kim, M.-S. (2017). Independent operation of implicit working memory under cognitive load. *Consciousness and Cognition*, *55*, 214–222. https://doi.org/10.1016/j.concog.2017.08.014

Jiménez, L., Vaquero, J. M. M., & Lupiáñez, J. (2006). Qualitative differences between implicit and explicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 475–490. https://doi.org/10.1037/0278-7393.32.3.475

Johansson, R., & Johansson, M. (2014). Look Here, Eye Movements Play a Functional Role in Memory Retrieval. *Psychological Science*, *25*(1), 236–242. https://doi.org/10.1177/0956797613498260

Johnsdorf, M., Kisker, J., Gruber, T., & Schöne, B. (2023). Comparing encoding mechanisms in realistic virtual reality and conventional 2D laboratory settings: Event-related potentials in a repetition suppression paradigm. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1051938

Johnson, J. D., McDuff, S. G. R., Rugg, M. D., & Norman, K. A. (2009). Recollection, Familiarity, and Cortical Reinstatement: A Multivoxel Pattern Analysis. *Neuron*, *63*(5), 697–708. https://doi.org/10.1016/j.neuron.2009.08.011

Junker, F. B., Schlaffke, L., Axmacher, N., & Schmidt- Wilcke, T. (2021). Impact of multisensory learning on perceptual and lexical processing of unisensory Morse code. *Brain Research*, *1755*, 147259. https://doi.org/10.1016/j.brainres.2020.147259

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*(1), 103–109. https://doi.org/10.3758/BF03197276

Karakaş, S. (2020). A review of theta oscillation and its functional correlates. *International Journal of Psychophysiology*, *157*, 82–99. https://doi.org/10.1016/j.ijpsycho.2020.04.008

Kayser, C., & Logothetis, N. K. (2007). Do early sensory cortices integrate cross-modal information? *Brain Structure and Function*, *212*(2), 121–132. https://doi.org/10.1007/s00429-007-0154-0

Keil, J., & Senkowski, D. (2018). Neural Oscillations Orchestrate Multisensory Processing. *The Neuroscientist*, *24*(6), 609–626. https://doi.org/10.1177/1073858418755352

Kempermann, G. (2019). Environmental enrichment, new neurons and the neurobiology of individuality. *Nature Reviews Neuroscience*, *20*(4), 235–245. https://doi.org/10.1038/s41583-019-0120-x

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

Kerrén, C., Reznik, D., Doeller, C. F., & Griffiths, B. J. (2025). Exploring the role of dimensionality transformation in episodic memory. *Trends in Cognitive Sciences*, *0*(0). https://doi.org/10.1016/j.tics.2025.01.007

Kerz, E., Wiechmann, D., & Riedel, F. B. (2017). IMPLICIT LEARNING IN THE CROWD: INVESTIGATING THE ROLE OF AWARENESS IN THE ACQUISITION OF L2 KNOWLEDGE. *Studies in Second Language Acquisition*, *39*(4), 711–734. https://doi.org/10.1017/S027226311700002X

Khader, P. H., & Rösler, F. (2011). EEG power changes reflect distinct mechanisms during long-term memory retrieval. *Psychophysiology*, *48*(3), 362–369. https://doi.org/10.1111/j.1469-8986.2010.01063.x

Khosla, M., Ngo, G. H., Jamison, K., Kuceyeski, A., & Sabuncu, M. R. (2021). Cortical response to naturalistic stimuli is largely predictable with deep neural networks. *Science Advances*, *7*(22), eabe7547. https://doi.org/10.1126/sciadv.abe7547

Kilteni, K., Groten, R., & Slater, M. (2012). The Sense of Embodiment in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, *21*(4), 373–387. https://doi.org/10.1162/PRES_a_00124

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

Kirkhart, M. W. (2001). The Nature of Declarative and Nondeclarative Knowledge for Implicit and Explicit Learning. *The Journal of General Psychology*, *128*(4), 447–461. https://doi.org/10.1080/00221300109598921

Kisker, J., Gruber, T., & Schöne, B. (2021a). Behavioral realism and lifelike psychophysiological responses in virtual reality by the example of a height exposure. *Psychological Research*, *85*(1), 68–81. https://doi.org/10.1007/s00426-019-01244-9

Kisker, J., Gruber, T., & Schöne, B. (2021b). Virtual reality experiences promote autobiographical retrieval mechanisms: Electrophysiological correlates of laboratory and virtual experiences. *Psychological Research*, *85*(7), 2485–2501. https://doi.org/10.1007/s00426-020-01417-x

Klimesch, W., Fellinger, R., & Freunberger, R. (2011). Alpha Oscillations and Early Stages of Visual Encoding. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00118

Klimesch, W., Freunberger, R., & Sauseng, P. (2010). Oscillatory mechanisms of process binding in memory. *Neuroscience & Biobehavioral Reviews*, *34*(7), 1002–1014. https://doi.org/10.1016/j.neubiorev.2009.10.004

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

Köster, M., Martens, U., & Gruber, T. (2019). Memory entrainment by visually evoked theta-gamma coupling. *NeuroImage*, *188*, 181–187. https://doi.org/10.1016/j.neuroimage.2018.12.002

Kota, S., Rugg, M. D., & Lega, B. C. (2020). Hippocampal Theta Oscillations Support Successful Associative Memory Formation. *Journal of Neuroscience*, *40*(49), 9507–9518. https://doi.org/10.1523/JNEUROSCI.0767-20.2020

Kragel, J. E., Ezzyat, Y., Lega, B. C., Sperling, M. R., Worrell, G. A., Gross, R. E., Jobst, B. C., Sheth, S. A., Zaghloul, K. A., Stein, J. M., & Kahana, M. J. (2021). Distinct cortical systems reinstate the content and context of episodic memories. *Nature Communications*, *12*(1), 4444. https://doi.org/10.1038/s41467-021-24393-1

Kragel, J. E., VanHaerents, S., Templer, J. W., Schuele, S., Rosenow, J. M., Nilakantan, A. S., & Bridge, D. J. (2020). Hippocampal theta coordinates memory processing during visual exploration. *eLife*, *9*, e52108. https://doi.org/10.7554/eLife.52108

Krokos, E., Plaisant, C., & Varshney, A. (2019). Virtual memory palaces: Immersion aids recall. *Virtual Reality*, *23*(1), 1–15. https://doi.org/10.1007/s10055-018-0346-3

Kuhl, B. A., & Chun, M. M. (2014). Successful Remembering Elicits Event-Specific Activity Patterns in Lateral Parietal Cortex. *Journal of Neuroscience*, *34*(23), 8051–8060. https://doi.org/10.1523/JNEUROSCI.4328-13.2014

Lakatos, P., Gross, J., & Thut, G. (2019). A New Unifying Account of the Roles of Neuronal Entrainment. *Current Biology*, *29*(18), R890–R905. https://doi.org/10.1016/j.cub.2019.07.075

Lee, H., Stirnberg, R., Stöcker, T., & Axmacher, N. (2017). Audiovisual integration supports face–name associative memory formation. *Cognitive Neuroscience*, *8*(4), 177–192. https://doi.org/10.1080/17588928.2017.1327426

Lega, B., Burke, J., Jacobs, J., & Kahana, M. J. (2016). Slow-Theta-to-Gamma Phase–Amplitude Coupling in Human Hippocampus Supports the Formation of New Episodic Memories. *Cerebral Cortex*, *26*(1), 268–278. https://doi.org/10.1093/cercor/bhu232

Lehmann, S., & Murray, M. M. (2005). The role of multisensory memories in unisensory object discrimination. *Cognitive Brain Research*, *24*(2), 326–334. https://doi.org/10.1016/j.cogbrainres.2005.02.005

Lehn, H., Steffenach, H.-A., Strien, N. M. van, Veltman, D. J., Witter, M. P., & Håberg, A. K. (2009). A Specific Role of the Human Hippocampus in Recall of Temporal Sequences. *Journal of Neuroscience*, *29*(11), 3475–3484. https://doi.org/10.1523/JNEUROSCI.5370-08.2009

Levine, B., Black, S. E., Cabeza, R., Sinden, M., Mcintosh, A. R., Toth, J. P., Tulving, E., & Stuss, D. T. (1998). Episodic memory and the self in a case of isolated retrograde amnesia. *Brain*, *121*(10), 1951–1973. https://doi.org/10.1093/brain/121.10.1951

Lewis, P. A., & Bendor, D. (2019). How Targeted Memory Reactivation Promotes the Selective Strengthening of Memories in Sleep. *Current Biology*, *29*(18), R906–R912. https://doi.org/10.1016/j.cub.2019.08.019

Li, J., & Deng, S. W. (2023). Facilitation and interference effects of the multisensory context on learning: A systematic review and meta-analysis. *Psychological Research*, *87*(5), 1334–1352. https://doi.org/10.1007/s00426-022-01733-4

Libby, L. A., Hannula, D. E., & Ranganath, C. (2014). Medial Temporal Lobe Coding of Item and Spatial Information during Relational Binding in Working Memory. *Journal of Neuroscience*, *34*(43), 14233–14242. https://doi.org/10.1523/JNEUROSCI.0655-14.2014

Lifanov-Carr, J., Griffiths, B. J., Linde-Domingo, J., Ferreira, C. S., Wilson, M., Mayhew, S. D., Charest, I., & Wimber, M. (2024). Reconstructing Spatiotemporal Trajectories of Visual Object Memories in the Human Brain. *Eneuro*, *11*(9), ENEURO.0091-24.2024. https://doi.org/10.1523/ENEURO.0091-24.2024

Linde-Domingo, J., Treder, M. S., Kerrén, C., & Wimber, M. (2019). Evidence that neural information flow is reversed between object perception and object reconstruction from memory. *Nature Communications*, *10*(1), 179. https://doi.org/10.1038/s41467-018-08080-2

Liu, C., Ye, Z., Chen, C., Axmacher, N., & Xue, G. (2022). Hippocampal Representations of Event Structure and Temporal Context during Episodic Temporal Order Memory. *Cerebral Cortex*, *32*(7), 1520–1534. https://doi.org/10.1093/cercor/bhab304

Liu, Z., Johansson, M., Johansson, R., & Bramão, I. (2024). The effects of episodic context on memory integration. *Scientific Reports*, *14*(1), 30159. https://doi.org/10.1038/s41598-024-82004-7

Lohnas, L. J., Duncan, K., Doyle, W. K., Thesen, T., Devinsky, O., & Davachi, L. (2018). Time-resolved neural reinstatement and pattern separation during memory decisions in human hippocampus. *Proceedings of the National Academy of Sciences*, *115*(31), E7418–E7427. https://doi.org/10.1073/pnas.1717088115

Lohnas, L. J., & Healey, M. K. (2021). The role of context in episodic memory: Behavior and neurophysiology. In *Psychology of Learning and Motivation* (Bd. 75, S. 157–199). Elsevier. https://doi.org/10.1016/bs.plm.2021.06.003

Lohnas, L. J., Healey, M. K., & Davachi, L. (2023). Neural temporal context reinstatement of event structure during memory recall. *Journal of Experimental Psychology. General*, *152*(7), 1840–1872. https://doi.org/10.1037/xge0001354

Lohnas, L. J., & Kahana, M. J. (2014). Compound cuing in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 12–24. https://doi.org/10.1037/a0033698

Loonis, R. F., Brincat, S. L., Antzoulatos, E. G., & Miller, E. K. (2017). A Meta-Analysis Suggests Different Neural Correlates for Implicit and Explicit Learning. *Neuron*, *96*(2), 521-534.e7. https://doi.org/10.1016/j.neuron.2017.09.032

Lüddecke, R., & Felnhofer, A. (2022). Virtual Reality Biofeedback in Health: A Scoping Review. *Applied Psychophysiology and Biofeedback*, *47*(1), 1–15. https://doi.org/10.1007/s10484-021-09529-9

Ma, Y., Brown, J. A., Chen, C., Ding, M., Wu, W., & Li, W. (2025). Alpha-Frequency Stimulation Enhances Synchronization of Alpha Oscillations with Default Mode Network Connectivity. *eNeuro*, *12*(3). https://doi.org/10.1523/ENEURO.0449-24.2025

Macaluso, E. (2006). Multisensory Processing in Sensory-Specific Cortical Areas. *The Neuroscientist*, *12*(4), 327–338. https://doi.org/10.1177/1073858406287908

Magosso, E., & Borra, D. (2024). The strength of anticipated distractors shapes EEG alpha and theta oscillations in a Working Memory task. *NeuroImage*, *300*, 120835. https://doi.org/10.1016/j.neuroimage.2024.120835

Makowski, D., Sperduti, M., Nicolas, S., & Piolino, P. (2017). “Being there” and remembering it: Presence improves memory encoding. *Consciousness and Cognition*, *53*, 194–202. https://doi.org/10.1016/j.concog.2017.06.015

Malmberg, K. J. (2008). Recognition memory: A review of the critical findings and an integrated theory for relating them. *Cognitive Psychology*, *57*(4), 335–384. https://doi.org/10.1016/j.cogpsych.2008.02.004

Manning, J. R., Polyn, S. M., Baltuch, G. H., Litt, B., & Kahana, M. J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings of the National Academy of Sciences*, *108*(31), 12893–12897. https://doi.org/10.1073/pnas.1015174108

Marks, W. D., Yokose, J., Kitamura, T., & Ogawa, S. K. (2022). Neuronal Ensembles Organize Activity to Generate Contextual Memory. *Frontiers in Behavioral Neuroscience*, *16*, 805132. https://doi.org/10.3389/fnbeh.2022.805132

Mather, M., & Sutherland, M. R. (2011). Arousal-Biased Competition in Perception and Memory. *Perspectives on Psychological Science*, *6*(2), 114–133. https://doi.org/10.1177/1745691611400234

Mathewson, K. E., Kuziek, J. P., Scanlon, J. E. M., & Robles, D. (2024). The moving wave: Applications of the mobile EEG approach to study human attention. *Psychophysiology*, *61*(9), e14603. https://doi.org/10.1111/psyp.14603

Medvedeva, A., Materassi, M., Neacsu, V., Beresford-Webb, J., Hussin, A., Khan, N., Newton, F., & Galli, G. (2019). Effects of Anodal Transcranial Direct Current Stimulation Over the Ventrolateral Prefrontal Cortex on Episodic Memory Formation and Retrieval. *Cerebral Cortex*, *29*(2), 657–665. https://doi.org/10.1093/cercor/bhx347

Menesse, G., & Torres, J. J. (2024). Information dynamics of in silico EEG Brain Waves: Insights into oscillations and functions. *PLOS Computational Biology*, *20*(9), e1012369. https://doi.org/10.1371/journal.pcbi.1012369

Michael, E., Covarrubias, L. S., Leong, V., & Kourtzi, Z. (2023). Learning at your brain's rhythm: Individualized entrainment boosts learning for perceptual decisions. *Cerebral Cortex*, *33*(9), 5382–5394. https://doi.org/10.1093/cercor/bhac426

Michelmann, S., Bowman, H., & Hanslmayr, S. (2016). The Temporal Signature of Memories: Identification of a General Mechanism for Dynamic Memory Replay in Humans. *PLOS Biology*, *14*(8), e1002528. https://doi.org/10.1371/journal.pbio.1002528

Michelmann, S., Bowman, H., & Hanslmayr, S. (2018). Replay of Stimulus-specific Temporal Patterns during Associative Memory Formation. *Journal of Cognitive Neuroscience*, *30*(11), 1577–1589. https://doi.org/10.1162/jocn_a_01304

Minarik, T., Berger, B., & Sauseng, P. (2018). The involvement of alpha oscillations in voluntary attention directed towards encoding episodic memories. *NeuroImage*, *166*, 307–316. https://doi.org/10.1016/j.neuroimage.2017.10.064

Monaro, M., Mazza, C., Colasanti, M., Colicino, E., Bosco, F., Ricci, E., Biondi, S., Rossi, M., & Roma, P. (2024). Testing memory of a VR environment: Comparison with the real environment and 2D pictures. *Virtual Reality*, *28*(2), 100. https://doi.org/10.1007/s10055-024-00999-w

Montoya, A. K. (2023). Selecting a Within- or Between-Subject Design for Mediation: Validity, Causality, and Statistical Power. *Multivariate Behavioral Research*, *58*(3), 616–636. https://doi.org/10.1080/00273171.2022.2077287

Mudrik, L., Hirschhorn, R., & Korisky, U. (2024). Taking consciousness for real: Increasing the ecological validity of the study of conscious vs. unconscious processes. *Neuron*, *112*(10), 1642–1656. https://doi.org/10.1016/j.neuron.2024.03.031

Mullet, E., & Chasseigne, G. (2018). Assessing information integration processes: A comparison of findings obtained with between-subjects designs versus within-subjects designs. *Quality & Quantity*, *52*(4), 1977–1988. https://doi.org/10.1007/s11135-017-0592-6

Mulligan, N. W. (2011). Conceptual implicit memory and environmental context. *Consciousness and Cognition*, *20*(3), 737–744. https://doi.org/10.1016/j.concog.2010.11.008

Mulligan, N. W., Spataro, P., & West, J. T. (2023). Memory and attention: A double dissociation between memory encoding and memory retrieval. *Cognition*, *238*, 105509. https://doi.org/10.1016/j.cognition.2023.105509

Murray, M. M., Foxe, J. J., & Wylie, G. R. (2005). The brain uses single-trial multisensory memories to discriminate without awareness. *NeuroImage*, *27*(2), 473–478. https://doi.org/10.1016/j.neuroimage.2005.04.016

Murray, M. M., Michel, C. M., Grave De Peralta, R., Ortigue, S., Brunet, D., Gonzalez Andino, S., & Schnider, A. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *NeuroImage*, *21*(1), 125–135. https://doi.org/10.1016/j.neuroimage.2003.09.035

Murray, M. M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., & Matusz, P. J. (2016). The multisensory function of the human primary visual cortex. *Neuropsychologia*, *83*, 161–169. https://doi.org/10.1016/j.neuropsychologia.2015.08.011

Nanni-Zepeda, M., DeGutis, J., Wu, C., Rothlein, D., Fan, Y., Grimm, S., Walter, M., Esterman, M., & Zuberer, A. (2024). Neural signatures of shared subjective affective engagement and disengagement during movie viewing. *Human Brain Mapping*, *45*(4), e26622. https://doi.org/10.1002/hbm.26622

Ngo, C. T., Horner, A. J., Newcombe, N. S., & Olson, I. R. (2019). Development of Holistic Episodic Recollection. *Psychological Science*, *30*(12), 1696–1706. https://doi.org/10.1177/0956797619879441

Nikolaev, A. R., Bramão, I., Johansson, R., & Johansson, M. (2023). Episodic memory formation in unrestricted viewing. *NeuroImage*, *266*, 119821. https://doi.org/10.1016/j.neuroimage.2022.119821

Noel, J., Blanke, O., & Serino, A. (2018). From multisensory integration in peripersonal space to bodily self-consciousness: From statistical regularities to statistical inference. *Annals of the New York Academy of Sciences*, *1426*(1), 146–165. https://doi.org/10.1111/nyas.13867

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, *110*(4), 611–646. https://doi.org/10.1037/0033-295X.110.4.611

Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences*, *97*(20), 11120–11124. https://doi.org/10.1073/pnas.97.20.11120

Nyhus, E., & Curran, T. (2010). Functional role of gamma and theta oscillations in episodic memory. *Neuroscience & Biobehavioral Reviews*, *34*(7), 1023–1035. https://doi.org/10.1016/j.neubiorev.2009.12.014

Okray, Z., Jacob, P. F., Stern, C., Desmond, K., Otto, N., Talbot, C. B., Vargas-Gutierrez, P., & Waddell, S. (2023). Multisensory learning binds neurons into a cross-modal memory engram. *Nature*, *617*(7962), 777–784. https://doi.org/10.1038/s41586-023-06013-8

Ostrowski, J., & Rose, M. (2024). Increases in pre-stimulus theta and alpha oscillations precede successful encoding of crossmodal associations. *Scientific Reports*, *14*(1), 7895. https://doi.org/10.1038/s41598-024-58227-z

Otten, L. J., Henson, R. N. A., & Rugg, M. D. (2002). State-related and item-related neural correlates of successful memory encoding. *Nature Neuroscience*, *5*(12), 1339–1344. https://doi.org/10.1038/nn967

Otten, L. J., Quayle, A. H., & Puvaneswaran, B. (2010). Prestimulus Subsequent Memory Effects for Auditory and Visual Events. *Journal of Cognitive Neuroscience*, *22*(6), 1212–1223. https://doi.org/10.1162/jocn.2009.21298

Ozubko, J. D., Sirianni, L. A., Ahmad, F. N., MacLeod, C. M., & Addante, R. J. (2021). Recallable but not recognizable: The influence of semantic priming in recall paradigms. *Cognitive, Affective, & Behavioral Neuroscience*, *21*(1), 119–143. https://doi.org/10.3758/s13415-020-00854-w

Pacheco Estefan, D., Sánchez-Fibla, M., Duff, A., Principe, A., Rocamora, R., Zhang, H., Axmacher, N., & Verschure, P. F. M. J. (2019). Coordinated representational reinstatement in the human hippocampus and lateral temporal cortex during episodic memory retrieval. *Nature Communications*, *10*(1), 2255. https://doi.org/10.1038/s41467-019-09569-0

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

Panichello, M. F., & Buschman, T. J. (2021). Shared mechanisms underlie the control of working memory and attention. *Nature*, *592*(7855), 601–605. https://doi.org/10.1038/s41586-021-03390-w

Park, H., & Kayser, C. (2019). Shared neural underpinnings of multisensory integration and trial-by-trial perceptual recalibration in humans. *eLife*, *8*, e47001. https://doi.org/10.7554/eLife.47001

Park, J. L., & Donaldson, D. I. (2019). Detecting the neural correlates of episodic memory with mobile EEG: Recollecting objects in the real world. *NeuroImage*, *193*, 1–9. https://doi.org/10.1016/j.neuroimage.2019.03.013

Parker, A., Dagnall, N., & Coyle, A.-M. (2007). Environmental context effects in conceptual explicit and implicit memory. *Memory*, *15*(4), 423–434. https://doi.org/10.1080/09658210701309834

Parsons, T. D. (2015). Virtual Reality for Enhanced Ecological Validity and Experimental Control in the Clinical, Affective and Social Neurosciences. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00660

Pecher, D., & Zeelenberg, R. (2022). Does multisensory study benefit memory for pictures and sounds? *Cognition*, *226*, 105181. https://doi.org/10.1016/j.cognition.2022.105181

Peelen, M. V., & Downing, P. E. (2023). Testing cognitive theories with multivariate pattern analysis of neuroimaging data. *Nature Human Behaviour*, *7*(9), 1430–1441. https://doi.org/10.1038/s41562-023-01680-z

Penaud, S., Yeh, D., Gaston-Bellegarde, A., & Piolino, P. (2023). The role of bodily self-consciousness in episodic memory of naturalistic events: An immersive virtual reality study. *Scientific Reports*, *13*(1), 17013. https://doi.org/10.1038/s41598-023-43823-2

Peney, T., & Skarratt, P. A. (2024). Increasing the immersivity of 360° videos facilitates learning and memory: Implications for theory and practice. *Educational Technology Research and Development*, *72*(6), 3103–3115. https://doi.org/10.1007/s11423-024-10402-9

Peteranderl, S., & Oberauer, K. (2018). Serial recall of colors: Two models of memory for serial order applied to continuous visual stimuli. *Memory & Cognition*, *46*(1), 1–16. https://doi.org/10.3758/s13421-017-0741-0

Petrucci, A. S., McCall ,Cade, Schofield ,Guy, Wardell ,Victoria, Safi ,Omran K., & and Palombo, D. J. (2025). The relationship between environmentally induced emotion and memory for a naturalistic virtual experience. *Cognition and Emotion*, *39*(1), 180–195. https://doi.org/10.1080/02699931.2024.2333067

Place, R., Farovik, A., Brockmann, M., & Eichenbaum, H. (2016). Bidirectional prefrontal-hippocampal interactions support context-guided memory. *Nature Neuroscience*, *19*(8), 992–994. https://doi.org/10.1038/nn.4327

Poch, C., Campo, P., & Barnes, G. R. (2014). Modulation of alpha and gamma oscillations related to retrospectively orienting attention within working memory. *European Journal of Neuroscience*, *40*(2), 2399–2405. https://doi.org/10.1111/ejn.12589

Poch, C., Valdivia, M., Capilla, A., Hinojosa, J. A., & Campo, P. (2018). Suppression of no-longer relevant information in Working Memory: An alpha-power related mechanism? *Biological Psychology*, *135*, 112–116. https://doi.org/10.1016/j.biopsycho.2018.03.009

Polyn, S. M., & Kahana, M. J. (2008). Memory search and the neural representation of context. *Trends in Cognitive Sciences*, *12*(1), 24–30. https://doi.org/10.1016/j.tics.2007.10.010

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), 129–156. https://doi.org/10.1037/a0014420

Pomper, U., Szaszkó, B., Pfister, S., & Ansorge, U. (2023). Cross-modal attentional effects of rhythmic sensory stimulation. *Attention, Perception, & Psychophysics*, *85*(3), 863–878. https://doi.org/10.3758/s13414-022-02611-2

Pooja, R., Ghosh, P., & Sreekumar, V. (2024). Towards an ecologically valid naturalistic cognitive neuroscience of memory and event cognition. *Neuropsychologia*, *203*, 108970. https://doi.org/10.1016/j.neuropsychologia.2024.108970

Portuguez-Castro, M., & Santos Garduño, H. (2024). Beyond Traditional Classrooms: Comparing Virtual Reality Applications and Their Influence on Students' Motivation. *Education Sciences*, *14*(9), Article 9. https://doi.org/10.3390/educsci14090963

Preston, A. 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

Pu, Y., Kong, X.-Z., Ranganath, C., & Melloni, L. (2022). Event boundaries shape temporal organization of memory by resetting temporal context. *Nature Communications*, *13*(1), 622. https://doi.org/10.1038/s41467-022-28216-9

Qu, J., Ma, K., & Hommel, B. (2021). Cognitive load dissociates explicit and implicit measures of body ownership and agency. *Psychonomic Bulletin & Review*, *28*(5), 1567–1578. https://doi.org/10.3758/s13423-021-01931-y

Rameson, L. T., Satpute, A. B., & Lieberman, M. D. (2010). The neural correlates of implicit and explicit self-relevant processing. *NeuroImage*, *50*(2), 701–708. https://doi.org/10.1016/j.neuroimage.2009.12.098

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., & Hsieh, L.-T. (2016). The hippocampus: A special place for time. *Annals of the New York Academy of Sciences*, *1369*(1), 93–110. https://doi.org/10.1111/nyas.13043

Rau, E. M. B., Fellner, M.-C., Heinen, R., Zhang, H., Yin, Q., Vahidi, P., Kobelt, M., Asano, E., Kim-McManus, O., Sattar, S., Lin, J. J., Auguste, K. I., Chang, E. F., King-Stephens, D., Weber, P. B., Laxer, K. D., Knight, R. T., Johnson, E. L., Ofen, N., & Axmacher, N. (2025). Reinstatement and transformation of memory traces for recognition. *Science Advances*, *11*(8), eadp9336. https://doi.org/10.1126/sciadv.adp9336

Raud, L., Sneve, M. H., Vidal-Piñeiro, D., Sørensen, Ø., Folvik, L., Ness, H. T., Mowinckel, A. M., Grydeland, H., Walhovd, K. B., & Fjell, A. M. (2023). Hippocampal-cortical functional connectivity during memory encoding and retrieval. *NeuroImage*, *279*, 120309. https://doi.org/10.1016/j.neuroimage.2023.120309

Reber, P. J., & Squire, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learning & Memory*, *1*(4), 217–229. https://doi.org/10.1101/lm.1.4.217

Repetto, C., Serino, S., Macedonia, M., & Riva, G. (2016). Virtual Reality as an Embodied Tool to Enhance Episodic Memory in Elderly. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01839

Richter, F. R., Chanales, A. J. H., & Kuhl, B. A. (2016). Predicting the integration of overlapping memories by decoding mnemonic processing states during learning. *NeuroImage*, *124*, 323–335. https://doi.org/10.1016/j.neuroimage.2015.08.051

Rissman, J., & Wagner, A. D. (2012). Distributed Representations in Memory: Insights from Functional Brain Imaging. *Annual Review of Psychology*, *63*(Volume 63, 2012), 101–128. https://doi.org/10.1146/annurev-psych-120710-100344

Ritchey, M., Wing, E. A., LaBar, K. S., & Cabeza, R. (2013). Neural Similarity Between Encoding and Retrieval is Related to Memory Via Hippocampal Interactions. *Cerebral Cortex*, *23*(12), 2818–2828. https://doi.org/10.1093/cercor/bhs258

Roberts, B. M., Clarke, A., Addante, R. J., & Ranganath, C. (2018). Entrainment enhances theta oscillations and improves episodic memory. *Cognitive Neuroscience*, *9*(3–4), 181–193. https://doi.org/10.1080/17588928.2018.1521386

Robertson, L. C. (2003). Binding, spatial attention and perceptual awareness. *Nature Reviews Neuroscience*, *4*(2), 93–102. https://doi.org/10.1038/nrn1030

Rose, M., Haider, H., & Buchel, C. (2010). The Emergence of Explicit Memory during Learning. *Cerebral Cortex*, *20*(12), 2787–2797. https://doi.org/10.1093/cercor/bhq025

Saarimäki, H. (2021). Naturalistic Stimuli in Affective Neuroimaging: A Review. *Frontiers in Human Neuroscience*, *15*. https://doi.org/10.3389/fnhum.2021.675068

Sabo, M., & Schneider, D. (2022). Pattern reinstatement and attentional control overlap during episodic long-term memory retrieval. *Scientific Reports*, *12*(1), 10739. https://doi.org/10.1038/s41598-022-14090-4

Sadeh, T., & Moscovitch, M. (2024). Retrieval of temporal structure at recall can occur automatically. *Cognition*, *242*, 105647. https://doi.org/10.1016/j.cognition.2023.105647

Sakaki, M., Fryer, K., & Mather, M. (2014). Emotion strengthens high priority memory traces but weakens low priority memory traces. *Psychological science*, *25*(2), 387–395. https://doi.org/10.1177/0956797613504784

Schechtman, E., Heilberg, J., & Paller, K. A. (2023). Memory consolidation during sleep involves context reinstatement in humans. *Cell Reports*, *42*(4), 112331. https://doi.org/10.1016/j.celrep.2023.112331

Schneider, S. L., & Rose, M. (2016). Intention to encode boosts memory-related pre-stimulus EEG beta power. *NeuroImage*, *125*, 978–987. https://doi.org/10.1016/j.neuroimage.2015.11.024

Schnotz, W., & Kürschner, C. (2007). A Reconsideration of Cognitive Load Theory. *Educational Psychology Review*, *19*(4), 469–508. https://doi.org/10.1007/s10648-007-9053-4

Scholz, S., Dutke, S., & Busch, N. A. (2021). Oscillatory Correlates of Intentional Forgetting: The Role of Theta and Alpha Power in Item-Method Directed Forgetting. *eNeuro*, *8*(5). https://doi.org/10.1523/ENEURO.0022-21.2021

Scholz, S., Schneider, S. L., & Rose, M. (2017). Differential effects of ongoing EEG beta and theta power on memory formation. *PLOS ONE*, *12*(2), e0171913. https://doi.org/10.1371/journal.pone.0171913

Schöne, B., Kisker, J., Sylvester, R. S., Radtke, E. L., & Gruber, T. (2023). Library for universal virtual reality experiments (luVRe): A standardized immersive 3D/360° picture and video database for VR based research. *Current Psychology*, *42*(7), 5366–5384. https://doi.org/10.1007/s12144-021-01841-1

Schreiner, T., Doeller, C. F., Jensen, O., Rasch, B., & Staudigl, T. (2018). Theta Phase-Coordinated Memory Reactivation Reoccurs in a Slow-Oscillatory Rhythm during NREM Sleep. *Cell Reports*, *25*(2), 296–301. https://doi.org/10.1016/j.celrep.2018.09.037

Schroeder, C. E., & Foxe, J. (2005). Multisensory contributions to low-level, ‘unisensory’ processing. *Current Opinion in Neurobiology*, *15*(4), 454–458. https://doi.org/10.1016/j.conb.2005.06.008

Seger, C. A. (1994). Implicit learning. *Psychological Bulletin*, *115*(2), 163–196. https://doi.org/10.1037/0033-2909.115.2.163

Seinfeld, S., Zhan, M., Poyo-Solanas, M., Barsuola, G., Vaessen, M., Slater, M., Sanchez-Vives, M. V., & de Gelder, B. (2021). Being the victim of virtual abuse changes default mode network responses to emotional expressions. *Cortex*, *135*, 268–284. https://doi.org/10.1016/j.cortex.2020.11.018

Senkowski, D., & Engel, A. K. (2024). Multi-timescale neural dynamics for multisensory integration. *Nature Reviews Neuroscience*, *25*(9), 625–642. https://doi.org/10.1038/s41583-024-00845-7

Senkowski, D., Schneider, T. R., Foxe, J. J., & Engel, A. K. (2008). Crossmodal binding through neural coherence: Implications for multisensory processing. *Trends in Neurosciences*, *31*(8), 401–409. https://doi.org/10.1016/j.tins.2008.05.002

Shahbaba, B., Li, L., Agostinelli, F., Saraf, M., Cooper, K. W., Haghverdian, D., Elias, G. A., Baldi, P., & Fortin, N. J. (2022). Hippocampal ensembles represent sequential relationships among an extended sequence of nonspatial events. *Nature Communications*, *13*(1), 787. https://doi.org/10.1038/s41467-022-28057-6

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, *12*(11), 411–417. https://doi.org/10.1016/j.tics.2008.07.006

Shams, L., Wozny, D. R., Kim, R., & Seitz, A. (2011). Influences of Multisensory Experience on Subsequent Unisensory Processing. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00264

Shin, Y. S., Masís-Obando, R., Keshavarzian, N., Dáve, R., & Norman, K. A. (2021). Context-dependent memory effects in two immersive virtual reality environments: On Mars and underwater. *Psychonomic Bulletin & Review*, *28*(2), 574–582. https://doi.org/10.3758/s13423-020-01835-3

Siapas, A. G., Lubenov, E. V., & Wilson, M. A. (2005). Prefrontal Phase Locking to Hippocampal Theta Oscillations. *Neuron*, *46*(1), 141–151. https://doi.org/10.1016/j.neuron.2005.02.028

Smallwood, J., Bernhardt, B. C., Leech, R., Bzdok, D., Jefferies, E., & Margulies, D. S. (2021). The default mode network in cognition: A topographical perspective. *Nature Reviews Neuroscience*, *22*(8), 503–513. https://doi.org/10.1038/s41583-021-00474-4

Smith, D. M., Yang, Y. Y., Subramanian, D. L., Miller, A. M. P., Bulkin, D. A., & Law, L. M. (2022). The limbic memory circuit and the neural basis of contextual memory. *Neurobiology of Learning and Memory*, *187*, 107557. https://doi.org/10.1016/j.nlm.2021.107557

Smith, S. A., & and Mulligan, N. W. (2021). Immersion, presence, and episodic memory in virtual reality environments. *Memory*, *29*(8), 983–1005. https://doi.org/10.1080/09658211.2021.1953535

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review*, *8*(2), 203–220. https://doi.org/10.3758/BF03196157

Sols, I., DuBrow, S., Davachi, L., & Fuentemilla, L. (2017). Event Boundaries Trigger Rapid Memory Reinstatement of the Prior Events to Promote Their Representation in Long-Term Memory. *Current Biology*, *27*(22), 3499-3504.e4. https://doi.org/10.1016/j.cub.2017.09.057

Song, S., Howard, J. H., & Howard, D. V. (2007). Implicit probabilistic sequence learning is independent of explicit awareness. *Learning & Memory*, *14*(3), 167–176. https://doi.org/10.1101/lm.437407

Staresina, B. P., & Davachi, L. (2009). Mind the Gap: Binding Experiences across Space and Time in the Human Hippocampus. *Neuron*, *63*(2), 267–276. https://doi.org/10.1016/j.neuron.2009.06.024

Staresina, B. P., Henson, R. N. A., Kriegeskorte, N., & Alink, A. (2012). Episodic Reinstatement in the Medial Temporal Lobe. *The Journal of Neuroscience*, *32*(50), 18150–18156. https://doi.org/10.1523/JNEUROSCI.4156-12.2012

Staresina, B. P., Michelmann, S., Bonnefond, M., Jensen, O., Axmacher, N., & Fell, J. (2016). Hippocampal pattern completion is linked to gamma power increases and alpha power decreases during recollection. *eLife*, *5*, e17397. https://doi.org/10.7554/eLife.17397

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

Staudigl, T., & Hanslmayr, S. (2019). Reactivation of neural patterns during memory reinstatement supports encoding specificity. *Cognitive Neuroscience*, *10*(4), 175–185. https://doi.org/10.1080/17588928.2019.1621825

Staudigl, T., Vollmar, C., Noachtar, S., & Hanslmayr, S. (2015). Temporal-Pattern Similarity Analysis Reveals the Beneficial and Detrimental Effects of Context Reinstatement on Human Memory. *The Journal of Neuroscience*, *35*(13), 5373–5384. https://doi.org/10.1523/JNEUROSCI.4198-14.2015

Stern, Y., Katz, R., & Sadeh, T. (2020). Explicit Sequence Memory in Recall of Temporally-structured Episodes. *Scientific Reports*, *10*(1), 2666. https://doi.org/10.1038/s41598-020-59472-8

Straube, B., Wroblewski, A., Jansen, A., & He, Y. (2018). The connectivity signature of co-speech gesture integration: The superior temporal sulcus modulates connectivity between areas related to visual gesture and auditory speech processing. *NeuroImage*, *181*, 539–549. https://doi.org/10.1016/j.neuroimage.2018.07.037

Strube, A., Rose, M., Fazeli, S., & Büchel, C. (2021). Alpha-to-beta- and gamma-band activity reflect predictive coding in affective visual processing. *Scientific Reports*, *11*(1), 23492. https://doi.org/10.1038/s41598-021-02939-z

Strunk, J., & Duarte, A. (2019). Prestimulus and poststimulus oscillatory activity predicts successful episodic encoding for both young and older adults. *Neurobiology of Aging*, *77*, 1–12. https://doi.org/10.1016/j.neurobiolaging.2019.01.005

Su, M., Hu, K., Liu, W., Wu, Y., Wang, T., Cao, C., Sun, B., Zhan, S., & Ye, Z. (2024). Theta Oscillations Support Prefrontal-hippocampal Interactions in Sequential Working Memory. *Neuroscience Bulletin*, *40*(2), 147–156. https://doi.org/10.1007/s12264-023-01134-6

Sun, R., Slusarz, P., & Terry, C. (2005). The Interaction of the Explicit and the Implicit in Skill Learning: A Dual-Process Approach. *Psychological Review*, *112*(1), 159–192. https://doi.org/10.1037/0033-295X.112.1.159

Swallow, K., Broitman, A. W., Riley, E., & Turker, H. B. (2022). Grounding the Attentional Boost Effect in Events and the Efficient Brain. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.892416

Swallow, K., Makovski, T., & Jiang, Y. (2009). The attentional boost effect. *Journal of Vision*, *9*(8), 183. https://doi.org/10.1167/9.8.183

Sweeney-Reed, C. M., Zaehle, T., Voges, J., Schmitt, F. C., Buentjen, L., Kopitzki, K., Richardson-Klavehn, A., Hinrichs, H., Heinze, H.-J., Knight, R. T., & Rugg, M. D. (2016). Pre-stimulus thalamic theta power predicts human memory formation. *NeuroImage*, *138*, 100–108. https://doi.org/10.1016/j.neuroimage.2016.05.042

Tambini, A., Rimmele, U., Phelps, E. A., & Davachi, L. (2017). Emotional brain states carry over and enhance future memory formation. *Nature Neuroscience*, *20*(2), 271–278. https://doi.org/10.1038/nn.4468

Tan, J. H., & Hsieh, P.-J. (2016). Low-level auditory and visual features can be decoded across early sensory cortices. *Journal of Vision*, *16*(12), 581. https://doi.org/10.1167/16.12.581

Tanaka, K. Z., Pevzner, A., Hamidi, A. B., Nakazawa, Y., Graham, J., & Wiltgen, B. J. (2014). Cortical Representations Are Reinstated by the Hippocampus during Memory Retrieval. *Neuron*, *84*(2), 347–354. https://doi.org/10.1016/j.neuron.2014.09.037

Tang, M., Salvatori, T., Millidge, B., Song, Y., Lukasiewicz, T., & Bogacz, R. (2023). Recurrent predictive coding models for associative memory employing covariance learning. *PLOS Computational Biology*, *19*(4), e1010719. https://doi.org/10.1371/journal.pcbi.1010719

Tang, X., Wu, J., & Shen, Y. (2016). The interactions of multisensory integration with endogenous and exogenous attention. *Neuroscience & Biobehavioral Reviews*, *61*, 208–224. https://doi.org/10.1016/j.neubiorev.2015.11.002

ten Oever, S., De Weerd, P., & Sack, A. T. (2020). Phase-dependent amplification of working memory content and performance. *Nature Communications*, *11*(1), 1832. https://doi.org/10.1038/s41467-020-15629-7

Thavabalasingam, S., O’Neil, E. B., Tay, J., Nestor, A., & Lee, A. C. H. (2019). Evidence for the incorporation of temporal duration information in human hippocampal long-term memory sequence representations. *Proceedings of the National Academy of Sciences*, *116*(13), 6407–6414. https://doi.org/10.1073/pnas.1819993116

Thelen, A., & Murray, M. M. (2013). The Efficacy of Single-Trial Multisensory Memories. *Multisensory Research*, *26*(5), 483–502. https://doi.org/10.1163/22134808-00002426

Thelen, A., Talsma, D., & Murray, M. M. (2015). Single-trial multisensory memories affect later auditory and visual object discrimination. *Cognition*, *138*, 148–160. https://doi.org/10.1016/j.cognition.2015.02.003

Thunell, E., Peter, M., Iravani, B., Porada, D. K., Prenner, K., Darki, F., & Lundström, J. N. (2025). Unisensory visual and auditory objects are processed in olfactory cortex, independently of odor association. *Cortex*, *186*, 74–85. https://doi.org/10.1016/j.cortex.2025.04.002

Tompary, A., Duncan, K., & Davachi, L. (2016). High-resolution investigation of memory-specific reinstatement in the hippocampus and perirhinal cortex. *Hippocampus*, *26*(8), 995–1007. https://doi.org/10.1002/hipo.22582

Torres-Morales, C., & Cansino, S. (2024). Brain representations of space and time in episodic memory: A systematic review and meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, *24*(1), 1–18. https://doi.org/10.3758/s13415-023-01140-1

Tovar, D. A., Murray, M. M., & Wallace, M. T. (2020). Selective Enhancement of Object Representations through Multisensory Integration. *The Journal of Neuroscience*, *40*(29), 5604–5615. https://doi.org/10.1523/JNEUROSCI.2139-19.2020

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

Uncapher, M. R., Otten, L. J., & Rugg, M. D. (2006). Episodic Encoding Is More than the Sum of Its Parts: An fMRI Investigation of Multifeatural Contextual Encoding. *Neuron*, *52*(3), 547–556. https://doi.org/10.1016/j.neuron.2006.08.011

Ursino, M., Cesaretti, N., & Pirazzini, G. (2023). A model of working memory for encoding multiple items and ordered sequences exploiting the theta-gamma code. *Cognitive Neurodynamics*, *17*(2), 489–521. https://doi.org/10.1007/s11571-022-09836-9

Ursino, M., Cuppini, C., & Magosso, E. (2014). Neurocomputational approaches to modelling multisensory integration in the brain: A review. *Neural Networks*, *60*, 141–165. https://doi.org/10.1016/j.neunet.2014.08.003

Ursino, M., & Pirazzini, G. (2024). Theta–gamma coupling as a ubiquitous brain mechanism: Implications for memory, attention, dreaming, imagination, and consciousness. *Current Opinion in Behavioral Sciences*, *59*, 101433. https://doi.org/10.1016/j.cobeha.2024.101433

Van De Ven, V., Kleuters, G., & Stuiver, J. (2021). *Multisensory integration of contextual boundaries affects temporal order memory, but not encoding or recognition*. https://doi.org/10.31234/osf.io/xnqh9

van Ede, F. (2018). Mnemonic and attentional roles for states of attenuated alpha oscillations in perceptual working memory: A review. *European Journal of Neuroscience*, *48*(7), 2509–2515. https://doi.org/10.1111/ejn.13759

van Atteveldt, N., Murray, M. M., Thut, G., & Schroeder, C. E. (2014). Multisensory Integration: Flexible Use of General Operations. *Neuron*, *81*(6), 1240–1253. https://doi.org/10.1016/j.neuron.2014.02.044

Vandierendonck, A. (2016). Modality independence of order coding in working memory: Evidence from cross-modal order interference at recall. *Quarterly Journal of Experimental Psychology*, *69*(1), 161–179. https://doi.org/10.1080/17470218.2015.1032987

Vaz, A. P., Wittig, J. H., Inati, S. K., & Zaghloul, K. A. (2020). Replay of cortical spiking sequences during human memory retrieval. *Science*, *367*(6482), 1131–1134. https://doi.org/10.1126/science.aba0672

Vedder, A., Smigielski, L., Gutyrchik, E., Bao, Y., Blautzik, J., Pöppel, E., Zaytseva, Y., & Russell, E. (2015). Neurofunctional Correlates of Environmental Cognition: An fMRI Study with Images from Episodic Memory. *PLOS ONE*, *10*(4), e0122470. https://doi.org/10.1371/journal.pone.0122470

Ventura, S., Brivio, E., Riva, G., & Baños, R. M. (2019). Immersive Versus Non-immersive Experience: Exploring the Feasibility of Memory Assessment Through 360° Technology. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02509

Virk, T., Letendre, T., & Pathman, T. (2024). The convergence of naturalistic paradigms and cognitive neuroscience methods to investigate memory and its development. *Neuropsychologia*, *196*, 108779. https://doi.org/10.1016/j.neuropsychologia.2023.108779

Wais, P. E., Arioli, M., Anguera-Singla, R., & Gazzaley, A. (2021). Virtual reality video game improves high-fidelity memory in older adults. *Scientific Reports*, *11*(1), 2552. https://doi.org/10.1038/s41598-021-82109-3

Waldhauser, G. T., Braun, V., & Hanslmayr, S. (2016). Episodic Memory Retrieval Functionally Relies on Very Rapid Reactivation of Sensory Information. *The Journal of Neuroscience*, *36*(1), 251–260. https://doi.org/10.1523/JNEUROSCI.2101-15.2016

Waldhauser, G. T., Johansson, M., & Hanslmayr, S. (2012). Alpha/Beta Oscillations Indicate Inhibition of Interfering Visual Memories. *The Journal of Neuroscience*, *32*(6), 1953–1961. https://doi.org/10.1523/JNEUROSCI.4201-11.2012

Wang, C., Rajagovindan, R., Han, S.-M., & Ding, M. (2016). Top-Down Control of Visual Alpha Oscillations: Sources of Control Signals and Their Mechanisms of Action. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00015

Wang, D. (2024). Rhythmic sensory stimulation as a noninvasive tool to study plasticity mechanisms in human episodic memory. *Current Opinion in Behavioral Sciences*.

Wang, H., Xu, X., Xu, X., Gao, J., & Zhang, T. (2020). Enriched Environment and Social Isolation Affect Cognition Ability via Altering Excitatory and Inhibitory Synaptic Density in Mice Hippocampus. *Neurochemical Research*, *45*(10), 2417–2432. https://doi.org/10.1007/s11064-020-03102-2

Warren, C. M., Murphy, P. R., & Nieuwenhuis, S. (2016). Cognitive control, dynamic salience, and the imperative toward computational accounts of neuromodulatory function. *The Behavioral and Brain Sciences*, *39*, e227. https://doi.org/10.1017/S0140525X15001983

Waskom, M. L., Okazawa, G., & Kiani, R. (2019). Designing and Interpreting Psychophysical Investigations of Cognition. *Neuron*, *104*(1), 100–112. https://doi.org/10.1016/j.neuron.2019.09.016

Weinberger, A. B., & Green, A. E. (2022). Dynamic development of intuitions and explicit knowledge during implicit learning. *Cognition*, *222*, 105008. https://doi.org/10.1016/j.cognition.2021.105008

Wen, T., Mitchell, D. J., & Duncan, J. (2020). The Functional Convergence and Heterogeneity of Social, Episodic, and Self-Referential Thought in the Default Mode Network. *Cerebral Cortex*, *30*(11), 5915–5929. https://doi.org/10.1093/cercor/bhaa166

Wessel, J. R., Haider, H., & Rose, M. (2012). The transition from implicit to explicit representations in incidental learning situations: More evidence from high-frequency EEG coupling. *Experimental Brain Research*, *217*(1), 153–162. https://doi.org/10.1007/s00221-011-2982-7

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*(3), 331–354. https://doi.org/10.1037/0033-2909.121.3.331

Wierzchoń, M., & Derda, M. (2019). Implicit learning under attentional load. In *Implicit Learning*. Routledge.

Williams, J. N. (2005). LEARNING WITHOUT AWARENESS. *Studies in Second Language Acquisition*, *27*(02). https://doi.org/10.1017/S0272263105050138

Willingham, D. B., & Goedert-Eschmann, K. (1999). The Relation Between Implicit and Explicit Learning: Evidence for Parallel Development. *Psychological Science*, *10*(6), 531–534. https://doi.org/10.1111/1467-9280.00201

Wilson, J. H., & Criss, A. H. (2017). The list strength effect in cued recall. *Journal of Memory and Language*, *95*, 78–88. https://doi.org/10.1016/j.jml.2017.01.006

Wing, E. A., Ritchey, M., & Cabeza, R. (2015). Reinstatement of Individual Past Events Revealed by the Similarity of Distributed Activation Patterns during Encoding and Retrieval. *Journal of Cognitive Neuroscience*, *27*(4), 679–691. https://doi.org/10.1162/jocn_a_00740

Wong, C. W., Olafsson, V., Plank, M., Snider, J., Halgren, E., Poizner, H., & Liu, T. T. (2014). Resting-State fMRI Activity Predicts Unsupervised Learning and Memory in an Immersive Virtual Reality Environment. *PLOS ONE*, *9*(10), e109622. https://doi.org/10.1371/journal.pone.0109622

Wu, S., Éltető, N., Dasgupta, I., & Schulz, E. (2023). Chunking as a rational solution to the speed–accuracy trade-off in a serial reaction time task. *Scientific Reports*, *13*(1), 7680. https://doi.org/10.1038/s41598-023-31500-3

Xiao, X., Dong, Q., Gao, J., Men, W., Poldrack, R. A., & Xue, G. (2017). Transformed Neural Pattern Reinstatement during Episodic Memory Retrieval. *The Journal of Neuroscience*, *37*(11), 2986–2998. https://doi.org/10.1523/JNEUROSCI.2324-16.2017

Xu, J., & Zhong, B. (2018). Review on portable EEG technology in educational research. *Computers in Human Behavior*, *81*, 340–349. https://doi.org/10.1016/j.chb.2017.12.037

Xu, X., Jiang, M., Yin, H., Wang, G., Colzato, L., Zhang, W., & Hommel, B. (2024). The impact of stimulus format and presentation order on social working memory updating. *Social Cognitive and Affective Neuroscience*, *19*(1), nsae067. https://doi.org/10.1093/scan/nsae067

Xue, G. (2022). From remembering to reconstruction: The transformative neural representation of episodic memory. *Progress in Neurobiology*, *219*, 102351. https://doi.org/10.1016/j.pneurobio.2022.102351

Yadav, N., Noble, C., Niemeyer, J. E., Terceros, A., Victor, J., Liston, C., & Rajasethupathy, P. (2022). Prefrontal feature representations drive memory recall. *Nature*, *608*(7921), 153–160. https://doi.org/10.1038/s41586-022-04936-2

Yaffe, R. B., Kerr, M. S. D., Damera, S., Sarma, S. V., Inati, S. K., & Zaghloul, K. A. (2014). Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. *Proceedings of the National Academy of Sciences*, *111*(52), 18727–18732. https://doi.org/10.1073/pnas.1417017112

Yaffe, R. B., Shaikhouni, A., Arai, J., Inati, S. K., & Zaghloul, K. A. (2017). Cued Memory Retrieval Exhibits Reinstatement of High Gamma Power on a Faster Timescale in the Left Temporal Lobe and Prefrontal Cortex. *Journal of Neuroscience*, *37*(17), 4472–4480. https://doi.org/10.1523/JNEUROSCI.3810-16.2017

Yang, H.-L., Chan, P.-T., Chang, P.-C., Chiu, H.-L., Sheen Hsiao, S.-T., Chu, H., & Chou, K.-R. (2018). Memory-focused interventions for people with cognitive disorders: A systematic review and meta-analysis of randomized controlled studies. *International Journal of Nursing Studies*, *78*, 44–51. https://doi.org/10.1016/j.ijnurstu.2017.08.005

Yang, J., & Li, P. (2012). Brain Networks of Explicit and Implicit Learning. *PLoS ONE*, *7*(8), e42993. https://doi.org/10.1371/journal.pone.0042993

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

Yoo, H. B., Umbach, G., & Lega, B. (2021). Neurons in the human medial temporal lobe track multiple temporal contexts during episodic memory processing. *NeuroImage*, *245*, 118689. https://doi.org/10.1016/j.neuroimage.2021.118689

Young, D., Lawlor, P. A., Leone, P., Dragunow, M., & During, M. J. (1999). Environmental enrichment inhibits spontaneous apoptosis, prevents seizures and is neuroprotective. *Nature Medicine*, *5*(4), 448–453. https://doi.org/10.1038/7449

Yu, H., Wang*, A., Li, Q., Liu, Y., Yang, J., Takahashi, S., Ejima, Y., Zhang, M., & Wu, J. (2021). Semantically Congruent Bimodal Presentation with Divided-Modality Attention Accelerates Unisensory Working Memory Retrieval. *Perception*, *50*(11), 917–932. https://doi.org/10.1177/03010066211052943

Zaehle, T., Rach, S., & Herrmann, C. S. (2010). Transcranial Alternating Current Stimulation Enhances Individual Alpha Activity in Human EEG. *PLOS ONE*, *5*(11), e13766. https://doi.org/10.1371/journal.pone.0013766

Zareian, B., Maboudi, K., Daliri, M. R., Abrishami Moghaddam, H., Treue, S., & Esghaei, M. (2020). Attention strengthens across-trial pre-stimulus phase coherence in visual cortex, enhancing stimulus processing. *Scientific Reports*, *10*(1), 4837. https://doi.org/10.1038/s41598-020-61359-7

Zazio, A., Ruhnau, P., Weisz, N., & Wutz, A. (2022). Pre-stimulus alpha-band power and phase fluctuations originate from different neural sources and exert distinct impact on stimulus-evoked responses. *European Journal of Neuroscience*, *55*(11–12), 3178–3190. https://doi.org/10.1111/ejn.15138

Zhang, W., Van Ast, V. A., Klumpers, F., Roelofs, K., & Hermans, E. J. (2018). Memory Contextualization: The Role of Prefrontal Cortex in Functional Integration across Item and Context Representational Regions. *Journal of Cognitive Neuroscience*, *30*(4), 579–593. https://doi.org/10.1162/jocn_a_01218

Zhao, J., Al-Aidroos, N., & Turk-Browne, N. B. (2013). Attention Is Spontaneously Biased Toward Regularities. *Psychological Science*, *24*(5), 667–677. https://doi.org/10.1177/0956797612460407

Zhou, C. Y., Talmi, D., Daw, N. D., & Mattar, M. G. (2025). Episodic retrieval for model-based evaluation in sequential decision tasks. *Psychological Review*, *132*(1), 18–49. https://doi.org/10.1037/rev0000505

Zhu, Y., Zhang, C., Poikonen, H., Toiviainen, P., Huotilainen, M., Mathiak, K., Ristaniemi, T., & Cong, F. (2020). Exploring Frequency-Dependent Brain Networks from Ongoing EEG Using Spatial ICA During Music Listening. *Brain Topography*, *33*(3), 289–302. https://doi.org/10.1007/s10548-020-00758-5

Appendix A: Study I

Enriched environments enhance the development of explicit memory in an incidental learning task. Jaap*, C., Maack*, M. C., Taesler, P., Steinicke, F., & Rose, M. (2022). *Scientific Reports*, *12*(1), 18717.

*Equal contributions

scientific reports

Enriched environments enhance the development of explicit memory in an incidental learning task

Carina Jaap[1,3], Marike C. Maack[1,3], Philipp Taesler[1], Frank Steinicke[2] & Michael Rose[1]✉

Learning, rendered in an implicit (unconscious) or explicit (conscious) way, is a crucial part of our daily life. Different factors, like attention or motivation, influence the transformation from implicit to explicit memory. Via virtual reality a lively and engaging surrounding can be created, whereby motivational processes are assumed to be a vital part of the transition from implicit to explicit memory. In the present study, we tested the impact of an enriched virtual reality compared to two conventional, non-enriched 2D-computer-screen based tasks on implicit to explicit memory transformation, using an audio-visual sequential association task. We hypothesized, that the immersive nature of the VR surrounding enhances the transfer from implicit to explicit memory. Notably, the overall amount of learned sequence pairs were not significantly different between experimental groups, but the degree of awareness was affected by the different settings. However, we observed an increased level of explicitly remembered pairs within the VR group compared to two screen-based groups. This finding clearly demonstrates that a near-natural experimental setting affects the transformation process from implicit to explicit memory.

Abbreviations

ITI	Intertrial interval
SEM	Standard error of the mean
VR	Virtual reality

Learning is influenced by multiple factors, like attention, and motivation, and is rendered in an implicit (unconscious) or explicit (conscious) way[1,2]. Implicit memory is usually encoded incidentally so that regularities in the environment are extracted without the actual intention to learn them, and thereby influence our behavior without awareness. Evidence suggests, that implicitly acquired knowledge can become explicit, allowing us to extract and use regularities from the environment without having learned them consciously[3–5]. This memory transformation, from the implicit to the explicit domain, represents a crucial mechanism as learning such regularities and complex rules facilitates the development of higher cognitive functions such as reasoning and language[6]. This way, implicit learning forms the basis to adapt to a complex and changing world and guide the decision making in our daily life.

The fluency hypothesis.

Whether implicit information is transformed to become explicit and thereby conscious, depends on a multitude of factors[6–8]. One prominent theory to explain the transformation from incidental learning to explicit memory is based on the ongoing evaluation of predictions[6,9]. In particular, implicit perceptual associations result in predictions for upcoming events and these predictions are assumed to be used to monitor the outcome of consecutive processing. To study this effect, a sequential task is often implemented[6,10]. During the presentation of a predictable sequence of stimuli, a discrepancy between the predicted and the actual processing speed can be detected by the participants. Hence, this detection process can trigger attention towards the cause of this discrepancy and due to the consecutive search processes, explicit memory is generated[4,11]. This is in line with the Unexpected-Event theory[11], which postulates that any metacognitive judgment, like unex-

[1]NeuroImage Nord, Department for Systems Neuroscience, University Medical Center Hamburg Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany. [2]Human-Computer Interaction, Department of Informatics, University of Hamburg, Vogt-Kölln-Str. 30, 22527 Hamburg, Germany. [3]These authors contributed equally: Carina Jaap and Marike C. Maack. ✉email: rose@uke.uni-hamburg.de

pected fluency or accuracy, can be an unexpected event and trigger attributive processes. It can be assumed that this fluency of processing is enhanced in computer-screen-related tasks because, in more near-natural settings, the evaluation of stimuli and responses are slowed down by several factors. For example, in VR applications, the participants can explore the virtual environment to different as well as variable amounts of time. Additionally, the use of a controller as a response device slows down the response in contrast to classical response devices. These factors directly reduce the fluency of stimulus processing and should therefore affect the emergence of explicit memory.

Following this, our implemented 2D-computer-screen application probably creates more fluency of the responses, related to different factors, like smoother transition of trials compared to trials within the VR task design and therefore enable a rhythmical stimulus processing. However, the fluency in the computer-screen application is favored by the time needed for stimulus evaluation and response times due to different response devices. Following these assumptions, we hypothesized to find an enhanced emergence of explicit memory in a conventional 2D-computer-screen based environment in case that the fluency of stimulus processing is a crucial part of implicit to explicit memory formation. In the following we call this assumption, the *fluency hypothesis*. The contrary hypothesis, the *enriched environmental hypothesis*, can be formulated based on theories regarding learning within a more near-natural environment, assuming that rather realistic stimuli and context directly enhance explicit learning processes[12,13].

The enriched environmental hypothesis. The current cognitive research aspires to apply experimental designs in real-world settings (i.e. audio-visual processing while walking or performing daily activities[14,15]), as it was demonstrated that the context of learning (and retrieval) processes are highly relevant for their outcome[8,12,13]. Despite their advantages, these experimental settings are, however, extremely vulnerable to uncontrollable variables. To control external variables and provide a realistic setting, VR environments have become an attractive option. Previous applications demonstrated that these VR environments enable a sufficient, near-natural experience using interactive elements and multisensory stimulation, resulting in multiple levels of excitement and engagement[16,17], which support an improved multi-sensory integration[18]. Moreover, VR facilitates a more salient processing of stimuli by dynamic engagement of the sensorimotor system, which provokes more naturalistic behavioral and physiological responses than abstract stimuli[19,20]. It was previously shown that learning in VR promotes better performance in an enriched environment[21,22] and the literature demonstrated that these enriched stimuli are responsible for the motivational significance of stimuli[1,23,24]. This motivational effect can also enhance the detection of violated predictions due to an increased level of attention instead of performance fluency. Notably, compared to desktop PC settings, VR Head-Mounted Displays like the HTC Vive induce greater feelings of being present in the VR experimental surrounding, and higher motivation to interact with the environment[16,25]. So far, VR has already been broadly used in the research of explicit episodic and spatial learning (for a review see[22]). Here it has been shown that in contrast to basic computer setups, episodic memory performance is increased in VR settings[26,27], which is most likely caused by the near-natural or more specifically immersive VR environments. Therefore, the benefit of using VR studies for explicit memory supports our assumption that VR can enhance the transformation from implicit to explicit memory. It is an open question whether the emergence of explicit memory during incidental learning can also benefit from a more near-natural experimental setting or if this phenomenon is exclusive to task designs, in which the volunteers we instructed to memorize the stimuli. We expect, that the detection of unexpected events within ongoing prediction evaluations, needed for the transfer from implicit to explicit knowledge, benefits from the near-natural experimental application in VR reflected in increased explicit memory performance. This hypothesis, the *enriched environmental hypothesis*, is based on the existing evidence from the implicit learning domain and potential effects within a VR experience (i.e., rising motivational and attentional as well as engaging processes).

The present study. In the present study, we employed a sequential-association task, which has been introduced in previous studies[6,28,29]. The task was used as a between-subjects design contrasting a conventional 2D-computer-based presentation with an enriched near-natural VR application. As the evaluation of stimuli and responses are probably slowed in the near-natural VR task, we tested our design within two different 2D-computer-screen groups, differing only in the number of trials, and one VR group. One 2D-computer-screen group practiced the identical amount of trials as presented in the VR condition, which operated as control group for the duration of VR condition and to exclude an influence of experimental length. Previous studies already demonstrated this smaller amount of trial in a 2D-computer-screen application lead to a partly generation of explicit memory in an incidental learning task[4,6]. Both 2D-computer-screen condition were control conditions for our fluency hypothesis.

The use of stimuli from different modalities (crossmodal) has been shown to be beneficial in learning paradigms because memory formation seems to benefit if encoded elements are derived from separate modalities, facilitating their integration (i.e., visual, and auditory[30–32]). The crossmodal stimuli implemented in the conventional 2D-computer-screen based task were modified from a previous study and consisted of simple tones and plain squares as visual stimuli[6] (see Condition 2 for further details). The sequential regularities consisted of the presentation of alternating visual and auditory stimuli, building a fixed eight-digit sequence in 85% of the trials. Only in 15% of all trials this sequence was violated. The sequential task structure was unknown to the participants, and the content could only be learned incidentally.

Importantly, to assess the degree of explicit memory for the embedded sequences, an identical completion task[33] and a free recall test were conducted at the end of all experimental conditions and always outside the VR. Thus, the results from these tasks can be directly compared between the different learning settings. Both post-tests were combined with a confidence rating[6] to identify participants' explicit knowledge[5,8,34,35]. Previous

Condition	N taking part in total	N of discarded datasets	N of datasets in the final analysis
VR	51	2	49
PC-short	22	1	21
PC-long	29	2	27

Table 1. Overview of number of participants taking part in each condition before and after discarding relevant datasets.

applications revealed that participants with explicit memory express their knowledge with high confidence. The correct responses under high confidence are an indicator of explicit memory and, hence, will be used to differentiate them from implicit memory[36,37]. To test for probable differences in memory formation, based on the latter mentioned evaluations of the implemented stimuli in VR and the 2D screen conditions, a stimulus value rating was performed as a last step of the post-experimental assessment.

Hypothesis. We hypothesized that both 2D-computer-screen groups perform similar, as it can be assumed that a *fluency* based unexpected event is detected in both variations of the experiments. Our modified version of the task for the VR contained images of different naturalistic landscapes and complex instrumental sounds. Furthermore, the enriched aspect of the VR was achieved by putting the participants directly into the naturalistic landscapes while performing the sequential task. Furthermore, we hypothesized that an enhanced explicit memory can be explained by two potential mechanisms. On the one hand, if *fluency* is an important factor in the transfer of implicit to explicit memory, we expected to find a boosted performance in explicit memory formation in the 2D-computer-screen application compared to the group performing the task in the VR condition. On the other hand, if the performance of explicit memory is greater within the VR condition, motivational and enriched environmental-based attention processes play a crucial role in naturalistic learning scenarios.

Study aim. The aim of the present study was a direct comparison of both the *fluency* and *enriched environmental hypothesis* regarding the mechanism of incidental perceptual learning processes concerning the transfer from implicit to explicit memory. Therefore, we have contrasted a simple 2D-computer-screen experiment and a complex and enriched VR experiment. The influence of both hypotheses can be compared between the different experimental surroundings by assessing the degree of implicit and explicit memory after learning within two tasks that were identical for all experimental conditions.

Material and methods

We tested a sequential association task within a VR environment as well as in two conventional 2D-computer-screen-based tasks. The conventional 2D-computer-screen based conditions will be called PC-short (260 trials) or condition 2 and PC-long (520 trials) or condition 3 in the following.

Participants. 102 (51 in cond. 1; 22 in cond. 2 and 29 in cond. 3) healthy participants with normal hearing and normal or corrected-to-normal vision took part in this study. Data of five participants had to be discarded as the participants did not complete the tutorial successfully (two in cond. 1, one in cond. 2 and two in cond. 3).

Data of 49 participants were part of the final analysis of the VR condition (29 Females, age M = 27.07 years). In the 2nd condition (PC-short), the datasets of 21 (11 Females, age M = 24.7 years) participants were included in the final analysis. In the 3rd condition (PC-long), the datasets of 27 (21 Females, age M = 27.6 years) participants were included in the final analysis (see Table 1). All experiment protocols were approved by the local Ethics Committee of the General Medical Council Hamburg (PV7022) and our methods were carried out in accordance with ethical guidelines and regulations. Before taking part in the experiments, all participants gave their written informed consent and were paid an expense allowance of 10 €/h.

Condition 1 (VR). *Apparatus.* Inside the VR. The visual stimuli were presented in a virtual surrounding, built with the Unity 3.0 engine, via a head-mounted display i.e., HTC Vive. The responses were tracked with a HTC Vive controller. The acoustic stimuli were presented via headphones. The volume was adjusted by the participants to a comfortable level. The tutorial, as well as the main experiment, took place within the VR environment.

Outside the VR. After the main experiment, a completion task, a free recall, and a stimulus value rating were performed on a computer screen (23″, ~ 1 m distance to the participant) using a standard computer mouse. The tones were presented via two loudspeakers (HD 201, Sennheiser, Germany) one on each side of the screen. The volume of the acoustic stimuli was adjusted by the participants to a comfortable level during the before mentioned post-tests.

Stimuli. The visual stimuli in the VR consisted of four different landscapes (ocean, desert, ice, and forest). The participants were teleported into a landscape and were able to get a 360° view of the landscape. The landscapes were designed with the Unity engine. So, it was a fully immersive experience and not just a 360° image of the landscapes. The acoustic stimuli consisted of four complex instrumental tones, with a duration of 1000 ms each.

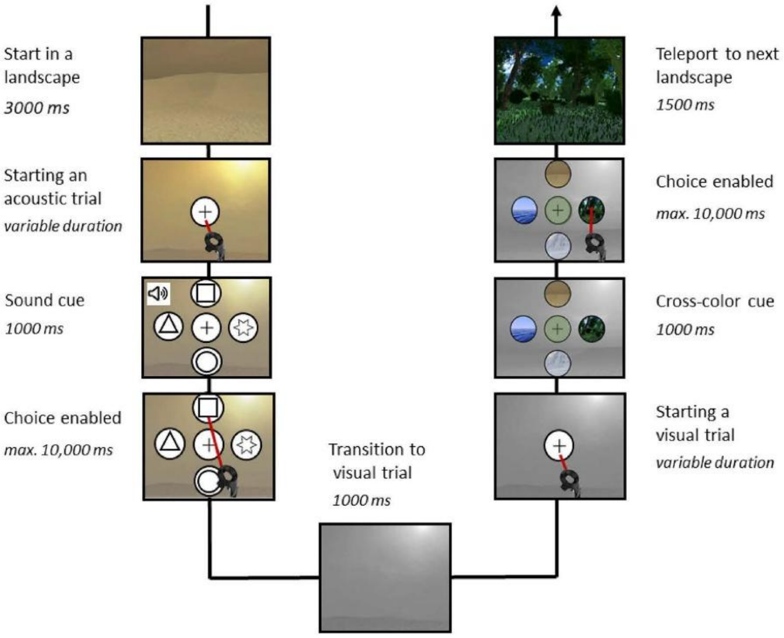


Figure 1. Timeline of a trial within the VR experiment (from upper left to the upper right). The participants were teleported to a virtual landscape where they had 3000 ms time for exploration. Afterwards, a white circle with a cross in the middle appeared in front of them, symbolizing the option to start a trial. The participants could autonomously start by pointing at the circle with the beam of their controller. Next, a sound was played for 1000 ms and four symbols occurred. The sound-symbol combinations were previously learned in the tutorial. Importantly, during the sound was played, participants could not respond. After the sound had ended, participants could choose one of the four symbols within 10,000 ms. After selecting the corresponding symbol, there was a transition from the acoustic to visual trial within 1000 ms during which the color of the current surrounding was desaturated. Afterwards, the visual trial started in the same way as the acoustic trial. Again, the participants could autonomously start the visual trial by pointing at the circle with the beam of their controller. Upon start, four visual stimuli symbolizing four landscapes appeared. Then the target circle in the middle changed from white to the dominant color of one of the four predefined landscapes within 1000 ms. During the color changing process, the participants could not give an answer. Next, participants had to select the icon representing the color of the target (i.e., if the target turned green, participants were expected to select the forest icon within 10,000 ms). After selecting a landscape icon, the participants were teleported to the next correct landscape, independent of their choice (i.e., if the participant chose the ocean landscape although the target was green, still the forest landscape was presented).

Each sound was paired with a symbol (e.g. the piano sound was resembled as a star). By matching each sound with a specific symbol, the participants could match the played sound with the respective symbol within the experiment (see Fig. 1).

Experimental design and procedure. Before entering the main experiment, each participant was introduced to the stimuli and the task instructions for the main experiment during a tutorial. First, the participants had to learn the correct combination of tones and their corresponding symbol in a familiarization task. For the visual trials, the participants learned to match a color with one of the four landscape icons within a familiarization task. Each color was chosen in accordance to the dominant color of the corresponding landscape (desert = yellow). The tutorial ended when less than 2 errors over the last 10 trials were generated by the participants for each trial type. In the main experiment the participants were instructed to answer as fast and precise as possible in each visual and acoustic trial. The VR condition consisted of 260 teleports which is the equivalent of 520 (50% acoustic and 50% visual) trials. To avoid motion sickness, the participants were teleported slowly into the next scenery. The participants were offered several breaks during the VR experiment.

For a detailed overview of the timeline of trials within a teleport, representing a set of an acoustic and visual stimulus presentation in the main experiment, see Fig. 1.

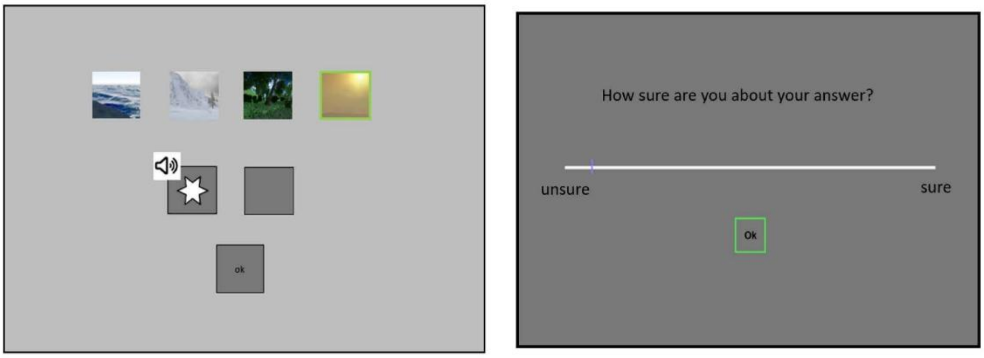


Figure 2. A completion task trial (left) and confidence rating (right) performed outside the VR. In the completion task, four stimuli were displayed above a given stimulus. The participants then had to match the correct visual or acoustic stimulus to the given one. Within the shown crossmodal completion task trial, the participants were asked to match the given acoustic stimulus with the surrounding they associate with it. The participants were asked to make a guess if necessary. After each trial in the completion task (as well as the free recall) the participants had to perform a confidence rating (left).

As we were interested in the learning behavior within the VR, the acoustic and visual trial presentations were part of a sequence consisting of eight digits formed by pairs of eight different stimuli. Within the sequence, each visual stimulus (one of four landscapes) was paired with a specific sound (one of four musical instruments) e.g. visual 1: Desert with acoustic 1: Piano sound. Therefore, pairs or even the whole sequence could be learned in principle. The sequence was interrupted by deviants (15%) with a maximum of three in a row. The sequential regularity of the trial presentations was not introduced to the participants. To avoid motor learning, the order of the visual response option was randomized in each trial[33].

Post-experimental assessment of knowledge and stimulus value rating outside the VR. A completion task (see Fig. 2), a free recall, and a stimulus rating followed the main VR experiment. All post-experimental tasks were retrospectively performed outside the VR on a computer screen. In each trial, one of eight stimuli was given and had to be completed with a stimulus that matched the given stimulus (see Fig. 2). The trials were either crossmodal, with a given stimulus in one modality and four choices given from the opposite modality, or unimodal in which the given stimulus and the choices of answers were of the same modality. The completion task consisted of 64 trials with 50% crossmodal trials (25% visual and 25% acoustic matching). After each trial, the participants had to choose if they were sure or unsure about the given answer (see Fig. 2). With this rating, we later could separate the given answers into implicit (correct answer rated as unsure) and explicit (correct answer rated as sure) knowledge about the presented sequence.

The completion task was followed by a free recall. At this point, the participants were told that there was an order in which the stimuli were presented most of the time during the experiment. Within the free recall, the participants were asked to choose an order for the eight given stimuli. The chosen order should resemble a sequence the participants most likely were presented within the main experiment. The order for the chosen stimuli was not limited. After the participants logged their choice in, by clicking on an "ok" button, the participants were asked if they were sure or unsure about their chosen order.

At the end of the latter mentioned memory assessment tasks, the participants performed a stimulus value rating of the visual stimuli. We implemented a stimulus value rating to test for potential motivational significance of stimuli between the stimuli used in the visually enriched VR and the non-enriched screen based experiments, as this contextual influence can probably correlate with learning effects due to motivational and attentional factors[23]. Each visual stimulus was presented once and the participants could rate it on a continuous scale with "I dislike it" (negative rating) on the left, "Neutral" in the middle and "I like it" (positive rating) on the right side of the scale.

Condition 2 (PC-short). *Apparatus.* The visual stimuli were presented on a 23″ screen (SyncMaster P2370; Samsung). The distance between screen and participant was approximately 1 m. The tones were presented via two loudspeakers (HD 201, Sennheiser, Germany) one on each side of the screen. The volume was adjusted by the participants. For recording the answer of the participant, we used a standard keyboard and computer mouse. The psychtoolbox on Matlab was used to present the experiment.

The visual stimuli consisted of four colored and easily distinguishable squares. The color of each square was chosen analogue to a landscape within the VR condition (blue = ocean, green = forest, yellow = desert, white = icy landscape). Four simple tones (sine waves: 120 Hz, 286 Hz, 389 Hz and 527 Hz) with a duration of 1000 ms, were used as auditory stimuli[38]. Black circles of diameters ranging from 20 to 80% of the size of the visual stimulus were displayed as a visual response option for the acoustic stimuli. The circle size represented the frequency height e.g., the biggest circle represented the tone with the lowest frequency.

Experimental design and procedure. The participants were instructed to respond as quickly and correctly as possible to the target stimulus which was presented in the centre of the screen for the visual stimuli or as a tone to which the participants had to match one of four circles. The participants underwent a training before entering the main experiment (see condition 1). The response options were displayed above the target. The last visual target was still present during an acoustic trial to keep it analogue to the setting in the VR experiment in which the participants remained in a landscape during the acoustic trial (see Fig. 1). After the participants gave their response, the trial ended. Answers had to be given within 2500 ms per trial.

Responses had to be made with the index and middle finger of both hands on a regular keyboard. The enabled keys were "y", "x", ", ; " ,and ". : ".

Condition 2 or PC-short consisted of 260 trials. The trials were part of a sequence consisting of eight stimuli, four in each condition starting with a visual stimulus (Sequence: V1 (e.g. blue square) A1 (e.g. 286 Hz), V2A2 V3A3 V4A4; for more details, see Condition 1).

Post-experimental assessment of knowledge and stimulus value rating. A completion task, a free recall and a stimulus rating followed the main experiment. All post-experimental tasks were performed on a computer screen and were identical, with an exception for the used visual and acoustic stimuli, to the tasks performed by the experimental group 1 (see cond. 1, VR).

Condition 3 (PC-long).

Apparatus. See cond. 2.

Stimuli. See cond. 2.

Experimental design and procedure. The procedure was the same as in condition 2 except that condition 3 consisted of 520 instead of 260 trials. See cond. 2 for further details.

Post-experimental assessment of knowledge and stimulus value rating. See cond. 1 and cond. 2.

Behavioral data analysis.

The important parameters, which can be compared between all experimental conditions, are the amount of memory expressed as implicit or explicit memory. These parameters can be taken from the completion task, as well as the free recall. The latter tasks were identical for all three conditions, except for a change in both visual and acoustic stimuli in the VR condition. All correct answers were taken into account for the analysis of task performance in the completion task, as well as the free recall. Within the main experimental conditions, specific acoustic and visual stimuli formed pairs within a sequence in 85% of the trials. We counted an answer as correct if the participant was able to match a given stimulus with a stimulus of the other modality that was either the following or the previous stimulus within the sequence. As the last visual stimulus was present, when the acoustic stimulus was presented and vice versa, we assumed, that not only forward but also backward learning within the sequence was feasible. Therefore, we accepted an answer in both directions. The amount of implicit memory was calculated as the percentage of "unsure", correct, answers from the sum of all possible answers per participant. The amount of explicit memory was calculated as the percentage of "sure", correct, answers from the sum of all possible trials per participant. This method was used for the assessment of the type of acquired knowledge (implicit; explicit) in the different conditions in both the completion task as well as in the free recall.

To test for differences in the amount as well as quality of gained knowledge, we performed an ANOVA with the factors Condition (VR; PC-short; PC-long) and Learning-Type (implicit; explicit) with performance in each learning type as the dependent variable within the completion task as well as the free recall. Furthermore, we tested for probable differences in the quality of gained knowledge between the two conventional 2D-screen-based tasks and performed an ANOVA with the factors PC-Conditions (PC-short; PC-long) and Learning-Type (implicit; explicit) with performance in each learning type as the dependent variable within the completion task as well as the free recall. For the analysis of the stimulus value rating, all given answers were taken into account. We controlled for potential outliers, i.e. participants selecting only "sure" and while showing constant errors, before we went on with the further analysis. For each rating, where a value between 0 (unpleasant) and 1 (pleasant) was possible, the absolute distance to the neutral rating (0.5) was calculated. A mean over these adjusted stimulus ratings of the four visual stimuli was calculated for each participant in each condition and used for the further analysis. To test for a probable impact of the enriched stimuli, used in the VR, on the stimulus value rating, we performed a two-sided two-sample t-test over stimulus ratings in VR versus PC (PC-short & PC-long merged) conditions. Furthermore, we tested for probable correlations of stimulus value ratings on implicitly and explicitly gained knowledge in each group. The statistical analysis was performed in R (4.0.5) and Matlab (2020b).

Results

Completion task.

Notably, we found no difference in the amounts of overall learned sequence pairs between conditions (see Table 2), but the degree of developed explicit memory was affected by the different settings in the three conditions.

The interaction effect of Conditions and Learning-Type ($F_{(2188)} = 27.3$, $p < 0.0001$; see Fig. 3) revealed more explicit memory in the VR group as compared to both PC versions. Overall, volunteers acquired more implicit than explicit memory (main effect of Learning-Type: $F_{(1188)} = 23.1$, $p < 0.0001$).

Type	Conditions					
	VR [%]	SEM [%]	PC-short [%]	SEM [%]	PC-long [%]	SEM [%]
Implicit	19.2	3.15	45.8	5.44	52	2.5
Explicit	36.7	4.2	14.4	4.4	4.62	1.5
Total learned	56	3.12	60.3	3.52	56.6	2.4

Table 2. Completion task performance over learning types in each condition. Mean and SEM in percent for the performance between and within conditions (VR = 49, PC-short = 21, PC-long = 27) for Learning-Type (implicit; explicit) and over all learned stimulus pairs.

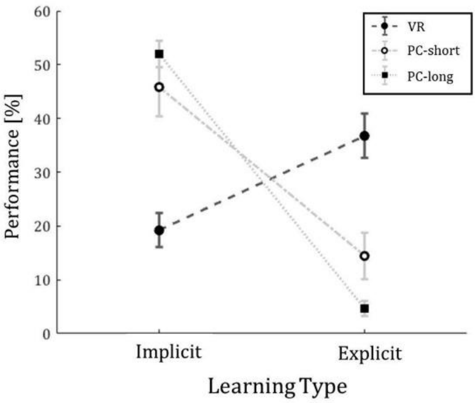


Figure 3. Completion task performance of learned pairs within the given sequence divided by Learning-Type (implicit; explicit) and Condition (VR, PC-short, PC-long). The light grey, dotted lines represent the performance of participants of PC-short (N = 21) and PC-long (N = 27) and the dark grey, dashed line represents the performance of VR (N = 49) participants in the completion task. The performance is divided into implicitly learned sequence pairs (left) and explicitly learned sequence pairs (right). The mean of the performance is visualized as black circle (VR), hollow circle (PC-short) and black square (PC-long). The error bars represent the SEM.

Free recall. Notably, we found no difference in the amounts of overall learned sequence pairs between conditions (see Fig. 4), but the degree of awareness was affected by the different settings between conditions.

In the free recall, the VR experimental group had a mean performance of 37.57% for implicit and 26.03% (SEM = 3.61%) for explicit memory. In contrast, the conventional, 2D screen based experimental groups resulted in an average performance of 46.73% implicit and 7.44% explicit memory (PC-short; SEM = 5.12%) and 65.06% implicit and 1.27% explicit memory (PC-long; SEM = 4.87%).

Overall, the volunteers acquired more implicit than explicit memory (main effect of Learning-Type: $F_{(1188)} = 55.9$, $p < 0.0001$). The interaction effect of Conditions and Learning-Types ($F_{(2188)} = 10.9$, $p < 0.0001$; see Fig. 4) revealed more explicit memory in the VR group as compared to both PC groups.

Gained knowledge in the short versus long version of the 2D-computer-screen based conditions. We tested for probable differences in the mean of gained explicit and implicit knowledge between a short and a long version of the 2D-computer-screen based sequential-association-task. Notably, we found no difference in the amounts of overall learned sequence pairs between the 2D-computer-screen based conditions. Overall, volunteers acquired more implicit than explicit memory in both the completion task (main effect of Learning-Type: $F_{(1,92)} = 82.94$, $p < 0.0001$; see Fig. 3) as well as the free recall (main effect of Learning-Type: $F_{(1,92)} = 265.2$, $p < 0.0001$; see Fig. 4). An interaction effect for the influence of Conditions PC-short versus PC-long on Learning-Type was significant for both completion task ($F_{(1,92)} = 12.44$, $p < 0.001$) and free recall ($F_{(1,92)} = 4.770$, $p < 0.0315$) (see Figs. 3 and 4). However, this effect is related to less explicit memory in the PC versions.

Stimulus value rating. We tested for probable differences in the mean of stimulus value ratings between an enriched visual environment in the VR task and non-enriched visual stimuli in both PC tasks.

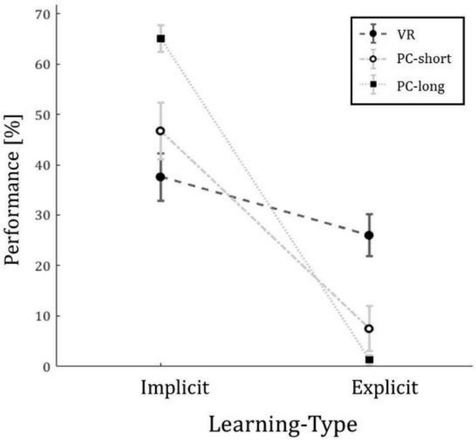


Figure 4. Free recall performance of learned pairs within the given sequence divided by Learning-Type (implicit; explicit) and Condition (VR, PC-short, PC-long). The light grey, dotted lines represent the performance of participants of PC-short (N = 21) and PC-long (N = 27) and the dark grey, dashed line represents the performance of VR (N = 49) participants in the free recall. The performance is divided into implicitly learned sequence pairs (left) and explicitly learned sequence pairs (right). The mean of the performance is visualized as black circle (VR), hollow circle (PC-short) and black square (PC-long). The error bars represent the SEM.

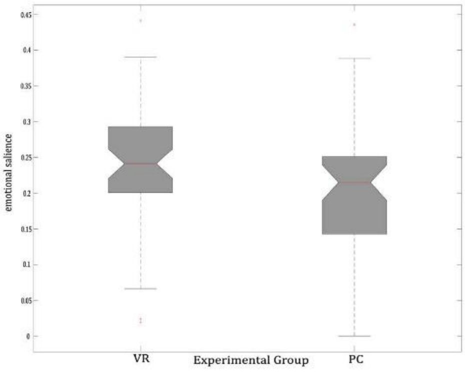


Figure 5. The stimulus value rating for the VR condition and the two merged 2D-computer-screen conditions. The boxplot shows for the VR condition a median of 0.24 (25-percentiles of 0.2 and 75-percentiles of 0.29) and 0.21 (25-percentiles of 0.14 and 75-percentiles of 0.25) for the computer-screen conditions.

We could not find a statistically significant difference in stimulus value ratings ($t_{(95)} = 1.82$, $p = 0.071$) between conditions of VR (Mean = 0.24; SEM = 0.012) compared to both conventional 2D-computer-screen based conditions taken together (Mean = 0.21; SEM = 0.015; see Fig. 5). Furthermore, we tested for associations of stimulus values and the amount of implicit and explicit knowledge separately for each task. We found no correlation of the stimulus value ratings and the performance of implicit memory, as well as no correlation between the stimulus value ratings and the performance of explicit memory (all $p > 0.05$).

We additionally calculated an ANCOVA for both post-experimental assessments thereby the stimulus value rating functions as the covariate to test whether the interaction effect is affected by the stimulus value. We gained comparable significant effects to the previous performed ANOVA. The ANCOVA for the completion task resulted in an interaction effect between Condition and Learning-Type (F (2188) = 2.82, $p < 0.0001$) and a main effect for Learning-Type: F (1188) = 12.21, $p < 0.0001$). The ANCOVA for the free recall resulted in an interaction effect between Condition and Learning-Type (F (2188) = 31.79, $p < 0.0001$) and a main effect for Learning-Type: F (1188) = 73.45, $p < 0.0001$).

Discussion

Although incidental learning is a fundamental process occurring in everyday life, its underlying mechanism is commonly investigated in artificial laboratory environments. In particular, the potential benefits for the generation of explicit memory during learning in naturalistic contexts remain elusive. In this study, we aimed on closing this gap by implementing a near-natural sequential association task in an enriched environment presented in VR. We compared the generated knowledge in the enriched task environment with the learning behavior we got when participants perform a sequential association task in a classical and non-enriched environment. Our results revealed that participants successfully acquired a comparable amount of memory in both enriched and non-enriched learning environments.

Enhanced explicit memory formation in an enriched task design. Although the amount of learned items was similar across all applications, two memory tests (completion and free recall task) revealed that the extent of later expressed explicit memory was enhanced in the enriched environment compared to the conventional non-enriched environment. These results support the view that an enriched setting is an essential factor that can explain an increased explicit memory in the VR application of the incidental sequential association task. Therefore, the *enriched environmental hypothesis* is supported by our results. We could not find a significant difference in the visual stimulus value rating used in the VR condition compared to the conventional 2D-computer-screen based applications, indicating that the stimulus alone outside the VR cannot explain the difference in the explicit memory generation. Therefore, the enhanced development of explicit memory can be related to various advantages of the VR application[19,21,22,39]. One possible factor includes that VR applications facilitate the transition of information by its general characteristics in mobility and natural movements like the upright body position, and not only arms but especially head movement freedom. That way, VR applications might reflect a motivating environment with the potential to increase the feeling of an embodiment, like being physically present in the scenario, by which encoding efficiency and the transition into explicit memory are increased[25,40]. Closely related, evidence from Smith & Mulligan (2021)[25] includes the concept of immersion, which represents the degree of natural features in the VR portrayed by the virtual environment's sensory and interactive properties (see[41,42]). Previous research showed that immersion strongly correlates with participants' attentional engagement level[43,44]. Hence, the VR application not only provides a less vulnerable setting to attention decreases (i.e., due to mind wandering), it might even enhance attentional processes[16,25,45]. These advantages are a potential explanation for the enhanced explicit memory in the VR condition and constitute to the higher enriched environmental aspects that are increased in more real-life settings. Besides the environment posing an essential factor, near-natural perception is also increased when stimulus material is enriched by context-related components, enabling superior memory formation[46–48]. These factors of the *enriched environmental hypothesis* have a clear impact on the transition from implicit to explicit memory.

The role of motivation in learning and why an enriched environment facilitates explicit memory formation by boosting motivational and attentional processes. In neuroscience, recent evidence has specifically connected the ventral striatum as a key player in the transition mechanism of implicit to explicit memory formation in incidental learning. The ventral striatum is mainly associated to motivational and reward processes[49] releasing dopamine in rewarding situations, consequently enhancing the generation of explicit memory. In a study by Clos et al. (2018)[10], which also included a sequential task, the dopaminergic level was pharmacologically modulated in human adults. It was shown that an increase in dopamine was directly connected to an increased transfer of information. This increase in information transfer was linked to the enhanced formation of explicit memory in an incidental sequential task. This way, successful predictions during task processing reflect an achievement within our neural system rewards. These rewarding processes encourage insight into hidden regularities and the emergence of explicit memory. The role of the dopaminergic system in the transfer from implicit to explicit memory may be related to the increased explicit memory rate in the present study as a consequence of the more rewarding and motivating experimental setting in an enriched task environment like we introduced to the participants in the VR surrounding.

The fluency hypothesis versus the enriched environmental hypothesis. To control for the feeling of task fluency (see the introduction for further details), we implemented two simple conventional 2D-computer-screen based versions, which differed in experiment length. The increased amount of implicit memory in the 2D-computer-screen condition indicated that the content is still learned, but is transferred to explicit memory only to a reduced degree. By directly comparing the effect of the *fluency* factor and the *enriched environmental* factor in the present experiment, the impact of the motivational and enriched experimental environment seems to have a larger effect on the emergence of explicit memory.

Conclusion. To summarize, the increased transition from implicit to explicit memory during incidental learning in the VR setting is related to an interaction of enhanced task processing, rewarding processes and attentional as well as motivational factors. These observations in the VR application have strong therapeutic implications for the rehabilitation of patients, who suffer from learning and memory impairments according to neurological diseases. We show that the transition from implicit to explicit memory is considerably influenced by the enriched environment, which includes context-enriched stimuli. In this work, we have obtained compelling evidence that the transition from implicit to explicit memory is primarily influenced by environment-engaging processes like attention, motivation, and presence.

Limitations of this study. Within this study, we gained evidence that the environment used for studying incidental learning in humans influences the outcome of the type of generated knowledge to a great extent. As this study is one of the first attempts in elucidating the complex processes such an enriched and near-natural study design, here presented in VR, can have on incidental learning, we cannot clearly state which of the factors mentioned within our discussion plays a key role that led to the enhanced transfer from implicit to explicit memory formation. Yet, the findings emphasize the relevance for investigating learning and memory processes in more near-natural scenarios.

Data availability

The datasets generated and/or analyzed during this study are available on request from the corresponding author.

Received: 21 June 2022; Accepted: 27 October 2022

References

1. Anderson, B. A. The attention habit: How reward learning shapes attentional selection. *Ann. N. Y. Acad. Sci.* **1369**(1), 24–39. https://doi.org/10.1111/nyas.12957 (2016).
2. Astleitner, H. & Wiesner, C. An integrated model of multimedia effects on learning. *J. Educ. Multimed. Hypermedia* **11**(2), 177–191 (2002).
3. Cleeremans, A., Jiménez, L. & French, R. M. Implicit learning and consciousness: A graded, dynamic perspective. *Implicit Learn Conscious* https://doi.org/10.4324/9780203989401-9 (2002).
4. Esser, S. & Haider, H. Action–effects enhance explicit sequential learning. *Psychol. Res.* **82**(6), 1113–1129. https://doi.org/10.1007/s00426-017-0883-5 (2018).
5. Haider, H., Eberhardt, K., Kunde, A. & Rose, M. Implicit visual learning and the expression of learning. *Conscious Cogn.* **22**(1), 82–98. https://doi.org/10.1016/j.concog.2012.11.003 (2013).
6. Rose, M., Haider, H. & Büchel, C. The emergence of explicit memory during learning. *Cereb. Cortex.* **20**(12), 2787–2797. https://doi.org/10.1093/cercor/bhq025 (2010).
7. Esser, S., Lustig, C. & Haider, H. What triggers explicit awareness in implicit sequence learning? Implications from theories of consciousness. *Psychol. Res.* **86**, 1442–1457. https://doi.org/10.1007/s00426-021-01594-3 (2021).
8. Goujon, A., Didierjean, A. & Poulet, S. The emergence of explicit knowledge from implicit learning. *Mem. Cogn.* **42**(2), 225–236. https://doi.org/10.3758/s13421-013-0355-0 (2014).
9. Rünger, D. & Frensch, P. A. How incidental sequence learning creates reportable knowledge: The role of unexpected events. *J. Exp. Psychol. Learn. Mem. Cogn.* **34**(5), 1011–1026. https://doi.org/10.1037/a0012942 (2008).
10. Clos, M., Sommer, T., Schneider, S. L. & Rose, M. Enhanced transformation of incidentally learned knowledge into explicit memory by dopaminergic modulation. *PLoS ONE* **13**(6), 1–15. https://doi.org/10.1371/journal.pone.0199013 (2018).
11. Lustig, C., Esser, S. & Haider, H. The interplay between unexpected events and behavior in the development of explicit knowledge in implicit sequence learning. *Psychol. Res.* https://doi.org/10.1007/s00426-021-01630-2 (2021).
12. Schwabe, L. & Wolf, O. T. The context counts: Congruent learning and testing environments prevent memory retrieval impairment following stress. *Cogn. Affect. Behav. Neurosci.* **9**(3), 229–236. https://doi.org/10.3758/CABN.9.3.229 (2009).
13. Smith, S. M. & Vela, E. Environmental context-dependent memory. *Psychon. Bull. Rev.* **8**(2), 203–220 (2001).
14. Hölle, D., Meekes, J. & Bleichner, M. G. Mobile ear-EEG to study auditory attention in everyday life: Auditory attention in everyday life. *Behav. Res. Methods* **53**(5), 2025–2036. https://doi.org/10.3758/s13428-021-01538-0 (2021).
15. Piñeyro Salvidegoitia, M. *et al.* Out and about: Subsequent memory effect captured in a natural outdoor environment with smartphone EEG. *Psychophysiology* **56**(5), 1–15. https://doi.org/10.1111/psyp.13331 (2019).
16. Lüddecke, R. & Felnhofer, A. Virtual reality biofeedback in health: A scoping review. *Appl. Psychophysiol. Biofeedback* **47**(1), 1–15. https://doi.org/10.1007/s10484-021-09529-9 (2021).
17. Maarsingh, B. M., Bos, J., Van Tuijn, C. F. J. & Renard, S. B. Changing stress Mindset through Stressjam: A virtual reality game using biofeedback. *Games Health J.* **8**(5), 326–331. https://doi.org/10.1089/g4h.2018.0145 (2019).
18. Talsma, D., Senkowski, D., Soto-Faraco, S. & Woldorff, M. G. The multifaceted interplay between attention and multisensory integration. *Trends Cogn. Sci.* **14**(9), 400–410. https://doi.org/10.1016/j.tics.2010.06.008 (2010).
19. Bohil, C. J., Alicea, B. & Biocca, F. A. Virtual reality in neuroscience research and therapy. *Nat. Rev. Neurosci.* https://doi.org/10.1038/NRN3122 (2011).
20. Schmidt, S., Bruder, G. & Steinicke, F. Effects of virtual agent and object representation on experiencing exhibited artifacts. *Comput. Graph.* **83**, 1–10. https://doi.org/10.1016/j.cag.2019.06.002 (2019).
21. Schomaker, J. Unexplored territory: Beneficial effects of novelty on memory. *Neurobiol. Learn. Mem.* **2019**(161), 46–50. https://doi.org/10.1016/j.nlm.2019.03.005 (2018).
22. Smith, S. A. Virtual reality in episodic memory research: A review. *Psychon. Bull. Rev.* **26**(4), 1213–1237. https://doi.org/10.3758/s13423-019-01605-w (2019).
23. Chelazzi, L., Perlato, A., Santandrea, E. & Della, L. C. Rewards teach visual selective attention. *Vis. Res.* **85**, 58–72. https://doi.org/10.1016/j.visres.2012.12.005 (2013).
24. Sanz, L. R. D., Vuilleumier, P. & Bourgeois, A. Cross-modal integration during value-driven attentional capture. *Neuropsychologia* **120**, 105–112. https://doi.org/10.1016/j.neuropsychologia.2018.10.014 (2018).
25. Smith, S. A. & Mulligan, N. W. Immersion, presence, and episodic memory in virtual reality environments. *Memory* **29**(8), 983–1005. https://doi.org/10.1080/09658211.2021.1953535 (2021).
26. Dehn, L. B. *et al.* Training in a comprehensive everyday-like virtual reality environment compared to computerized cognitive training for patients with depression. *Comput. Hum. Behav.* **79**, 40–52. https://doi.org/10.1016/j.chb.2017.10.019 (2018).
27. Ruddle, R. A., Volkova, E., Mohler, B. & Bülthoff, H. H. The effect of landmark and body-based sensory information on route knowledge. *Mem. Cogn.* **39**(4), 686–699. https://doi.org/10.3758/s13421-010-0054-z (2011).
28. Kemény, F. & Lukács, Á. Sequence in a sequence: Learning of auditory but not visual patterns within a multimodal sequence. *Acta Psychol. (Amst.)* https://doi.org/10.1016/j.actpsy.2019.102905 (2019).
29. Zhao, F., Gaschler, R., Nöhring, D. O., Röttger, E. & Haider, H. Sequential modulation of across-task congruency in the serial reaction time task. *Acta Psychol. (Amst.)* **205**, 103043. https://doi.org/10.1016/j.actpsy.2020.103043 (2020).
30. Bavelier, D. & Neville, H. J. Cross-modal plasticity: Where and how?. *Nat. Rev. Neurosci.* **3**(6), 443–452. https://doi.org/10.1038/nrn848 (2002).

31. Butler, A. J. & James, K. H. Cross-modal versus within-modal recall: Differences in behavioral and brain responses. *Behav. Brain Res.* **224**(2), 387–396. https://doi.org/10.1016/J.BBR.2011.06.017 (2011).
32. Shams, L. & Seitz, A. R. Benefits of multisensory learning. *Trends Cogn. Sci.* **12**(11), 411–417. https://doi.org/10.1016/j.tics.2008.07.006 (2008).
33. Rose, M., Haider, H., Salari, N. & Büchel, C. Functional dissociation of Hippocampal mechanism during implicit learning based on the domain of associations. *J. Neurosci.* **31**(39), 13739–13745. https://doi.org/10.1523/JNEUROSCI.3020-11.2011 (2011).
34. Haider, H., Eberhardt, K., Esser, S. & Rose, M. Implicit visual learning: How the task set modulates learning by determining the stimulus-response binding. *Conscious Cogn.* **26**(1), 145–161. https://doi.org/10.1016/j.concog.2014.03.005 (2014).
35. Jablonowski, J. & Rose, M. The functional dissociation of posterior parietal regions during multimodal memory formation. 1–17. https://doi.org/10.1002/hbm.25861 (2022).
36. Esser, S. & Haider, H. The emergence of explicit knowledge in a serial reaction time task: The role of experienced fluency and strength of representation. *Front. Psychol.* **8**(502), 1–18. https://doi.org/10.3389/fpsyg.2017.00502 (2017).
37. Dienes, Z. & Seth, A. Gambling on the unconscious: A comparison of wagering and confidence ratings as measures of awareness in an artificial grammar task. *Conscious Cogn.* **19**(2), 674–681. https://doi.org/10.1016/j.concog.2009.09.009 (2010).
38. Conway, C. M. & Christiansen, M. H. Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychol. Sci.* **17**(10), 905–912. https://doi.org/10.1111/j.1467-9280.2006.01801.x (2006).
39. Plancher, G., Barra, J., Orriols, E. & Piolino, P. The influence of action on episodic memory: A virtual reality study. *Q. J. Exp. Psychol.* **66**(5), 895–909. https://doi.org/10.1080/17470218.2012.722657 (2013).
40. Kilteni, K., Groten, R. & Slater, M. The sense of embodiment in virtual reality. *Presence Teleoper. Virtual Environ.* **21**(4), 373–387. https://doi.org/10.1162/PRES_a_00124 (2012).
41. Fuchs, P., Moreau, G. & Guitton, P. *Virtual Reality: Concepts and Technologies* (CRC Press, Florida, 2011).
42. Slater, M. A note on presence terminology. *Presence Connect.* **3**(3), 1–5 (2003).
43. Darken, R. P., Bernatovich, D., Lawson, J. P. & Peterson, B. Quantitative measures of presence in virtual environments: The roles of attention and spatial comprehension. *Cyberpsychol. Behav. Impact Internet Multimed. Virtual Real Behav. Soc.* **2**(4), 337–347. https://doi.org/10.1089/cpb.1999.2.337 (1999).
44. Lessiter, J., Freeman, J., Keogh, E. & Davidoff, J. A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence Teleoper. Virtual Environ.* **10**(3), 282–297. https://doi.org/10.1162/105474601300343612 (2001).
45. Dehaene, S. & Changeux, J. P. Experimental and theoretical approaches to conscious processing. *Neuron* **70**(2), 200–227. https://doi.org/10.1016/j.neuron.2011.03.018 (2011).
46. Eggermont, J. J. Multisensory processing. *Hear Loss* https://doi.org/10.1016/b978-0-12-805398-0.00003-7 (2017).
47. Schmid, C., Büchel, C. & Rose, M. The neural basis of visual dominance in the context of audio-visual object processing. *Neuroimage* 55(1), 304–311. https://doi.org/10.1016/j.neuroimage.2010.11.051 (2011).
48. Van Atteveldt, N., Murray, M. M., Thut, G. & Schroeder, C. E. Multisensory integration: Flexible use of general operations. *Neuron* **81**(6), 1240–1253. https://doi.org/10.1016/j.neuron.2014.02.044 (2014).
49. Pool, E. R. *et al.* Differential contributions of ventral striatum Subregions to the motivational and hedonic components of the affective processing of reward. *J. Neurosci.* **42**(13), 2716–2728. https://doi.org/10.1523/JNEUROSCI.1124-21.2022 (2022).

Acknowledgements

The present research was supported by the German Research Foundation (DFG; SFB/Transregio 169, Project B3, C8 and DFG RO 2653/9-1). We thank Theresia Vordermark and Celia Kumpmann for helping measuring the data.

Author contributions

M.R. and F.S. designed the study. C.J. and P.T recorded the data, C.J., M.M. and M.R. analyzed the data and C.J., M.M. and M.R. wrote the manuscript. All authors reviewed the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.R.

Reprints and permissions information is available at www.nature.com/reprints.

Appendix B: Study II

Disentangling the Functional Roles of Pre-Stimulus Oscillations in Crossmodal Associative Memory Formation via Sensory Entrainment. Maack*, M. C., Ostrowski*, J., & Rose, M. C. (2025). *bioRxiv*, 2025-06.

*Equal contributions

Disentangling the Functional Roles of Pre-Stimulus Oscillations in Crossmodal Associative Memory Formation via Sensory Entrainment

Jan Ostrowski[1†], Marike C. Maack[1†], Michael Rose[1*]

1 *Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany;* j.ostrowski@uke.de, m.maack@uke.de, rose@uke.de

† Equal contributions

** Correspondence:* rose@uke.de

16-digit ORCID:
JO: 0000-0001-9928-272X
MM: 0000-0002-6960-4465
MR: 0000-0002-9789-7066

Impact Statement

By using sensory entrainment of pre-stimulus oscillations we could show thatalpha-band stimulation in particular enhanced crossmodal memory. These findings reveal a frequency-specific functional dissociation and highlight the potential of targeting preparatory brain rhythms to improve crossmodal memory formation.

Abstract

The state of neural dynamics prior to the presentation of an external stimulus significantly influences its subsequent processing. This neural preparatory mechanism might be of particular importance for crossmodal memory formation. The integration of stimuli across different sensory modalities is a fundamental mechanism underlying the formation of episodic memories. However, the causal role of pre-stimulus neural activity in this process remains largely unclear. In this preregistered study, we investigate the direct relationship between transient brain states induced by sensory entrainment and crossmodal memory encoding. Participants (*n* = 105) received rhythmic visual stimuli at theta (5 Hz) or alpha (9 Hz) frequencies to evoke specific brain states. EEG recordings confirmed successful entrainment, with sustained increases in neural activity within the stimulated frequency bands persisting until stimulus onset. Notably, induced alpha oscillatory activity enhanced recognition memory performance reflected by increased sensitivity, and suggesting that alpha oscillations prepare the brain for optimal multisensory integration. These findings highlight the functional significance of distinct oscillatory brain states in facilitating memory encoding by increasing cortical excitability before stimulus presentation. Overall, our results emphasize the importance of pre-stimulus brain states in shaping the efficiency of memory formation across sensory modalities and shed light on how dynamic neural preparations support learning.

Keywords

Pre-Stimulus Entrainment, Visual Sensory Stimulation, Crossmodal Learning, Associative Memory Formation

Introduction

Multisensory learning is fundamental for human cognition, enabling the encoding and retrieval of complex environmental information. In daily life, individuals continuously integrate sensory information from multiple modalities, such as visual and auditory stimuli, to enhance memory performance. This ability to form crossmodal associations supports essential cognitive functions, especially episodic memory formation (Dickerson & Eichenbaum, 2010; Gasser & Davachi, 2023). Given the relevance of multisensory learning, understanding the underlying neural mechanisms has become a key objective in cognitive neuroscience. Brain oscillations play a critical role in coordinating neural activity during multisensory learning. Theta oscillations (3–7 Hz) have been widely implicated in the formation of episodic memory, particularly in binding disparate elements of experience into coherent memories (Klimesch et al., 2011; Rudoler et al., 2023; Staudigl & Hanslmayr, 2013). Research suggests that theta rhythms support the temporal organization of information, facilitating associative encoding across modalities (Buzsáki & Moser, 2013; Herweg et al., 2020; Terada et al., 2017). Additionally, alpha oscillations (8–12 Hz) have been associated with attentional selection, serving as a gating mechanism to suppress irrelevant sensory input and enhance task-relevant processing (Foxe & Snyder, 2011; Jensen & Mazaheri, 2010; Waldhauser et al., 2012). However, the precise role of pre-stimulus theta and alpha dynamics of memory formation during multisensory learning remains unclear.

A growing body of evidence highlights the importance of pre-stimulus neural activity in shaping subsequent cognitive processing (Lindenbaum et al., 2023; Roberts et al., 2018; Salari & Rose, 2016; Taesler & Rose, 2022; Van Dijk et al., 2008; Zazio et al., 2022). Pre-stimulus theta and alpha power fluctuations have been linked to successful memory formation (Addante et al., 2011; Schneider & Rose, 2016; Scholz et al., 2017; Sweeney-Reed et al., 2016; Winterling et al., 2019), suggesting that oscillatory states before stimulus presentation may serve a preparatory function (Cruzat et al., 2021; Strunk & Duarte, 2019; Zoefel & VanRullen, 2017). In particular, we were able to support this notion in a previous investigation, where participants were required to memorize audiovisual pairs in a Subsequent Memory Effects task (SME; Ostrowski & Rose, 2024). We could demonstrate that theta and alpha oscillations have a significant impact on memory encoding during the pre-stimulus phase, as increases in theta (3–7 Hz) and alpha power (8–12 Hz) observed before stimulus presentation were associated with enhanced memory performance. Specifically, higher pre-stimulus theta and alpha activity has been linked to better recognition of crossmodal associations between stimuli, such as visual and auditory inputs.

These findings propose that pre-stimulus oscillations might optimize encoding conditions (Amil et al., 2024; Salari & Rose, 2016), aligning neural activity with upcoming information (Schneider & Rose, 2016; Terporten et al., 2019; Winterling et al., 2019; Yeh & Rose, 2019). However, a causal link between pre-stimulus oscillatory activity and successful learning has not yet been demonstrated. One promising approach is the modulation of pre-stimulus frequencies through entrainment. These methods, such as transcranial alternating current stimulation (tACS) and rhythmic sensory stimulation, provide the means to modulate oscillatory activity in a non-invasive manner (Bree et al., 2021; Neuling et al., 2015, 2017; Veniero et al., 2015). The application of external rhythmic stimulation can synchronize endogenous neural rhythms at targeted frequencies (Duecker et al., 2024; Notbohm et al., 2016; Notbohm & Herrmann, 2016; Thut et al., 2011), thereby affecting cognitive processes, and subsequently behavior (Bree et al., 2021; Michael et al., 2022; Wang et al., 2024). In sensory entrainment, neural oscillations are modified by an external visual or auditory stimulus during encoding. Depending on the sensory domain, either luminance or amplitude oscillate in a specific frequency, leading to increases in oscillatory power. As the brain synchronizes with these external rhythms, it may become more aligned at integrating sensory details into structured memories (Grover et al., 2021; Köster & Gruber, 2022; Singer, 1993; Wälti et al., 2020). Given the evidence that pre-stimulus oscillatory activity can affect memory performance, investigating whether externally applied rhythmic stimulation can modulate these oscillatory states to enhance learning is crucial to reveal a direct functional role of this neural mechanism. Furthermore, this might allow researchers to determine in a causal framework whether the potential enhancement of multisensory memory formation stems from improved temporal binding (theta) or more effective suppression of irrelevant information (alpha). However, studies investigating sensory entrainment in the context of multisensory learning have yielded mixed results so far (Hanslmayr et al., 2019; Wälti et al., 2020; Wang et al., 2018).

This pre-registered study aims to examine a direct link between pre-stimulus states of theta and alpha oscillations and multisensory memory formation by using visual sensory entrainment, while also addressing existing challenges in sensory entrainment through an optimized experimental paradigm. Using a between-subjects design, participants were required to memorize and later recognize pairs of visual and auditory stimuli. Visual sensory entrainment was presented immediately before each stimulus from the encoding task at either 5 Hz (theta group) or 9 Hz (alpha group). The choice of entrainment frequencies was based on observed effects from prior research where the same SME paradigm was used (Ostrowski & Rose, 2024). Arrhythmic stimulation was used as a control condition in which the entrainment oscillations were derived randomly from frequencies between 13 and 24 Hz. This approach extends previous work through a refined experimental design, allowing us to test whether pre-stimulus sensory entrainment might influence neural oscillations and memory performance.

Building upon prior research, the current study aims to replicate and extend previous findings through a refined experimental design that systematically manipulates brain oscillations before stimulus onset. First, we first expected that using an oscillating image as a stimulus for sensory entrainment will be successful in modifying oscillatory activity and hypothesized that it will lead to increased oscillatory power within the entrained frequency ranges (H1). Importantly, we hypothesized that both theta (H2a) and alpha (H2b) entrainment would enhance memory performance as compared to controls. Moreover, we hypothesized that theta and alpha entrainment might affect memory performance to a different degree, resulting in potential differences between the two conditions (H2c). Additionally, we expected that both theta and alpha entrainment might lead to improved memory performance as compared to no entrainment, which we assessed through a statistical comparison with the dataset from the previous study (H3).

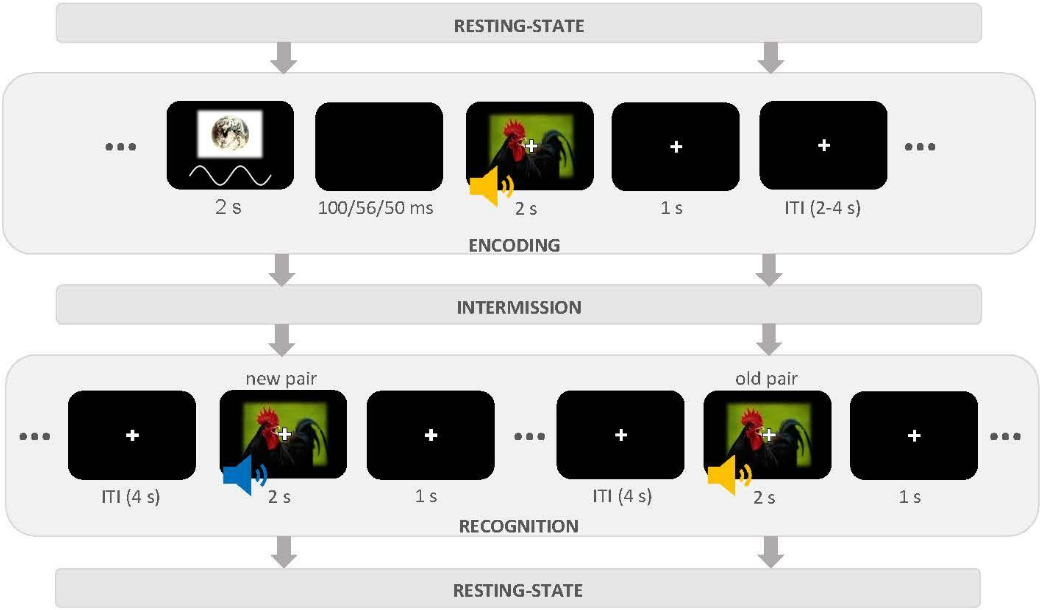


Figure 1. Schematic overview of one experimental run of the SME task. Each encoding trial began with a frequency-specific visual entrainment stimulus that lasted 2 s, followed by a black screen whose presentation duration differed depending on the group (theta: 100 ms; alpha: 56 ms; control: 50 ms). During entrainment, the luminance varied in a sinusoidal manner, resulting in a rhythmic oscillation of 5 Hz in the theta group and 9 Hz in the alpha group. Then, an image–sound pair was presented for 2 s, followed by a fixation cross for 1 s. Taken together, this comprised the possible window for responses. Participants judged whether both stimuli represented animals and were instructed to memorize each audiovisual combination. A fixation cross remained visible on the screen during the subsequent inter-trial interval of 2 to 4 s. In the recognition task, participants were presented with previously shown (old) and recombined (new) audiovisual pairs and indicated whether they remembered the particular combination of image and sound or not. Stimuli were shown for 2 s, and responses were recorded up to 3 s after stimulus onset. The inter-trial interval was fixed at 4 s, during which a fixation cross was shown. In the intermission between each encoding and recognition task, participants were presented with a short distraction task.

Results

In this study, participants (n = 105) performed in a sequential memory encoding and recognition task across three experimental runs, each containing audiovisual pairings that the participants were instructed to memorize **(Figure 1)**. Prior to each pairing in the encoding phase, participants were exposed to rhythmic visual stimulation at either theta (5 Hz) or alpha (9 Hz) frequencies, or exposed to arrhythmic stimulation (control) for two seconds. Each encoding run was followed by a short distractor task and a recognition phase, in which previously seen pairs were randomly intermixed with recombined lures. Participants indicated whether each pair was old or new via button press. We implemented an open-ended sequential design for gathering evidence, taking advantage of the Bayesian statistical framework. The data collection concluded either when at least moderate evidence had been gathered to accept or reject the null hypothesis for the respective contrast or when group size reached $k = 35$ for each group.

Successful pre-stimulus visual stimulation modified targeted frequency

The focus of this study was to test whether visual sensory entrainment before the presentation of a stimulus would affect its subsequent encoding and thus result in changes in memory performance. As a prerequisite, we needed to make sure that the entrainment procedure would increase oscillatory power in the frequency bands corresponding to the entrainment frequencies (H1). To that end, oscillatory power in the late entrainment period (-1.1 s to -0.1 s relative to stimulus onset) from the theta and alpha entrainment groups was contrasted with the recorded activity from the control group. Two-tailed independent-samples *t*-tests were used on sample level, with a cluster-based permutation approach to account for multiple comparisons. The analysis was conducted for a frequency range of 1 to 40 Hz across the entire channel space. Comparing activity from the theta group with the control group, our analysis revealed a positive cluster ranging from 3 to 7 Hz and spanning the entire late entrainment period ($p < .025$, corrected, **Supplementary Figure S1A**), demonstrating the successful entrainment of pre-stimulus theta activity. Simultaneously, a negative cluster was observed that ranged from 13 Hz to 40 Hz ($p < .025$, corrected), covering most of the beta as well as lower gamma bands. In the comparison between the alpha group and control group, a positive cluster was found in the range of 6 to 10 Hz that spanned the whole analysis window ($p < .025$, corrected), also showing the specific entrainment of alpha band oscillations before the onset of the stimulus pair. Furthermore, the analysis revealed a negative cluster in the high beta/low gamma band ranging from 29 to 34 Hz, spanning the whole analysis window as well ($p < .025$, corrected). Generally, the entrainment seemed to be centered around occipital and parieto-occipital electrodes , and the effects in all entrainment groups were observed only in the pre-stimulus period (**Figure 2**), since our analysis revealed that oscillatory activity after stimulus onset did not differ between the theta, alpha, and control groups ($p = .069$; **Supplementary Figure S2**) In addition, comparing pre-stimulus power from both entrainment groups with oscillatory activity from the NE group using identical analysis parameters revealed similar patterns. Specifically, we found a significant positive cluster in the theta frequency range (theta vs NE; $p < .025$, corrected) as well as in the alpha band (alpha vs NE; $p < .025$, corrected). The common effects found in the control as well as in the NE contrast are shown in **Figure 2B** and **2C**, demonstrating the specificity of the different entrainment protocols (for visualizations of individual contrasts, see **Supplementary S2**). These results suggest that the entrainment of 5 Hz in the theta group and 9 Hz in the alpha group successfully modified oscillatory activity in the pre-stimulus window and the targeted frequency selectively and consistently. As no alpha modification was observed in the theta group, and no theta modification in the alpha group, the observations support our hypothesis that sensory visual entrainment can selectively modify ongoing oscillations (H1).

In some studies, the entrainment frequency is individually tailored to match participants' endogenous rhythms (Duecker et al., 2024; Zaehle et al., 2010). This approach is particularly common in alpha entrainment research, where resting-state EEG is used to identify the individual alpha frequency (IAF) as a target for stimulation (Janssens et al., 2022; Kasten et al., 2019; Klimesch, 2012; Stecher et al., 2017). To explore whether the match between stimulation frequency and endogenous alpha rhythms modulated entrainment strength in the present study, we computed the absolute difference between each participant's IAF and the stimulation frequency in the alpha group (9 Hz). The IAF was extracted from resting-state EEG recorded prior to the main experiment by calculating power spectra using a multitaper fast Fourier transform (1–40 Hz). It was defined as the frequency showing the maximum power within the 8–12 Hz range, averaged across posterior electrodes (Pz, POz, Oz, O1, O2). Thus, we correlated this IAF distance with the maximal relative change in alpha power during the recognition phase, as an index of entrainment strength. This analysis was restricted to participants in the alpha entrainment condition. The correlation was not statistically significant, $r(43) = -0.034$, $p = .849$, indicating that the individual distance from the stimulation frequency did not predict the strength of neural entrainment as indexed by maximal alpha power modulation. Surprisingly, the same analysis for theta revealed a significant negative correlation between the individual theta frequency (ITF) distance and the maximal relative change in theta power during the encoding phase, $r(55) = -0.296$, $p = 0.028$. This suggests that a smaller difference between the individual theta frequency and the entrainment frequency might be associated with greater increases in theta power.

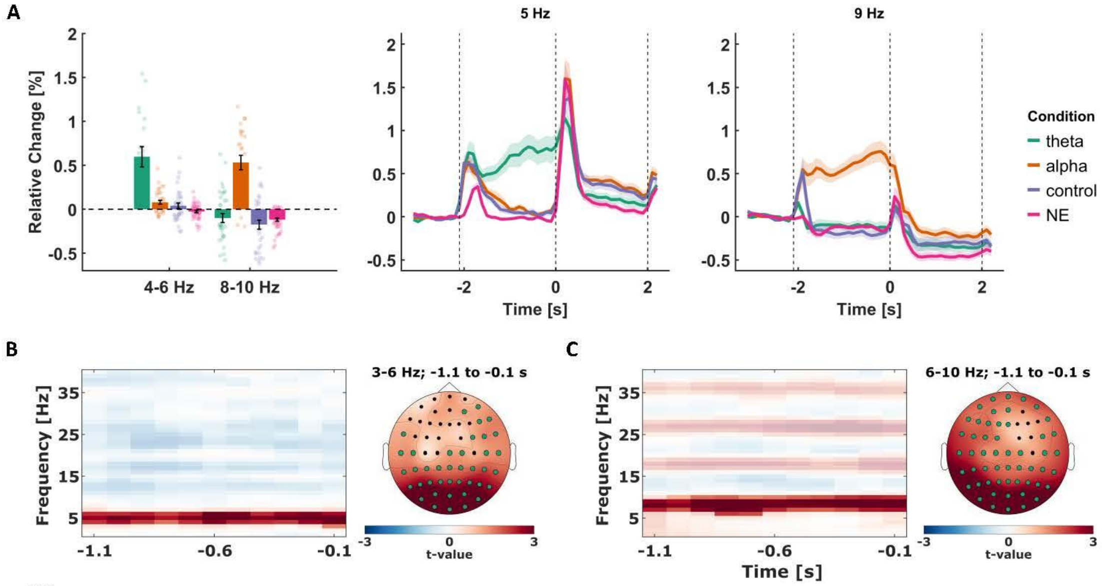


Fig. 2. Effects of visual sensory entrainment on recorded EEG activity. (A) *Left* shows the average relative change of oscillatory power in reference to the baseline period for every group. Specifically, the average change is displayed for the mean activity from the 5 Hz and 9 Hz envelopes (± 1 Hz). Each data point shows the respective mean value for one participant, and the black error bars represent the standard error of means. *Right* shows the average time courses of relative change in power for the 5 Hz and 9 Hz narrow bands across the whole trial period for every group. The shadings around the lines represent the standard error of means across participants. All three figures depict relative change derived from the mean of the occipital electrodes *O1, O2,* and *Oz*. The first dashed vertical line marks the onset of the entrainment stimulus, while the other two mark the stimulus presentation window. **(B)** Visualization of the common effects of visual entrainment that were found in the contrast of the theta group with the control group, as well as in the contrast with the NE group. The time-frequency plot (*left*) shows the dimension of the common cluster along the time and frequency dimensions, depicting the average *t*-values across all contributing electrodes. Positive *t*-values signify greater relative change in the theta group, while opaque data points mark a significant difference at $p < .025$ (corrected). On the *right*, the topographical distribution of the common effects is shown, with electrodes contributing to the cluster marked in green. **(C)** Same as in (B) but for the alpha entrainment group.

Alpha but not theta entrainment enhances memory performance

In this study, we aimed to determine whether changes in oscillatory activity during the pre-stimulus interval could causally influence an individual's ability to encode audiovisual associations. First, the analysis of performance in the categorization task during encoding yielded moderate evidence in support of the null hypothesis, $BF_{10} = 0.381$, suggesting no significant differences in accuracy across conditions. These findings suggest that participants consistently adhered to task demands throughout the experiment, supporting the validity of subsequent analyses on oscillatory activity and memory performance (see **Supplementary 3** for further details). Notably, memory performance in the recognition task as measured by the sensitivity index *d'* was significantly enhanced in the alpha entrainment group (*M* = 1.46, *SD* = 0.60) as compared to the control group (*M* = 1.18, *SD* = 0.50; **Figure 3A**). An independent-samples *t*-test yielded a Bayes factor of BF_{10} = 3.29, providing moderate evidence for the alternative hypothesis and suggesting that increased alpha-band activity induced by visual entrainment may facilitate the formation of audiovisual associations (H2b). Further analysis revealed that this effect was primarily driven by a measurable increase in hit rate in the alpha group (M = 61.498%, SD = 14.443%) as compared to the control group (M = 49.792%, SD = 15.029%), BF_{10} = 22.5742. Simultaneously, no differences were observed in the false positive rate between the groups (alpha: M = 13.9%, SD = 6.67%; control: M = 13%, SD = 5.73%), BF_{10} = 0.2852 (**Figure 3B**). This suggests that participants in the alpha group were more likely to correctly recognize an old stimulus pair compared to those in the control condition. In contrast, the comparison of sensitivity between the theta group (*M* = 1.28, *SD* = 0.55) and the control group yielded a Bayes factor of BF_{10} = 0.49, indicating weak evidence for the null-hypothesis. Similarly, the direct comparison between the theta and alpha groups resulted in a Bayes factor of BF_{10} = 0.53, further suggesting weak support for the null hypothesis. Although these results do not support our hypotheses H2a and H2c, they provide evidence that any effect of theta entrainment on encoding performance may be smaller or more variable than anticipated. Together, these findings point to a potentially specific role of alpha oscillations in enhancing audiovisual memory encoding, highlighting the importance of frequency-specific mechanisms in pre-stimulus neural dynamics.

To address our preregistered hypothesis (H3), we compared memory performance in the entrainment groups to that of participants from a previous study who were not exposed to any rhythmic stimulation during the pre-stimulus interval but instead viewed a static fixation cross (NE group; M = 1.27, SD = 0.53). The results of the Bayesian *t*-test showed that neither the theta group nor the control group differed significantly from the NE group, BF_{10} = 0.26 and BF_{10} = 0.16, respectively, indicating moderate-to-strong evidence for the absence of an effect. The comparison between the alpha group and the NE group yielded a Bayes factor of BF_{10} = 1.03, indicating that the data did not provide conclusive evidence for either hypothesis. However, we found moderate evidence suggesting that the hit rate measured in the alpha group still differed from the hit rate in the NE group (M = 52.662%, SD = 13.724%), BF_{10} = 3.8563. Again, we also found moderate evidence that the false positive rate from the NE group (M = 12.8%, SD = 6.1%) was likely not statistically different from the false positive rate in the alpha group. This indicates that participants receiving alpha band stimulation correctly remembered old stimuli more often than participants who were not stimulated at all. In addition, we investigated whether sensory entrainment might have affected how memory performance changed across the experiment. While we found that memory performance generally increased over the course of the experiment, this effect was not modulated by the pre-stimulus condition, BF = 0.0343, indicating that the improvement was consistent across entrainment conditions (see **Supplementary 4** for further details).

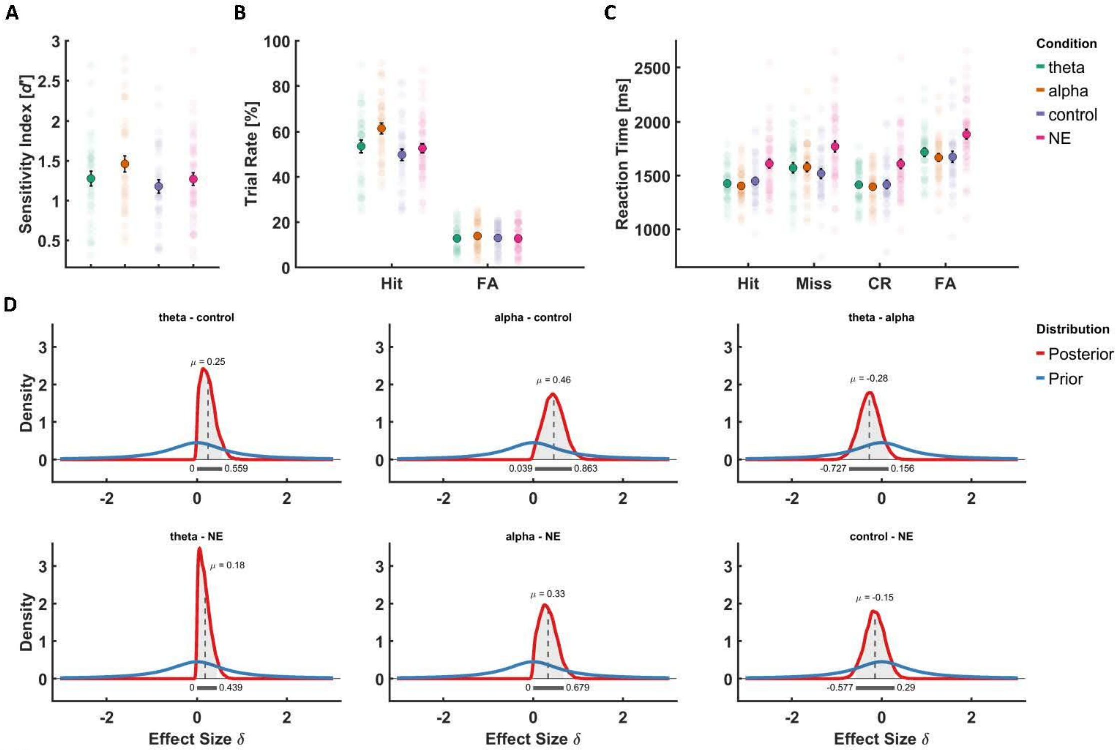


Figure 3. Effects of neural entrainment on recognition memory performance. (A) Mean sensitivity index (*d'*) with standard errors of the mean (SEM) across the three entrainment conditions **(B)** Average trial rates depicting the proportion of remembered old pairs (Hit) and new pairs erroneously categorized as old (false alarms; FA) across individuals for each group. Black error bars mark the SEM. **(C)** Group means of reaction times (RTs) for all response categories from the recognition task illustrating differences in processing speed between the groups. In addition to Hits and FAs, the figure includes RTs for not recognized old pairs (Miss) and correctly rejected new pairs (CR). Individual group means are reported in **Supplementary 5**. **(D)** Prior and posterior distributions for the individual group contrasts for sensitivity indices. The Highest Density Interval (HDI) is marked by the grey bar and the light grey shaded area under the curve of the posterior distribution. The mean effect size μ is marked by the dashed line.

Visual entrainment accelerates response times

Next, we explored the effect of visual entrainment on response times. First, we compared group means of reaction times from the categorization task during encoding using the Bayesian version of a one-way ANOVA with the factor *pre-stimulus condition* (theta, alpha, control, NE). The analysis revealed moderate evidence in favor of the alternative hypothesis, $BF_{10} = 5.0737$, suggesting relevant differences in response times between the groups during encoding. Further analysis revealed that participants in the NE group responded significantly slower as compared to the entrainment groups (see **Supplementary 3** for details). Next, differences in reaction times from the recognition task were assessed using the same statistical approach but conducted separately for all four response categories (hits, misses, correct rejections, false alarms). Importantly, we found strong evidence that response times differed significantly between conditions for every response category (hit: $BF_{10} = 78.0847$; miss: $BF_{10} = 8.0419$; correct rejections: $BF_{10} = 49.3906$; false alarms: $BF_{10} = 18.2397$). Subsequent analyses revealed that the entrainment groups did not differ in response times, regardless of which response category was tested. The NE group, however, displayed consistently slower reaction times than the other groups for every response category (**Figure 3C;** see **Supplementary 5** for more information on the individual group contrasts). The results suggest that participants who received visual stimulation during the encoding phase responded faster in the recognition phase than participants who were not stimulated. In addition, we investigated whether the entrainment condition would affect the discrepancy in response times between recognition trials with correct and incorrect responses. To that end, response times from all response categories were grouped according to the correctness of the corresponding trial and submitted to a Bayesian mixed-design ANOVA. The analysis yielded strong evidence in favor of the null hypothesis, $BF_{10} = 0.1161$, indicating that the difference in response times between correct and incorrect trials was not modulated by the entrainment condition.

No lasting effect of entrainment condition on resting state activity and salience reports

To complement our pre-registered analysis, we conducted exploratory investigations examining resting-state EEG data before and after the experiment, as well as differences between groups in subjective salience reports. Comparing the difference in resting-state activity from before and after the experiment across the theta, alpha and control groups revealed no notable differences ($p = .2972$, corrected). This suggests that any changes in baseline activity due to the experiment were not dependent on the entrainment condition and appeared to be a general effect instead (see also **Supplementary 6**). To control for the subjective experience of the visual entrainment, participants rated its pleasantness and salience as well as their own perceived attention and fatigue at the end of every encoding task. Although we observed noticeable decreases in perceived attention as well as increases in perceived fatigue over the course of the experiment, $BF_{\mathbf{attention}} = 1.855 \times 10^{9}$, $BF_{\mathbf{fatigue}} = 0.3415 \times 10^{15}$, these effects did not interact with the entrainment condition, $BF_{\mathbf{attention}} = 0.195$, $BF_{\mathbf{fatigue}} = 0.0552$ (**Supplementary 7**). Overall, these results suggest that subjective perceptions and task engagement were comparable across groups, reducing the likelihood of confounds influencing the behavioral outcomes.

Discussion

This study aimed to investigate the causal role of pre-stimulus oscillations in the encoding of crossmodal associations. Using visual sensory entrainment targeted at theta (5 Hz) and alpha (9 Hz) frequencies, we aimed to modulate neural rhythmic activity prior to stimulus presentation and assess its impact on memory performance. Our main findings demonstrated successful frequency-specific entrainment of pre-stimulus oscillations. Notably, alpha-band entrainment before stimulus presentation resulted in significantly improved recognition memory, as evidenced by increased sensitivity driven by higher hit rates. In contrast, pre-stimulus theta entrainment did not produce measurable behavioral effects. These results support a functional dissociation of pre-stimulus alpha and theta oscillations in relation to memory encoding, with alpha activity playing a more prominent role in facilitating successful associative memory formation. Importantly, these effects were driven solely by transient pre-stimulus modulation, with no evidence for lasting entrainment effects during stimulus presentation or changes in post-experiment resting-state activity, enabling the differentiation of their respective functional contributions.

Previous research has demonstrated that pre-stimulus brain activity significantly influences episodic memory formation (Addante et al., 2011; Salari & Rose, 2016; Schneider & Rose, 2016; Scholz et al., 2017; Sweeney-Reed et al., 2016; Winterling et al., 2019). Oscillatory activity, particularly within the theta (3–7 Hz) and alpha (8–12 Hz) bands, has been frequently associated with successful encoding processes (Cruzat et al., 2021; Ostrowski & Rose, 2024). Elevated pre-stimulus theta power has been linked to enhanced binding of contextual information and more accurate source memory, suggesting a preparatory role for subsequent memory performance (Addante et al., 2011). Similarly, increases in alpha oscillations prior to stimulus onset are thought to reflect a state of attentional preparation, facilitating the encoding of complex audiovisual associations (Strunk & Duarte, 2019). Importantly, attentional engagement modulates these oscillatory patterns, indicating that intentional focus can create neural conditions beneficial for memory formation (Schneider & Rose, 2016; Uncapher et al., 2011).

Our findings build upon this established framework, demonstrating that externally induced pre-stimulus alpha entrainment boosts audiovisual associative encoding. The observed increase in alpha power in our experimental condition aligns with theories that posit alpha oscillations as essential for sensory anticipation and attentional gating (Foxe & Snyder, 2011; Kizuk & Mathewson, 2017; Leske et al., 2025; Morrow et al., 2023). This externally driven alpha synchronization likely enhanced preparatory attentional states, enabling more efficient inhibition of irrelevant information and promoting engagement of memory-related neural networks such as parietal and hippocampal regions (Klimesch et al., 2011; Palva et al., 2010; Parish et al., 2018; Raud et al., 2023; Tian et al., 2021). Behaviorally, this facilitation translated into higher recognition sensitivity, driven primarily by increased hit rates, while false alarm rates remained unaffected. These findings support the hypothesis that alpha oscillations modulate sensory preparation and attentional gating during encoding via bottom-up processes, and thereby improve associative memory performance.

In contrast, pre-stimulus theta entrainment did not produce significant behavioral benefits, despite successfully increasing theta power. This suggests that power enhancement alone may be not sufficient to influence memory performance within this paradigm. A key factor could be the temporal specificity of theta's role in encoding. While theta oscillations were shown to be critical for episodic memory and associative binding (Buzsáki & Moser, 2013; Herweg et al., 2020; Terada et al., 2017), their effectiveness appears to depend heavily on activity during stimulus processing (Hsieh & Ranganath, 2014; Nyhus & Curran, 2010). Prior studies demonstrating memory improvements with theta entrainment typically targeted the period during stimulus presentation, likely optimizing engagement of memory networks (Hanslmayr et al., 2019; Herweg et al., 2020; Köster et al., 2019). These findings indicate that the contribution of theta activity to memory encoding may be more dynamic, occurring during active processing rather than as a preparatory state alone. Our targeting of pre-stimulus activity likely aimed to set a preparatory neural state that may not have directly engaged the neural mechanisms necessary for effective multisensory binding, although oscillatory power was modulated. Furthermore, the role of theta oscillations in memory encoding often involves activity across widespread and synchronized networks such as hippocampal-cortical circuits (Boran et al., 2019; Etter et al., 2023; Nyhus & Curran, 2010), which may not have been fully engaged through unimodal occipital stimulation alone. While the stimulation successfully increased theta power, the lack of phase coherence or cross-regional synchronization may have limited its influence on encoding. These findings emphasize that the contribution of theta activity to memory may be more context-dependent and particularly crucial during active processing phases, rather than solely during pre-stimulus intervals. This aligns with prior studies emphasizing the importance of timing and phase alignment in theta-mediated memory processes.

In addition to the effects of entrainment on primary performance measures we observed a secondary effect on response times. Specifically, participants that received visual stimulation before the encoding of audiovisual pairs responded consistently faster during encoding and, most importantly, during the subsequent recognition task as compared to participants who did not undergo entrainment, while no differences were observed between the three entrainment groups. Given this pattern of results, it is plausible to assume that the faster response times from the entrainment groups could be interpreted as an effect of increased alertness during the encoding period due to general visual stimulation before stimulus onset. Visual stimulation in general has been associated with an increase of alertness before (Figueiro et al., 2018; Golmohammadi et al., 2021; Lok et al., 2018), while alertness, in turn, has been shown to decrease RTs in tasks recruiting executive control systems (Nieuwenhuis & de Kleijn, 2013; Weinbach & Henik, 2012). One could argue that visual information processing might benefit from an improved inhibition of peripheral information (Poirel et al., 2014). This indicates that visual stimulation might have enhanced a preparatory mechanism that is independent from the specific cognitive demand of encoding information but might rather point towards an increased ability to remain vigilant and maintain attention throughout the task despite increasing subjective feelings of fatigue.

Although the entrainment procedure applied in this study led to a reliable modification of pre-stimulus theta and alpha activity, our analyses revealed that both the theta and alpha group exhibited significantly lower beta band power as compared to controls. While we cannot rule out completely that the behavioral effects presented here could also be attributed to modifications of beta band oscillations, it's plausible to assume that the observed difference was caused by increases in beta activity in the control group rather than decreases in the theta and alpha groups. This is supported by the fact that we found no negative clusters in the beta band when comparing activity from the theta and alpha groups with the NE group, and that the effects common to both the control and NE group contrasts are centered around the respective entrainment frequencies. Instead, the arhythmic stimulation in the control group might have modified pre-stimulus beta-oscillations due to potential additive effects of single-frequency cycles randomly chained together. As individual arhythmic luminance functions were computed for every participant in the control group, individual cycles of the same frequency that ended up at the same time point could have had an amplifying effect during averaging procedures, resulting in what seemed as beta power enhancement. However, this does not invalidate the usefulness of arythmic stimulation, as it plays a complementary role in validating and specifying the precision of entrainment procedures. With this, the present work is in line with previous studies using arhythmic stimulation as an additional control mechanism to ensure that oscillatory responses to the entrainment actually arise from the rhythmicity of a specific frequency (Albouy et al., 2017; Michael et al., 2022; Notbohm & Herrmann, 2016; Thut et al., 2011).

In sum, our results highlight the distinct functional roles of alpha and theta oscillations in multisensory learning and memory. Alpha oscillations appear to serve as a gating mechanism that can be externally modulated to optimize sensory processing and attentional filtering (Foxe & Snyder, 2011; Waldhauser et al., 2012), with our findings providing causal evidence that externally driven alpha rhythms prior to encoding facilitate associative memory performance. In contrast, the unsuccessful behavioral impact of theta entrainment highlights the importance of time specificity and multisensory synchronization for the mnemonic functions of theta oscillations (Herweg et al., 2020; Wang et al., 2018). These insights contribute to a nuanced understanding of how tailored oscillatory modulation can differentially influence neural states underpinning successful memory formation, emphasizing the potential of targeted neurostimulation techniques, personalized cognitive interventions, and novel therapeutic approaches for memory disorders. This is highlighting the significant clinical potential of utilizing specific oscillatory pathways to enhance learning and memory.

Methods

Participants

In total, 176 healthy young adults were recruited for this pre-registered study (http://osf.io/5gprt). Participants were required to have normal or corrected-to-normal vision and hearing ability. We had to exclude several participants from the analysis due to unsuccessful entrainment (21.59%). Further exclusions were the result of false positive rates above the predetermined threshold (14.77%). An additional 3.98% were excluded for both failure to entrain and high false positive rate. Taken together, a sample of *n* = 105 (72.38% female) participant data sets were submitted to the analysis, with a group size of *k* = 35 for each experimental group. On average, participants were 24.8 years old (SD = 4.17), with the age ranging from 18 to 35 years. All participants gave their informed consent and received either financial reimbursement or course credit for taking part in the study, which was approved by the ethics committee of the Hamburg Medical Council (PV5893). We confirm that all experiments were performed in accordance with relevant guidelines and regulations.

Experimental design

The Subsequent Memory Effects task (SME) implemented in this study is a slight variation from the design used in Ostrowski & Rose (2024). The pre-stimulus interval in the encoding task was modified to accommodate the entrainment procedure, while the recognition task remained the same. Participants received the same instructions as in the previous study, with the addition that they were made aware of the presence of an oscillating image. For this study, a between-subjects design was employed, with *entrainment condition* serving as the independent variable with three groups: 5 Hz (theta group), 9 Hz (alpha group), and arhythmic (control group). To take advantage of the Bayesian framework, we implemented an open-ended sequential design for gathering evidence, but added the additional constraint of a maximum group size *k* (Schönbrodt & Wagenmakers, 2018). Thus, data collection was carried out evenly between the groups until a group size with *k* = 15 usable data sets was reached. Subsequent statistical hypothesis testing was conducted incrementally for each additional usable data set, using the pre-registered dependent variable (sensitivity index), with changes in evidence being continuously monitored across all groups. Data collection would stop either when statistical testing showed moderate support for either the alternative or null hypothesis ($BF_{10} > 3$ or $BF_{10} < 1/3$; Jeffreys, Harold, 1998; Lee & Wagenmakers, 2014) or when group size reached *k* = 35 for each group. This resulted in group sizes of *k* = 35 for the theta, alpha, as well as the control group.

Stimulus material

Stimulus pairs consisting of one image and one sound were selected randomly from an internal database, and the selection was unique for each experimental run. All images featured a resolution of 640 x 480 pixels and a 24-bit color depth. Each image depicted a photograph of either natural or man-made scenes. An additional neutral image depicting a photograph of Earth in space was chosen as the entrainment stimulus to be shown in every trial. We inverted the colors of the entrainment stimulus to increase contrast, thereby increasing the intensity of the stimulation. According to the principle of the Arnold tongue (Pikovsky et al., 2003; Tass et al., 1998), higher stimulation intensity might compensate for a slight frequency mismatch between the entraining signal and the ongoing oscillations in the brain, thus increasing the probability of a successful entrainment. The sounds were real-life recordings of either sounds from nature (e.g. animal calls) or from man-made or artificial environments (e.g. a honk of a car). All sounds were cropped to a duration of 2 s, and featured a bit rate of 1411 kBit/s. All pairings were created in a manner so that no effects of semantic congruency would arise (Parise & Spence, 2012). While it was possible that e.g. animal images could be paired with animal sounds, pairings containing an image of an animal and the corresponding sound of that animal were excluded.

Sensory entrainment

Sensory stimulation was used during the pre-stimulus intervals of encoding trials to manipulate narrow-band oscillatory activity and investigate its effects on subsequent encoding. The entrainment stimulus, which was the same for all participants and all groups, was presented for 2 s before stimulus onset. Specifically, its luminance varied in a sinusoidal manner, resulting in a rhythmic oscillation of 5 Hz in the theta group and 9 Hz in the alpha group. To achieve high temporal resolution of the luminance sine curve, we used a monitor with a frame rate of 240 Hz (Alienware 27 AW2723DF, Dell Technologies, Round Rock, USA). This enabled us to change luminance every 4.2 ms, resulting in luminance change that closely followed a sine curve instead of a box car function. The frequencies of the entrainment signal were determined based on evidence from our previous study (Ostrowski & Rose, 2024), where the peak subsequent memory effects in the pre-stimulus interval were found at 5 Hz in the theta range, as well as 9 Hz in the alpha range. While the luminance in the theta and alpha groups was kept at a steady rhythm in the respective frequencies, the luminance waveforms in the control condition were arhythmic. The waveforms consisted of single cycles of differing frequencies pulled randomly from the interval of 13 to 24 Hz. Importantly, we excluded frequencies of 15 Hz, 18 Hz, and 20 Hz, as these are harmonic frequencies of the entrainment frequencies in the other entrainment conditions. Each participant in the control group was presented with a unique arhythmic waveform with a duration comparable to the 2 s of entrainment in the other groups (M = 1.976 s , SD = 0.022). Notably, the luminance waveform for every group always started and ended at zero luminance (image not visible). To ensure that the stimulus pair would be presented in line with the entrainment rhythm, we implemented a gap of 100 ms between the end of stimulation and stimulus onset in the theta group (56 ms in the alpha group, respectively), which constitutes half of a cycle in the entrainment frquency. In the control condition, this gap was set to 50 ms.

Task and procedure

The experimental procedure was the same regardless of experimental group. After giving informed consent and receiving a short introduction by the experimenter, participants were seated in a sound-attenuated chamber. The experimental session started with a recording of 3.5 minutes of resting-state activity, during which participants were told to fixate a fixation cross on the screen. This was followed by the SME task, which consisted of a short training session and three experimental runs that only differed in the stimulation material presented to the participants. Each experimental run included an encoding phase, an intermission, and a subsequent recognition phase (see **Figure 1**). One encoding phase consisted of 47 trials. During each trial, participants were simultaneously presented with an image and a sound for 2 s. A white fixation cross was visible during stimulus presentation and remained on the screen for 3 to 5 s after stimulus offset. Before stimulus onset, the entrainment stimulus was presented in the respective frequency. Participants were instructed to memorize the combination of image and sound from every trial. Furthermore, participants should indicate whether both the image and sound represented an animal (right mouse button) or not (left mouse button). Button presses were registered as a valid response during the first 3 s after stimulus onset but were otherwise counted as a missed response. The experimental trials were followed by four survey questions measuring the participants' perception of the entrainment procedure. Specifically, the questions measured salience, attention, fatigue, and distractive qualities in relation to the entrainment procedure. During the subsequent intermission of approximately 3 minutes, the participants were asked to count down aloud from 100 (115 and 125 in the second and third run, respectively) in steps of 7 (9 and 13 in the other runs, respectively).

In the recognition phase, the 47 audiovisual pairings from the preceding encoding phase were presented again but intermixed with 47 new pairings, which were created by randomly shuffling the original ones. Note that the individual images and sounds used for the combinations remained the same within each experimental run. All stimulus pairs were again presented for 2 s, with a small white fixation cross layered on top of the image. The fixation cross remained on the screen after stimulus onset. The participants were asked to indicate via button-press whether the current pair had already been presented in the preceding encoding phase (left mouse button) or not (right mouse button). They were further encouraged to press the right mouse button when they felt highly uncertain about a stimulus pair. As in the encoding phase, valid responses were recorded up to 3 s after stimulus onset, and otherwise labeled as a missed response trial. The subsequent inter-trial interval was set to 4 s, during which the white fixation cross was visible on the screen. Across all three experimental runs, participants were presented with 141 unique encoding trials and 282 recognition trials. At the end of the experimental session, resting-state activity was measured again for 3.5 minutes while participants fixated the middle of the screen.

EEG data acquisition and preprocessing

We used a 64-channel electrode setup (ActiCap, BrainProducts, Gilching, Germany) to record EEG. Four of those electrodes were placed on the left and right temple, as well as above and below the left eye, to record vertical and horizontal EOG. The signal was referenced online to *FCz* and re-referenced offline to a common average. The ground electrode was placed at *Iz* below *Oz*, and electrode impedences were kept below 10 kΩ. The signal was amplified with a low cut-off frequency of 0.53 Hz (0.3 s time constant) and recorded at a sampling rate of 500 Hz. EEG activity was recorded during all encoding and recognition phases, but not during intermissions. These settings were used for resting-state recordings as well as for the recordings during the SME task.

Offline preprocessing was done using the Fieldtrip (Oostenveld et al., 2011) and EEGLAB (Delorme & Makeig, 2004) toolboxes for MATLAB (Release 2023a, The Mathworks Inc., Natick, Massachusetts, USA). For the data from the encoding task, an automated approach was used to epoch and clean the data for further processing. The raw data was divided into segments from -3.4 s to 2.5 s relative to the onset of the stimulus pair. A bandpass filter was used to filter out all frequencies outside the range of 0.5 Hz to 40 Hz. Next, trials containing temporally distinct artifacts based on muscular activity or related to electronics were rejected in an automated pipeline using the *ft_artifact_zvalue* function from Fieldtrip. The trial data was filtered, z-transformed, and averaged over channels. An accumulated z-score was computed for each trial based on the types of artifacts. The cutoff value was set to z = 60 for jump artifacts and z = 30 for artifacts caused by phasic muscular activity. Trials were then rejected if the accumulated z-score was larger than the corresponding threshold value. The resulting data was submitted to an automated Independent Component Analysis (ICA) to remove underlying noise from muscular activity as well as artifacts resulting from blinks and eye movements using the *ICLabel* plugin for EEGLAB (Pion-Tonachini et al., 2019). Components that showed at least a probability of 80% of being related to eye-movements, noise caused by muscular activation, or line noise were flagged for removal. On average, 6.37 (SD = 3.8) independent components were removed from the data of the theta group, 6.17 (SD = 3.99) for the alpha group, and 6.63 (SD = 3.32) for the control group. The data was then re-referenced again to the common average. After preprocessing, 4.74 trials (SD = 4.9) out of 141 encoding trials were removed from data sets in the theta group. In the alpha group, an average of 5.14 trials (SD = 3.84) was rejected per participant, while 3.77 trials (SD = 2.65) were rejected in the control group. As the data from the previous study were also analyzed again in the context of the present investigation (Ostrowski & Rose, 2024), all corresponding EEG data were submitted to the same processing pipeline to ensure comparability. After ICA, 4.79 (SD = 2.19) independent components were rejected from the data on average per participant. After preprocessing, an average of 6.42 trials (SD = 4.67) per participant was removed from the data.

From the 3.5 minutes of recorded resting-state activity before and after the experiment, the first and last 15 seconds were omitted for offline processing. Pre- and post-experiment data were processed separately. From this point, we will refer to the data as RestPre and RestPost, respectively. We used a bandpass filter to remove activity below 0.5 Hz and above 40 Hz from the remaining 3-minute interval. The data were then divided into 90 epochs with a length of 2 s each and cleaned from temporally distinct artifacts with the same automated pipeline that was used with the experimental data. Epochs containing artifacts were then removed from the data. The results were submitted to an automated ICA using the same parameters as for the experimental data. On average, 1.11 epochs (SD = 1.3) were removed from RestPre data in the theta group per participant (alpha: 0.91, SD = 1.07; control: 1.2, SD = 1.45). After ICA, 2.91 (SD = 1.8) independent components were rejected per participant in the theta group (alpha: 2.97, SD = 2.35; control: 3.17, SD = 2.16). For RestPost, an average of 1.4 epochs (SD = 2.24) were removed in the theta group per participant (alpha: 0.89, SD = 1.47; control: 1.14, SD = 1.57). On average, 3.74 (SD = 2.78) independent components were rejected (alpha: 3.54, SD = 2.76; control: 5.51, SD = 3.78).

Entrainment validation

As a first step, the pre-processed experimental data from the theta and alpha groups were decomposed into the time-frequency domain. We chose a frequency range of 1 to 40 Hz with frequency bins of 1 Hz, and a time interval of -3.1 s to 2.2 s relative to the onset of the stimulus pairs. Fieldtrip's *mtmconvol* method (Oostenveld et al., 2011) was used in conjunction with a Hanning window of 500 ms and a step size of 100 ms. The additional 300 ms before and after the chosen time interval that were retained during preprocessing served as padding to avoid edge artifacts from the decomposition process. After conducting the decomposition for every trial, the resulting oscillatory power was then averaged over trials for every participant. Next, the data was normalized using a measure of change percentage relative to baseline activity that was defined as the activity from -3.1 s to -2.1 s before stimulus onset. For every individual data set from the theta and alpha group, an average was computed from the data of occipital electrodes (O1, O2, and Oz). As the sensory entrainment took place in the visual domain, the most prominent response should be expected in the electrodes adjacent to the visual cortex. Separate frequency envelopes were chosen for the theta group (5 Hz ± 1 Hz) and the alpha group (9 ± 1 Hz), with a common time interval of interest ranging from -1.1 s to - 0.1 s relative to stimulus onset. We used the latter half of the entrainment interval to estimate entrainment success, as phase alignment and entrainment typically develop over time and tend to plateau after an initial adjustment period (Riecke et al., 2015; Wacker et al., 2011). Assessing the full interval may underestimate entrainment strength due to lower power at the beginning of the stimulation. The entrainment was deemed successful if a relative change in power of at least 10 % could be observed for at least 500 ms within the time interval of interest.

Statistical analysis

Behavioral data

We used a Bayesian framework to test the hypotheses relating to behavioral effects, utilizing the *BayesFactor* package for *R* (v.4.3.3). In line with the signal detection theory (Pastore & Scheirer, 1974; Stanislaw & Todorov, 1999), four percentage measures were extracted for every participant from the recognition data: Correctly remembered old pairings (*hits*), not remembered old pairings (*misses*), new pairings correctly rejected as new (*correct rejections*), and new pairings seemingly remembered as old (*false alarms*). Our main dependent variable, the sensitivity index *d'*, was computed by calculating the difference between the z-transformed hit and false alarm rates for every participant. When group size reached k = 15, we used the Bayesian version of a t-test to statistically compare memory performance between groups. Testing was then repeated every time k increased by one for each group. Specifically, one-sided tests were computed to compare performance between the theta group and controls (H2a), as well as between the alpha group and controls (H2b). For estimating differences between both entrainment groups (H2c), a two-sided test was performed. In all cases, a Cauchy distribution of medium width was used as prior, i.e. with an r scale of $\sqrt{2}/2$.

For the comparison of the experimental groups from the current study with the data from the previous investigation (H3), the NE group data was processed in the same manner. To keep in line with our exclusion criteria, participants with a false positive rate > 25% were not considered in the analysis, resulting in a sample size of k = 45 for the NE group. Due to the difference in group size, a sampling approach was chosen in which a subsample was randomly pulled from the NE data set that matched the group size of the entrainment groups. The average sensitivity was calculated from that subsample and compared to the mean of the full NE sample. This procedure was repeated 50 times. Ultimately, we chose the subsample where the difference in means was minimal, ensuring that the subsample would be representative of the original NE sample. We then conducted one-sided Bayesian t-tests to compare memory performance between both entrainment groups and the previous data set using the same settings as in testing for H2.

Bayesian statistical approaches were further used to explore differences between groups in secondary behavioral variables. A Bayesian one-way ANOVA with the factor *entrainment condition* (theta, alpha, control, NE) was used to assess differences in accuracy and response times from the categorization task during encoding. Changes in memory sensitivity over the course of the experiment were investigated using a mixed-design Bayesian ANOVA with the factors *entrainment condition* and *experimental run* (a, b, c). Furthermore, Bayesian one-way ANOVAs with the factor *entrainment condition* were used to assess group differences for every response category of reaction times. To investigate, whether the entrainment condition modulated the discrepancy in response time between correct and incorrect trials, a mixed-design Bayesian ANOVA with the factors *entrainment condition* and *correctness* was utilized. Finally, Bayesian mixed-design ANOVAs with the factors *entrainment condition* and *experimental run* were used to assess differences in the subjective perception of the entrainment procedure, as well as state of attention and fatigue. The analyses were conducted separately for each survey item. To estimate the relative likelihood of the interactions in the these analyses, the ratio of Bayes factors corresponding to the full model and the model containing only the main effects was computed. For all analyses, a Cauchy distribution with an r scale of $\sqrt{2}/2$ was used as prior.

EEG data

To statistically test the success of entrainment, we compared oscillatory activity from the theta and alpha groups with activity from the control group. Specifically, the baseline-normalized time-frequency data was restricted to the latter half of the entrainment period (-1.1 s to -0.1 s relative to stimulus onset), and the frequency range was set to 1 to 40 Hz. We used a non-parametric permutation testing approach with a cluster-based correction for multiple comparisons as implemented in Fieldtrip (Oostenveld et al., 2011). Independent-samples t-tests were computed for every data point across participants from the channel-time-frequency space. Data points that showed significant differences between conditions ($p < .05$) were organized into clusters based on temporal, spatial, and spectral proximity. For each cluster, statistical values were summed to yield a cluster-level statistic, and the highest of these sums was selected as the principal test statistic for condition comparisons. To construct a reference distribution, a Monte Carlo approach was employed: all trials from both conditions were merged into a single dataset and randomly split into two groups. Statistical testing was performed again at the level of individual data points within these shuffled groups, and cluster-level statistics were recalculated. This randomization process was repeated 4000 times. During each iteration, the largest cluster-level statistics were recorded to generate the null distribution, separately for positive and negative clusters. The final p-value for condition differences was obtained by determining the proportion of randomizations that produced a test statistic greater than that observed in the original data. This method was applied across all detected clusters, yielding a p-value for each cluster's comparison between conditions. The same statistical approach was used for the comparison of pre-stimulus activity from the theta and alpha groups to activity from the NE group. In addition, this approach was also used to assess differences between the entrainment groups (theta, alpha, control) in brain activity during stimulus presentation. However, an independent-samples *F*-test was used on sample level in this case.

A similar statistical approach was used for the exploration of resting-state data. For every entrainment group, the preprocessed RestPre and RestPost data were decomposed into the frequency domain by using the Fast Fourier Transform on single epochs for a frequency range of 1 to 40 Hz. All epoch spectra were then averaged to a subject-specific mean frequency spectrum. This was done separately for RestPre and RestPost. For each of the entrainment groups, we compared RestPre and RestPost with a paired-samples *t*-test on sample level. To further assess whether these differences varied between groups, difference scores were computed between RestPre and RestPost data and submitted to cluster-based permutation testing using an independent-samples *F*-test on sample level. Thresholds for *p*-values were kept as laid out above.

To determine each participant's Individual Alpha Frequency (IAF) and Individual Theta Frequency (ITF), we analyzed resting state EEG data acquired prior to the experimental task. Preprocessed data underwent spectral analysis using a multitaper Fast Fourier Transform (FFT) approach implemented in FieldTrip. Spectral power was computed across frequencies from 1 to 40 Hz in 1 Hz steps, with a 2 Hz smoothing kernel applied. For each participant, we extracted the mean power spectrum across a set of posterior electrodes (Pz, POz, Oz, O1, O2, P3, P4, PO3, PO4). The IAF was identified as the frequency within the alpha range (8–12 Hz) exhibiting the maximum spectral power across these electrodes. Conversely, the ITF was defined as the frequency within the theta range (3–7 Hz) with the highest spectral power. We calculated the absolute differences between each participant's IAF and ITF and their corresponding entrainment frequencies (theta: 5 Hz; alpha: 9 Hz). To assess whether the proximity of an individual's intrinsic frequencies to the stimulation frequencies influenced entrainment efficacy, we conducted Pearson correlation analyses between these frequency-distance measures and the maximum relative change in spectral power at the individual peak channel during stimulation.

Data availability

The raw EEG and behavioral data underlying our findings have been uploaded to an open repository of the University of Hamburg for accessibility (https://doi.org/10.25592/uhhfdm.17620).

Acknowledgements

The present research was supported by the German Research Foundation (DFG; SFB/Transregio 169, Project B3 and DFG RO 2653/9-1). We thank Michele Frerichs and Lennart May for assistance during data acquisition.

Author Contributions

M.R., J.O. and M.M. designed the study. J.O. and M.M. performed data acquisition. J.O. and M.M. analyzed the data. M.R. acquired funding, conceptualized, and supervised the project. J.O., M.M. and M.R. wrote the original manuscript. J.O., M.M., and M.R. reviewed and edited the manuscript and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional Information

Correspondence and requests for materials should be addressed to M.R.

References

Addante, R. J., Watrous, A. J., Yonelinas, A. P., Ekstrom, A. D., & Ranganath, C. (2011). Prestimulus theta activity predicts correct source memory retrieval. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(26), 10702–10707. https://doi.org/10.1073/pnas.1014528108

Albouy, P., Weiss, A., Baillet, S., & Zatorre, R. J. (2017). Selective Entrainment of Theta Oscillations in the Dorsal Stream Causally Enhances Auditory Working Memory Performance. *Neuron*, *94*(1), 193-206.e5. https://doi.org/10.1016/j.neuron.2017.03.015

Amil, A. F., Albesa-González, A., & Verschure, P. F. M. J. (2024). Theta oscillations optimize a speed-precision trade-off in phase coding neurons. *PLOS Computational Biology*, *20*(12), e1012628. https://doi.org/10.1371/journal.pcbi.1012628

Boran, E., Fedele, T., Klaver, P., Hilfiker, P., Stieglitz, L., Grunwald, T., & Sarnthein, J. (2019). Persistent hippocampal neural firing and hippocampal-cortical coupling predict verbal working memory load. *Science Advances*, *5*(3), eaav3687. https://doi.org/10.1126/sciadv.aav3687

Bree, S. van, Sohoglu, E., Davis, M. H., & Zoefel, B. (2021). Sustained neural rhythms reveal endogenous oscillations supporting speech perception. *PLOS Biology*, *19*(2), e3001142. https://doi.org/10.1371/journal.pbio.3001142

Buzsáki, G., & Moser, E. I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature Neuroscience 2013 16:2*, *16*(2), 130–138. https://doi.org/10.1038/nn.3304

Cruzat, J., Torralba, M., Ruzzoli, M., Fernández, A., Deco, G., & Soto-Faraco, S. (2021). The phase of Theta oscillations modulates successful memory formation at encoding. *Neuropsychologia*, *154*, 107775. https://doi.org/10.1016/j.neuropsychologia.2021.107775

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

Dickerson, B. C., & Eichenbaum, H. (2010). The Episodic Memory System: Neurocircuitry and Disorders. *Neuropsychopharmacology*, *35*(1), 86–104. https://doi.org/10.1038/npp.2009.126

Duecker, K., Doelling, K. B., Breska, A., Coffey, E. B. J., Sivarao, D. V., & Zoefel, B. (2024). Challenges and Approaches in the Study of Neural Entrainment. *Journal of Neuroscience*, *44*(40). https://doi.org/10.1523/JNEUROSCI.1234-24.2024

Etter, G., Carmichael, J. E., & Williams, S. (2023). Linking temporal coordination of hippocampal activity to memory function. *Frontiers in Cellular Neuroscience*, *17*. https://doi.org/10.3389/fncel.2023.1233849

Figueiro, M., Nagare, R., & Price, L. (2018). Non-visual effects of light: How to use light to promote circadian entrainment and elicit alertness. *Lighting Research & Technology*, *50*(1), 38–62. https://doi.org/10.1177/1477153517721598

Foxe, J. J., & Snyder, A. C. (2011). The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00154

Gasser, C., & Davachi, L. (2023). Cross-Modal Facilitation of Episodic Memory by Sequential Action Execution. *Psychological Science*, *34*(5), 581–602. https://doi.org/10.1177/09567976231158292

Golmohammadi, R., Yousefi, H., Safarpour Khotbesara, N., Nasrolahi, A., & Kurd, N. (2021). Effects of Light on Attention and Reaction Time: A Systematic Review. *Journal of Research in Health Sciences*, *21*(4), e00529. https://doi.org/10.34172/jrhs.2021.66

Grover, S., Nguyen, J. A., & Reinhart, R. M. G. (2021). Synchronizing Brain Rhythms to Improve Cognition. *Annual Review of Medicine*, *72*(Volume 72, 2021), 29–43. https://doi.org/10.1146/annurev-med-060619-022857

Hanslmayr, S., Axmacher, N., & Inman, C. S. (2019). Modulating Human Memory via Entrainment of Brain Oscillations. *Trends in Neurosciences*, *42*(7), 485–499. https://doi.org/10.1016/j.tins.2019.04.004

Herweg, N. A., Solomon, E. A., & Kahana, M. J. (2020). Theta Oscillations in Human Memory. *Trends in Cognitive Sciences*, *24*(3), 208–227. https://doi.org/10.1016/j.tics.2019.12.006

Hsieh, L. T., & Ranganath, C. (2014). Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. *NeuroImage*, *85*, 721–729. https://doi.org/10.1016/j.neuroimage.2013.08.003

Janssens, S. E. W., Sack, A. T., Ten Oever, S., & de Graaf, T. A. (2022). Calibrating rhythmic stimulation parameters to individual electroencephalography markers: The consistency of individual alpha frequency in practical lab settings. *European Journal of Neuroscience*, *55*(11–12), 3418–3437. https://doi.org/10.1111/ejn.15418

Jeffreys, Harold. (1998). *Theory of probability* (3rd ed.). Oxford University Press.

Jensen, O., & Mazaheri, A. (2010). Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition. *Frontiers in Human Neuroscience*, *4*. https://doi.org/10.3389/fnhum.2010.00186

Kasten, F. H., Duecker, K., Maack, M. C., Meiser, A., & Herrmann, C. S. (2019). Integrating electric field modeling and neuroimaging to explain inter-individual variability of tACS effects. *Nature Communications*, *10*(1), 5427. https://doi.org/10.1038/s41467-019-13417-6

Kizuk, S. A. D., & Mathewson, K. E. (2017). Power and Phase of Alpha Oscillations Reveal an Interaction between Spatial and Temporal Visual Attention. *Journal of Cognitive Neuroscience*, *29*(3), 480–494. https://doi.org/10.1162/jocn_a_01058

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. https://doi.org/10.1016/j.tics.2012.10.007

Klimesch, W., Fellinger, R., & Freunberger, R. (2011). Alpha oscillations and early stages of visual encoding. *Frontiers in Psychology*, *2*(118). https://doi.org/10.3389/fpsyg.2011.00118

Köster, M., & Gruber, T. (2022). Rhythms of human attention and memory: An embedded process perspective. *Frontiers in Human Neuroscience*, *16*. https://doi.org/10.3389/fnhum.2022.905837

Köster, M., Martens, U., & Gruber, T. (2019). Memory entrainment by visually evoked theta-gamma coupling. *NeuroImage*, *188*, 181–187. https://doi.org/10.1016/j.neuroimage.2018.12.002

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.

Leske, S., Endestad, T., Volehaugen, V., Foldal, M. D., Blenkmann, A. O., Solbakk, A.-K., & Danielsen, A. (2025). Beta oscillations predict the envelope sharpness in a rhythmic beat sequence. *Scientific Reports*, *15*(1), 3510. https://doi.org/10.1038/s41598-025-86895-y

Lindenbaum, L., Steppacher, I., Mehlmann, A., & Kissler, J. M. (2023). The effect of neural pre-stimulus oscillations on post-stimulus somatosensory event-related potentials in disorders of consciousness. *Frontiers in Neuroscience*, *17*. https://doi.org/10.3389/fnins.2023.1179228

Lok, R., Smolders, K. C. H. J., Beersma, D. G. M., & de Kort, Y. A. W. (2018). Light, Alertness, and Alerting Effects of White Light: A Literature Overview. *Journal of Biological Rhythms*, *33*(6), 589–601. https://doi.org/10.1177/0748730418796443

Michael, E., Covarrubias, L. S., Leong, V., & Kourtzi, Z. (2022). Learning at your brain's rhythm: Individualized entrainment boosts learning for perceptual decisions. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhac426

Morrow, A., Elias, M., & Samaha, J. (2023). Evaluating the Evidence for the Functional Inhibition Account of Alpha-band Oscillations during Preparatory Attention. *Journal of Cognitive Neuroscience*, *35*(8), 1195–1211. https://doi.org/10.1162/jocn_a_02009

Neuling, T., Ruhnau, P., Fuscà, M., Demarchi, G., Herrmann, C. S., & Weisz, N. (2015). Friends, not foes: Magnetoencephalography as a tool to uncover brain dynamics during transcranial alternating current stimulation. *NeuroImage*, *118*, 406–413. https://doi.org/10.1016/j.neuroimage.2015.06.026

Neuling, T., Ruhnau, P., Weisz, N., Herrmann, C. S., & Demarchi, G. (2017). Faith and oscillations recovered: On analyzing EEG/MEG signals during tACS. *NeuroImage*, *147*, 960–963. https://doi.org/10.1016/j.neuroimage.2016.11.022

Nieuwenhuis, S., & de Kleijn, R. (2013). The impact of alertness on cognitive control. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(6), 1797–1801. https://doi.org/10.1037/a0033980

Notbohm, A., & Herrmann, C. S. (2016). Flicker Regularity Is Crucial for Entrainment of Alpha Oscillations. *Frontiers in Human Neuroscience*, *10*. https://www.frontiersin.org/articles/10.3389/fnhum.2016.00503

Notbohm, A., Kurths, J., & Herrmann, C. S. (2016). Modification of Brain Oscillations via Rhythmic Light Stimulation Provides Evidence for Entrainment but Not for Superposition of Event-Related Responses. *Frontiers in Human Neuroscience*, *10*. https://www.frontiersin.org/articles/10.3389/fnhum.2016.00010

Nyhus, E., & Curran, T. (2010). Functional role of gamma and theta oscillations in episodic memory. *Neuroscience & Biobehavioral Reviews*, *34*(7), 1023–1035. https://doi.org/10.1016/J.NEUBIOREV.2009.12.014

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

Ostrowski, J., & Rose, M. (2024). Increases in pre-stimulus theta and alpha oscillations precede successful encoding of crossmodal associations. *Scientific Reports*, *14*(1), 7895. https://doi.org/10.1038/s41598-024-58227-z

Palva, S., Monto, S., & Palva, J. M. (2010). Graph properties of synchronized cortical networks during visual working memory maintenance. *NeuroImage*, *49*(4), 3257–3268. https://doi.org/10.1016/j.neuroimage.2009.11.031

Parise, C. V., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: A study using the implicit association test. *Experimental Brain Research*, *220*(3), 319–333. https://doi.org/10.1007/s00221-012-3140-6

Parish, G., Hanslmayr, S., & Bowman, H. (2018). The Sync/deSync Model: How a Synchronized Hippocampus and a Desynchronized Neocortex Code Memories. *Journal of Neuroscience*, *38*(14), 3428–3440. https://doi.org/10.1523/JNEUROSCI.2561-17.2018

Pastore, R. E., & Scheirer, C. J. (1974). Signal detection theory: Considerations for general application. *Psychological Bulletin*, *81*(12), 945–958. https://doi.org/10.1037/h0037357

Pikovsky, A., Rosenblum, M., & Kurths, J. (2003). Synchronization: A Universal Concept in Nonlinear Sciences. *Synchronization, by Arkady Pikovsky and Michael Rosenblum and Jürgen Kurths, Pp. 432. ISBN 052153352X. Cambridge, UK: Cambridge University Press, May 2003.*, *2*. https://doi.org/10.4249/scholarpedia.1459

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Poirel, N., Krakowski, C. S., Sayah, S., Pineau, A., Houdé, O., & Borst, G. (2014). Do You Want to See the Tree? Ignore the Forest. *Experimental Psychology*. https://econtent.hogrefe.com/doi/10.1027/1618-3169/a000240

Raud, L., Sneve, M. H., Vidal-Piñeiro, D., Sørensen, Ø., Folvik, L., Ness, H. T., Mowinckel, A. M., Grydeland, H., Walhovd, K. B., & Fjell, A. M. (2023). Hippocampal-cortical functional connectivity during memory encoding and retrieval. *NeuroImage*, *279*, 120309. https://doi.org/10.1016/j.neuroimage.2023.120309

Riecke, L., Sack, A. T., & Schroeder, C. E. (2015). Endogenous Delta/Theta Sound-Brain Phase Entrainment Accelerates the Buildup of Auditory Streaming. *Current Biology*, *25*(24), 3196–3201. https://doi.org/10.1016/j.cub.2015.10.045

Roberts, B. M., Clarke, A., Addante, R. J., & Ranganath, C. (2018). Entrainment enhances theta oscillations and improves episodic memory. *Cognitive Neuroscience*, *9*(3–4), 181–193. https://doi.org/10.1080/17588928.2018.1521386

Rudoler, J. H., Herweg, N. A., & Kahana, M. J. (2023). Hippocampal Theta and Episodic Memory. *Journal of Neuroscience*, *43*(4), 613–620. https://doi.org/10.1523/JNEUROSCI.1045-22.2022

Salari, N., & Rose, M. (2016). Dissociation of the functional relevance of different pre-stimulus oscillatory activity for memory formation. *NeuroImage*, *125*, 1013–1021. https://doi.org/10.1016/j.neuroimage.2015.10.037

Schneider, S. L., & Rose, M. (2016). Intention to encode boosts memory-related pre-stimulus EEG beta power. *NeuroImage*, *125*, 978–987. https://doi.org/10.1016/J.NEUROIMAGE.2015.11.024

Scholz, S., Schneider, S. L., & Rose, M. (2017). Differential effects of ongoing EEG beta and theta power on memory formation. *PLOS ONE*, *12*(2), e0171913. https://doi.org/10.1371/journal.pone.0171913

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*(1), 128–142. https://doi.org/10.3758/s13423-017-1230-y

Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annual Review of Physiology*, *55*, 349–374.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

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

Stecher, H. I., Pollok, T. M., Strüber, D., Sobotka, F., & Herrmann, C. S. (2017). Ten Minutes of α-tACS and Ambient Illumination Independently Modulate EEG α-Power. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00257

Strunk, J., & Duarte, A. (2019). Prestimulus and poststimulus oscillatory activity predicts successful episodic encoding for both young and older adults. *Neurobiology of Aging*, *77*, 1–12. https://doi.org/10.1016/j.neurobiolaging.2019.01.005

Sweeney-Reed, C. M., Zaehle, T., Voges, J., Schmitt, F. C., Buentjen, L., Kopitzki, K., Richardson-Klavehn, A., Hinrichs, H., Heinze, H.-J., Knight, R. T., & Rugg, M. D. (2016). Pre-stimulus thalamic theta power predicts human memory formation. *NeuroImage*, *138*, 100–108. https://doi.org/10.1016/j.neuroimage.2016.05.042

Taesler, P., & Rose, M. (2022). Multivariate prediction of pain perception based on pre-stimulus activity. *Scientific Reports*, *12*(1), 1–9. https://doi.org/10.1038/s41598-022-07208-1

Tass, P., Rosenblum, M. G., Weule, J., Kurths, J., Pikovsky, A., Volkmann, J., Schnitzler, A., & Freund, H.-J. (1998). Detection of N:m Phase Locking from Noisy Data: Application to Magnetoencephalography. *Physical Review Letters*, *81*(15), 3291–3294. https://doi.org/10.1103/PhysRevLett.81.3291

Terada, S., Sakurai, Y., Nakahara, H., & Fujisawa, S. (2017). Temporal and Rate Coding for Discrete Event Sequences in the Hippocampus. *Neuron*, *94*(6), 1248-1262.e4. https://doi.org/10.1016/J.NEURON.2017.05.024

Terporten, R., Schoffelen, J.-M., Dai, B., Hagoort, P., & Kösem, A. (2019). The Relation between Alpha/Beta Oscillations and the Encoding of Sentence induced Contextual Information. *Scientific Reports*, *9*(1), 20255. https://doi.org/10.1038/s41598-019-56600-x

Thut, G., Schyns, P., & Gross, J. (2011). Entrainment of Perceptually Relevant Brain Oscillations by Non-Invasive Rhythmic Stimulation of the Human Brain. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00170

Tian, Y., Zhou, H., Zhang, H., & Li, T. (2021). Research on Differential Brain Networks before and after WM Training under Different Frequency Band Oscillations. *Neural Plasticity*, *2021*(1), 6628021. https://doi.org/10.1155/2021/6628021

Uncapher, M. R., Hutchinson, J. B., & Wagner, A. D. (2011). Dissociable Effects of Top-Down and Bottom-Up Attention during Episodic Encoding. *Journal of Neuroscience*, *31*(35), 12613–12628. https://doi.org/10.1523/JNEUROSCI.0152-11.2011

Van Dijk, H., Schoffelen, J. M., Oostenveld, R., & Jensen, O. (2008). Prestimulus Oscillatory Activity in the Alpha Band Predicts Visual Discrimination Ability. *Journal of Neuroscience*, *28*(8), 1816–1823. https://doi.org/10.1523/JNEUROSCI.1853-07.2008

Veniero, D., Vossen, A., Gross, J., & Thut, G. (2015). Lasting EEG/MEG Aftereffects of Rhythmic Transcranial Brain Stimulation: Level of Control Over Oscillatory Network Activity. *Frontiers in Cellular Neuroscience*, *9*. https://doi.org/10.3389/fncel.2015.00477

Wacker, M., Galicki, M., Putsche, P., Milde, T., Schwab, K., Haueisen, J., Ligges, C., & Witte, H. (2011). A Time-Variant Processing Approach for the Analysis of Alpha and Gamma MEG Oscillations During Flicker Stimulus Generated Entrainment. *IEEE Transactions on Biomedical Engineering*, *58*(11), 3069–3077. https://doi.org/10.1109/TBME.2011.2160640

Waldhauser, G. T., Johansson, M., & Hanslmayr, S. (2012). Alpha/Beta Oscillations Indicate Inhibition of Interfering Visual Memories. *Journal of Neuroscience*, *32*(6), 1953–1961. https://doi.org/10.1523/JNEUROSCI.4201-11.2012

Wälti, M. J., Woolley, D. G., & Wenderoth, N. (2020). Assessing Rhythmic Visual Entrainment and Reinstatement of Brain Oscillations to Modulate Memory Performance. *Frontiers in Behavioral Neuroscience*, *14*. https://doi.org/10.3389/fnbeh.2020.00118

Wang, D., Clouter, A., Chen, Q., Shapiro, K. L., & Hanslmayr, S. (2018). Single-trial phase entrainment of theta oscillations in sensory regions predicts human associative memory performance. *Journal of Neuroscience*, *38*(28), 6299–6309. https://doi.org/10.1523/JNEUROSCI.0349-18.2018

Wang, D., Marcantoni, E., Clouter, A., Shapiro, K. L., & Hanslmayr, S. (2024). Rhythmic sensory stimulation as a noninvasive tool to study plasticity mechanisms in human episodic memory. *Current Opinion in Behavioral Sciences*, *58*, 101412. https://doi.org/10.1016/j.cobeha.2024.101412

Weinbach, N., & Henik, A. (2012). The relationship between alertness and executive control. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1530–1540. https://doi.org/10.1037/a0027875

Winterling, S. L., Shields, S. M., & Rose, M. (2019). Reduced memory-related ongoing oscillatory activity in healthy older adults. *Neurobiology of Aging*, *79*, 1–10. https://doi.org/10.1016/j.neurobiolaging.2019.03.012

Yeh, N., & Rose, N. S. (2019). How Can Transcranial Magnetic Stimulation Be Used to Modulate Episodic Memory?: A Systematic Review and Meta-Analysis. *Frontiers in Psychology*, *10*, 993. https://doi.org/10.3389/fpsyg.2019.00993

Zaehle, T., Rach, S., & Herrmann, C. S. (2010). Transcranial Alternating Current Stimulation Enhances Individual Alpha Activity in Human EEG. *PLOS ONE*, *5*(11), e13766. https://doi.org/10.1371/journal.pone.0013766

Zazio, A., Ruhnau, P., Weisz, N., & Wutz, A. (2022). Pre-stimulus alpha-band power and phase fluctuations originate from different neural sources and exert distinct impact on stimulus-evoked responses. *European Journal of Neuroscience*, *55*(11–12), 3178–3190. https://doi.org/10.1111/ejn.15138

Zoefel, B., & VanRullen, R. (2017). Oscillatory Mechanisms of Stimulus Processing and Selection in the Visual and Auditory Systems: State-of-the-Art, Speculations and Suggestions. *Frontiers in Neuroscience*, *11*(MAY). https://doi.org/10.3389/FNINS.2017.00296

Disentangling the Functional Roles of Pre-Stimulus Oscillations in Crossmodal Associative Memory Formation via Sensory Entrainment

Supplementary material

Jan Ostrowski[1†], Marike C. Maack[1†], Michael Rose[1*]

1 *Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany;* j.ostrowski@uke.de, m.maack@uke.de, rose@uke.de

† Equal contributions

** Correspondence:* rose@uke.de

16-digit ORCID:
JO: 0000-0001-9928-272X
MM: 0000-0002-6960-4465
MR: 0000-0002-9789-7066

Supplementary 1. Contrasting pre-stimulus activity from theta and alpha groups to activity from the control and NE groups

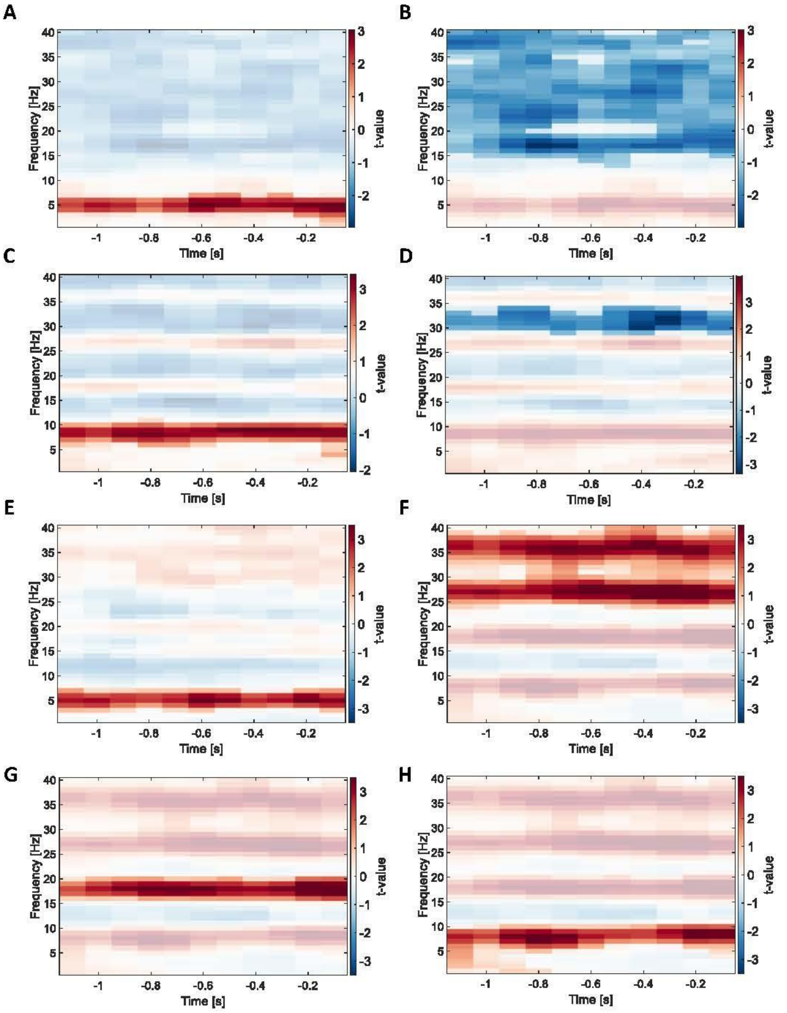


Figure S1. Results of EEG data contrasts of entrainment groups with the control and NE groups. The figure shows time-frequency plots depicting the results of the statistical comparison of relative change in pre-stimulus activity. **(A)** and **(B)** show the two significant clusters from the comparison of the theta group with the control group. **(C)** and **(D)** depict the two significant clusters resulting from contrasting the alpha group with the control group. **(E)** shows the positive cluster revealed by comparing activity from the theta group with the NE group. **(F) – (G)** depict the statistical results comparing the alpha group with the NE group. In all time-frequency plots, positive *t*-values signify greater relative change in the theta or alpha groups, respectively. Opaque data points show the extent of a statistically significant cluster ($p < .025$, corrected). Each subplot shows one distinct cluster and depicts the *t*-values averaged over the electrodes comprising the cluster.

Oscillatory power in the late entrainment period (-1.1 s to -0.1 s relative to stimulus onset) from the theta and alpha group was each contrasted with the activity from the NE group in the same time period. We used a cluster-based permutation approach to account for multiple comparisons, and two-tailed independent-samples *t*-tests on sample level. The frequency range for this analysis was set to 1 to 40 Hz, and all electrodes were included. Comparing the theta entrainment condition with the NE group, the analysis yielded one significant positive cluster in the frequency range of 3 to 7 Hz, spanning the whole analysis window ($p < .025$, corrected). This suggests significantly increased oscillatory power in the envelope around 5 Hz for the theta group as compared to the NE group **(Figure S2E)**. Contrasting activity from the alpha group with the NE group revealed three distinct positive clusters, each spanning the whole analysis window. The clusters covered the frequency ranges of 24 to 40 Hz and 16 to 20 Hz. Importantly, the third cluster ranged from 1 to 10 Hz up until -0.8 s relative to stimulus onset and was centered one the 9 Hz envelope for the remaining part of the analysis time window (**Figure S2H**).

Supplementary 2. Analysis of oscillatory activity in the stimulus presentation window during encoding

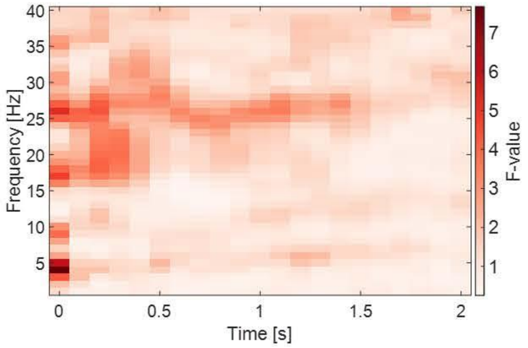


Figure S2. Difference in post-stimulus activity between entrainment groups. The figure shows a time-frequency plot of the stimulus presentation time window from the encoding task over a frequency range of 1 to 40 Hz. The color dimension displays the F-values from the independent-samples F-test.

In order to assess potential differences in oscillatory activity during stimulus presentation between groups, we compared oscillatory power from the post-stimulus interval (0 s to 2s relative to stimulus onset) among the entrainment groups (theta, alpha, and control) using an independent-samples *F*-test on sample level. Data was included for a frequency range of 1 to 40 Hz and all electrodes, and cluster-based permutation was used for multiple-comparison corrections. However, the analysis showed only a tendency for a significant cluster in the electrode-frequency-time space, suggesting that there are no significant differences in post-stimulus power between the entrainment groups (p = .069, corrected).

Supplementary 3. Analysis of categorization task performance during encoding

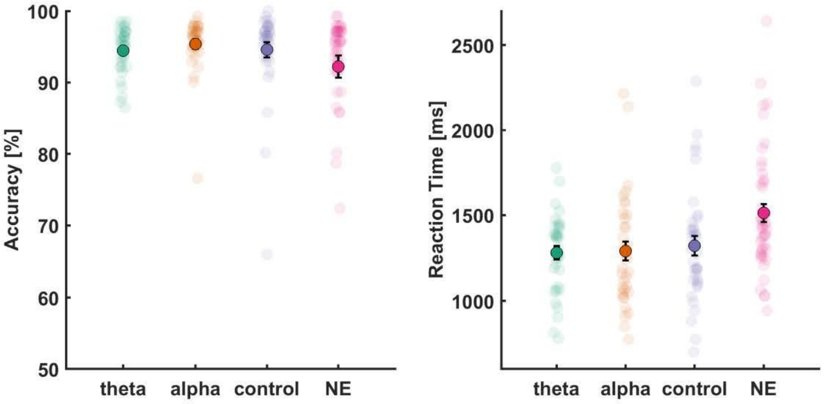


Figure S3. Performance in the categorization task during encoding. *Left*: Accuracy values for the categorization task during encoding for each group. *Right:* Average response time for the categorization task during encoding. Transparent data points mark individual task accuracy, and the black error bars signify the standard error of means.

In the categorization task from the encoding phase, participants showed high accuracy across all entrainment conditions, 94.45% (SD = 3.23) for theta, 95.38% (SD = 3.88) for alpha, and 94.57% (SD = 6.27) for control condition. Accuracy in the NE condition as an additional control was slightly lower at 92.21% (SD = 10.43). The overall accuracy across all four conditions was 94.02% (SD = 6.97). Using a Bayesian one-way ANOVA model yielded a Bayes factor of $BF_{10} = 0.381$, suggesting moderate evidence in support of the null hypothesis of no significant differences among the groups. These results suggest that participants in the current study maintained high compliance with task demands throughout the experiment, which was essential for accurately assessing the subsequent impact of oscillatory activity on memory performance. Reaction times showed a similar pattern. Participants responded fastest in the entrainment conditions (theta: 1279.4 ms, SD = 230.3; alpha: 1288.9 ms, SD = 327.6; control: 1320.4 ms, SD = 343.7), with slower responses in the NE condition (1512.6 ms, SD = 355.3). The overall average reaction time was 1361.11 ms (SD = 333.1). However, a Bayesian one-way ANOVA indicated moderate evidence for the alternative hypothesis, $BF_{10} = 5.0737$, indicating measurable differences among the groups. Individual group contrasts revealed that there was likely no difference in response times between the entrainment groups (theta vs control: $BF_{10} = 0.2849$; alpha vs control: $BF_{10} = 0.2628$; theta vs alpha: $BF_{10} = 0.248$). However, the evidence suggests a moderate-to-strong effect for differences between the theta and alpha groups and the NE group (theta vs NE: $BF_{10} = 19.1684$; alpha vs NE: $BF_{10} = 5.7022$). The comparison between the control and the NE group yielded only weak evidence for significant difference, $BF_{10} = 2.3366$.

Supplementary 4. ***Changes in performance over the course of the experiment***

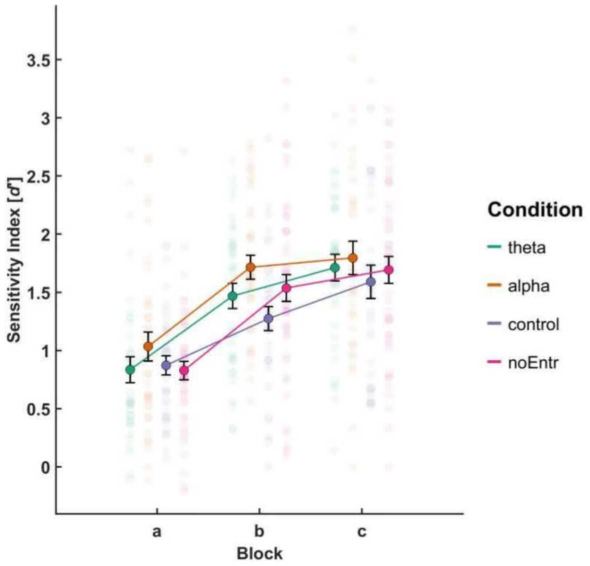


Figure S4. Changes in sensitivity indices across experimental runs. The plot depicts mean sensitivity indices (*d'*) over participants for every group and the three experimental runs. Every transparent data point marks the individual sensitivity index of one participant. Black error bars indicate the standard error of means.

To assess changes in behavioral performance over time, we conducted a mixed-design Bayesian ANOVA with the within-subjects factor *block* (a, b, c) and the between-subjects factor *entrainment condition* (theta, alpha, control, NE), both as fixed factors. The participant ID was included as a random effect. The best-supported model included *block* and the participant ID, BF = 1.58×10^{35}, indicating extreme evidence for a main effect of *block*. Adding *entrainment condition* reduced model support by a factor of approximately 4.6 (BF = 3.43×10^{34}), while including the *block* x *entrainment condition* interaction further reduced support by a factor of approximately 135, BF = 1.17×10^{33}. The model with only *entrainment condition* and participant ID was 7.3 times less likely than the null model (BF = 0.14), providing strong evidence against a main effect of *entrainment condition*. This indicates that the improvement in memory performance over the course of the experiment was consistent across entrainment conditions and was not modified by the type of entrainment.

Supplementary 5. ***Individual group and variable contrasts for response times during recognition***

Table S1

Individual group comparisons of RTs for all response categories

Response category	Group 1	M [ms]	SD	Group 2	M [ms]	SD	*BF*^{a}
hit	alpha	1405	165	control	1450	206	0.3789
	alpha	1405	165	NE	1612	272	47.4767**
	alpha	1405	165	theta	1427	156	0.2825
	control	1450	206	NE	1612	272	4.279*
	control	1450	206	theta	1427	156	0.2773
	NE	1612	272	theta	1427	156	18.9926**
miss	alpha	1580	257	control	1519	271	0.3652
	alpha	1580	257	NE	1771	344	3.3256*
	alpha	1580	257	theta	1573	283	0.2472
	control	1519	271	NE	1771	344	19.4755**
	control	1519	271	theta	1573	283	0.326
	NE	1771	344	theta	1573	283	3.3623*
CR	alpha	1397	174	control	1417	236	0.265
	alpha	1397	174	NE	1609	296	31.4731**
	alpha	1397	174	theta	1414	185	0.2642
	control	1417	236	NE	1609	296	6.9499*
	control	1417	236	theta	1414	185	0.2463
	NE	1609	296	theta	1414	185	13.1874**
FA	alpha	1669	222	control	1675	309	0.2469
	alpha	1669	222	NE	1885	310	37.8379**
	alpha	1669	222	theta	1718	230	0.3493
	control	1675	309	NE	1885	310	9.4932*
	control	1675	309	theta	1718	230	0.2961
	NE	1885	310	theta	1718	230	5.2933*

Note. This table shows the results from the statistical analysis of reaction times based on entrainment group differences. CR = correct rejection, FA = false alarm, M = mean (arithmetic), SD = standard deviation, df = degrees of freedom

a The depicted Bayes factor values are equivalent to BF_{10}, estimating the evidence for the alternative hypothesis (statistical difference) relative to the null hypothesis (no difference). A Cauchy distribution of medium width was used as prior for each comparison.

* at least moderate evidence for H_1, ** at least strong evidence for H_1

Supplementary 6. *Resting State*

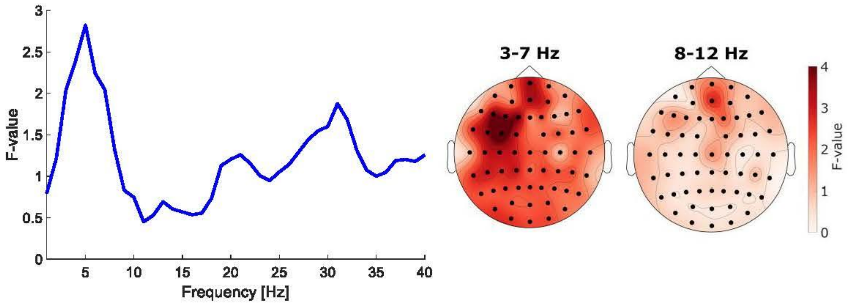


Figure S5. Group contrast for differences in resting-state activity before and after the experiment. *Left*: F-values for the analysis window of 1 to 40 Hz from a cluster-based permutation test with an independent-samples *F*-test on sample level assessing group differences in pre- and post-experiment resting-state discrepancies. . *Right*: Topographical distribution of *F*-values averaged across the 3–7 Hz (theta) and 8–12 Hz (alpha) bands. No significant clusters were observed in this analysis (p = .2972, corrected).

As some studies report lingering oscillatory effects due to entrainment procedures **[Kasten & Herrmann, 2022; Gallina et al., 2023)**, we explored differences in resting-state spectra that were recorded once before (RestPre) and once after the SME task (RestPost) to determine whether traces of the entrainment could be observed even after the experiment. For the analysis, we used power spectra in the frequency range of 1 to 40 Hz and submitted the data to a cluster-based permutation test with two-tailed paired-samples t-tests on the sample level. Note that this analysis was conducted separately for the theta group, alpha group, as well as the control group. For the theta group, one negative cluster was found in the alpha band (8 to 12 Hz), indicating increased power after the experiment (p < .025, corrected). Similarly, a negative cluster ranging from 7 to 18 Hz was observed for the comparison in the alpha group (p < .025, corrected), while the analysis in the control revealed a negative cluster in the alpha band (8 - 12 Hz, p < .025, corrected). Interestingly, the control group analysis yielded a second negative cluster in the beta band, ranging from 17 to 33 Hz (p < .025, corrected). As the effect in the alpha band and, to a certain degree, in the beta band was observed in all three groups, we were interested in whether the effect magnitude differed between the groups.

Supplementary 7. ***No differences in subjective perception of entrainment***

Survey items (translated from German into English):

I1 (pleasentness): How pleasant did you find the flickering of the image?

I2 (distraction): To what extent did you feel distracted by the flickering of the image while trying to remember the pairs?

I3 (attention): How would you rate your level of attention during the task?

I4 (fatigue): How exhausted do you feel at the moment?

Participants rated on a scale from 0 (*not at all*) to 5 (*very much*) in steps of 0.5.

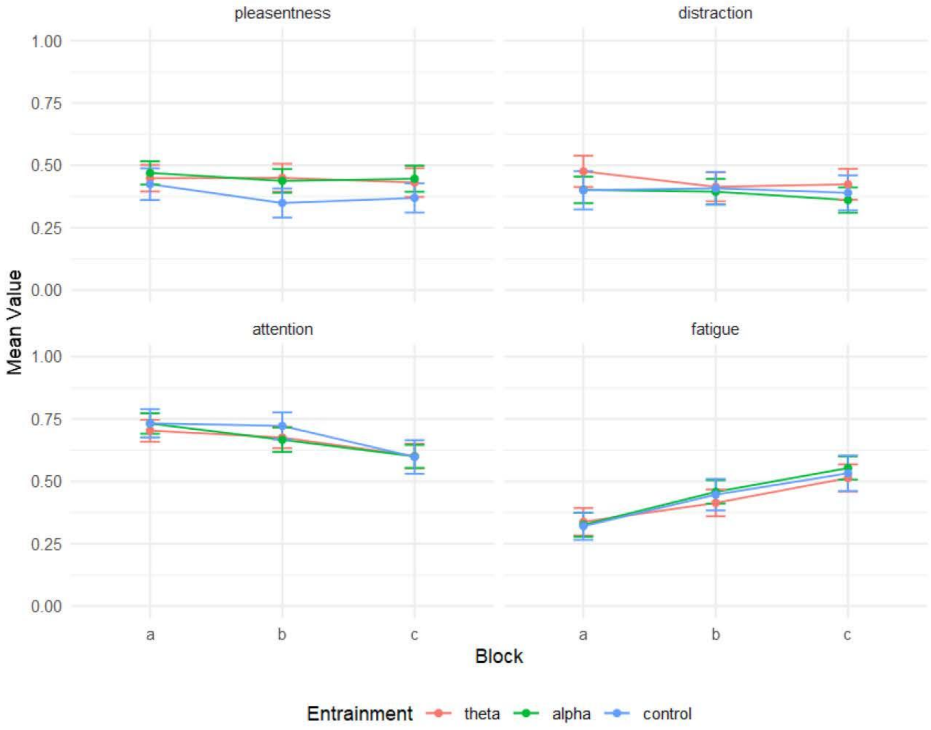


Figure S6. Subjective ratings of task-related experience across blocks and entrainment conditions. The data points depict group and block averages across participants. The error bars mark the standard error of means.

As individual perception qualities of images with oscillating luminance may vary, we investigated whether the subjective perception of the sensory stimulation might differ between the entrainment groups to control for salience effects. Participants received four survey items after each encoding phase, measuring the *pleasantness* and the *distracting qualities* of the entrainment procedure, as well as *attention* and *fatigue*. We conducted a Bayesian mixed-design ANOVA for every item, with a between-subjects factor *pre-stimulus condition* (theta, alpha, control) and a within-subjects factor *block* (A, B, C). Scores from every item did not significantly differ between levels of *pre-stimulus condition* ($BF_{\textbf{pleasentness}} = 0.3099$, $BF_{\textbf{distraction}} = 0.3853$, $BF_{\textbf{attention}} = 0.1132$, $BF_{\textbf{fatigue}} = 0.1279$), indicating that the type of entrainment procedure had no differential effect. However, the analysis revealed an effect of *block* for the variables *distraction, attention,* and *fatigue* ($BF_{\textbf{distraction}} = 7.8834$, $BF_{\textbf{attention}} = 1.855 \times 10^{9}$, $BF_{\textbf{fatigue}} = 1.3415 \times 10^{15}$). No interactions of *pre-stimulus condition* and *block* were observed ($BF_{\textbf{pleasentness}} = 0.0732$, $BF_{\textbf{distraction}} = 0.0578$, $BF_{\textbf{attention}} = 0.195$, $BF_{\textbf{fatigue}} = 0.0552$. Participants felt less distracted by the entrainment in block C of the experiment (M = 0.382, SD = 0.224) than in block A (M = 0.45, SD = 0.243). Conversely, participants rated their level of attention in block C (M = 0.573, SD = 0.203) consistently lower than in block A (M = 0.694, SD = 0.186). This was accompanied by increased fatigue ratings in block C (M = 0.528, SD = 0.223) as compared to block A (M = 0.338, SD = 0.197). In sum, evidence from the survey data indicates that the entrainment procedures were received equally pleasant and distracting, suggesting no confound of the behavioral results due to subjective perception.

Appendix C: Study III

The order of multisensory associative sequences is reinstated as context feature during successful recognition. Maack, M. C., Ostrowski, J., & Rose, M. (2025). *Scientific Reports*, 15(1), 1-20.

www.nature.com/scientificreports

scientific reports

OPEN

The order of multisensory associative sequences is reinstated as context feature during successful recognition

Marike Christiane Maack, Jan Ostrowski & Michael Rose

The ability of the human brain to encode and recognize sequential information from different sensory modalities is key to memory formation. The sequence in which these modalities are presented during encoding critically affects recognition. This study investigates the encoding of sensory modality sequences and its neural impact on recognition using multivariate pattern analysis (MVPA) of oscillatory EEG activity. We examined the reinstatement of multisensory episode-specific sequences in *n* = 32 participants who encoded sound-image associations (e.g., the image of a ship with the sound of a frog). Images and sounds were natural scenes and 2-second real-life sounds, presented sequentially during encoding. During recognition, stimulus pairs were presented simultaneously, and classification was used to test whether the modality sequence order could be decoded as a contextual feature in memory. Oscillatory results identified a distinct neural signature during successful retrieval, associated with the original modality sequence. Furthermore, MVPA successfully decoded neural patterns of different modality sequences, hinting at specific memory traces. These findings suggest that the sequence in which sensory modalities are encoded forms a neural signature, affecting later recognition. This study provides novel insights into the relationship between modality encoding and recognition, with broad implications for cognitive neuroscience and memory research.

The ability to remember episodes from the past is a cornerstone of human memory. Episodes consist of multiple features that may stem from different modalities (e.g., visual, auditory), reaching us in specific sequences. In line, it has been shown that humans encode not only the semantic content but also the temporal order (sequence) of features, which is crucial for recalling the flow of a past episode[1,2]. The capacity to encode and retrieve sequential information allows us to mentally rebuild the dynamic structure of events, highlighting the role of temporal context in episodic memory[1,3]. Recalling the order that features of an episode were originally encoded in, is based on sequential reinstatement[4].

Episodic memory relies on the integration of contextual information during encoding, with the hippocampus playing a key role in binding event features such as sensory modality and spatial-temporal context[5–7]. Prior research has investigated how we remember the temporal order of events, including the role of unimodal cueing[8–10]. Evidence suggests that episodic memory involves temporal compression and event segmentation, where the hippocampus supports memory organization by structuring event sequences and contextual boundaries[9,11]. Temporal compression refers to the tendency of episodic memory to condense events during recall, influenced by event segmentation and goal-directed actions[12–15]. Event segmentation, in turn, affects how temporal order is remembered, as events chunked at perceptual boundaries enhance object-context binding but may reduce precise temporal order memory[9,16]. The hippocampus further supports the encoding and retrieval of event sequences, integrating spatial and temporal contexts essential for remembering event order[17,18]. In line, episodic memory retrieval is shaped by the availability of contextual information at encoding, with reinstatement of encoding context enhancing recognition in providing characteristic cues that mitigate interference[19–21]. Sensory modality sequences, as part of contextual information, contribute to the organization of memory representations[22,23]. The hippocampus is crucial for integrating these contextual elements, facilitating recognition by reactivating modality-based associations rather than strictly reconstructing event sequences in order[18,24,25].

While previous studies have explored unimodal cueing and temporal order memory, the current study specifically investigates whether the modality sequence acts as a contextual feature that influences encoding and

Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Martinistr. 52, Building W34, 20248 Hamburg, Germany. email: rose@uke.de

recognition rather than sequential recall per se. Contextual cues, including sensory modality sequences, can mitigate interference effects and improve memory performance by reinstating elements of the original encoding context[21,26,27]. Conversely, mismatches between encoding and retrieval contexts may lead to competition between overlapping memory traces, impairing recall[4,28,29]. Both physical and mental reinstatement of contextual information facilitate episodic retrieval, with mental reconstruction yielding comparable benefits to direct environmental cues[30,31]. Furthermore, cognitive control plays a role in sustaining contextual reinstatement, as individuals with higher working memory capacity are better able to maintain contextual associations through strategic memory processes[32].

Beyond unimodal sequences, results from the animal, as well as human studies, suggest that sequential information plays a critical role in recalling multisensory episodes, where encoded features from different sensory modalities (e.g., an image-sound pair) are reinstated during retrieval[33–36]. Multisensory features enable the brain to store and retrieve information across various perceptual domains, enhancing our ability to recall past experiences and make informed decisions[37–39]. As such, the human memory has evolved to function optimally under multisensory conditions[40]. Interestingly, multisensory memories can be cued by unimodal features upon remembering, suggesting that multisensory encoding enhances subsequent recognition[39,41]. This means that one modality may serve as a cue to retrieve the complete episode, even if it includes features from different modalities[42]. It is, however, unclear so far whether and how one modality (i.e., auditory) relates to the reinstatement of another (i.e., visual). The Scene Reconstruction Theory suggests that especially the hippocampus helps to reconstruct memories by integrating various sensory details associated with an event[4]. This integration is supposed to allow richer, and more detailed, memory formation, as the brain can draw upon multiple sensory inputs to create a coherent narrative of the event[11,43].

While neural plasticity allows multisensory learning[44], understanding the neural reinstatement of sequential information across sensory modalities remains crucial for elucidating the broader mechanisms of remembering. Recall generally involves the activation and reconstruction of neural pathways tied to previously encoded details[45]. This process is influenced by retrieval cues and familiarity with the material[46]. Beyond sequential and multisensory reinstatement, the recollection of contextual details surrounding past events, such as where and when they occurred, significantly contribute to the liveliness and specificity of the memory representation[47–50]. The Context Maintenance and Retrieval (CMR) model here provides a framework to understand how the brain organizes memories around contextual and temporal cues, facilitating the accurate retrieval of episode-specific feature sequences[51,52], propelling the reconstruction of the correct order and updating of associations. In line, previous research has highlighted that the process of episodic remembering involves not only recalling specific items but also reinstating the contextual details of the original event[53–55]. These studies have demonstrated that the success of memory retrieval is closely associated with the reactivation of the encoding-related memory trace[52,56–58]. Interestingly, it has been shown that during memory reactivation, not only episode-specific features but also contextual features, that were not directly related to the current memory task, are reinstated[29,59–65]. However, the precise neural mechanisms underlying this sequential reinstatement, especially in multisensory contexts, remain elusive.

EEG is a powerful tool to track these reinstatement processes including multisensory episode-specific feature sequences. Here, especially multivariate temporal-pattern analysis has emerged as the gold-standard to examine how neural activity during retrieval reflects (sequential) reinstatement. Moreover, it represents a tool for investigating the role of context reinstatement in memory processes, revealing the (beneficial) effects of reinstating neural encoding patterns in memory retrieval[30,64]. Accordingly, multivariate pattern analysis (MVPA) has been used to decode oscillatory activity patterns during memory retrieval, successfully classifying specific neural signatures tied to remembering[30,64]. Here, low-frequency activity (e.g., beta (13–30 Hz) and theta (3–7 Hz) oscillations) have been shown to be particularly important for episodic memory processes, facilitating successful retrieval[66–70]. Multisensory inputs from different modalities as well as their sequential encoding thereby enrich the formation of stable memory traces[39,41]. As multivariate approaches have shown that neural pattern reinstatement is indeed associated with episode-specific feature sequences, this study aims to investigate the oscillatory mechanisms underlying the retrieval of sequential information in human memory, focusing on whether the modality sequence in which information was presented during encoding is reinstated during recognition. Specifically, we utilize EEG and MVPA to classify the neural patterns associated with modality sequence reinstatement, providing new insights into the role of oscillatory activity in organizing and retrieving sequential memories across different sensory modalities. Importantly, the current study aims to demonstrate how modality sequence functions as a contextual feature during encoding and recognition (but not temporal reinstatement in the sense of reactivating the sequential order during retrieval; i.e., first visual, second auditory). This design offers a novel perspective on context memory in representing a sharp contrast to previous designs, which employed a parallel presentation of modalities during encoding[71,72].

Methods

Participants

Thirty-six healthy participants were recruited for the experiment. Data from four participants were excluded due to a high number of missing trials ($n = 1$) and poor behavioral performance ($n = 3$). Outliers in terms of memory performance (d') were identified and excluded if they exceeded ± 3 absolute deviations from the median (MAD;[73]. Therefore, the final sample included $n = 32$ participants (19 females, 52.77% female) with a mean age of 24.25 years (SD = 3.34), ranging from 18 to 33 years. All participants had normal or corrected-to-normal vision and hearing ability and reported no neurological or psychiatric diseases. They gave written informed consent and received financial reimbursement for participating in the study. The Hamburg Medical Council ethics committee (PV5893) approved this investigation. We confirm that all research was performed in accordance with relevant guidelines and regulations.

Task and procedure

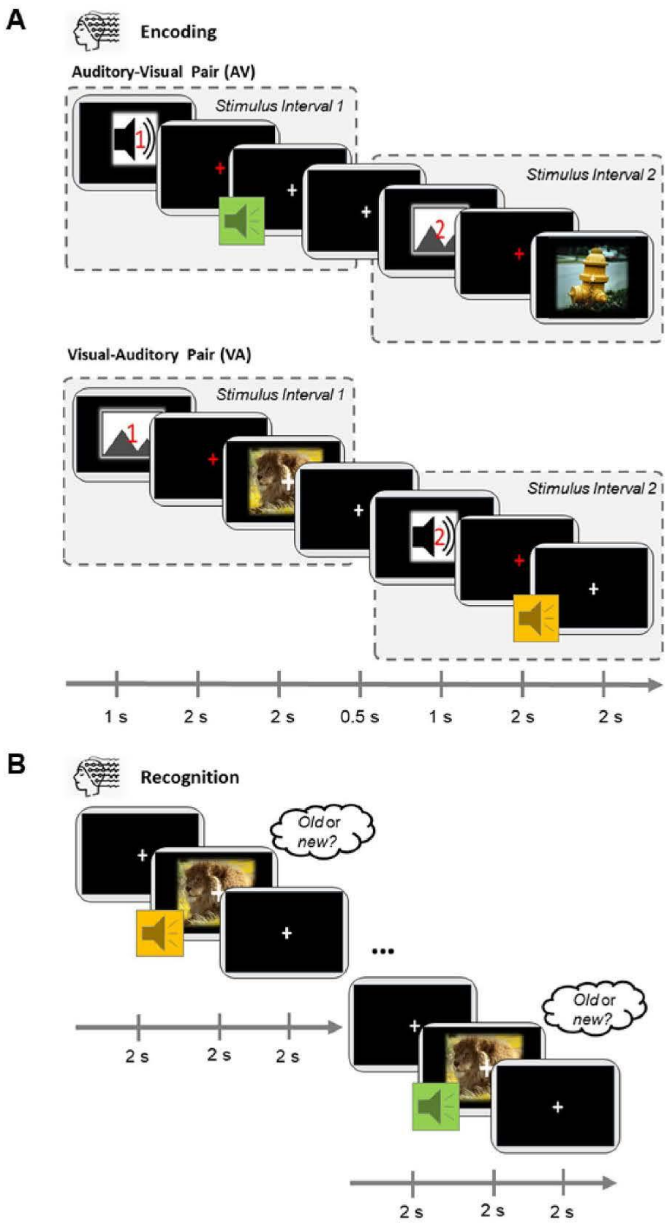
We implemented an explicit sequential associative memory task consisting of an encoding task, a short intermission, and a subsequent recognition task. In order to measure associative multisensory memory, participants were presented sequential image-sound pairs during the encoding task. The images (resolution: 640 × 480 pixels; 24-bit color depth) and sounds (length: 2 s; bitrate: 1411kBit/s) were randomly selected from an internal stimulus database, presenting real-life objects, animals, and landscapes. The individual stimuli were paired pseudo-randomly (semantically congruent stimuli pairs were excluded). Each pair was presented once during the encoding task. In each encoding trial, the stimulus pairs were presented sequentially. Each trial started with a 1-second *modality cue* indicating whether a visual or an auditory stimulus would be presented as the first pair component (Fig. 1). The *modality cue* was represented as an icon, cueing the following modality (image or tone). Additionally, the icon involved a red number, stating whether the next stimulus would be the first or second stimulus of the pair. The *modality cue* was followed by a red fixation cross for 2 s, which served as a visual cue for the upcoming stimulus. Afterwards, the first stimulus was presented, followed by a 500 ms Inter-Stimulus-Interval (ISI). Complementary to the first stimulus presentation, a second *modality cue* indicated whether a visual or an auditory stimulus would be presented as the second stimulus. Again, a red fixation cross served as a visual cue for the upcoming stimulus followed by the announced stimulus and the inter-trial interval (ITI). The ITI was jittered between 3 s and 5 s. Trials were counterbalanced for *modality-order*, such that the encoding task consisted of the same number of visual-auditory (VA) as auditory-visual (AV) trials. We applied a modality cueing procedure, including both modality-specific cues and numerical indicators, to clearly signal the upcoming auditory and visual pairings, thereby facilitating robust and explicit encoding of the stimulus associations. All participants were explicitly instructed to memorize the stimulus pairs, and not to focus on the individual images or sounds. This ensured that subsequent memory would later reflect associative memory, but not item memory. We divided the experiment into three consecutive blocks, each consisting of an encoding and subsequent recognition task. During each encoding task, 47 stimulus pairs were presented. After the encoding task, a short 3-minute intermission followed, during which participants were asked to count down aloud from 100 (115 and 125 in the second and third run, respectively) in steps of 7 (9 and 13 in the second and third run, respectively). In the subsequent recognition task, participants were presented with the 47 stimulus pairs shown during the previous encoding phase, as well as 47 new pairs. Both components of a pair were presented in parallel, in sharp contrast to the encoding task, where the components were shown sequentially. New pairs consisted of the same individual components that rendered the pairs from the encoding task, but were shuffled to create 47 new pairs. The participants were asked to indicate via button press whether the presented pairs were already known from the encoding task or not. Stimulus presentation lasted for 2 s, and no cue was used. The inter-trial interval was fixed to 5 s. The recognition task was followed by a short break of 3 to 5 min, followed by the encoding task of the next run. Across the three blocks, 141 stimulus pairs were presented during encoding, and 282 were presented during recognition in total.

EEG data acquisition

EEG data were collected using a 64-channel Ag/AgCl active electrode system (ActiCap64; BrainProducts, Gilching, Germany), arranged in accordance with the extended 10–20 system[74]. Sixty electrodes were positioned at the most central scalp locations. To facilitate offline artifact removal, a bidirectional bipolar electrooculogram (EOG) was concurrently recorded using the remaining four electrodes. These bipolar EOG electrode pairs were positioned above and below the left eye, as well as at the lateral ends of the bicanthal plane. FCz served as the reference electrode for data acquisition, while the ground electrode was situated at position Iz. Signals were digitized at a sampling rate of 500 Hz and was amplified with a low cut-off frequency of 0.53 (0.3 s time constant). Impedances were maintained below 10 kΩ throughout the recording session.

EEG preprocessing and time-frequency decomposition

The acquired EEG data were preprocessed offline using the FieldTrip toolbox[75] in MATLAB (Release 2022a, The MathWorks Inc., Natick, Massachusetts, USA). For each participant, the encoding and the recognition task were analyzed separately. For the recognition task, epochs were extracted from – 2500 ms to 3500 ms relative to stimulus onset, resulting in a trial duration of 5 s. A high-pass filter at 0.5 Hz was applied to remove extreme low-frequency fluctuations. The data were visually inspected, and trials containing artifacts, such as high-frequency noise indicating muscular activity or spikes resembling poor electrode connections, were removed. Independent Component Analysis (ICA) was used to identify components corresponding to blinks and other ocular activity, and the data were corrected accordingly. On average, 4.8 components (SD = 3.31) were removed per subject. The data were then visually inspected again, and trials with artefacts were excluded. The remaining trials for each subject were split into correct old (remembered; mean = 83.11; SD = 26.40), incorrect old (forgotten; mean = 46.94; SD = 24.08), correctly rejected (mean = 112.46; SD = 17.05) and false alarm (mean = 18.60; SD = 14.78) trials. The correct old trials were additionally split for the "visual-auditory" (VA) and "auditory-visual" (AV) conditions according to modality order in the encoding task, resulting in a similar number of trials in each condition (44 trials on average in each condition; AV: mean = 44.75, SD = 10.79; VA: mean = 44.41, SD = 11.02). Time-frequency decomposition was performed from – 1 to 2 s relative to stimulus onset, covering the 1 to 40 Hz frequency range. This was achieved using a multitaper convolution approach with a sliding Hanning window of 500 ms and a 100 ms step size. Although a 500 ms window has an intrinsic frequency resolution of 2 Hz (1/0.5 s), we applied zero-padding to the maximum trial length, which interpolated the spectrum to a 1 Hz grid, thereby effectively balancing temporal precision with frequency resolution. No baseline correction was applied for the recognition tasks, as the primary focus was on within-subject differences in oscillatory power between the remembered and forgotten pairs. We did not apply baseline correction to preserve potential pre-stimulus effects, which have been implicated in learning and memory processes. Additionally, baseline correction assumes a proportional scaling



of oscillatory and non-oscillatory (1/f) activity, an assumption that may not always hold. As a result, the reported effects should be interpreted in terms of relative power changes rather than absolute polarity shifts. Additionally, for the data of the recognition task, the resulting frequency spectra (electrode × frequency × time) from stimulus onset onward, reflecting the processing phase, were used to predict the modality order in the encoding task for each participant and condition. Only trials with successfully remembered items were included in the analysis, while pre-stimulus activity was not included. For the multivariate analyses, the Matlab toolbox for classification and regression of multi-dimensional data (MVPA-Light;[76] was used. A support vector machine (SVM) with a

◀**Fig. 1**. Schematic overview of the encoding task in both sequence variations as well as the following recognition task. (**A**) The two different modality order sequences in the encoding task. Stimuli were presented sequentially. The trial started with a cue, indicating which stimulus modality is presented first followed by a red fixation cross. The red fixation cross indicated that a stimulus would be presented in the next 2 s. Afterwards, a complex natural image was presented. The presentation of the second stimulus of the pair followed the same procedure. We utilized two different encoding sequence orders: either the sound was present first and the image afterwards (auditory-visual; AV-condition or the image was presented first and the sound afterwards (visual-auditory; VA-condition). For the analysis of the encoding task, the task was subdivided into Stimulus Intervals 1 and 2, indicated by the dashed boxes. (**B**) One example trial from the recognition task. Unlike the encoding phase, stimulus pairs were presented simultaneously. Participants had to identify whether the pair **had** been presented during encoding (old) or not (new) within 4 s. The recognition task pairs were composed of the stimuli from the encoding resulting in new and old pairs.

k = 5-fold cross-validation was used for classification with a five-time repetition. The classifier was trained on all electrodes using single-trial frequency spectra.

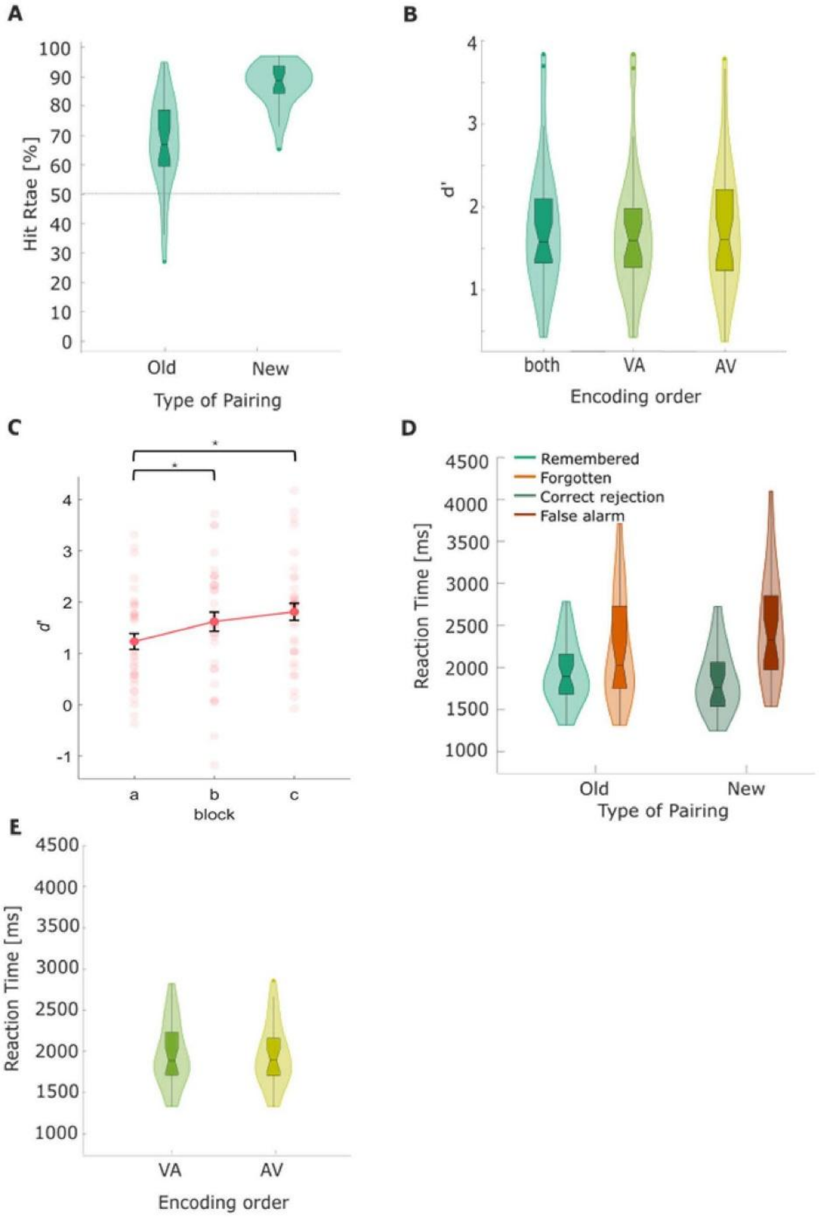
The preprocessing routine and the time-frequency analysis for the encoding task were the same as for the recognition task, with deviations during epoching. The deviation in epoching procedure resulted from the difference in presentation mode. While sound-image associations were presented simultaneously in the recognition task, the associations were presented sequentially in the encoding task. Epochs were extracted from the first modality cue until 2 s after the onset of the second stimulus, resulting in a trial duration of 10.5 s. Using ICA, on average 4.8 independent components were rejected from each individual data set (SD = 2.61). After preprocessing, the remaining trials for each subject were split into a "visual-auditory" and "auditory-visual" condition according to modality order in the encoding task, resulting in a similar number of trials in each condition. These trials were then used in the subsequent analyses of later remembered and forgotten trials (55 trials on average in each condition). After time-frequency decomposition, data were averaged separately for remembered and forgotten trials for each participant.

Statistical analysis

In the recognition task, the percentages of remembered old pairs (hits), correctly rejected new pairs, forgotten old pairs, and falsely remembered new pairs (false alarms) were extracted. In order to index memory performance, we utilized the sensitivity measure d', which is the difference between the z-transformed hit and false alarm rates[77–79]. A one-sample t-test against zero was conducted to probe associative memory formation with *d'* as dependent variable. Additionally, a repeated-measures ANOVA was used to analyze reaction times during the recognition task, with the within-subject factors *Pairing* (old vs. new) and *Correctness* (correct: remembered or correct rejection vs. incorrect: forgotten or false alarm).

Statistical analysis of the time-frequency EEG data acquired during the recognition phase of the experiment was conducted to explore the potential effect of remembered and forgotten trials within the low-frequency spectrum. This analysis was further differentiated by modality-independent sequence order during encoding, as well as sequential auditory-visual and visual-auditory presentations. Employing a non-parametric permutation testing approach with cluster-based correction for multiple comparisons, as implemented in the FieldTrip toolbox[75], we statistically compared time-frequency representations corresponding to remembered trials against those of forgotten trials. The cluster-based permutation test defines a cluster as a set of contiguous significant points in a three-dimensional space comprising electrode location (spatial dimension), frequency, and time. A significant point is identified when the test statistic at a specific electrode, frequency, and time point surpasses a predefined threshold ($p < .01$, uncorrected). These significant points are then grouped into clusters based on their adjacency. This means that a cluster can extend across multiple electrodes, frequency bands, and time points, rather than being restricted to a single dimension. We used the electrode neighbourhood structure defined by Fieldtrip function ft_prepare_neighbours to determine adjacency, ensuring that spatially close electrodes are considered neighbors. In the frequency and time dimensions, adjacency is defined by consecutive frequency bins and time points. To be classified as a cluster, an effect needed to span at least two neighbouring electrodes., preventing isolated effects in one electrode from being classified as clusters. Multiple comparison correction was applied using the cluster-based permutation test (*cfg.correctm = 'cluster', cfg.method = 'montecarlo'*), which controls the family-wise error rate (FWER). This means that while individual points initially pass a cluster-forming threshold ($p < .01$, uncorrected), the final significance of a cluster is determined via a permutation-based correction ($p < .05$ two-tailed, cluster-corrected). Thus, only clusters that survived this multiple comparison correction are reported as significant. Subsequently, Monte Carlo method was utilized to generate a distribution of t-values[80].

The main focus of the study was to probe neuronal reinstatement of the stimulus modality order (visual-auditory, auditory-visual) during subsequent recognition via MVPA. To investigate this effect, we employed MVPA in the time-frequency domain of the recognition task. The SVM classifier distinguished between the two different encoding stimulus modality orders (AV/VA) based on the EEG data from the recognition task in which the previously encoded associations were presented simultaneously. To comprehensively capture neural processing during the recognition phase, we applied MVPA to the entire a priori defined dataset, analyzing the full trial period during recognition from 0 to 2 s relative to stimulus onset, across the 1–40 Hz frequency range and all 60 electrodes. Classification accuracy was assessed using a single–subject k = 5-fold cross-validation procedure, ensuring that model training and testing were performed on separate data splits within each subject to reduce overfitting and improve generalization. We conducted the MVPA on remembered trials of the



recognition task to classify the two different modality orders in which the stimulus pairs were presented during the encoding task (*Modality Sequence Classifier*). To assess the statistical significance of the accuracy achieved by the *Modality Sequence Classifier*, we conducted a one-sample *t*-test (one-tailed) comparing the overall classifier accuracy against chance-level (50%) The classifier accuracy was determined by averaging the individual classifier performance within the entire analysis window. This methodology enabled us to assess whether the performance of the *Modality Sequence Classifier* significantly exceeds the chance classification level, thus providing insight into the presence of meaningful patterns associated with memory retrieval as opposed to random classification. Subsequently, to probe statistically significant accuracy of the *Modality Sequence Classifier* across the entire

◀**Fig. 2.** Behavioral results from the recognition task. (**A**) Hit-rate in the recognition task for old and newly rearranged pairs. Within the boxplots, the horizontal lines indicate the median of the subset, while the notch around the median represents its 95% confidence interval. The upper and lower edges indicate quartiles 1 and 3. (**B**) The distribution of memory performance (*d'*) overall, as well as separated between the two modality-order sequences during the encoding task (VA = visual-auditory, AV = auditory-visual). (C) The block effect for memory performance (d'). Asterisks indicate significant differences between blocks. $p < .05(*)$. (**D**) Distribution of reaction times for the respective response categories from the recognition task, split for remembered, forgotten, correct rejection and false alarm trials. (**E**) Distribution of reaction times for the respective response categories from the recognition task, split between the two modality-order sequences during the encoding task (VA = visual-auditory, AV = auditory-visual) for the remembered trials.

analysis window, we performed a cluster-based permutation t-test, comparing the accuracy values from the *Modality Sequence Classifier* to chance level (50%). To further examine frequency-specific effects, accuracy values were averaged within predefined frequency bands: theta (3–7 Hz), alpha (8–13 Hz), low beta (13–21 Hz), high beta (22–32 Hz), beta (13–32 Hz), and gamma (32–40 Hz). We determined the electrode with the highest mean t-value for each band and computed individual mean accuracy values at these electrodes.

Furthermore, a correlational analysis was conducted to examine the relationship between the oscillatory power contrast between remembered and forgotten trials during encoding task and d' (memory performance) estimates from the recognition task. This analysis is expected to yield valuable insights into multisensory processing during associative learning and memory. Drawing on previous research[81,82], we expected power differences to correlate with the memory performance within the low-frequency bands during the presentation of the first and second stimuli. Specifically, these power differences were expected to vary according to the modality sequence order (AV vs. VA condition). To accomplish this objective, we employed a non-parametric cluster-based permutation technique for the correlation analysis to control for alpha-error inflation. Here, we calculated the neuronal activity power differences (remembered > forgotten; AV/VA condition) and correlated these with the behavioural measure d'. Cluster-based permutation t-tests were conducted for each time-frequency data point across channels and participants. Significant differences between conditions ($p < .05$) resulted in adjacent data points being grouped into clusters based on temporal, spatial, and spectral criteria.

Results

Successful acquisition of sequentially encoded multisensory associative pairs

Participants completed one recognition task after each encoding task, which consisted of 47 multisensory associative pairs (Fig. 1). In the recognition task, participants were presented pairs of images and sounds in parallel, which were sequentially presented during previous encoding. This task included previously presented pairs (*old*) and tested memory specificity by presenting newly formed pairs, consisting of old stimuli elements (*new*). Participants had to indicate whether they remembered the presented stimulus pair from the encoding task (old) or whether it was a newly rearranged pair, consisting of an old image with a sound previously paired with another image (new). Overall, participants performed very well in remembering *old* pairs, with an average hit rate of M(SD) = 67.19 (± 15.29%; Fig. 2A). Although the *new* pairs consisted of images and sounds from encoding that were now rearranged, the false alarm rate was low (M(SD) = 12.12 (± 7.68%)). Accordingly, signal detection theory-based analysis confirmed robust learning, expressed by an average associative d' of M(SD) = 1.76 (± 0.76) independent of stimulus modality order ($t_{(30)} = 12.69$, $p < .001$, *cohen's* $d = 0.59$). The d' estimates of auditory-visual (AV; M(SD) = 1.74 ± 0.76) and visual-auditory stimuli (VA; M(SD) = 1.77 ± 0.77) did not differ significantly ($t_{(30)}$ = -0.14, $p = .884$, see Fig. 2B). Taken together, behavioral results confirmed the successful acquisition of sequentially encoded multisensory associative pairs, with no differences in performance due to modality order.

Our study was intentionally designed as an explicit learning and memory paradigm, where participants were specifically instructed to remember the pairs. Given the experimental structure, it was expected that participants adapt their strategies over time, mainly as they were aware that a recognition test followed each block. Crucially, due to the design, these improvements were likely to occur consistently across participants, regardless of their overall performance. Therefore, this effect should not be seen as a systematic bias but rather as an inherent characteristic of explicit learning. To assess whether behavioural performance changed across blocks, we conducted a repeated measures ANOVA on the performance measure d′, which revealed a significant main effect of Block, $F_{(2, 70)} = 8.92, p < .001, \eta^2 = 0.05$. Given that Mauchly's test indicated a violation of sphericity ($W = 0.92$, $p = .239$), we applied the Greenhouse-Geisser correction ($\varepsilon = 0.925$), yielding a corrected significance value of $p < .001$. Post hoc tests further revealed that performance significantly differed between Block A and Block B ($p = .025$) as well as Block A and Block C ($p = .002$; Fig. 2C).

Next, we compared participants' reaction times during recognition as an index of memory confidence[83,84]. A two-way repeated-measures ANOVA revealed a significant main effect of the factor Correctness ($F_{(1,31)} = 69.84$, $p < .001$, $\eta^2 = 0.69$; Fig. 2D), indicating faster responses during remembered pairs, compared to forgotten. Furthermore, this analysis revealed a main effect of Pairing ($F_{(1,31)} = 9.24$, $p = .005$, $\eta^2 = 0.23$; Fig. 2D), indicating faster responses to old pairs, compared to new. Critically, we observed a significant Pairing × Correctness interaction ($F_{(1,31)} = 15.56$, $p = .001$, $\eta^2 = 0.33$; see Table 1). The post-hoc *t*-test revealed a significant decrease in reaction times for correctly recognised old pairs ($t_{(126)}$ = -0.90, $p < .001$, *Cohen's* $d = -0.44$), indicating an increase in memory confidence in light of correctly retrieved associations compared to forgotten. The reaction time corresponding to auditory-visual (AV; RT = 1944.6 ± 381.7 ms) and visual-auditory stimuli (VA; RT = 1959.4 ± 385.9 ms) did not differ significantly ($t_{(62)} = 0.15$, $p = .877$, see Fig. 2E). Interestingly, participants correctly rejected new pairs significantly faster compared to mistakenly categorizing them as old (false alarm;

Effect	F(df)	p	Partial η^2
Pairing (Old vs. New)	9.23(1,31)	0.005	0.23
Correctness (Correct vs. Incorrect)	69.84(1,31)	< 0.001	0.69
Pairing × Correctness	15.56(1,31)	< 0.001	0.33

Table 1. Repeated-Measures ANOVA results for reaction times.

Comparison	t(df)	p	p_{corr}
Remembered vs. forgotten	-2.26(62)	0.027	0.162
Remembered vs. correct rejected	0.81(62)	0.421	1
Remembered vs. false alarm	-4.30(62)	< 0.001***	< 0.001***
Correct rejected vs. forgotten	-2.83(62)	0.006**	0.036*
Correct rejected vs. false alarm	-4.88(62)	< 0.001***	< 0.001***
Forgotten vs. false alarm	-1.63(62)	0.110	0.660

Table 2. Post-hoc t-test results for reaction times. p_{corr} relates to p-values after Bonferroni correction. $p < .05$(*), $p < .01$(**), $p < .001$ (***).

$t_{(62)} = -4.88$, $p < .001$), supporting the idea that decision confidence might influence RTs in recognition memory. Also, the response in correctly rejected trials were significantly faster as compared to response in forgotten trials, $t_{(62)} = -2.84$, $p = .006$, suggesting that correctly detecting novel information is easier than failing to recognize old pairs. Responses to forgotten trials did not significantly differ in reaction times as compared to false alarm trials ($t_{(62)} = -1.63$, $p = .110$; Table 2).

Successful recognition of multisensory associations relies on alpha/beta oscillations

In the next step, we investigated the oscillatory power differences between remembered and forgotten multisensory pairs within the recognition task, which presented the previously sequentially encoded pairs simultaneously. Here, participants had to indicate whether the presented pair was previously shown in sequential order during encoding. In the following analysis, we split the recognition trials according to their modality-sequence during encoding (VA/AV), and conducted time-frequency analyses. In VA associations (remembered > forgotten) we observed a significant negative cluster of oscillatory activity covering the high theta to low beta frequency range (0.8–1.2 s after stimulus onset; 7–23 Hz; negative cluster: $p < .003$, SD = 0.001). This indicates that remembered pairs, which were represented in a VA sequence order during encoding, were associated with alpha and beta power during memory retrieval as compared to forgotten pairs. This effect was primarily driven by activity differences in frontotemporal and lateral-occipital areas (0.8–1.2 s after stimulus onset, 7–23 Hz; Fig. 3A). Interestingly, AV trials showed a different pattern, including a negative cluster in the theta and alpha range (0.5–1.9 s; 6–13 Hz; negative cluster 1: $p < .002$, SD = 0.009; negative cluster 2: $p < .045$, SD = 0.005), indicating that remembered pairs, which were represented in an AV sequence order during encoding, were associated with theta and alpha power during memory retrieval as compared to forgotten pairs. This effect was primarily driven by activity differences in parietal-occipital areas (0.5–1.9 s after stimulus onset, 6–13 Hz; Fig. 3B). The results indicate differential processes concerning the oscillatory processing of modality-sequences during recognition and were used to restrict the following MVPA analysis. When analyzing both conditions (combining AV and VA) together as an independent modality-sequence condition, we observed similar significant neuronal activity effects, further reinforcing the underlying processing patterns across modalities. The results revealed a negative cluster covering the theta (median cluster size = 10), alpha (median cluster size = 27), and beta (median cluster size = 18) bands, occurring 0.9 to 1.8 s after stimulus onset (negative cluster: $p < .001$, SD = 0.004; see Fig. 3C). This indicates that remembered pairs were associated with lower theta, alpha, and beta power during memory retrieval as compared to forgotten pairs. This effect was primarily driven by activity differences in frontotemporal and lateral-central areas (1.0 to 1.7 s after stimulus onset, 3–7 Hz) and in lateral parietal regions (0.9 to 1.7 s after stimulus onset, 8–18 Hz; see Fig. 3C).

Modality-sequences are reinstated as context-features during recognition

The neuronal signature within the recognition task indicated differential processing of the stimulus modality order, which we hypothesized to also be expressed as context-specific features of the underlying memory trace (i.e. the temporal sequence of the stimulus from different modalities). Accordingly, our main hypothesis stated that the neural signature during the recognition task would reflect the modality sequence in which the pairs were presented during the encoding task. To probe this effect, we employed MVPA, moving beyond the univariate comparisons of oscillatory power. The *Modality Sequence Classifier* distinguished between the two different encoding stimulus modality orders (AV/VA) based on the EEG data from the recognition task in which the previously encoded associations were presented simultaneously. Classification accuracy was assessed using a cross-validation (k = 5) procedure. The overall mean classifier performance for the *Modality Sequence Classifier* was 52.26%, which significantly exceeded the chance level of 50% ($t_{(31)} = 4.28$, $p < .001$, *cohen's d* = 0.76).

For the more detailed classification analyses, we calculated the classification accuracy over all electrodes for each frequency band within the significant data points, as well as the average classification accuracy for the electrode with the maximum mean t-value of the entire time within the specific frequency range. The cluster-based permutation test comparing MVPA accuracy values for remembered trials when decoding the two modality orders against chance level (50%) revealed eight significant positive clusters (most prominent cluster: $p < .001$, cluster-level statistic $t = 29.069$, SD = 0.0003, CI range = 0.0006). This cluster encompassed 45.302 data points out of a total of 50.400 (60 channels × 40 frequency × 21 time), covering approximately 89.88% of the analyzed search space. This significant cluster extended from stimulus onset to 2 s post-stimulus, spanning frequencies from 1 to 40 Hz, and was distributed across all electrodes, with the lowest representation in P7 (84.28%) and the highest in FC4 (93.69%). The remaining clusters had p-values ranging from 0.017 to 0.047, with their respective cluster statistics and confidence intervals indicating robust effects across multiple frequency bands and time points. No significant negative clusters were detected.

To further interpret the decoding performance, average accuracy values were computed over the significant time points identified by the cluster-based permutation test, focusing on time points and electrodes with the highest mean t-values. To identify the electrodes exhibiting the strongest effects, we computed the mean t-value for each electrode by averaging all significant t-values ($p < .05$) across frequency bands and time points. This metric served as an index of the relative effect size at each electrode, highlighting regions that consistently demonstrated robust neural discrimination effects in the cluster-based permutation test. By focusing on electrodes with the highest mean t-values, we aimed to characterize the spatial distribution of the most pronounced neural decoding effects during the recognition task. The overall mean accuracy across all frequencies (1–40 Hz) was 52.26%, with the highest mean t-value observed at electrode F4 ($t = 4.18$). The mean individual accuracy at this electrode was 52.32%, while the highest individual accuracy value reached 62.38% at 0.8 s and 30 Hz. The maximal t-value was reached at F4 ($t = 6.58$).

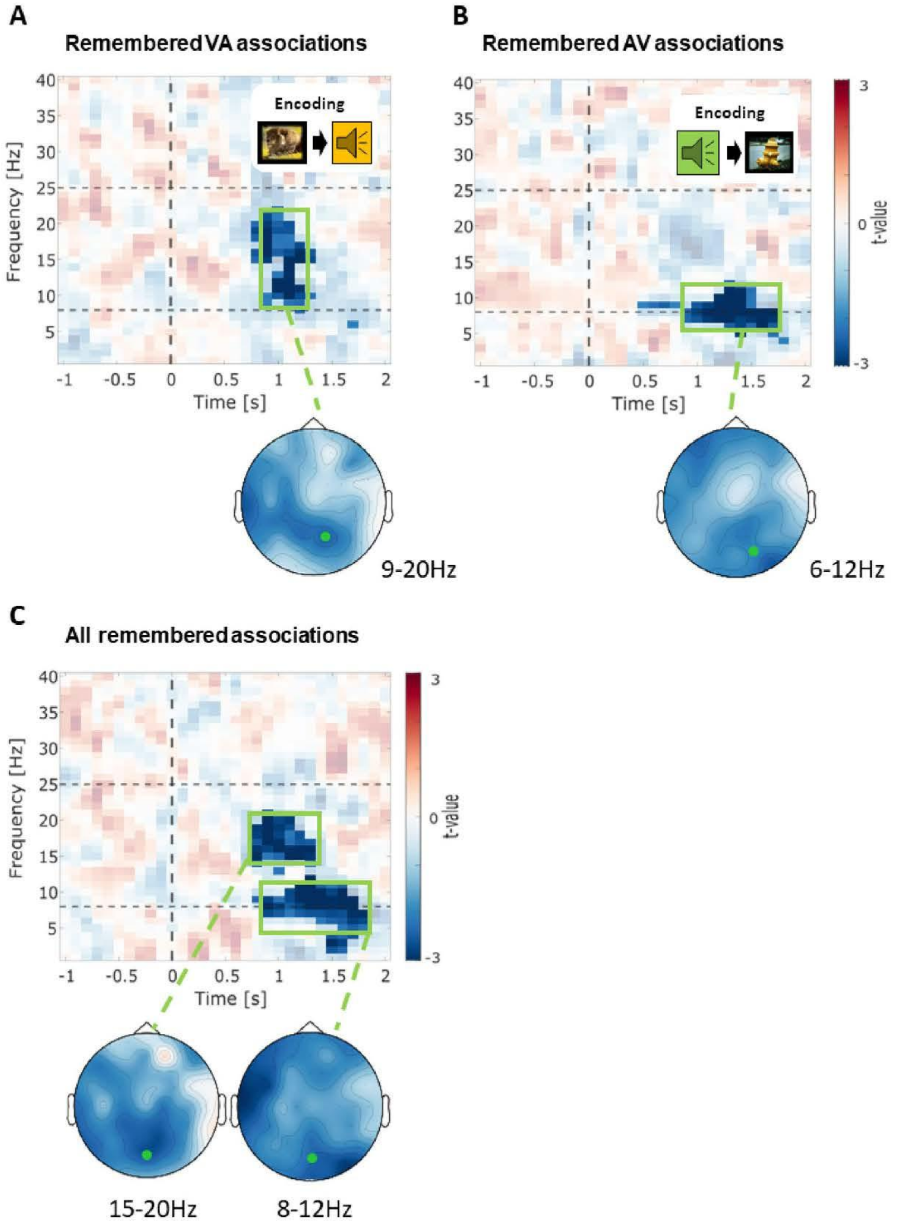
When examining specific frequency bands, the theta range (3–7 Hz) yielded the highest mean accuracy of 52.49%, with the strongest effect at electrode AF7 ($t = 4.22$), and an individual peak accuracy of 58.80% at 0.4 s and 7 Hz. The maximal t-value was reached at AF7 ($t = 6.29$). In the alpha band (8–13 Hz), the mean accuracy was 52.40%, with the highest t-value recorded at electrode Fpz ($t = 3.92$), and an individual maximum accuracy of 58.67% at 0.1 s and 11 Hz. The maximal t-value was reached at F4 ($t = 6.58$). The low beta range (13–21 Hz) showed a mean accuracy of 52.29%, with the most significant effect at electrode Fpz ($t = 3.80$) and a peak accuracy of 58.63% at 0.4 s and 21 Hz. The maximal t-value was reached at FC2 ($t = 5.93$). Similarly, the high beta range (22–32 Hz) demonstrated a mean accuracy of 52.22%, with the strongest effect at electrode P1 ($t = 4.27$) and a peak accuracy of 58.29% at 1.9 s and 24 Hz. The maximal t-value was reached at CP2 ($t = 6.31$).

When considering the full beta range (13–32 Hz), the mean accuracy was 52.24%, with the highest t-value at P1 ($t = 4.07$) and an individual maximum accuracy of 61.00% at 1.9 s and 24 Hz. The maximal t-value was reached at CP2 ($t = 6.31$). The mean accuracy in the gamma band (32–40 Hz) was 52.23%, with the effect at electrode P2 ($t = 4.69$) and an individual peak accuracy of 52.98% at 0.1 s and 33 Hz. The maximal t-value was reached at PO8 ($t = 6.18$). All results are summarised in Table 3. These findings suggest that decoding performance was significantly above chance level across multiple time points, frequencies, and electrode locations, with particularly strong effects in the theta, alpha, and beta bands.

To further explore the spatial distribution of significant effects, we visualized topographical maps of t-values obtained from the cluster-based permutation test for different frequency bands (theta, alpha, low beta, high beta, and gamma) over the trial time course (Fig. 4). Interestingly, the higher beta and gamma frequency range (21–40 Hz) exhibited pronounced discriminative power between the two conditions during the entire stimulus presentation over the centro-parietal electrodes, specifically strongest at the beginning (0.4 to 1.3 s) and the end (1.7 to 2.0 s). Additionally, also early (0 to 1.0 s to stimulus onset) centro-frontal electrode cluster in the lower frequency range (3 to 20 Hz) showed a significant classification performance. The results demonstrate a convergence between the neural activity patterns during successful recognition and retrospective discrimination between the modality order of the sequential encoding during recognition. In sum, these findings highlight the ability to encode and differentiate VA from AV sequences during the retrieval task, as evidenced by the distinct neural patterns observed in the EEG data during subsequent recognition. In addition to the averaged accuracy values from electrodes with the highest mean t-values, time-resolved decoding accuracy (MVPA) for remembered trials is shown at exemplary electrodes across different frequency bands (Fig. 5).

Successful recognition of multisensory associations relies on low-frequency oscillations during encoding

As an explorative analysis, we analyzed the dynamics between the neuronal activity from the encoding task and the recognition performance. Thus, we asked whether specific oscillations during encoding propel successful memory formation. First, we computed the differences in oscillatory power between later remembered and forgotten trials separately for the presentation of the first and second stimulus of each pair. Therefore, we focused on the 2-s period before and after each stimulus onset for the VA and AV conditions, resulting in two analysis time windows (Stimulus Interval 1 and Stimulus Interval 2; Fig. 1). Initially, we analyzed oscillatory data independent of the stimulus modality order, dividing Stimulus intervals 1 (SI1) and 2 (SI2). The differential time-frequency spectra between remembered and forgotten pairs were computed in each participant. These difference values were correlated with associative d' values from the recognition task. Results revealed significant correlations within SI2 for both encoding sequence conditions. During the visual stimulus of AV pairs, two clusters of significant negative correlation between associative d' and average power differences (remembered > forgotten) were revealed at multiple electrodes in the parietal and central region in the alpha band (pre-stimulus positive cluster: $p < .042$, SD < 0.005; post-stimulus negative cluster: $p < .013$, SD = 0.003, Fig. 6A) at multiple electrodes in the central-parietal region (Fig. 6A) within the alpha range. In VA pairs, however, during the presentation of



the auditory stimulus, we observed a significant positive post-stimulus correlation ($p < .006$, SD = 0.002; Fig. 6B) and negative pre-stimulus correlation ($p < .018$, SD = 0.003) at multiple electrodes in the parietal-occipital region (Fig. 6B) within the alpha and beta range.

Discussion

Successful retrieval of events is strongly bound to the context of encoding[55,85,86]. While it is well established that context can be reflected as the surrounding environment[31,87,88], it can also be represented differentially, i.e. as the sequence of episode-specific features. In line, context feature reinstatement (of sequential information)

◀**Fig. 3**. Memory effects on time-frequency power before and during recognition for electrodes with maximal t-value for corresponding encoding stimulus modality order. (**A**) Time-frequency plot of the statistical comparison of REMEMBERED > FORGOTTEN oscillatory power at P2 for recognized associations, which were presented in the visual-auditory sequence during encoding. Opaque data points show a significant difference at $p < .05$ (corrected). The lower panel shows the topographical distribution within the significant cluster during stimulus presentation in the theta-alpha range (0.9 to 1.2 s; left). The green marker illustrates P2. (**B**) Time-frequency plot of the statistical comparison of REMEMBERED > FORGOTTEN oscillatory power at PO4 for recognized associations, which were presented in the auditory-visual sequence during encoding. Opaque data points show a significant difference at $p < .05$ (corrected). The lower panel shows the topographical distribution within the significant cluster during stimulus in the alpha/beta range (0.8 to 1.4 s). The green marker illustrates P2. (**C**) Top: Time-frequency plot of the statistical comparison of REMEMBERED > FORGOTTEN oscillatory power at POz, independent of modality order presentation during encoding. The vertical line marks the stimulus onset, and the horizontal lines mark the frequency bins alpha and beta (8 Hz and 25 Hz). Opaque data points show a significant difference at $p < 0.\ 5$ (corrected). Negative t-values signify higher power in FORGOTTEN trials. The lower panel shows the topographical distribution within the significant cluster during stimulus in the beta range (0.8 to 1.3 s; left) and the significant cluster in the alpha range (0.9 to 1.8 s, right). The marker illustrates POz.

Frequency range [Hz]	Mean Accuracy [%]	Peak mean t-value Electrode	Time [s]	Frequency [Hz]	Peak mean t-value Max. [%]
Overall [1 40]	52.26	F4	0.8	30	62.38
Theta [3 7]	52.49	AF7	0.4	7	58.80
Alpha [8 13]	52.40	Fpz	0.1	11	58.67
Beta [13 21]	52.29	Fpz	0.4	21	58.63
Beta [22 32]	52.22	P1	1.9	24	58.29
Beta [13 32]	52.24	P1	1.9	24	61.00
Gamma [32 40]	52.23	P2	0.1	33	52.98

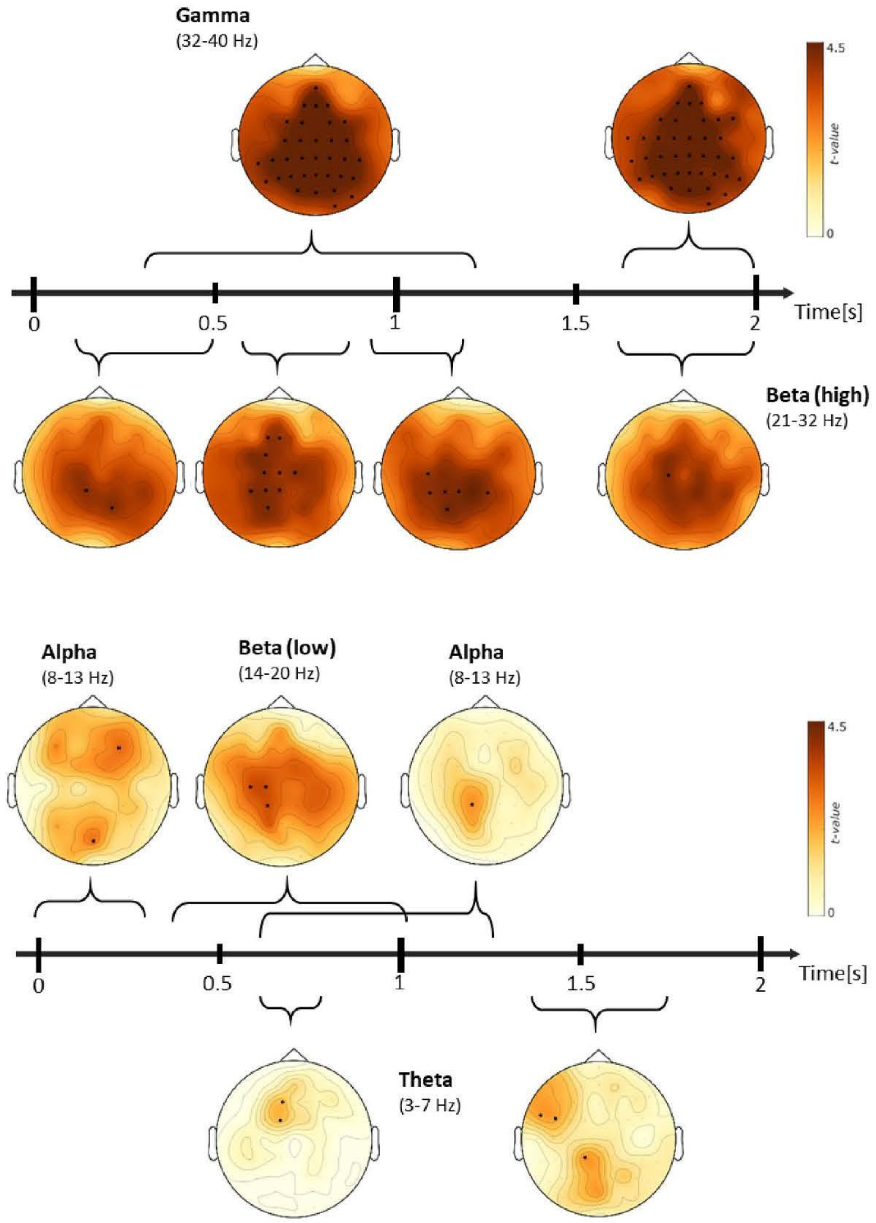
Table 3. Average Accuracy for each frequency range and at the electrode with the maximummean t-value.

has been shown to be of central importance for memory encoding and retrieval processes[51,88–91]. The role of contextual reinstatement of multisensory features that stem, i.e. from the auditory and visual domain, remains so far unexplored. Here we aimed to shed light on the oscillatory mechanisms underlying the recognition of sequentially encoded multisensory episodes. Our findings show that modality sequences are (incidentally) encoded within the memory trace and serve as a context feature that drives recognition based on theta, alpha and beta frequency pattern reinstatement.

It is well established that context reinstatement plays a crucial role in memory retrieval, allowing the brain to access the temporal and environmental cues associated with past events[51,52]. However, previous research has largely focused on unimodal stimuli, rendering our understanding about how complex multisensory sequences are encoded and retrieved incomplete. In the current study, participants acquired and retrieved the image-sound/sound-image associations overall very well, with recognition performance as measured by d' being comparable to similar study designs[72,92–94]. Importantly, recognition performance did not differ when comparing AV (auditory-visual) vs. VA (visual-auditory) pairs, suggesting that the order of features itself did not influence memory formation. While we did not test for incidental stimulus acquisition, several studies report reinstatement of encoding specific features in memory tasks[29,95,96].

MVPA of EEG data revealed distinct neural signatures depending on the modality sequence presented during encoding, even though overall memory performance for both conditions was the same. This suggests that the brain encodes the order of multisensory episodes as part of the contextual memory trace, which aids in the retrieval process. This finding is consistent with prior work demonstrating context-specific temporal patterns during both encoding and retrieval processes[97]. Furthermore, it aligns with the Context Maintenance and Retrieval (CMR) model, which states that temporal and contextual features of episodes are essential components of the memory trace[52]. Our study extends this model by showing that modality order, as a contextual feature, can be decoded from oscillatory activity during memory retrieval.

Our multivariate results from EEG recordings suggest that decoding performance went significantly above chance level across multiple time points, frequencies, and electrode locations, with particularly strong effects in the theta, alpha, and beta bands. Theta oscillations have been associated with the binding of information into coherent memory traces and are crucial for organizing sequentially ordered working memory items[60,68,98]. These oscillations are thought to coordinate neural activity across different brain regions, facilitating the binding of sensory inputs into a coherent memory trace[97–99], acting as the "glue"[100]. Alpha oscillations have been generally related to the inhibition of irrelevant information and are involved in processing incoming information relevant to memory[101,102]. Decreases in alpha power during memory tasks have been associated with enhanced memory performance, particularly in semantic encoding tasks[103,104], while beta oscillations have been linked to memory formation, with elevated pre-stimulus beta power associated with successful memory encoding[105,106]. This activity is thought to reflect a memory-promoting state, possibly moderated by attentional or inhibitory processes[105]. In sum, theta, alpha, and beta oscillations play distinct yet interconnected roles in



memory processing[107,108]. Taken together, the fact that the MVPA analysis revealed significant classification of modality order across different oscillatory bands provides strong evidence that sequential information in the form of a context feature is retained and reinstated during retrieval rather than merely reflecting general associative activation. Thus, our results suggest that encoding processes are sequence-specific, with VA pairs potentially engaging greater anticipatory processing due to the nature of auditory stimulus processing. One may speculate, that the reinstatement of the observed oscillatory patterns may facilitate the synchronization of neural activity across sensory processing regions, ensuring that the original modality sequence is represented during recognition within the specific memory trace, which has been suggested by several human studies[4,11,82,109,110].

◀**Fig. 4**. Topographical plots of *t*-values from the cluster-based permutation test comparing MVPA accuracy for remembered trials against the chance level of 50%. The plots illustrate the distribution of significant effects across the scalp for different frequency bands (theta, alpha, low beta, high beta, and gamma) over the trial time course. Highlighted electrodes indicate regions with the highest number of neighbouring significant data points), reflecting areas with the strongest decoding effects. The gamma band showed significant effects between 400–1300 ms and 1700–2000 ms, with strong activations in centro-posterior regions. The high beta band displayed significant clusters in multiple time windows (100–500 ms, 600–800 ms, 900–1200 ms, and 1700–2000 ms), predominantly in centro-parietal areas. The low beta band exhibited significant effects between 400–1000 ms, mainly over central electrodes. The alpha band showed spatially distributed effects in early (0–300 ms) and later (600–1300 ms) time windows, particularly in partial-occipital and fronto-central sites. The theta band revealed significant clusters in two discrete time windows (600–700 ms and 1400–1700 ms), with the strongest effects observed in frontal and central regions. The results suggest that significant decoding effects are not uniformly distributed across the scalp but are concentrated in specific electrode regions, particularly in central and parietal areas in the beta and gamma range and frontal areas in the theta and alpha range.

Considering the potential cognitive processes elicited by the experimental design, one could argue that the initial stimulus might evoke visual or auditory mental imagery, creating an expectation of the second stimulus. In fact, there is evidence showing that mental imagery could also influence associative memory retrieval by engaging both modality-specific and modality-independent neural networks[111,112], thereby aiding in overcoming potential modality mismatches during encoding induced by incongruent association pairs[113]. This process involves the activation of sensory-specific regions, such as the visual and auditory cortices, alongside a modality-independent core network, including the default mode network, which supports imagery across different sensory domains[111,112,114]. The overlap between brain regions involved in mental imagery and those supporting retrieval suggests that successful retrieval relies on the same neural mechanisms that facilitate imagery[111,115]. Furthermore, encoding specificity plays a critical role in remembering, as the reactivation of encoding-related neural patterns benefits retrieval when there is a match between encoding and retrieval modalities but can impair memory under mismatch conditions[116]. However, individuals can flexibly employ mental imagery to compensate for mismatches, generating and maintaining mental representations even when encoding involves incongruent audiovisual information[117]. Moreover, imagery-based strategies, such as integrating items into interactive mental images, have been shown to enhance associative memory, emphasizing the functional significance of mental imagery in retrieval processes[118].

While these findings demonstrate that mental imagery is connected to memory encoding and retrieval, prior research also suggests that multisensory, sequential encoding can enhance memory through encoding variability, introducing competition effects that alter retrieval dynamics[28,29]. This interpretation gains support from our univariate oscillatory findings, suggesting a modality-specific influence of alpha and beta oscillations during the encoding of sequentially presented audiovisual stimuli. Specifically, for auditory-visual (AV) pairs, increased pre-stimulus alpha and beta power before the visual stimulus and a subsequent decrease during stimulus presentation may indicate a preparatory state followed by active sensory processing. In contrast, the pattern is reversed for visual-auditory (VA) pairs: pre-stimulus decreases in alpha and beta before the auditory stimulus, and increased power during its presentation suggests a shift in processing demands between modalities. Previous research has linked decreased alpha-band activity in the prefrontal and occipital cortex to successful visual encoding, indicating that lower alpha power facilitates visual information processing and enhances memory formation[119–122]. Similarly, increased pre-stimulus beta power has been associated with improved memory formation, potentially reflecting attentional or inhibitory processes that aid in binding stimulus components[105,123]. The observed pre-stimulus increases in beta power in AV pairs may, therefore, indicate an anticipatory mechanism supporting visual encoding, whereas the decrease in VA pairs might reflect a shift in sensory dominance from vision to audition. Alpha and beta oscillations have further been implicated in the processing of expectations and prediction errors. An alpha-to-beta desynchronization (ERD) has been linked to expected stimulus valence, suggesting that these frequency bands contribute to prediction mechanisms that influence encoding efficiency[124]. In audiovisual tasks, alpha oscillations modulate sensory processing and attention, influencing the temporal integration of stimuli[125]. However, alpha activity does not solely predict auditory stimulus detection consistently due to its interaction with broadband neural activity[126]. Beta oscillations, on the other hand, are associated with top-down control processes and enhance memory formation across sensory modalities, including auditory processing[105]. Taken together, these findings suggest that alpha and beta oscillations in sequential encoding are modality-dependent rather than purely memory-driven. The observed pre-stimulus shifts in power may reflect preparatory mechanisms that optimize encoding by modulating attention and sensory processing demands across modalities. Specifically, lower alpha power in occipital-parietal regions has been associated with improved perceptual sensitivity and the enhancement of stimulus representations[127], supporting the idea that modality-dependent oscillatory changes may reflect the differential engagement of sensory and integrative processes during retrieval. Finally, theta oscillations play a critical role in cross-modal binding, supporting associative memory by synchronizing neural activity across sensory regions[128]. Given the angular gyrus' role in multimodal integration[129] and the contribution of multisensory cues to episodic retrieval[130–132], it is likely that power changes across frequency bands reflect both modality-specific processing and memory-related mechanisms in an interactive manner. Thus, the observed oscillatory dynamics align with well-established mechanisms of sensory reactivation, cortical excitability, and associative memory retrieval.

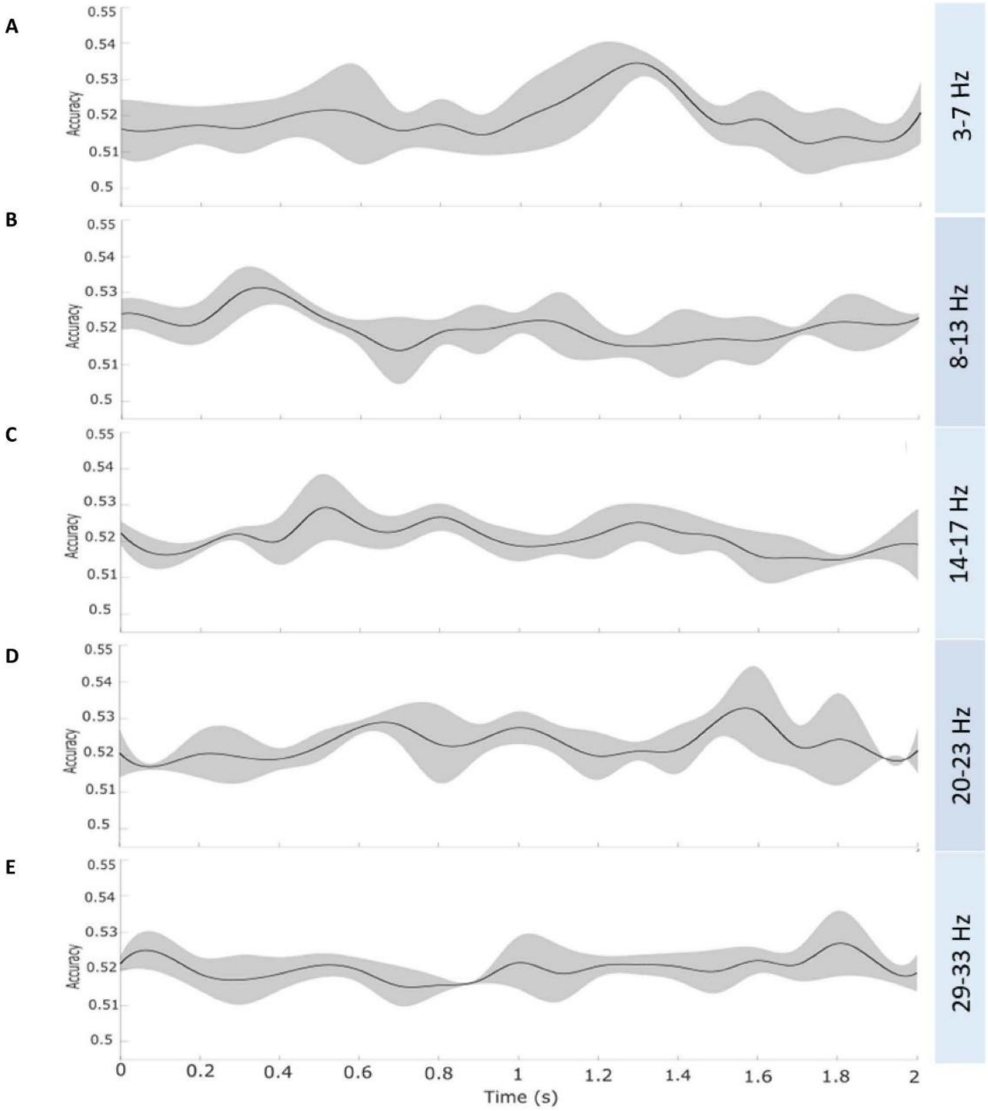
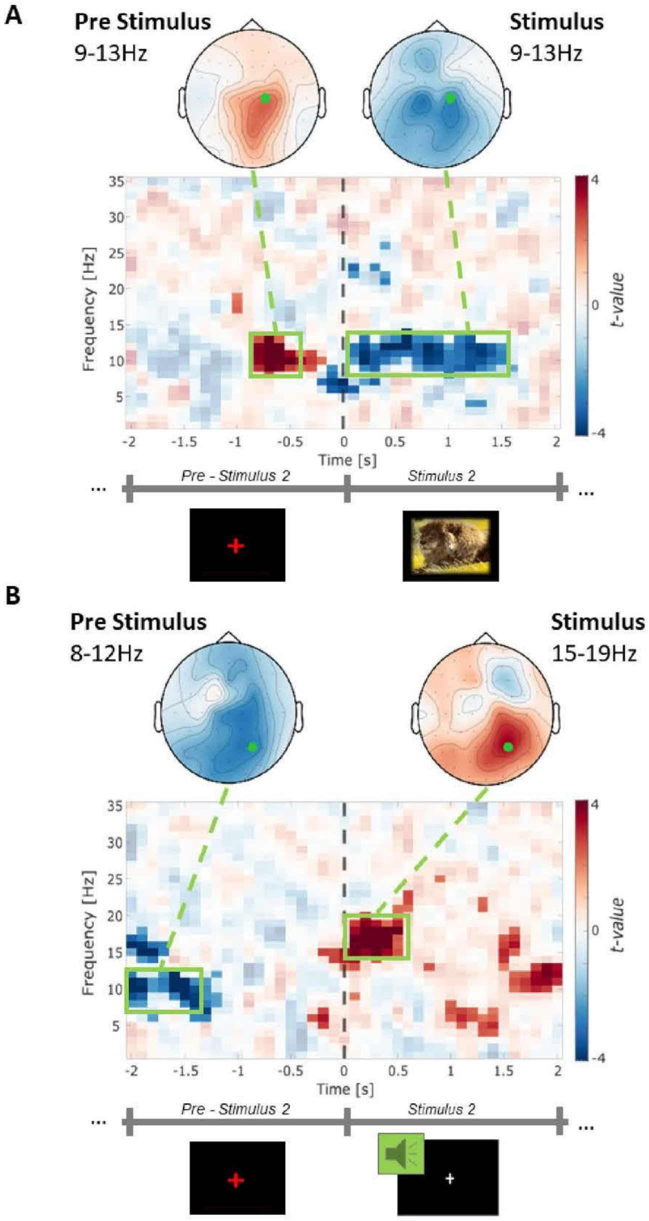


Fig. 5. Full-time courses of accuracy values from exemplary electrodes for each frequency range. Time-resolved decoding accuracy (MVPA) for remembered trials to classify modality order from 0 to 2 s relative to stimulus onset is shown at exemplary electrodes. Accuracy values, averaged within the respective frequency bands, range from 51–54%. (**A**) Theta (3–7 Hz) at Cz exhibits a subtle peak around 1.3 s. (**B**) Alpha (8–13 Hz) at F4 shows a slight peak around 0.3 s. (**C**) Decoding accuracy in the low beta (14–17 Hz) at Fz, (**D**) high beta (20–23 Hz) at C3, and (**E**) gamma (29–33 Hz) at F4 bands fluctuates over time but did not reveal distinct classification peaks. Caution should be exercised, as the averaging of accuracy estimates over multiple time points and across participants may have contributed to an overall reduction in classification performance.



The role of modality order in retrieval is further supported by the finding that above-chance decoding of modality order from EEG activity during retrieval indicates its integration into the memory trace. If one modality had dominated encoding, successful decoding of modality order would not be expected. Prior research supports the idea that encoding modality order contributes to retrieval by providing structured cues that facilitate reconstruction of past experiences[133]. Contextual information is a well-established component of episodic memory, with hippocampal mechanisms playing a crucial role in binding sensory details into coherent memory representations[22,85,103,134,135]. Our study extends this body of work by demonstrating that modality order, as a contextual feature, can be decoded from oscillatory activity during retrieval. While our data do not conclusively establish a causal link between modality order encoding and retrieval success, they provide novel

◀ **Fig. 6**. Correlation between the power difference of remembered vs forgotten trials in Stimulus interval 2 and the memory performance (d') from the recognition task. (**A**) The correlation between the power differences and the memory performance as measured by d' during visual stimuli presentation (AV pair) at C2. Topographical distribution within the significant cluster in the pre stimulus (-0.8 to -0.1 s; left) and the significant cluster in the stimulus presentation (0 to 1.5 s) in the alpha range (9 to 13 Hz, right) during visual stimuli presentation. The marker illustrates C2. (**B**) The correlation between the power differences and the memory performance as measured by d' during auditory stimulus presentation (VA pair) at P4. Topographical distributions are shown for the pre-stimulus cluster (-2 to -1.3 s; 8–12 Hz) and stimulus presentation cluster (0–0.5 s; 15–19 Hz). The green marker illustrates P4.

insights into the neural dynamics supporting multisensory sequential memory. Although no behavioural data confirm explicit retrieval of order information, the significant decoding results indicate that modality order was included in the memory trace and reinstated during retrieval. Future studies should explore the extent to which such reinstatement contributes to explicit order memory and whether implicit representations influence retrieval performance.

Finally, we observed modality effects in parietal locations during both, encoding, and recognition tasks. This might reflect multisensory association processes[136,137], as the parietal cortex is crucial for integrating information from various sensory modalities[138–141]. Interestingly, our findings align well with recent work around multisensory processing, which shows that information from different sensory modalities is integrated within several cortical regions (e.g. the parietal lobe;[139,141,142]). While the intraparietal sulcus (IPS) is known to process multisensory information[143–145], the angular gyrus has been shown to be centrally involved in binding information into coherent narratives[131]. In our results, classifier accuracy derived from MVPA was highest at centro-parietal electrodes (e.g., CP1), covering the superior parietal lobe. This might suggest a distinct role of the superior parietal cortex in multisensory sequential reinstatement processes, which supports the hypothesis that the neural systems for sequence encoding and multisensory integration are closely linked to facilitate the binding of presented items, thereby forming an episode.

In conclusion, our study provides new insights into the neural mechanisms underlying multisensory memory retrieval. The findings of this study yield important implications for our understanding of memory processes. First, the (incidental) encoding of modality sequence order as a context feature suggests that the brain actively integrates temporal and sensory information during memory formation. This has important implications for models of episodic memory, particularly those that emphasize the role of context[52]. Our findings suggest that the neural mechanisms underlying context-feature retrieval are not limited to unimodal tasks[4,51,64,146]. Instead, they extend to more complex multisensory episodes, during which we encode and integrate sequences of different sensory modalities. Crucially, the sequence of modalities as a contextual feature within the memory trace directly affects memory retrieval, cognitive control, and learning processes[147–150]. Prior research suggests that different sensory modalities contribute uniquely to encoding and retrieval mechanisms, with auditory and visual sequences influencing attentional engagement and memory consolidation in distinct ways (e.g.,[151,152]. Auditory sequences, for instance, have been linked to more durable temporal structuring, while visual sequences often benefit from spatial organization[71]. Recognizing modality sequence as a contextual feature allows us to investigate how the structure of sensory input shapes memory representations rather than focusing solely on content-based associations. Understanding these effects can help optimize learning and memory strategies by leveraging the strengths of different modalities. This might offer insights into educational and rehabilitative applications where multimodal integration plays a key role. Our findings provide new insights into how the brain encodes and retrieves complex episodic memories, particularly those that involve multisensory information.

Data availability

The raw EEG and behavioral data underlying our findings have been uploaded to an open repository (https://www.fdr.uni-hamburg.de/; https://www.fdr.uni-hamburg.de/record/17120) for accessibility.

Received: 28 October 2024; Accepted: 14 May 2025

References

1. Dannenberg, H., Alexander, A. S., Robinson, J. C. & Hasselmo, M. E. The role of hierarchical dynamical functions in coding for episodic memory and cognition. *J. Cogn. Neurosci.* **31**, 1271–1289 (2019).
2. Sheldon, S., Farb, N., Palombo, D. J. & Levine, B. Intrinsic medial Temporal lobe connectivity relates to individual differences in episodic autobiographical remembering. *Cortex* **74**, 206–216 (2016).
3. Sadeh, S. & Clopath, C. Theory of neuronal perturbome in cortical networks. *Proc. Natl. Acad. Sci.* **117**, 26966–26976 (2020).
4. Kim, J. S. & Lee, S. A. Hippocampal orchestration of associative and sequential memory networks for episodic retrieval. *Cell. Rep.* **42**, 112989 (2023).
5. Yonelinas, A. P., Ranganath, C., Ekstrom, A. D. & Wiltgen, B. J. A contextual binding theory of episodic memory: systems consolidation reconsidered. *Nat. Rev. Neurosci.* **20**, 364–375 (2019).
6. Pronier, É., Morici, J. F. & Girardeau, G. The role of the hippocampus in the consolidation of emotional memories during sleep. *Trends Neurosci.* **46**, 912–925 (2023).
7. Cooper, R. A. & Ritchey, M. Progression from Feature-Specific brain activity to hippocampal binding during episodic encoding. *J. Neurosci.* **40**, 1701–1709 (2020).
8. Crystal, J. D. Temporal foundations of episodic memory. *Learn. Behav.* **52**, 35–50 (2024).
9. Clewett, D. & Davachi, L. The ebb and flow of experience determines the Temporal structure of memory. *Curr. Opin. Behav. Sci.* **17**, 186–193 (2017).

10. Horner, A. J., Bisby, J. A., Wang, A., Bogus, K. & Burgess, N. The role of Spatial boundaries in shaping long-term event representations. *Cognition* **154**, 151–164 (2016).
11. Eichenbaum, H. Time (and space) in the hippocampus. *Curr. Opin. Behav. Sci.* **17**, 65–70 (2017).
12. Leroy, N., Majerus, S. & D'Argembeau, A. Working memory capacity for continuous events: the root of Temporal compression in episodic memory? *Cognition* **247**, 105789 (2024).
13. Jeunehomme, O., Folville, A., Stawarczyk, D., Van Der Linden, M. & D'Argembeau, A. Temporal compression in episodic memory for real-life events. *Memory* **26**, 759–770 (2018).
14. Jeunehomme, O. & D'Argembeau, A. Event segmentation and the Temporal compression of experience in episodic memory. *Psychol. Res.* **84**, 481–490 (2020).
15. D'Argembeau, A., Jeunehomme, O. & Stawarczyk, D. Slices of the past: how events are temporally compressed in episodic memory. *Memory* **30**, 43–48 (2022).
16. Riegel, M., Granja, D., Amer, T., Vuilleumier, P. & Rimmele, U. Opposite effects of emotion and event segmentation on Temporal order memory and object-context binding. *Cogn. Emot.* **39**, 117–135 (2025).
17. Thavabalasingam, S., O'Neil, E. B. & Lee, A. C. H. Multivoxel pattern similarity suggests the integration of Temporal duration in hippocampal event sequence representations. *NeuroImage* **178**, 136–146 (2018).
18. Liu, C., Ye, Z., Chen, C., Axmacher, N. & Xue, G. Hippocampal representations of event structure and Temporal context during episodic Temporal order memory. *Cereb. Cortex.* **32**, 1520–1534 (2022).
19. Rey, L. et al. Episodic memory and recognition are influenced by cues' sensory.
20. Tamminen, J. & Mebude, M. Reinstatement of odour context cues veridical memories but not false memories. *Memory* **27**, 575–579 (2019).
21. Sorokowska, A., Nord, M., Stefańczyk, M. M. & Larsson, M. Odor-based context-dependent memory: influence of olfactory cues on declarative and nondeclarative memory indices. *Learn. Mem.* **29**, 136–141 (2022).
22. Marks, W. D., Yokose, J., Kitamura, T. & Ogawa, S. K. Neuronal ensembles organize activity to generate contextual memory. *Front. Behav. Neurosci.* **16**, 805132 (2022).
23. Noyce, A. L., Cestero, N., Shinn-Cunningham, B. G. & Somers, D. C. Short-term memory stores organized by information domain. *Atten. Percept. Psychophys.* **78**, 960–970 (2016).
24. Dimsdale-Zucker, H. R. et al. Representations of complex contexts: A role for Hippocampus. *J. Cogn. Neurosci.* **35**, 90–110 (2022).
25. Bird, C. M. The role of the hippocampus in recognition memory. *Cortex* **93**, 155–165 (2017).
26. Robin, J. & Moscovitch, M. Familiar real-world Spatial cues provide memory benefits in older and younger adults. *Psychol. Aging.* **32**, 210–219 (2017).
27. Wälti, M. J., Woolley, D. G. & Wenderoth, N. Assessing rhythmic visual entrainment and reinstatement of brain oscillations to modulate memory performance. *Front. Behav. Neurosci.* **14**, 118 (2020).
28. Zhang, M. & Hupbach, A. The effects of variable encoding contexts on item and source recognition. *Mem. Cognit.* **51**, 391–403 (2023).
29. Bramão, I., Jiang, J., Wagner, A. D. & Johansson, M. Encoding contexts are incidentally reinstated during competitive retrieval and track the Temporal dynamics of memory interference. *Cereb. Cortex.* **32**, 5020–5035 (2022).
30. Bramão, I. & Johansson, M. Neural Pattern Classification Tracks Transfer-Appropriate Processing in Episodic Memory. *eneuro* 5, ENEURO.0251-18.2018 (2018).
31. Wheeler, R. L. & Gabbert, F. Using Self-Generated cues to facilitate recall: A narrative review. *Front. Psychol.* **8**, 1830 (2017).
32. Unsworth, N. & Spillers, G. J. Variation in working memory capacity and episodic recall: the contributions of strategic encoding and contextual retrieval. *Psychon Bull. Rev.* **17**, 200–205 (2010).
33. Versace, R. et al. Act-In: an integrated view of memory mechanisms. *J. Cogn. Psychol.* **26**, 280–306 (2014).
34. Duarte, S. E., Ghetti, S. & Geng, J. J. Object memory is multisensory: Task-irrelevant sounds improve recollection. *Psychon Bull. Rev.* **30**, 652–665 (2023).
35. Versace, R., Labeye, É., Badard, G. & Rose, M. The contents of long-term memory and the emergence of knowledge. *Eur. J. Cogn. Psychol.* **21**, 522–560 (2009).
36. Juan, C. et al. The variability of multisensory processes of natural stimuli in human and non-human primates in a detection task. *PLOS ONE.* **12**, e0172480 (2017).
37. Lehmann, S. & Murray, M. M. The role of multisensory memories in unisensory object discrimination. *Cogn. Brain Res.* **24**, 326–334 (2005).
38. Moran, J. K. et al. Multisensory processing can compensate for Top-Down attention deficits in schizophrenia. *Cereb. Cortex.* **31**, 5536–5548 (2021).
39. Thelen, A., Talsma, D. & Murray, M. M. Single-trial multisensory memories affect later auditory and visual object discrimination. *Cognition* **138**, 148–160 (2015).
40. Murray, C. A. & Shams, L. Crossmodal interactions in human learning and memory. *Front. Hum. Neurosci.* **17**, 1181760 (2023).
41. Tovar, D. A., Murray, M. M. & Wallace, M. T. Selective enhancement of object representations through multisensory integration. *J. Neurosci.* **40**, 5604–5615 (2020).
42. Benini, E., Koch, I., Mayr, S., Frings, C. & Philipp, A. M. Contextual features of the cue enter episodic bindings in task switching. *J. Cogn.* **5**, 29 (2022).
43. Karacsony, S. & Abela, M. R. L. Stimulating sense memories for people living with dementia using the Namaste care programme: what works, how and why? *J. Clin. Nurs.* **31**, 1921–1932 (2022).
44. Shams, L. & Seitz, A. R. Benefits of multisensory learning. *Trends Cogn. Sci.* **12**, 411–417 (2008).
45. Dudai, Y. Molecular bases of long-term memories: a question of persistence. *Curr. Opin. Neurobiol.* **12**, 211–216 (2002).
46. Sridhar, S., Khamaj, A. & Asthana, M. K. Cognitive neuroscience perspective on memory: overview and summary. *Front. Hum. Neurosci.* **17**, 1217093 (2023).
47. Brainerd, C. J., Gomes, C. F. A. & Nakamura, K. Dual recollection in episodic memory. *J. Exp. Psychol. Gen.* **144**, 816–843 (2015).
48. Chang, M. et al. Spatial context scaffolds long-term episodic richness of weaker real-world autobiographical memories in both older and younger adults. *Memory* **32**, 431–448 (2024).
49. Dobbins, I. G. & Wagner, A. D. Domain-general and Domain-sensitive prefrontal mechanisms for recollecting events and detecting novelty. *Cereb. Cortex.* **15**, 1768–1778 (2005).
50. King, D. R., De Chastelaine, M., Elward, R. L., Wang, T. H. & Rugg, M. D. Recollection-Related increases in functional connectivity predict individual differences in memory accuracy. *J. Neurosci.* **35**, 1763–1772 (2015).
51. Lohnas, L. J., Healey, M. K. & Davachi, L. Neural Temporal context reinstatement of event structure during memory recall. *J. Exp. Psychol. Gen.* **152**, 1840–1872 (2023).
52. Polyn, S. M., Norman, K. A. & Kahana, M. J. A context maintenance and retrieval model of organizational processes in free recall. *Psychol. Rev.* **116**, 129–156 (2009).
53. Clewett, D., DuBrow, S. & Davachi, L. Transcending time in the brain: how event memories are constructed from experience. *Hippocampus* **29**, 162–183 (2019).
54. Davachi, L. & DuBrow, S. How the hippocampus preserves order: the role of prediction and context. *Trends Cogn. Sci.* **19**, 92–99 (2015).
55. DuBrow, S. & Davachi, L. Temporal binding within and across events. *Neurobiol. Learn. Mem.* **134**, 107–114 (2016).

56. Gordon, A. M., Rissman, J., Kiani, R. & Wagner, A. D. Cortical reinstatement mediates the relationship between Content-Specific encoding activity and subsequent recollection decisions. *Cereb. Cortex.* **24**, 3350–3364 (2014).
57. Jafarpour, A., Horner, A. J., Fuentemilla, L., Penny, W. D. & Duzel, E. Decoding oscillatory representations and mechanisms in memory. *Neuropsychologia* **51**, 772–780 (2013).
58. Staresina, B. P. & Davachi, L. Mind the Gap: binding experiences across space and time in the human Hippocampus. *Neuron* **63**, 267–276 (2009).
59. Folkerts, S., Rutishauser, U. & Howard, M. W. Human episodic memory retrieval is accompanied by a neural contiguity effect. *J. Neurosci.* **38**, 4200–4211 (2018).
60. Herweg, N. A., Solomon, E. A. & Kahana, M. J. Theta oscillations in human memory. *Trends Cogn. Sci.* **24**, 208–227 (2020).
61. Koen, J. D. & Rugg, M. D. Memory reactivation predicts resistance to retroactive interference: evidence from multivariate classification and pattern similarity analyses. *J. Neurosci.* **36**, 4389–4399 (2016).
62. Leiker, E. K. & Johnson, J. D. Pattern reactivation co-varies with activity in the core recollection network during source memory. *Neuropsychologia* **75**, 88–98 (2015).
63. Miller, J. E., Carlson, L. A. & McAuley, J. D. When what you hear influences when you see: listening to an auditory rhythm influences the Temporal allocation of visual attention. *Psychol. Sci.* **24**, 11–18 (2013).
64. Staudigl, T., Vollmar, C., Noachtar, S. & Hanslmayr, S. Temporal-Pattern similarity analysis reveals the beneficial and detrimental effects of context reinstatement on human memory. *J. Neurosci.* **35**, 5373–5384 (2015).
65. Yu, J. Y., Liu, D. F., Loback, A., Grossrubatscher, I. & Frank, L. M. Specific hippocampal representations are linked to generalized cortical representations in memory. *Nat. Commun.* **9**, 2209 (2018).
66. Ratcliff, R., Sederberg, P. B., Smith, T. A. & Childers, R. A single trial analysis of EEG in recognition memory: tracking the neural correlates of memory strength. *Neuropsychologia* **93**, 128–141 (2016).
67. Rugg, M. D. & Wilding, E. L. Retrieval processing and episodic memory. *Trends Cogn. Sci.* **4**, 108–115 (2000).
68. Staudigl, T., Hanslmayr, S. & Bäuml, K. H. T. Theta oscillations reflect the dynamics of interference in episodic memory retrieval. *J. Neurosci.* **30**, 11356–11362 (2010).
69. Tan, R. J., Rugg, M. D. & Lega, B. C. Direct brain recordings identify hippocampal and cortical networks that distinguish successful versus failed episodic memory retrieval. *Neuropsychologia* **147**, 107595 (2020).
70. Waldhauser, G. T., Braun, V. & Hanslmayr, S. Episodic memory retrieval functionally relies on very rapid reactivation of sensory information. *J. Neurosci.* **36**, 251–260 (2016).
71. Cowan, N. & Guitard, D. Encoding colors and tones into working memory concurrently: A developmental investigation. *Dev. Sci.* **27**, e13552 (2024).
72. Ostrowski, J. & Rose, M. Increases in pre-stimulus theta and alpha oscillations precede successful encoding of crossmodal associations. *Sci. Rep.* **14**, 7895 (2024).
73. Leys, C., Ley, C., Klein, O., Bernard, P. & Licata, L. Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* **49**, 764–766 (2013).
74. Klem, G. H., Lüders, H. O. & Elger, C. *The ten-twenty Electrode System of the International Federation* (The International Federation of Clinical Neurophysiology, 1999).
75. Oostenveld, R., Fries, P., Maris, E., Schoffelen, J. M. & FieldTrip Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput. Intell. Neurosci.* 1–9 (2011). (2011).
76. Treder, M. S. & MVPA-Light A classification and regression toolbox for Multi-Dimensional data. *Front. Neurosci.* **14**, 289 (2020).
77. Artuso, C., Palladino, P. & Belacchi, C. Sensitivity detection in memory recognition: interference control as index of taxonomic memory development? *Memory* **28**, 187–195 (2020).
78. Haatveit, B. C. et al. The validity of *d* prime as a working memory index: results from the Bergen *n* -back task. *J. Clin. Exp. Neuropsychol.* **32**, 871–880 (2010).
79. Shermohammed, M., Davidow, J. Y., Somerville, L. H. & Murty, V. P. Stress impacts the fidelity but not strength of emotional memories. *Brain Cogn.* **133**, 33–41 (2019).
80. Marinari, E. Optimized Monte Carlo methods. in Advances in Computer Simulation (eds Kertész, J. & Kondor, I.) vol. 501 50–81 (Springer Berlin Heidelberg, (1998).
81. Kalafatovich, J., Lee, M. & Lee, S. W. Decoding declarative memory process for predicting memory retrieval based on source localization. *PLOS ONE.* **17**, e0274101 (2022).
82. Osipova, D. et al. Theta and gamma oscillations predict encoding and retrieval of declarative memory. *J. Neurosci.* **26**, 7523–7531 (2006).
83. Sikora-Wachowicz, B. et al. False recognition in Short-Term Memory – Age-Differences in confidence. *Front. Psychol.* **10**, 2785 (2019).
84. Weidemann, C. T. & Kahana, M. J. Assessing recognition memory using confidence ratings and response times. *R Soc. Open. Sci.* **3**, 150670 (2016).
85. Staudigl, T. & Hanslmayr, S. Theta oscillations at encoding mediate the Context-Dependent nature of human episodic memory. *Curr. Biol.* **23**, 1101–1106 (2013).
86. Sols, I., DuBrow, S., Davachi, L. & Fuentemilla, L. Event boundaries trigger rapid memory reinstatement of the prior events to promote their representation in Long-Term memory. *Curr. Biol.* **27**, 3499–3504e4 (2017).
87. Schwabe, L. & Wolf, O. T. The context counts: congruent learning and testing environments prevent memory retrieval impairment following stress. *Cogn. Affect. Behav. Neurosci.* **9**, 229–236 (2009).
88. Bramão, I., Karlsson, A. & Johansson, M. Mental reinstatement of encoding context improves episodic remembering. *Cortex* **94**, 15–26 (2017).
89. Lu, Y., Wang, C., Chen, C. & Xue, G. Spatiotemporal neural pattern similarity supports episodic memory. *Curr. Biol.* **25**, 780–785 (2015).
90. Sutterer, D. W., Foster, J. J., Serences, J. T., Vogel, E. K. & Awh, E. Alpha-band oscillations track the retrieval of precise Spatial representations from long-term memory. *J. Neurophysiol.* **122**, 539–551 (2019).
91. Xiao, X. et al. Transformed neural pattern reinstatement during episodic memory retrieval. *J. Neurosci.* **37**, 2986–2998 (2017).
92. Jablonowski, J. & Rose, M. The functional dissociation of posterior parietal regions during multimodal memory formation. *Hum. Brain Mapp.* **43**, 3469–3485 (2022).
93. Marian, V., Hayakawa, S. & Schroeder, S. R. Cross-Modal interaction between auditory and visual input impacts memory retrieval. *Front. Neurosci.* **15**, 661477 (2021).
94. Pillai, A. S., Gilbert, J. R. & Horwitz, B. Early sensory cortex is activated in the absence of explicit input during crossmodal item retrieval: evidence from MEG. *Behav. Brain Res.* **238**, 265–272 (2013).
95. Healey, M. K. Temporal contiguity in incidentally encoded memories. *J. Mem. Lang.* **102**, 28–40 (2018).
96. Mundorf, A. M. D., Lazarus, L. T. T., Uitvlugt, M. G. & Healey, M. K. A test of retrieved context theory: dynamics of recall after incidental encoding. *J. Exp. Psychol. Learn. Mem. Cogn.* **47**, 1264–1287 (2021).
97. Michelmann, S., Bowman, H. & Hanslmayr, S. The Temporal signature of memories: identification of a general mechanism for dynamic memory replay in humans. *PLOS Biol.* **14**, e1002528 (2016).
98. Rudoler, J. H., Herweg, N. A. & Kahana, M. J. Hippocampal Theta and episodic memory. *J. Neurosci.* **43**, 613–620 (2023).
99. Karakaş, S. A review of theta Oscillation and its functional correlates. *Int. J. Psychophysiol.* **157**, 82–99 (2020).

100. Clouter, A., Shapiro, K. L. & Hanslmayr, S. Theta phase synchronization is the glue that binds human associative memory. *Curr. Biol.* **27**, 3143–3148e6 (2017).
101. Schroeder, S. C. Y., Ball, F. & Busch, N. A. The role of alpha oscillations in distractor Inhibition during memory retention. *Eur. J. Neurosci.* **48**, 2516–2526 (2018).
102. Hall, M. C. et al. Oscillatory activity in bilateral prefrontal cortices is altered by distractor strength during working memory processing. *NeuroImage* **301**, 120878 (2024).
103. Griffiths, B. J., Martín-Buro, M. C., Staresina, B. P. & Hanslmayr, S. Disentangling neocortical alpha/beta and hippocampal theta/gamma oscillations in human episodic memory formation. *NeuroImage* **242**, 118454 (2021).
104. Vogelsang, D. A., Gruber, M., Bergström, Z. M., Ranganath, C. & Simons, J. S. Alpha oscillations during incidental encoding predict subsequent memory for new foil information. *J. Cogn. Neurosci.* **30**, 667–679 (2018).
105. Scholz, S., Schneider, S. L. & Rose, M. Differential effects of ongoing EEG beta and theta power on memory formation. *PLOS ONE.* **12**, e0171913 (2017).
106. Winterling, S. L., Shields, S. M. & Rose, M. Reduced memory-related ongoing oscillatory activity in healthy older adults. *Neurobiol. Aging.* **79**, 1–10 (2019).
107. Elmers, J., Yu, S., Talebi, N., Prochnow, A. & Beste, C. Neurophysiological effective network connectivity supports a threshold-dependent management of dynamic working memory gating. *iScience* **27**, 109521 (2024).
108. Siebenhühner, F., Wang, S. H., Palva, J. M. & Palva, S. Cross-frequency synchronization connects networks of fast and slow oscillations during visual working memory maintenance. *eLife* 5, e13451 (2016).
109. Engel, A. K. & Singer, W. Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* **5**, 16–25 (2001).
110. Pennartz, C. M. A. Identification and integration of sensory modalities: neural basis and relation to consciousness. *Conscious. Cogn.* **18**, 718–739 (2009).
111. Huijbers, W., Pennartz, C. M. A., Rubin, D. C. & Daselaar, S. M. Imagery and retrieval of auditory and visual information: neural correlates of successful and unsuccessful performance. *Neuropsychologia* **49**, 1730–1740 (2011).
112. Caenegem et al. Multisensory approach in mental imagery: ALE meta-analyses comparing motor, visual and auditory imagery. *Neurosci Biobehav Rev* **167**.
113. Tian, X., Zarate, J. M. & Poeppel, D. Mental imagery of speech implicates two mechanisms of perceptual reactivation. *Cortex* **77**, 1–12 (2016).
114. Daselaar, S. M., Porat, Y., Huijbers, W. & Pennartz, C. M. A. Modality-specific and modality-independent components of the human imagery system. *NeuroImage* **52**, 677–685 (2010).
115. Spagna, A. et al. Visual mental imagery: evidence for a heterarchical neural architecture. *Phys. Life Rev.* **48**, 113–131 (2024).
116. Staudigl, T. & Hanslmayr, S. Reactivation of neural patterns during memory reinstatement supports encoding specificity. *Cogn. Neurosci.* **10**, 175–185 (2019).
117. Umar, H., Mast, F. W., Cacchione, T. & Martarelli, C. S. The prioritization of visuo-spatial associations during mental imagery. *Cogn. Process.* **22**, 227–237 (2021).
118. Sahadevan, S. S., Chen, Y. Y. & Jeremy, B. C. Imagery-based strategies for memory for associations. *Memory* **29**:10, 1275–1295 .
119. Yin, Q. et al. Direct brain recordings reveal occipital cortex involvement in memory development. *Neuropsychologia* **148**, 107625 (2020).
120. Xie, Y., Feng, Z., Xu, Y., Bian, C. & Li, M. The different Oscillation patterns of alpha band in the early and later stages of working memory maintenance. *Neurosci. Lett.* **633**, 220–226 (2016).
121. Staudigl, T., Hartl, E., Noachtar, S., Doeller, C. F. & Jensen, O. Saccades are phase-locked to alpha oscillations in the occipital and medial Temporal lobe during successful memory encoding. *PLOS Biol.* **15**, e2003404 (2017).
122. Romei, V. et al. Spontaneous fluctuations in posterior -Band EEG activity reflect variability in excitability of human visual areas. *Cereb. Cortex.* **18**, 2010–2018 (2008).
123. Schneider, S. L. & Rose, M. Intention to encode boosts memory-related pre-stimulus EEG beta power. *NeuroImage* **125**, 978–987 (2016).
124. Strube, A., Rose, M., Fazeli, S. & Büchel, C. Alpha-to-beta- and gamma-band activity reflect predictive coding in affective visual processing. *Sci. Rep.* **11**, 23492 (2021).
125. Ronconi, L., Busch, N. A. & Melcher, D. Alpha-band sensory entrainment alters the duration of Temporal windows in visual perception. *Sci. Rep.* **8**, 11810 (2018).
126. Cunningham, E., Zimnicki, C. & Beck, D. M. The influence of prestimulus 1/f-Like versus Alpha-Band activity on subjective awareness of auditory and visual stimuli. *J. Neurosci.* **43**, 6447–6459 (2023).
127. Zhou, Y. J., Iemi, L., Schoffelen, J. M., De Lange, F. P. & Haegens, S. Alpha oscillations shape sensory representation and perceptual sensitivity. *J. Neurosci.* **41**, 9581–9592 (2021).
128. Wang, D., Clouter, A., Chen, Q., Shapiro, K. L. & Hanslmayr, S. Single-Trial phase entrainment of Theta oscillations in sensory regions predicts human associative memory performance. *J. Neurosci.* **38**, 6299–6309 (2018).
129. Yazar, Y., Bergström, Z. M. & Simons, J. S. Reduced multimodal integration of memory features following continuous theta burst stimulation of angular gyrus. *Brain Stimulat.* **10**, 624–629 (2017).
130. Tibon, R., Fuhrmann, D., Levy, D. A., Simons, J. S. & Henson, R. N. Multimodal integration and vividness in the angular gyrus during episodic encoding and retrieval. *J. Neurosci.* **39**, 4365–4374 (2019).
131. Grob, A. M., Heinbockel, H., Milivojevic, B., Doeller, C. & Schwabe, L. Causal role of the angular gyrus in insight-driven memory reconfiguration. *Preprint At.* https://doi.org/10.7554/eLife.91033.1 (2023).
132. Bonnici, H. M., Richter, F. R., Yazar, Y. & Simons, J. S. Multimodal feature integration in the angular gyrus during episodic and semantic retrieval. *J. Neurosci.* **36**, 5462–5471 (2016).
133. Pierce, B. H. & Gallo, D. A. Encoding modality can affect memory accuracy via retrieval orientation. *J. Exp. Psychol. Learn. Mem. Cogn.* **37**, 516–521 (2011).
134. Thakral, P. P., Benoit, R. G. & Schacter, D. L. Characterizing the role of the hippocampus during episodic simulation and encoding. *Hippocampus* **27**, 1275–1284 (2017).
135. Foudil, S. A., Pleche, C. & Macaluso, E. Memory for spatio-temporal contextual details during the retrieval of naturalistic episodes. *Sci. Rep.* **11**, 14577 (2021).
136. Rohe, T. & Noppeney, U. Distinct computational principles govern multisensory integration in primary sensory and association cortices. *Curr. Biol.* **26**, 509–514 (2016).
137. Sereno, M. I. & Huang, R. S. Multisensory maps in parietal cortex. *Curr. Opin. Neurobiol.* **24**, 39–46 (2014).
138. Andersen, R. A. Multimodal integration for the representation of space in the posterior parietal cortex. (1997).
139. Kayser, C. & Logothetis, N. K. Do early sensory cortices integrate cross-modal information? *Brain Struct. Funct.* **212**, 121–132 (2007).
140. Lin, J. Sensory inputs guiding cognitive behaviors and decision making. *Highlights Sci. Eng. Technol.* **74**, 1399–1404 (2023).
141. Porada, D. K., Regenbogen, C., Freiherr, J., Seubert, J. & Lundström, J. N. Trimodal processing of complex stimuli in inferior parietal cortex is modality-independent. *Cortex* **139**, 198–210 (2021).
142. Lewis, J. W. & Van Essen, D. C. Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *J. Comp. Neurol.* **428**, 112–137 (2000).
143. Anderson, K. L., Rajagovindan, R., Ghacibeh, G. A., Meador, K. J. & Ding, M. Theta oscillations mediate interaction between prefrontal cortex and medial Temporal lobe in human memory. *Cereb. Cortex.* **20**, 1604–1612 (2010).

144. Regenbogen, C. et al. The intraparietal sulcus governs multisensory integration of audiovisual information based on task difficulty. *Hum. Brain Mapp.* **39**, 1313–1326 (2018).
145. Sours, C. et al. Structural and functional integrity of the intraparietal sulcus in moderate and severe traumatic brain injury. *J. Neurotrauma.* **34**, 1473–1481 (2017).
146. DuBrow, S. & Davachi, L. The influence of context boundaries on memory for the sequential order of events. *J. Exp. Psychol. Gen.* **142**, 1277–1286 (2013).
147. Schmidt, M., Frings, C. & Tempel, T. Context-Dependent memory of motor sequences. *J. Cogn.* **4**, 15 (2021).
148. Kelber, P., Mackenzie, I. G. & Mittelstädt, V. Cognitive control in cross-modal contexts: abstract feature transitions of task-related but not task-unrelated stimuli modulate the congruency sequence effect. *J. Exp. Psychol. Learn. Mem. Cogn.* **50**, 902–919 (2024).
149. Pazdera, J. K. & Kahana, M. J. Modality effects in free recall: A retrieved-context account. *J. Exp. Psychol. Learn. Mem. Cogn.* **49**, 866–888 (2023).
150. Song, G. & Tan, X. Real-world Cross-modal retrieval via sequential learning. *IEEE Trans. Multimed.* **23**, 1708–1721 (2021).
151. Zhu, B. et al. Multiple interactive memory representations underlie the induction of false memory. *Proc. Natl. Acad. Sci.* **116**, 3466–3475 (2019).
152. Harrison, N. R. & Woodhouse, R. Modulation of auditory Spatial attention by visual emotional cues: differential effects of attentional engagement and disengagement for pleasant and unpleasant cues. *Cogn. Process.* **17**, 205–211 (2016).

Acknowledgements

The present research was supported by the German Research Foundation (DFG; SFB/Transregio 169, Project B3 and DFG RO 2653/9-1). We thank Michele Frerichs and Carla Mourkojannis for assistance during data acquisition.

Author contributions

M.R. and M.M. designed the study. M.M. performed data acquisition, and J.O. provided parts of the scripts for the behavioural and univariate analysis analysis. M.M. and M.R. analyzed the data. M.R. acquired funding, conceptualized, and supervised the project. M.M. and M.R. wrote the original manuscript. M.M., J.O., and M.R. reviewed and edited the paper and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.R.

Reprints and permissions information is available at www.nature.com/reprints.

Zusammenfassung auf Deutsch

Wie gelingt es dem menschlichen Gehirn, aus multisensorischen Erfahrungen kohärente episodische Erinnerungen zu bilden? Während klassische Gedächtnismodelle vorrangig semantische und räumliche Informationen hervorheben, rückt die neuere Forschung zunehmend kontextuelle Merkmale in den Fokus – darunter Umweltfaktoren, internale neuronale Zustände und zeitliche Strukturen. Diese Dissertation untersucht, wie solche Kontextmerkmale in multisensorische Assoziationen eingebunden werden und inwiefern sie bei der späteren Erinnerung reaktiviert werden können.

Diese Arbeit umfasst drei empirische Studien, die sich um die zentrale Fragestellung gruppieren lassen, wie kontextuelle Bindung die multisensorische Gedächtnisbildung und -wiedergabe unterstützt. In der ersten Studie lernten Teilnehmende audiovisuelle Assoziationen entweder in einer natürlichen virtuellen Umgebung oder in einem minimalistisch gestalteten Computer Experiment. Die behavioralen Ergebnisse zeigen, dass eine reizreiche Lernumgebung die Wiedererkennungsleistung verbessert, vermutlich durch eine stärkere perzeptuelle Kohärenz und semantische Einbettung. Die zweite Studie untersuchte die Rolle prä-stimulusbezogener Gehirnzustände bei der Gedächtniskodierung. Mittels rhythmischer visueller Stimulation wurden vor der Enkodierung Alpha- und Theta-Oszillationen gezielt moduliert. EEG-Daten zeigten, dass insbesondere eine erhöhte Alpha-Oszillationen mit einer verbesserten Gedächtnisleistung einherging. Dies weist daraufhin, dass internale Zustände das Gehirn für erfolgreiche Enkodierung vorbereiten können. In der dritten Studie lernten Teilnehmende audiovisuelle Reize in zwei festgelegten Modalitätsreihenfolgen (auditorisch-visuell oder visuell-auditorisch), wobei die Abfragephase eine simultane Präsentation verwendete. Multivariate Pattern Analysen der EEG-Daten zeigten, dass die ursprüngliche Modalitätsreihenfolge während des Abrufs neurononal reaktiviert wurde. Dies spricht dafür, dass selbst die zeitliche Struktur multisensorischer Episoden als kontextuelles Merkmal mit in die Gedächtnisspur aufgenommen wird.

Zusammengefasst zeigen die Studien, dass kontextuelle Bindung ein dynamischer, multidimensionaler Prozess ist, in dem externe Umweltmerkmale, internale neuronale Zustände und zeitliche Ordnungen flexibel in episodische Gedächtnisrepräsentationen integriert werden. Die drei Studien liefern komplementäre Evidenz für die kontextabhängige Organisation multisensorischer Erinnerungen und eröffnen neue Ansätze für die Optimierung von Lern- und Gedächtnisprozessen in alltagsnahen, pädagogischen und klinischen Kontexten.

Zusammenfassung auf Englisch

How do humans form coherent episodic memories from multisensory experiences? While traditional models have emphasised semantic and spatial information, growing evidence suggests that contextual features, which include environmental richness, internal neural states, and temporal structure, play a crucial role in how memories are encoded, stored, and retrieved. This dissertation examines how such contextual information becomes embedded in multisensory associative memory and influences its reactivation during retrieval.

The thesis includes three empirical studies that, although initially independent in design, converge on the central question of how contextual binding supports multisensory memory. In the first study, participants learned audiovisual associations in either a naturalistic virtual environment or a minimalist, artificial setting. Behavioural results showed improved recognition performance in the enriched condition, suggesting that environmental richness can strengthen encoding by enhancing perceptual coherence and semantic depth. The second study investigated the role of pre-stimulus oscillatory states in memory formation. Using rhythmic visual stimulation to entrain alpha and theta oscillations before encoding, this EEG study found that increased alpha power facilitated recognition performance, highlighting the importance of internal brain states in preparing the system for encoding. In the third study, participants encoded audiovisual stimulus pairs presented in either auditory–visual or visual–auditory order and later retrieved them under simultaneous presentation. Multivariate pattern analysis of EEG data during recognition revealed reinstatement of the original modality order, suggesting that the temporal structure itself had been integrated into the memory trace and was accessible during retrieval.

Taken together, these findings suggest that contextual binding is a flexible and dynamic process that combines external, internal, and temporal cues into episodic memory representations. Although the studies were not originally designed under a common theoretical framework, they collectively support the view that memory is shaped by the interaction between environmental input, neural preparatory states, and temporal regularities, providing new insights into the mechanisms of episodic memory, and highlighting potential options for improving memory performance in educational and clinical contexts.

Declaration of own contribution

Marike C. Maack*, Carina Jaap*, Phillipp Taesler., Frank Steinicke, & Michael Rose. (2022). Enriched environments enhance the development of explicit memory in an incidental learning task. **Scientific Reports**

Marike Christiane Maack analyzed the data and wrote the manuscript in collaboration with Carina Jaap.

Marike Maack*, Jan Ostrowski* & Michael Rose. Crossmodal Associative Memory Modulations by Visually Evoked Entrainment (**bioRxiv**, submitted)

Marike Christiane Maack designed, programmed and conducted the experiment, analyzed the data and wrote the manuscript, all in cooperation with Jan Ostowski.

Marike Maack, Jan Ostrowski & Michael Rose. (2025). The order of multisensory associative sequences is reinstated as context-feature during successful recognition. **Scientific Reports**

Marike Christiane Maack designed and programmed the experiment, conducted the experiment, analyzed the data and wrote the manuscript.

---

*Co-first authorship

Acknowledgments

This thesis marks the end of a long and winding road, one I could never have walked alone.

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Dr. Michael Rose, for his academic guidance, trust, and encouragement throughout my doctoral journey. I am also thankful to Prof. Dr. Patrik Bruns and Prof. Dr. Gregor Leicht for serving as my examination committee and for their valuable feedback on my work.

A special thanks goes to Prof. Dr. med. Christian Büchel and the entire Institute for Systems Neuroscience for providing a vibrant, diverse, and intellectually stimulating environment. In times when I questioned my path, the openness and support of this community helped me to refocus and regain confidence. I am especially grateful to my lab members—Christoph, Maren, Carina, and Jan—for your motivation, open ears, advice, and the laughter we shared. I'm incredibly happy that many of you became more than colleagues—true friends. I would also like to thank Carina Ufer and Maria Brutsch for always being there to listen to my never-ending monologues and emotional rollercoasters during this journey.

Looking back to where it all began, I feel incredibly lucky to have studied in Oldenburg, where my passion for neuroscience truly took root. I want to thank Prof. Christoph Herrmann for teaching us the hard way—hand-coding Fourier transforms without books or shortcuts—and for making us truly think about what we were doing. Prof. Stefan Debener showed me that no task was too small for a professor and taught us how to conduct EEG research with both precision and joy. Dr. Martin Bleichner and Prof. Stefan Scherbaum played a crucial role in my academic path, especially Stefan, who planted the first seed of an academic career during my bachelor's studies. At that time, I was certain I would never pursue a PhD or publish a paper, but I am thankful to have been proven wrong.

To the incredible group of friends from my master's studies—Daniel, Julia, Alina, Ariane, Björn, Merle, Elena, Johanna, and Paul—you shaped not just my academic but also my personal journey. From late-night MATLAB debugging sessions (some lasting 16 hours!) to lake swims, kitchen gatherings, and even training for a triathlon, your companionship made those years unforgettable.

To Hendrik—thank you for walking beside me every step of the way. Your calm advice, scientific insight, and unwavering belief in me, especially during difficult moments, were invaluable. I owe you more than words can express.

My family has been my unwavering anchor. Thank you for your endless love, trust, and belief that I would always find my way—even when I wasn't sure myself. That faith, quietly and consistently given, has been one of the greatest gifts in my life.

To Eva and Merle—there are no words needed. Thank you for understanding the strange overlap of academic chaos and horse-related sanity. You knew exactly when I needed a break or a hug (or both).

Finally, there are far too many people to name individually. However, I am grateful to everyone who has been part of this journey, whether through a small act of kindness, a scientific discussion, or simply sharing coffee and laughter along the way.

Curriculum Vitae

Lebenslauf aus datenschutzrechtlichen Gründen nicht enthalten.

Tools

ChatGPT version 4, OpenAI (https://uhhgpt.uni-hamburg.de), grammarly was used for the following purposes:

- Improving sentences and passages for their readability and flow
- Grammar and spelling checks

Eidesstattliche Versicherung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe, insbesondere ohne entgeltliche Hilfe von Vermittlungs- und Beratungsdiensten, verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe. Das gilt insbesondere auch für alle Informationen aus Internetquellen.

Soweit beim Verfassen der Dissertation KI-basierte Tools („Chatbots“) verwendet wurden, versichere ich ausdrücklich, den daraus generierten Anteil deutlich kenntlich gemacht zu haben. Die „Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft (DFG) zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG“ aus September 2023 wurde dabei beachtet.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich damit einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

08.07.2025

Datum Unterschrift: