



Published in final edited form as:

Neurobiol Learn Mem. 2020 November ; 175: 107317. doi:10.1016/j.nlm.2020.107317.

Generalization and the hippocampus: More than one story?

Dagmar Zeithamova, Caitlin R. Bowman

University of Oregon

Abstract

Memory-based cognition depends on both the ability to remember specific details of individual experiences and the ability to combine information across experiences to generalize and derive new knowledge. A hippocampal role in rapid encoding of specific events is long established. More recent research also demonstrates hippocampal contributions to generalization, but their nature is still debated. The current review provides an overview of hippocampal-based generalization in two lines of research—episodic inference and categorization—and discusses evidence for four candidate mechanisms and representational schemes that may underpin such generalization. We highlight evidence showing that the hippocampus contributes specific memories to generalization decisions, but also form generalized representations that integrate information across experiences. Multiple views are currently plausible of how such generalized representations form and relate to specific memories. Future research that uses behavioral and neural indices of both generalization and specificity may help resolve between the candidate generalization mechanisms, with the possibility that more than one view of hippocampal-based generalization may be valid. Importantly, all views share the emphasis on the broader role of the hippocampus in cognition that goes beyond remembering the past.

Keywords

hippocampus; episodic memory; memory integration; concept learning; categorization; inference

Memory is fundamental to all aspects of our lives. It stores specific details of individual events we encounter so we can differentiate one event from other similar ones. Memory specificity has been long known to rely on the episodic memory system, implemented by a network of regions centered on the hippocampus (Scoville & Milner, 1957; Squire & Zola, 1998). Importantly, memory also allows us to accumulate and combine information across time, extending beyond direct experience to guide decisions in novel situations. Yet, how we generalize from prior experience to make new decisions and how generalization relates to memory specificity is still actively debated.

Corresponding author: Dagmar Zeithamova, 1227 University of Oregon, Eugene, OR, 97403, dasa@uoregon.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declarations of competing interest: none.

Beyond the traditional multiple memory systems view

The multiple memory systems view posits that distinct memory representations, formed by distinct brain regions, support specificity and generalization (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; McClelland, McNaughton, & O'Reilly, 1995; Packard, Hirsh, & White, 1989; Poldrack & Packard, 2003; Tulving, 1987). A prevalent distinction has been between regions with different learning rates, or different neural plasticity. The hippocampus, with its unique synaptic plasticity, is well suited for rapid encoding of individual specific events. The hippocampus is then contrasted with other memory systems, such as the striatum (Poldrack & Foerde, 2008; Poldrack & Packard, 2003) or cortex (McClelland et al., 1995; O'Reilly & Norman, 2002) that learn slowly and incrementally to accumulate statistical regularities that are generalizable across experiences. Support for such a division of labor comes from a number of studies, including animal lesion and neuropsychological studies that show impaired specificity but relatively spared performance in generalization tasks following hippocampal damage (Bozoki, Grossman, & Smith, 2006; Knowlton, Mangels, & Squire, 1996; Knowlton & Squire, 1993; O'Connell et al., 2016; Packard et al., 1989).

While the multiple memory systems view has been widely accepted and fruitful, it is unlikely to be a complete picture of memory generalization. Notably, several studies from the last two decades have blurred this division of labor between the hippocampus and other memory systems, indicating that the episodic memory system, including the hippocampus itself, may form representations that integrate related experiences in service of generalization (e.g., Dusek & Eichenbaum, 1997; Heckers, Zalesak, Weiss, Ditman, & Titone, 2004; Preston, Shrager, Dudukovic, & Gabrieli, 2004; Shohamy & Wagner, 2008; Zeithamova, Dominick, & Preston, 2012). While each experience is unique, events share elements, which provide a means of linking individual experiences into more complex knowledge structures that represent derived information not present in any individual event. In contrast to slow statistical learning, such memory integration allows for new knowledge to be derived rapidly, by combining information across a small number of events (Schlichting, Mumford, & Preston, 2015; Shohamy & Wagner, 2008; Zeithamova, Dominick, et al., 2012; Zeithamova & Preston, 2010). Some of the most widely-used episodic integration tasks that show hippocampal involvement are acquired equivalence, associative inference and transitive inference (for a review, see Zeithamova, Schlichting, & Preston, 2012). Acquired equivalence refers to the spontaneous tendency to assume that if two stimuli share one property, they also share another (Figure 1; e.g., Honey & Hall, 1989; Shohamy & Wagner, 2008). Associative inference refers to the ability to derive new information by forming links between related associations, such as when a relationship is inferred between a man and a woman after they've been both associated with the same house (Figure 1; e.g., Preston et al., 2004; Zeithamova, Dominick, et al., 2012; Zeithamova & Preston, 2010). Transitive inference tasks involve learning a set of relationships (e.g., $A > B$, $B > C$, $C > D$, $D > E$), then testing whether the subjects integrated the pairwise relationships into a hierarchy by probing an untrained pair ($B ? D$; e.g., Heckers et al., 2004; Ryan et al., 2016; Zalesak & Heckers, 2009). Importantly, many studies have concluded that the hippocampal contribution to generalization and inference in these tasks cannot be reduced to its role in storing individual learning episodes, as we will discuss below.

Learning a new category is another situation in which individuals link across related experiences to form new, generalizable knowledge. In this case, generalization occurs when individuals extract a category structure from experiences with individual category members (Figure 2A) and then apply it to new members of the category (Figure 2B). Category learning has been traditionally studied in a separate line of research from episodic memory and attributed to other memory systems (Ashby et al., 1998; Ashby & Maddox, 2005; Poldrack & Foerde, 2008; Poldrack & Packard, 2003). However, several recent studies have implicated the hippocampus in category learning (Bowman & Zeithamova, 2018; Davis, Love, & Preston, 2012; Mack, Love, & Preston, 2016; Zeithamova, Maddox, & Schnyer, 2008). Moreover, theoretical models have started to outline how the hippocampus may contribute to both episodic memory and the formation of conceptual knowledge (Mack, Love, & Preston, 2018; Schapiro, Turk-Browne, Botvinick, & Norman, 2017). Thus, episodic inference and concept learning may rely on a similar rapid generalization mechanism—memory integration across related experiences—that involve the episodic memory system (Bowman & Zeithamova, 2018).

In this review, we focus on this rapid generalization that has become of interest across episodic generalization and category generalization tasks, and consider the computations and representations that may support it. The review of existing work will point to multiple ways in which the brain may represent related memories to support rapid generalization. We will argue that current evidence supports several plausible coding schemes, and that none of these coding schemes is likely to account for all existing data. Rather than demonstrating *the* mechanism of generalization, individual studies likely demonstrate *a* mechanism of generalization. Highlighting that more than one view of the hippocampal role in generalization may be valid, our review aims to facilitate a shift of focus from asking *which view is correct* to mapping out the telltale signs of underlying representations and factors that affect them, taking a step towards a more comprehensive account of generalization and its relation to memory specificity.

Do we need multiple memory representations?

The existence of multiple memory systems, forming distinct kinds of memory representations to serve distinct memory functions, has been the dominant view for some time. Nevertheless, some still doubt the need for multiple representations: the single-system view posits that multiple memory decisions could be made based on separate representations of individual experiences, without the need for generalized representations (Hintzman, 1986; Kumaran & McClelland, 2012a; Nosofsky & Johansen, 2000). In the following sections, we first expand on some of the models that embody the single-system view and demonstrate their strengths. We then transition to the evidence for the existence of generalized representations and the multiple ways such generalized representations may relate to specific memory representations.

Specific representations alone can support generalization

According to the single-system view, people form separate memories of specific events and make generalization decisions on demand, based on those specific memories (Curtis &

Jamieson, 2018; Hintzman & Ludlam, 1980; Kinder & Shanks, 2001; Zaki, Nosofsky, Jessup, & Unverzagt, 2003). Such a view provides a parsimonious explanation for the hippocampal role in both memory specificity and generalization: the hippocampus encodes individual events that can be flexibly accessed to make either type of judgment. Computational models of the single-system view have been proposed for both concept generalization (Hintzman, 1984; Kruschke, 1992; Nosofsky, 1988) and episodic inference tasks (Kumaran & McClelland, 2012a).

Formalizing the single-system view in categorization, exemplar models argue that there is no need for “generalized” concept representations to exist. Instead, concepts can be represented by specific category examples and generalization to new items can be accomplished by consideration of specific exemplars from all relevant categories (Kruschke, 1992; Medin & Schaffer, 1978; Nosofsky & Johansen, 2000). This view contrasts with the prototype models of categorization that posit that people extract the central tendency, the prototype, across category exemplars, which then guides categorization decisions (Posner & Keele, 1968, 1970; Smith & Minda, 1998). The two models (illustrated in Figure 3) have well-developed mathematical formalizations that allow for estimation of an individual’s representational strategy from their responses.

Exemplar models provide good fit to behavior (Heit, 1992; Nosofsky, 1987) and two studies have argued for the single-system view based on neural evidence (Mack, Preston, & Love, 2013; Nosofsky, Little, & James, 2012), although neither of the neuroimaging studies found any evidence for exemplar-specific representations in the hippocampus itself. Nosofsky and colleagues (2012) compared categorization, recognition, and “laxed” recognition (do not miss any old items, false alarms are acceptable). Brain activation during lax recognition was similar to categorization, which they interpreted as evidence that a single (exemplar) model supports both recognition and categorization. However, as lax recognition instructions did not ask subjects to discriminate old items from highly similar new items, one may argue that subjects were making similarity judgments rather than recognition judgments. More direct neural evidence for the exemplar model of generalization comes from a study by Mack et al. (2013) that used quantitative indices derived from the exemplar and prototype models in fMRI data analysis. Brain states during categorization were consistent with exemplar and not prototype model predictions. The exemplar network included lateral occipital, posterior parietal and lateral prefrontal cortices, which includes regions known to support memory specificity (Badre, Poldrack, Paré-Blagoev, Insler, & Wagner, 2005; Dennis, Bowman, & Vandekar, 2012; Slotnick & Schacter, 2004; Vilberg & Rugg, 2007). Thus, people may rely on memory for specific events even in traditional generalization tasks, and we should be careful when making assumptions regarding underlying representations and computations based only on whether the task includes generalization judgments.

Outside categorization, the REMERGE model was developed as a formal single-system model to account for generalization in episodic inference tasks (Banino, Koster, Hassabis, & Kumaran, 2016; Kumaran, 2012; Kumaran & McClelland, 2012a). The REMERGE model proposes that recurrent associative connections in the hippocampus easily support generalization even if individual experiences are encoded separately. For example, after

encoding a man-house association and a woman-house association, the network can efficiently make the man-woman inference as activation of a node representing the man activates a house node, which in turn activates a woman node. The co-activation of the two separate memories allows subjects to respond that the man and the woman are associated, without the need to form an integrated representation. Thus, this model provides a compelling mechanism that can account for the hippocampal role in generalization through its well-established role in forming separate representations of individual experiences.

Evidence that the hippocampal role in generalization and inference goes beyond storing specific memories

We agree that the hippocampal role in representing specific events can explain its apparent involvement in generalization to some degree (Banino et al., 2016; Kumaran, 2012; Kumaran & McClelland, 2012a; Nosofsky, Denton, Zaki, Murphy-Knudsen, & Unverzagt, 2012; Zaki, 2004). A comprehensive account of the hippocampal role in generalization will need to take the contribution of specific memories to generalization decisions into consideration. However, there are compelling findings from several generalization paradigms indicating that the hippocampal role goes beyond storing specific memories to include the formation of generalized representations that span experiences (Bowman & Zeithamova, 2018; Dusek & Eichenbaum, 1997; Preston et al., 2004; Shohamy & Wagner, 2008). A seminal study by Shohamy and Wagner (2008) used an acquired equivalence task (Figure 1) to argue that the hippocampus supports generalization beyond direct experience by integrating related memories into a combined memory representation. As evidence, they found that encoding (but not retrieval) activation in the hippocampus and midbrain differentiated good and poor generalizers. Furthermore, in good generalizers, reaction times on generalization judgments were as fast as retrieval of learned associations, indicating that links across events had already been formed at encoding.

Studies on associative inference (Figure 1) have also posited that the hippocampus contributes to generalization by integrating information across events. While two related events may be stored as separate memories and still support retrieval-based inference on demand (Figure 4, blue), overlap between events may lead to their integration in memory (Figure 4, red). Mechanistically, the overlap across events can serve as a retrieval cue, causing reactivation of related memory. As a result, one may encode an integrated memory that combines elements across the current and reactivated events. Several pieces of evidence support this account. Repetition suppression in the hippocampus, medial temporal lobe and midbrain is reduced when repeating events are related to one another, indicating that overlap with existing memories changes the way current events are encoded (Zeithamova, Manthuruthil, & Preston, 2016). Brain decoding approaches have demonstrated that prior related experiences are reactivated during encoding of overlapping events, with the degree of reactivation predicting subsequent novel inference across events (Richter, Chanales, & Kuhl, 2016; Zeithamova, Dominick, et al., 2012; Zeithamova & Preston, 2017). Similar to findings from acquired equivalence (Shohamy and Wagner, 2008), hippocampal encoding activation predicts subsequent inference success above and beyond memory for directly experienced events (Zeithamova, Dominick, et al., 2012), as also seen in transitive inference paradigms (Heckers et al., 2004). Thus, as new events are encoded in the context of prior knowledge,

the encoded memories may integrate together elements from current events and past memories, forming memory representations that transcend direct experiences.

In the area of categorization, it is also unlikely that neural category representations are only exemplar-based, despite the success of exemplar models in accounting for category generalization behavior. For example, an exemplar model MINERVA (Hintzman, 1984) can account for several prototype-like behaviors even though the underlying representation is purely exemplar-based. However, as a prototype representation emerges as an output of this exemplar network, it would be available in the brain as an input to downstream regions. Depending on regional connectivity, it could also be encoded back into the initially exemplar network itself, with the prototype becoming an increasingly robust part of the exemplar representation. The idea that reminding reactivates prior memories that are then re-encoded as a part of a new memory was promoted by Hintzman himself to account for behavioral phenomena, such as recognition confidence and judgments of frequency (Hintzman, 2004). At the neural level, recent work has shown that hippocampal outputs may be encoded back into the hippocampus via recurrent connections with cortex (Koster et al., 2018), providing a potential mechanism for prototype formation within the hippocampus. Thus, we argue that in biological networks that contain many layers and recurrent connections, a generalized (prototype) representation is likely to emerge from the representation of specific exemplars at one point in the hierarchy or another.

Empirical evidence also supports the idea that people do not always rely on memory for specific instances and indeed form generalized concept representations. One limitation to studies supporting the exemplar view of category generalization is that they often use category structures with a low coherence among category examples. When stimuli within a category have few things in common or some stimuli from one category are more similar to the central tendency of the other category, it can make the formation of a generalized category representation across such distinct items difficult (Bowman & Zeithamova, 2020; Rouder & Ratcliff, 2004). Thus, some category structures may be represented by specific exemplars, but others may not (Minda & Smith, 2001). Using more coherent categories, our recent model-based fMRI study showed better fit of the prototype compared to the exemplar model in both brain and behavior. Importantly, we found prototype-consistent signals in the anterior hippocampus, in addition to ventromedial prefrontal cortex (VMPFC) (Bowman & Zeithamova, 2018), and replicated these findings in a separate study (Bowman, Iwashita, & Zeithamova, 2020). The evidence of prototype representation in the hippocampus indicates that it may store abstract concept representation derived across exemplars.

Notably, the neural mechanisms of rapid generalization implicated in the categorization research parallel those implicated in episodic inference. Specifically, several categorization studies have shown contributions of both the hippocampus and the VMPFC to concept generalization (Bowman et al., 2020; Bowman & Zeithamova, 2018; Zeithamova et al., 2008), including evidence that these regions functionally interact (Frank, Bowman, & Zeithamova, 2019). The same VMPFC-hippocampal interactions have been also implicated in studies investigating generalization through episodic inference (Bunsey & Eichenbaum, 1996; DeVito, Lykken, Kanter, & Eichenbaum, 2010; Schlichting et al., 2015; Spalding et al., 2018; Zeithamova, Dominick, et al., 2012). For example, the strength of hippocampal-

VMPFC connectivity increased in an associative inference task with repeated presentations of overlapping information, suggesting that these regions may work together to integrate related experiences (Zeithamova, Dominick, et al., 2012). That these regions tend to also interact during concept generalization suggests that this integration mechanism contributes to multiple memory domains.

What is the relationship between specific and generalized representations?

How we encode experiences so we can both remember specific details and form generalizable knowledge is a fundamental question in memory research. The single system view aimed to resolve this question by pointing out how a single type of representations – separate representations of specific events – can be flexibly used for both specific and generalization judgments. The multiple memory systems view aimed to resolve this question by pointing out evidence for a division of labor among memory systems, with the rapidly-learning episodic systems supporting memory specificity and slow-learning systems supporting generalization. The studies discussed so far provided evidence that the hippocampus may contribute to generalization through its role in storing specific memories, but it may also form generalized representations that span experiences. Thus, these findings highlight elements from both the single system view and the traditional multiple memory systems view, but do not precisely align with either one of them.

The finding of generalized representations in the hippocampus also calls for revisiting the question of how memory can support both generalization and specificity. If related events become integrated in the hippocampus into representations that combine elements across events, can we still retain details of individual events? Or is the loss of specific details an inevitable downside of hippocampal-based generalization? We will argue that the jury is still out on this question and there may not be a single answer. Below, we discuss three views of the relationship between integrated vs. separated representations of related events, with distinct consequences for the relationship between memory specificity and generalization. Figure 5 provides a schematic depiction of these views, together with the single system view.

Integration and separation may be competing representational strategies for related memories

One possibility, often assumed in studies on episodic memory integration, is that related events are *either* encoded as separated memory representations *or* integrated into a combined representation, with dissociable mnemonic processes determining which type of representation individuals form during learning (Richter et al., 2016; Zeithamova & Preston, 2017). Compelling evidence for this view comes from studies on neuronal memory allocation in rodents. Rashid et al. (2016) and Cai et al. (2016) showed that neuronal engrams for related memories are sometimes overlapping and sometimes distinct, with distinct consequences for generalization. When two similar spatial contexts were encountered close in time, they became linked and represented through a shared neural assembly, resulting in generalization of behavior from one context the other. When they were encountered further apart, they were represented by distinct sets of neurons and behavior learned in one context did not transfer to the other. Consistent with this work,

Zeithamova and Preston (2017) used neuroimaging in humans and showed better and faster inferences across overlapping associations encountered on the same day compared to across days, accompanied by a greater multivariate integration signature during encoding of events that overlapped with same-day memories than prior-day memories. At test, fMRI activation patterns differed depending on whether inferences were made across events encountered on the same day versus those encountered across days, presumably reflecting whether participants were retrieving integrated memories or making on-demand inferences from separate memories. Collectively, this work indicates that new information related to existing memories can be linked to those existing memories or encoded as a separate memory trace, with distinct consequences for behavior. Both types of representations can support generalization and inference across events, but inferences based on separate representations are slower and less accurate (Schlichting, Zeithamova, & Preston, 2014; Shohamy & Wagner, 2008).

If representations of related events are either integrated or separated, one would expect a distinct pattern of advantages and disadvantages for each memory code. Most notably, one would expect a trade-off between remembering specific episodic details of an experience versus the ability to generalize and integrate across experiences. Integrated representations may facilitate new inferences but it may be difficult to retrieve details of a specific event without also activating elements from the related event, leading to source confusion or false memories (Gershman, Schapiro, Hupbach, & Norman, 2013; Roediger et al., 1995; Varga, Gaugler, & Talarico, 2019). A fear response in a neutral context that was experienced close in time to a context in which a mouse received a shock (Cai et al., 2016) could be viewed as an example of such false memory resulting from integration. In humans, an across-subject relationship between generalization and false memory was alluded to in the previously described study of acquired equivalence by Shohamy and Wagner (2008). From an informal debriefing, the authors noted that many subjects were unaware of being tested on new, untrained relationships and thought they had observed rather than inferred them. The few who did notice were those who generalized poorly. Presumably, most subjects formed integrated memories, helping them generalize but at a cost to source memory specificity, such as a false memory for the inferred relationship. Others encoded related associations as separate memories, leading to slower and weaker generalization but retention of specificity. However, this trade-off evidence was only anecdotal as source memory was not formally tested and was not observed in another acquired equivalence study that tested source memory directly (de Araujo Sanchez & Zeithamova, 2020). Bowman and Zeithamova (2020) added a recognition test to a category learning paradigm to more formally test the trade-off idea in concept generalization. They indeed found that participants who generalized better tended to have more false memories during an old/new recognition (Bowman & Zeithamova, 2020). However, the false memory-concept generalization relationship was only marginally significant and the recognition performance was overall quite poor. Thus, it is possible that there are individual differences in the tendency to integrate vs. separate related events, with distinct consequences for memory specificity and generalization, but more conclusive evidence is yet to be established.

On-demand generalization from specific memories can lead to the formation of integrated representations

Even if events are initially stored as separate representations, those representations may be affected in response to generalization demands. For example, the previously discussed computational model REMERGE (Kumaran, 2012; Kumaran & McClelland, 2012b) postulates that separated representations can easily support new inferences through recurrent associative connections. In our example, separate man-house and woman-house associations provide a pathway through which to activate the representation of the woman after being cued with the man (Figure 4, blue). However, it is plausible that once indirectly related elements are co-activated through recurrent connections, they can become associated directly (for example through Hebbian learning). A neural implementation of such a model has been recently described by Koster and colleagues (2018). A loop recurrence exists within the medial temporal lobe, where entorhinal layers that receive output from the hippocampus are connected with the entorhinal layers that provide input to the hippocampus. This provides a mechanism of how the hippocampal output to the entorhinal cortex can be recirculated as a new input back into the hippocampus. Thus, once the indirect association (the woman) is generated by the network in response to the cue (the man), the man-woman association may then be encoded by the hippocampus, leading to the formation of a memory representation that integrates elements across events. This idea also aligns with our prior discussion that generalized concept representations would be expected to emerge in the brain at one point of the hierarchy, even if we assume exemplar representations are at the base of the hierarchy (Hintzman, 1984; Hintzman & Curran, 1994). The recurrent models take this idea one step further, providing a mechanism for how initially separated memories may become integrated in the hippocampus itself. Integrated memories would form from separate ones as long as the separated memories become co-activated, which may happen spontaneously over time or more rapidly in response to task demands.

Behavioral data consistent with the idea that related memories can start out as separated representations but become co-activated and linked in response to task demands (Koster et al., 2018; Kumaran & McClelland, 2012a) have been recently reported in two episodic inference paradigms. In acquired equivalence, de Araujo Sanchez and Zeithamova (2020) observed reliable generalization learning during test even though test trials did not include feedback. Little evidence for generalization of preferences from one face to another (Figure 1A) was found the first time it was tested, but generalization tendency increased with repeated testing. Presumably, as related experiences were repeatedly jointly co-activated in response to task-demands, the associations between co-activated indirectly related elements were re-encoded as new memories, increasing the probability of generalization with each subsequent test. In associative inference, Carpenter and Schacter (Carpenter & Schacter, 2017) had subjects learn overlapping person-object associations, overlaid on a background indoor scene. For example, there may be a man with a toy in a room with a white couch and the overlapping associations may be a child with the same toy in a different room with a brown couch. Tests probed memory for direct associations (e.g., the man and the toy), inferred associations (e.g., the man and the child), as well as specific context details (e.g., was the man near a white, brown or green couch?). Results showed reduced memory for specific contextual details (such as increased false attribution of the brown couch rather than

the white couch to the man) but *only after successful inference*. The authors proposed that initially separate memories of related events became integrated after making a generalization judgment, which led to the loss of specific details that were distinguishing between related events (see also, Carpenter & Schacter, 2018a). Thus, the need to make inferences across experiences may facilitate a fast transformation from specific to generalized memories, similar to what is assumed to happen more slowly over time through consolidation (McClelland et al., 1995; McKenzie & Eichenbaum, 2011; Norman, 2010; O'Reilly & Norman, 2002; Winocur, Moscovitch, & Sekeres, 2007).

Experiences may be simultaneously represented at multiple levels of specificity

Does linking related information in service of generalization—whether spontaneously at encoding or in response to task demands—inevitably come at the cost to memory specificity? Or can both specific and generalized representations form in parallel, despite both involving the hippocampus? Generalization and false memory are often seen as flip sides of the same coin, both driven by memory integration (Carpenter & Schacter, 2017, 2018a; Roediger et al., 1995; Shohamy & Wagner, 2008; Varga et al., 2019; Zeithamova, Schlichting, et al., 2012). However, a relatively limited number of studies have explicitly tested the generalization-specificity tradeoff for the same stimuli, in the same subjects. Furthermore, not all studies that focused on such a potential tradeoff found it. As one example, Banino and colleagues (2016) tested generalization and source memory in the associative inference paradigm, asking participants after each test trial whether it was directly learned (AB, BC) or indirect (AC). AC inference success was positively related to the source memory for AB and BC trials, indicating that memory specificity and generalization go hand in hand. However, testing the relationship between AC inference and AC source memory would provide a more direct test of the tradeoff hypothesis. Second, anecdotal evidence from Shohamy and Wagner (2008) indicated that generalization of preferences from one face to another may be associated with false memory for that inferred association, but de Araujo Sanchez and Zeithamova (2020) did not find such relationship when using an explicit source memory probe. Finally, following up on their work on generalization-false memory relationship in young adults (Carpenter & Schacter, 2017, 2018a), Carpenter and Schacter found no evidence of it when testing a group of older adults (2018b). In a more applied domain, Chang and colleagues (2019) trained math fluency (known to increase hippocampal engagement) and found better transfer of trained math problems to novel ones to be associated with *better* discrimination of novel from trained problems. Thus, behavioral evidence does not rule out that it is possible to form generalizable knowledge *and* retain specific memories at the same time.

One challenge with determining whether specific and generalized memories may co-exist using purely behavioral methods is the potential contribution of specific memories to generalization decisions. In other words, no behavioral tradeoff between generalization and source memory may be observed because generalized representations may form alongside specific ones, but also because participants relied to a greater degree on separate memories instead of generalized representations (Banino et al., 2016; de Araujo Sanchez & Zeithamova, 2020). The degree of a trade-off (or lack thereof) observed in a given study may then indicate the degree of reliance on specific vs. generalized representations in that

particular study. Thus, it is important to also consider converging evidence from the implementation level. Is there neural evidence to help us determine whether specific and generalized representations may co-exist in parallel, even when both types of representations involve the hippocampus?

Notably, at least two mechanisms have been proposed for how the hippocampus may represent events at multiple levels of specificity in parallel. One mechanism stems from differences in the computational properties across hippocampal subfields (e.g., Schapiro, 2017). The dentate gyrus and CA3 are thought to maintain sparse representations of specific episodes that can be reinstated from partial input. In contrast, CA1 receives both retrieved memory input from CA3 and current perceptual input from the cortex (Steward, 1976). This makes CA1 well positioned to form representations spanning across experiences, linking elements from current events with reinstated prior memories. A biologically plausible computational model of the hippocampus that implements these dissociations has been shown to be able to learn the specifics of individual experiences while also extracting regularities across those experiences (Schapiro et al., 2017). The model is able to account for both episodic memory and generalization learning through separate anatomical pathways within the hippocampus: the monosynaptic pathways from cortex to CA1 can account for generalization learning while the trisynaptic pathway from cortex through dentate gyrus and CA3 to CA1 can maintain specific representations. Both rodent work (Clelland et al., 2009; Leutgeb, Leutgeb, Moser, & Moser, 2007; Nakashiba, Young, McHugh, Buhl, & Tonegawa, 2008) and human neuroimaging work (Schlichting et al., 2014; Zeithamova et al., 2016) indicate functional differences among hippocampal subfields that align with such specificity/generalization distinction, with CA3 supporting detailed representations of individual events and CA1 supporting integration across events.

Functional dissociations that map well onto the specificity/generalization dissociation also exist along the long axis of the hippocampus (Poppenk, Evensmoen, Moscovitch, & Nadel, 2013). Animal work shows an anterior-posterior gradient in receptive field size (Kjelstrup et al., 2008), with larger receptive fields in anterior (ventral in rodents) and smaller in posterior (dorsal in rodents) hippocampus. This indicates that the same information is simultaneously represented at multiple spatial scales, such as perhaps coding of one's position within a room, position of the room within a house, and position of the house in the wider spatial context. Consistent with the gradient in rodents, a recent human fMRI study has shown larger overlap in spatial and temporal representations in anterior compared to posterior hippocampus (Brunec et al., 2018). Tentatively, the overlapping spatial and temporal signals in anterior hippocampus may help link across similar contexts while the more distinct representations in posterior hippocampus may help with discrimination between similar contexts.

The anterior-posterior hippocampal gradient is also apparent in memory paradigms that are not purely spatial in nature. Schlichting and colleagues (2015) used neural pattern similarity to index hippocampal representations of novel objects after participants learned overlapping sets of object associations (e.g., object A with B, object B with C). Patterns in the anterior hippocampus were consistent with an integrated coding scheme for the overlapping associations (A and C becoming similarly represented following learning) while neural

patterns in posterior hippocampus were consistent with separated representations (A and C becoming less similarly represented following learning). Bowman and Zeithamova (2018, 2020) found generalized concept representations (prototypes) to be unique to anterior rather than posterior hippocampus. Collin and colleagues (Collin, Milivojevic, & Doeller, 2015) had participants watch movie scenes in which two seemingly unrelated scenes were connected by a third scene to form a larger narrative. Posterior hippocampal activation patterns were more closely linked with small-scale information while anterior hippocampal patterns were more closely linked to large-scale narratives that bridged across scenes and included inferred relationships. Thus, rather than representing events by a single engram, the hippocampus may represent events concurrently at multiple levels of specificity, with coarse, big picture representations providing broader context for detailed representations of individual events. Whether at the level of hippocampal subfields or through a representational gradient along the long hippocampal axis, representing information at multiple levels of specificity and multiple timescales would provide a mechanism for forming generalizable knowledge while still remembering specific details.

Notably, while the hippocampus may support both specificity and generalization, it appears to interact with distinct sets of cortical memory regions to support each function. As noted previously, research on both associative inference and concept generalization highlighted the role of hippocampal-VMPFC interactions in memory integration (Bowman & Zeithamova, 2018; Frank et al., 2019; Schlichting et al., 2015; Zeithamova, Dominick, et al., 2012). More broadly, the VMPFC is a region implicated by a large number of studies in schema-related memory (Baldassano, Hasson, & Norman, 2018; Brod, Lindenberger, & Shing, 2017; Ghosh, Moscovitch, Melo Colella, & Gilboa, 2014; Romero, Barense, & Moscovitch, 2019; Spalding, Jones, Duff, Tranel, & Warren, 2015; Tse et al., 2011; van Kesteren, Fernandez, Norris, & Hermans, 2010). Although not reported in episodic inference studies, lateral temporal cortex is also thought to represent some forms of generalized memories, such as semantic knowledge (Mummery et al., 2000; Renoult, Irish, Moscovitch, & Rugg, 2019) and gist representations (Dennis, Bowman, & Peterson, 2014; Dennis, Kim, & Cabeza, 2008; Turney & Dennis, 2017). Thus, the VMPFC, and perhaps lateral temporal cortices, may play a broader role in highlighting commonalities across experiences.

In contrast, other cortical memory regions, such as lateral prefrontal and lateral parietal regions, are instead implicated in differentiating similar memories to prevent interferences and maintain specificity (Badre & Wagner, 2005; Hutchinson, Uncapher, & Wagner, 2009; Kuhl & Chun, 2014). Interestingly, regions implicated in exemplar-based vs. prototype-based category representations align quite well with these cortical representational differences: generalized (prototype) concept representations have been found in VMPFC and lateral temporal cortices (Bowman & Zeithamova, 2018) while specific exemplar representations have been found in lateral parietal and lateral prefrontal cortices (Mack et al., 2013). A recent study found that both prototype and exemplar representations can form across these distinct regions within the same task (Bowman et al., 2020). Finally, Frank and colleagues (2019) found greater resting state and background connectivity of putative generalization regions (VMPFC, lateral temporal cortex) with the anterior hippocampus and putative specificity regions (lateral prefrontal, lateral parietal cortex) with the posterior hippocampus. These findings provide one mechanism how the hippocampus may form

distinct types of representations along the anterior-posterior axis through interactions with distinct cortical memory regions, differentially implicated in specificity and generalization. Because of differential neural mechanisms within the hippocampus and across cortex, specific and generalized representations can in principal co-exist rather than trade-off, to inform a range of judgments.

Conclusions

Reconciling how the brain is able to form representations that support two fundamental memory functions, specificity and generalization, has been a long-standing issue in memory research. In the current review, we first highlighted research that has blurred the division of labor between distinct memory systems tailored for each function, emphasizing the contribution of the hippocampus to many forms of generalization. We then outlined evidence in support of several potential representational schemes that may account for such hippocampal-based generalization, as schematically summarized in Figure 5. The evidence comes from tasks that span memory domains typically thought to rely on distinct memory and neural systems – episodic inference and category learning – suggesting they are likely domain general and not linked to a particular type of generalization task. Most broadly, the hippocampus likely contributes specific memories to generalization decisions, but also forms generalized representations that span experiences, indicating a hippocampal role in generalization that goes beyond remembering specific events. Furthermore, such generalized representations may form *instead* of specific ones, *from* specific ones, or perhaps *in addition to* them. While one of these views may become dominant in future research, current evidence suggests that more than one view may be valid, perhaps under different conditions or in different individuals. For example, under some conditions, individuals may form primarily separated representations of related events during learning and use those separate representations to generalize on demand during retrieval. In some of those cases, on-demand generalization may cause those separated representations to become integrated, causing loss of memory specificity. Other times, integrated representations may form without replacing or disrupting memory for individual events, allowing both specificity and generalization judgments without trade-off. The multi-layered and interconnected nature of the brain should be sufficient to support each of these types of representations.

Many studies to date have focused on demonstrating the existence of a specific mechanism, such as the existence of generalized representations in the hippocampus, the contribution of specific memories to generalization decisions, or demand-driven integration of initially separated memories. Importantly, taken together, the results indicate that any given study is likely uncovering *a* mechanism of generalization rather than *the* mechanism of generalization. We should not necessarily dismiss one view because a particular study found evidence for another. The question then becomes, not which of these representational schemes is correct, but rather what conditions lead memories to be represented one way versus another? Some conditions that promote one type of representational scheme over another have already been a focus of investigation and discussed in the current review, such as the coherence of a category structure or the temporal proximity of related events. Other conditions that promote each type of representation will be the focus of future research. Different representational schemes have different consequences for behavior, which will be

important for fundamental understanding of memory as well as applications of memory research to areas such as education. Using cognitive models, measuring both specificity and generalization in the same study, and leveraging neural indices to answer cognitive questions will be increasingly important for fully characterizing the memory representations underlying the multiple functions of memory.

Acknowledgments

Funding: This work was supported by the National Institutes of Health [grant numbers R01-NS112366 (DZ) and F32-AG054204 (CRB)];

References

- Ashby FG, Alfonso-Reese LA, Turken AU, & Waldron EM (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, 105(3), 442–481. 10.1037/0033-295X.105.3.442 [PubMed: 9697427]
- Ashby FG, & Maddox WT (2005). Human Category Learning. *Annual Review of Psychology*, 56(1), 149–178. 10.1146/annurev.psych.56.091103.070217
- Badre D, Poldrack RA, Paré-Blagoev EJ, Insler RZ, & Wagner AD (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*. 10.1016/j.neuron.2005.07.023
- Badre D, & Wagner AD (2005). Frontal lobe mechanisms that resolve proactive interference. *Cerebral Cortex*, 15(12), 2003–2012. 10.1093/cercor/bhi075 [PubMed: 15788702]
- Baldassano C, Hasson U, & Norman KA (2018). Representation of Real-World Event Schemas during Narrative Perception. *The Journal of Neuroscience*. 10.1523/jneurosci.0251-18.2018
- Banino A, Koster R, Hassabis D, & Kumaran D (2016). Retrieval-Based Model Accounts for Striking Profile of Episodic Memory and Generalization. *Scientific Reports*, 6 10.1038/srep31330
- Bowman CR, Iwashita T, & Zeithamova D (2020). Model-based fMRI reveals co-existing specific and generalized concept representations. *BioRxiv*.
- Bowman CR, & Zeithamova D (2018). Abstract memory representations in the ventromedial prefrontal cortex and hippocampus support concept generalization. *The Journal of Neuroscience*. 10.1523/JNEUROSCI.2811-17.2018
- Bowman CR, & Zeithamova D (2020). Training set coherence and set size effects on concept generalization and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(8), 142–1464.
- Bozoki A, Grossman M, & Smith EE (2006). Can patients with Alzheimer’s disease learn a category implicitly? *Neuropsychologia*, 44(5), 816–827. 10.1016/j.neuropsychologia.2005.08.001 [PubMed: 16229868]
- Brod G, Lindenberger U, & Shing YL (2017). Neural activation patterns during retrieval of schema-related memories: differences and commonalities between children and adults. *Developmental Science*. 10.1111/desc.12475
- Brunec IK, Bellana B, Ozubko JD, Man V, Robin J, Liu ZX, ... Moscovitch M (2018). Multiple Scales of Representation along the Hippocampal Anteroposterior Axis in Humans. *Current Biology*, 28(13), 2129–2135.e6. 10.1016/j.cub.2018.05.016 [PubMed: 29937352]
- Bunsey M, & Eichenbaum H (1996). Conservation of hippocampal memory function in rats and humans. *Nature*, 379(6562), 255–257. 10.1038/379255a0 [PubMed: 8538790]
- Carpenter AC, & Schacter DL (2017). Flexible retrieval: when true inferences produce false memories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 43(3), 335–349.
- Carpenter AC, & Schacter DL (2018a). False memories, false preferences: Flexible retrieval mechanisms supporting successful inference bias novel decisions. *Journal of Experimental Psychology: General*, 147(7), 988–1004. 10.1037/xge0000391 [PubMed: 29419307]

- Carpenter AC, & Schacter DL (2018b). Flexible retrieval mechanisms supporting successful inference produce false memories in younger but not older adults. *Psychology and Aging*, 33(1), 134–143. 10.1037/pag0000210 [PubMed: 29494184]
- Chang H, Rosenberg-Lee M, Qin S, & Menon V (2019). Faster learners transfer their knowledge better: Behavioral, mnemonic, and neural mechanisms of individual differences in children’s learning. *Developmental Cognitive Neuroscience*, 40 10.1016/j.dcn.2019.100719
- Clelland CD, Choi M, Romberg C, Clemenson GD, Fragniere A, Tyers P, ... Bussey TJ (2009). A functional role for adult hippocampal neurogenesis in spatial pattern separation. *Science (New York, N.Y.)*. 10.1126/science.1173215
- Collin SHP, Milivojevic B, & Doeller CF (2015). Memory hierarchies map onto the hippocampal long axis in humans. *Nature Neuroscience*, 18(11), 1562–1564. 10.1038/nn.4138 [PubMed: 26479587]
- Curtis ET, & Jamieson RK (2018). Computational and empirical simulations of selective memory impairments: Converging evidence for a single-system account of memory dissociations. *Quarterly Journal of Experimental Psychology*.
- Davis T, Love BC, & Preston AR (2012). Learning the exception to the rule: Model-based fMRI reveals specialized representations for surprising category members. *Cerebral Cortex*. 10.1093/cercor/bhr036
- de Araujo Sanchez MA, & Zeithamova D (2020). Generalization and source memory in acquired equivalence. *PsyArXiv*.
- Dennis NA, Bowman CR, & Peterson KM (2014). Age-related differences in the neural correlates mediating false recollection. *Neurobiology of Aging*, 35(2), 395–407. [PubMed: 24094578]
- Dennis NA, Bowman CR, & Vandekar SN (2012). True and phantom recollection: An fMRI investigation of similar and distinct neural correlates and connectivity. *NeuroImage*, 59(3), 2982–2993. [PubMed: 22001165]
- Dennis NA, Kim H, & Cabeza R (2008). Age-related differences in brain activity during true and false memory retrieval. *Journal of Cognitive Neuroscience*, 20(8), 1390–1402. 10.1162/jocn.2008.20096 [PubMed: 18303982]
- DeVito LM, Lykken C, Kanter BR, & Eichenbaum H (2010). Prefrontal cortex: Role in acquisition of overlapping associations and transitive inference. *Learning & Memory*, 17(3), 161–167. 10.1101/lm.1685710 [PubMed: 20189961]
- Dusek JA, & Eichenbaum H (1997). The hippocampus and memory for orderly stimulus relations. *Proc Natl Acad Sci U S A*, 94(13), 7109–7114. 10.1073/pnas.94.13.7109 [PubMed: 9192700]
- Frank LE, Bowman CR, & Zeithamova D (2019). Differential Functional Connectivity along the Long Axis of the Hippocampus Aligns with Differential Role in Memory Specificity and Generalization. *Journal of Cognitive Neuroscience*. 10.1162/jocn_a_01457
- Gershman SJ, Schapiro AC, Hupbach A, & Norman K. a. (2013). Neural context reinstatement predicts memory misattribution. *The Journal of Neuroscience*, 33(20), 8590–8595. 10.1523/JNEUROSCI.0096-13.2013 [PubMed: 23678104]
- Ghosh VE, Moscovitch M, Melo Colella B, & Gilboa A (2014). Schema representation in patients with ventromedial PFC lesions. *The Journal of Neuroscience*, 34(36), 12057–12070. 10.1523/JNEUROSCI.0740-14.2014 [PubMed: 25186751]
- Heckers S, Zalesak M, Weiss AP, Ditman T, & Titone D (2004). Hippocampal activation during transitive inference in humans. *Hippocampus*, 14(2), 153–162. 10.1002/hipo.10189 [PubMed: 15098721]
- Heit E (1992). Categorization using chains of examples. *Cognitive Psychology*. 10.1016/0010-0285(92)90011-P
- Hintzman DL (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, 16(2), 96–101. 10.3758/BF03202365
- Hintzman DL (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review*, 93(4), 411–428. 10.1037/0033-295X.93.4.411
- Hintzman DL (2004). Judgment of frequency versus recognition confidence: Repetition and recursive reminding. *Memory and Cognition*. 10.3758/BF03196863

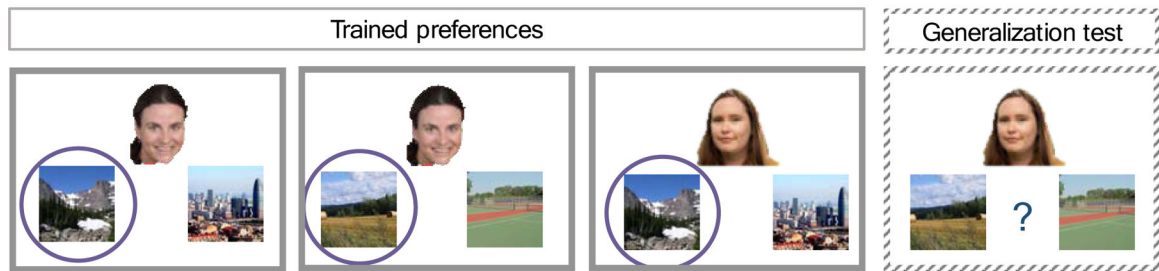
- Hintzman DL, & Curran T (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language*, 10.1006/jmla.1994.1001
- Hintzman DL, & Ludlam G (1980). Differential forgetting of prototypes and old instances: simulation by an exemplar-based classification model. *Memory & Cognition*, 8(4), 378–382. 10.3758/BF03198278 [PubMed: 7421579]
- Honey RC, & Hall G (1989). Acquired Equivalence and Distinctiveness of Cues. *Journal of Experimental Psychology: Animal Behavior Processes*. 10.1037/0097-7403.15.4.338
- Hutchinson JB, Uncapher MR, & Wagner AD (2009). Posterior parietal cortex and episodic retrieval: convergent and divergent effects of attention and memory. *Learning & Memory (Cold Spring Harbor, N.Y.)*, 16(6), 343–356. 10.1101/lm.919109
- Kinder A, & Shanks DR (2001). Amnesia and the declarative/nondeclarative distinction: A recurrent network model of classification, recognition, and repetition priming. *Journal of Cognitive Neuroscience*, 13(5), 648–669. 10.1162/089892901750363217 [PubMed: 11506662]
- Kjelstrup KB, Solstad T, Brun VH, Hafting T, Leutgeb S, Witter MP, ... Moser M-B (2008). Finite Scale of Spatial Representation in the Hippocampus. *Science*, 321(5885), 140–143. 10.1126/science.1157086 [PubMed: 18599792]
- Knowlton BJ, Mangels JA, & Squire LR (1996). A neostriatal habit learning system in humans. *Science (New York, N.Y.)*, 273(5280), 1399–1402. 10.1126/science.273.5280.1399
- Knowlton BJ, & Squire LR (1993). The learning of categories: parallel brain systems for item memory and category knowledge. *Science (New York, N.Y.)*, 262(5140), 1747–1749. 10.1126/science.8259522
- Koster R, Chadwick MJ, Chen Y, Berron D, Banino A, Düzel E, ... Kumaran D (2018). Big-Loop Recurrence within the Hippocampal System Supports Integration of Information across Episodes. *Neuron*, 99(6), 1342–1354. 10.1016/j.neuron.2018.08.009 [PubMed: 30236285]
- Kruschke JK (1992). ALCOVE: an exemplar-based connectionist model of category learning. *Psychological Review*, 99(1), 22–44. 10.1037/0033-295X.99.1.22 [PubMed: 1546117]
- Kuhl BA, & Chun MM (2014). Successful Remembering Elicits Event-Specific Activity Patterns in Lateral Parietal Cortex. *Journal of Neuroscience*, 34(23), 8051–8060. 10.1523/JNEUROSCI.4328-13.2014 [PubMed: 24899726]
- Kumaran D (2012). What representations and computations underpin the contribution of the hippocampus to generalization and inference? *Frontiers in Human Neuroscience*, 6(June), 157 10.3389/fnhum.2012.00157 [PubMed: 22675298]
- Kumaran D, & McClelland JL (2012a). Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychological Review*, 119(3), 573–616. 10.1037/a0028681 [PubMed: 22775499]
- Kumaran D, & McClelland JL (2012b). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, 119(3), 573–616. 10.1037/a0028681 [PubMed: 22775499]
- Leutgeb JK, Leutgeb S, Moser M-B, & Moser EI (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science (New York, N.Y.)*, 315(5814), 961–966. 10.1126/science.1135801
- Mack ML, Love BC, & Preston AR (2016). Dynamic updating of hippocampal object representations reflects new conceptual knowledge. *Proceedings of the National Academy of Sciences*. 10.1073/pnas.1614048113
- Mack ML, Love BC, & Preston AR (2018). Building concepts one episode at a time: The hippocampus and concept formation. *Neuroscience Letters*, 680, 31–38. 10.1016/j.neulet.2017.07.061 [PubMed: 28801273]
- Mack ML, Preston AR, & Love BC (2013). Decoding the brain's algorithm for categorization from its neural implementation. *Current Biology*, 23(20), 2023–2027. 10.1016/j.cub.2013.08.035 [PubMed: 24094852]
- McClelland JL, McNaughton BL, & O'Reilly RC (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of

- connectionist models of learning and memory. *Psychological Review*, 102(3), 419–457. 10.1037/0033-295X.102.3.419 [PubMed: 7624455]
- McKenzie S, & Eichenbaum H (2011). Consolidation and Reconsolidation: Two Lives of Memories? *Neuron*. 10.1016/j.neuron.2011.06.037
- Medin DL, & Schaffer MM (1978). Context theory of classification learning. *Psychological Review*, 85(3), 207–238. 10.1037/0033-295X.85.3.207
- Minda JP, & Smith JD (2001). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal Of Experimental Psychology-Learning Memory And Cognition*, 27(3), 775–799. 10.1037/0278-7393.27.3.775
- Mummery CJ, Patterson K, Price CJ, Ashburner J, Frackowiak RSJ, & Hodges JR (2000). A voxel-based morphometry study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*. 10.1002/1531-8249(200001)47:1<36::AID-ANA8>3.0.CO;2-L
- Nakashiba T, Young JZ, McHugh TJ, Buhl DL, & Tonegawa S (2008). Transgenic inhibition of synaptic transmission reveals role of CA3 output in hippocampal learning. *Science*. 10.1126/science.1151120
- Norman KA (2010). How hippocampus and cortex contribute to recognition memory: Revisiting the complementary learning systems model. *Hippocampus*, 20(11), 1217–1227. 10.1002/hipo.20855 [PubMed: 20857486]
- Nosofsky RM (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 13(1), 87–108. 10.1037/0278-7393.13.1.87
- Nosofsky RM (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(4), 700–708. 10.1037/0278-7393.14.4.700
- Nosofsky RM, Denton SE, Zaki SR, Murphy-Knudsen AF, & Unverzagt FW (2012). Studies of implicit prototype extraction in patients with mild cognitive impairment and early Alzheimer's disease. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 38(4), 860–880. 10.1037/a0028064
- Nosofsky RM, & Johansen MK (2000). Exemplar-based accounts of "multiple-system" phenomena in perceptual categorization. *Psychonomic Bulletin & Review*, 7(3), 375–402. [PubMed: 11082849]
- Nosofsky RM, Little DR, & James TW (2012). Activation in the neural network responsible for categorization and recognition reflects parameter changes. *Proceedings of the National Academy of Sciences*. 10.1073/pnas.1111304109
- O'Connell G, Myers CE, Hopkins RO, McLaren RP, Gluck MA, & Wills AJ (2016). Amnesic patients show superior generalization in category learning. *Neuropsychology*. 10.1037/neu0000301
- O'Reilly RC, & Norman KA (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences*. 10.1016/S1364-6613(02)02005-3
- Packard MG, Hirsh R, & White NM (1989). Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. *J Neurosci*, 9(5), 1465–1472. <https://doi.org/0270-6474/89/051465-08> [PubMed: 2723738]
- Poldrack RA, & Foerde K (2008). Category learning and the memory systems debate. *Neuroscience and Biobehavioral Reviews*. 10.1016/j.neubiorev.2007.07.007
- Poldrack RA, & Packard MG (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, 41(3), 245–251. 10.1016/S0028-3932(02)00157-4 [PubMed: 12457750]
- Poppenk J, Evensmoen HR, Moscovitch M, & Nadel L (2013). Long-axis specialization of the human hippocampus. *Trends in Cognitive Sciences*. 10.1016/j.tics.2013.03.005
- Posner MI, & Keele SW (1968). On the Genesis of Abstract Ideas. *Journal of Experimental Psychology*, 77(3, Pt.1), 353–363. 10.1037/h0025953 [PubMed: 5665566]
- Posner MI, & Keele SW (1970). The retention of abstract ideas. *Journal of Experimental Psychology*, 83(2), 304–308. 10.1037/h0028558

- Preston AR, Shrager Y, Dudukovic NM, & Gabrieli JDE (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus*, 14, 148–152. [PubMed: 15098720]
- Renoult L, Irish M, Moscovitch M, & Rugg MD (2019). From Knowing to Remembering: The Semantic–Episodic Distinction. *Trends in Cognitive Sciences*. 10.1016/j.tics.2019.09.008
- Richter FR, Chanals AJH, & Kuhl BA (2016). Predicting the integration of overlapping memories by decoding mnemonic processing states during learning. *NeuroImage*, 124, 323–335. 10.1016/j.neuroimage.2015.08.051 [PubMed: 26327243]
- Roediger HL, McDermott KB, Hintzman DL, Lindsay S, Rajaram S, & Tulving E (1995). Creating False Memories: Remembering Words Not Presented in Lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 803–814. 10.1037/0278-7393.21.4.803
- Romero K, Barense MD, & Moscovitch M (2019). Coherence and congruency mediate medial temporal and medial prefrontal activity during event construction. *NeuroImage*. 10.1016/j.neuroimage.2018.12.047
- Rouder JN, & Ratcliff R (2004). Comparing Categorization Models. *Journal of Experimental Psychology: General*. 10.1037/0096-3445.133.1.63
- Ryan JD, D’Angelo MC, Kamino D, Ostreicher M, Moses SN, & Rosenbaum RS (2016). Relational learning and transitive expression in aging and amnesia. *Hippocampus*, 26(2), 170–184. 10.1002/hipo.22501 [PubMed: 26234960]
- Schapiro AC, Turk-Browne NB, Botvinick MM, & Norman KA (2017). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372 10.1098/rstb.2016.0049
- Schlichting ML, Mumford JA, & Preston AR (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, 6, 8151 10.1038/ncomms9151
- Schlichting ML, Zeithamova D, & Preston AR (2014). CA1 subfield contributions to memory integration and inference. *Hippocampus*, 24(10), 1248–1260. 10.1002/hipo.22310 [PubMed: 24888442]
- Scoville WB, & Milner B (1957). Loss of recent memory after bilateral hippocampal lesions. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 20(11), 11–21. 10.1136/jnnp.20.1.11
- Shohamy D, & Wagner AD (2008). Integrating Memories in the Human Brain: Hippocampal-Midbrain Encoding of Overlapping Events. *Neuron*, 60(2), 378–389. 10.1016/j.neuron.2008.09.023 [PubMed: 18957228]
- Slotnick SD, & Schacter DL (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience*, 7(6), 664–672. 10.1038/nn1252 [PubMed: 15156146]
- Smith JD, & Minda JP (1998). Prototypes in the Mist: The Early Epochs of Category Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(6), 1411–1436. 10.1037/h0090333
- Spalding KN, Jones SH, Duff MC, Tranel D, & Warren DE (2015). Investigating the Neural Correlates of Schemas: Ventromedial Prefrontal Cortex Is Necessary for Normal Schematic Influence on Memory. *J Neurosci*, 35(47), 15746–15751. 10.1523/JNEUROSCI.2767-15.2015 [PubMed: 26609165]
- Spalding KN, Schlichting ML, Zeithamova D, Preston AR, Tranel D, Duff MC, & Warren DE (2018). Ventromedial prefrontal cortex is necessary for normal associative inference and memory integration. *The Journal of Neuroscience*. 10.1523/JNEUROSCI.2501-17.2018
- Squire LR, & Zola SM (1998). Episodic memory, semantic memory, and amnesia. *Hippocampus*. 10.1002/(SICI)1098-1063(1998)8:3<205::AID-HIPO3>3.0.CO;2-I
- Steward O (1976). Topographic organization of the projections from the entorhinal area to the hippocampal formation of the rat. *Journal of Comparative Neurology*. 10.1002/cne.901670303
- Tse D, Takeuchi T, Kakeyama M, Kajii Y, Okuno H, Tohyama C, ... Morris RGM (2011). Schema-dependent gene activation and memory encoding in neocortex. *Science*, 333(August), 891–895. 10.1126/science.1205274 [PubMed: 21737703]
- Tulving E (1987). Multiple memory systems and consciousness. *Human Neurobiology*.

- Turney IC, & Dennis NA (2017). Elucidating the neural correlates of related false memories using a systematic measure of perceptual relatedness. *NeuroImage*, 146, 940–950. 10.1016/j.neuroimage.2016.09.005 [PubMed: 27608601]
- van Kesteren MTR, Fernandez G, Norris DG, & Hermans EJ (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences*. 10.1073/pnas.0914892107
- Varga NL, Gaugler T, & Talarico J (2019). Are mnemonic failures and benefits two sides of the same coin?: Investigating the real-world consequences of individual differences in memory integration. *Memory & Cognition*, 47(3), 496–510. 10.3758/s13421-018-0887-4 [PubMed: 30617747]
- Vilberg KL, & Rugg MD (2007). Dissociation of the neural correlates of recognition memory according to familiarity, recollection, and amount of recollected information. *Neuropsychologia*, 45(10), 2216–2225. 10.1016/j.neuropsychologia.2007.02.027 [PubMed: 17449068]
- Winocur G, Moscovitch M, & Sekeres M (2007). Memory consolidation or transformation: Context manipulation and hippocampal representations of memory. *Nature Neuroscience*. 10.1038/nn1880
- Zaki SR (2004). Is categorization performance really intact in amnesia? A meta-analysis. *Psychonomic Bulletin & Review*, 11(6), 1048–1054. 10.3758/BF03196735 [PubMed: 15875974]
- Zaki SR, Nosofsky RM, Jessup NM, & Unverzagt FW (2003). Categorization and recognition performance of a memory-impaired group: evidence for single-system models. *Journal of the International Neuropsychological Society : JINS*, 9(3), 394–406. 10.1017/S1355617703930050 [PubMed: 12666764]
- Zalesak M, & Heckers S (2009). The role of the hippocampus in transitive inference. *Psychiatry Research - Neuroimaging*, 172(1), 24–30. 10.1016/j.psychresns.2008.09.008
- Zeithamova D, Dominick ALL, & Preston ARR (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron*, 75(1), 168–179. 10.1016/j.neuron.2012.05.010 [PubMed: 22794270]
- Zeithamova D, Maddox WT, & Schnyer DM (2008). Dissociable prototype learning systems: evidence from brain imaging and behavior. *Journal of Neuroscience*, 28(49), 13194–13201. 10.1523/JNEUROSCI.2915-08.2008 [PubMed: 19052210]
- Zeithamova D, Manthuruthil C, & Preston AR (2016). Repetition suppression in the medial temporal lobe and midbrain is altered by event overlap. *Hippocampus*, 26(11), 1464–1477. 10.1002/hipo.22622 [PubMed: 27479864]
- Zeithamova D, & Preston AR (2010). Flexible memories: differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 30(44), 14676–14684. 10.1523/JNEUROSCI.3250-10.2010 [PubMed: 21048124]
- Zeithamova D, & Preston AR (2017). Temporal proximity promotes integration of overlapping events. *Journal of Cognitive Neuroscience*, 29(8), 1311–1323. 10.1162/jocn_a_01116 [PubMed: 28253077]
- Zeithamova D, Schlichting ML, & Preston AR (2012). The hippocampus and inferential reasoning: building memories to navigate future decisions. *Frontiers in Human Neuroscience*, 6(March), 1–14. 10.3389/fnhum.2012.00070 [PubMed: 22279433]

Acquired equivalence paradigm



Associative inference paradigm



Figure 1. Example episodic inference tasks.

Episodes often share elements. Episodic inference tasks test the degree to which new knowledge can be inferred by linking information across different learning episodes.

Acquired equivalence. Through feedback-based learning, participants learn a correct choice for a set of cues. Circles (not visible to the participants while they are making their guesses) denote the correct choice for a given clue. For example, when presented with the first face, participants should choose the mountain image over the city image, indicating that Face 1 prefers mountains over city. Generalization through acquired equivalence—the tendency to assume that Faces that share one association also share another—is tested. For example, Face 1 and Face 2 both prefer the mountains. Face 1 also prefers fields over tennis courts. The tendency to choose fields over tennis court for Face 2, although that preference was never trained, would indicate generalization of preferences across the two faces. **Associative inference.** Participants are asked to remember pairs of images (paired associate learning). Some pairs share elements, providing an opportunity to link elements from events experienced at different times. For example, a woman and a man are presented on different trials with the same house. Inference of indirect relationship (here, the man and the woman live together) is then tested.

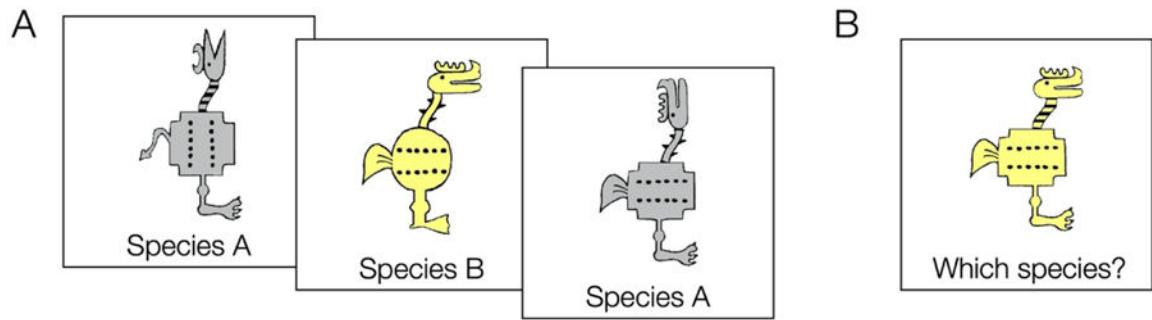


Figure 2. Example categorization task.

A. In training, participants learn category labels (here, species) for a set of category members (here cartoon animals). Training can be observational, where the label is presented with the stimulus, or feedback-based, where participants guess on each trial and then receive feedback. **B.** In a generalization test, participants categorize new animals that were not part of the training set, typically without feedback.

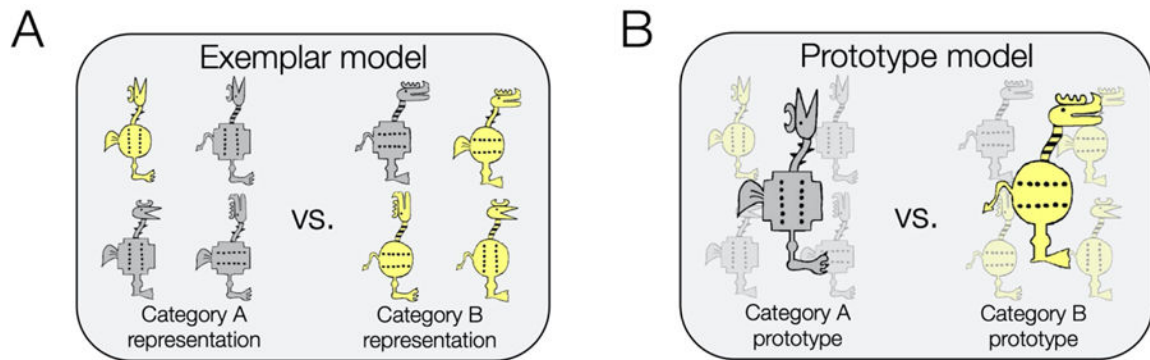


Figure 3. Formal categorization models.

A. Exemplar models posit that a category is represented by its specific exemplars. In this example, taken from Bowman and Zeithamova, 2018, participants were trained to categorize four members of category A and four members of category B, all cartoon animals with 8 binary dimensions. The trained exemplars are thought to form the representations of the two categories. Joint consideration of all exemplars from both categories underlies generalization judgments. **B.** Prototype models posit that a category is represented by its central tendency (prototype) generalized from specific instances. In this example, category prototypes combine all features that are characteristic for each category. Categorization judgments are then based on comparison of a stimulus to the two category prototypes.

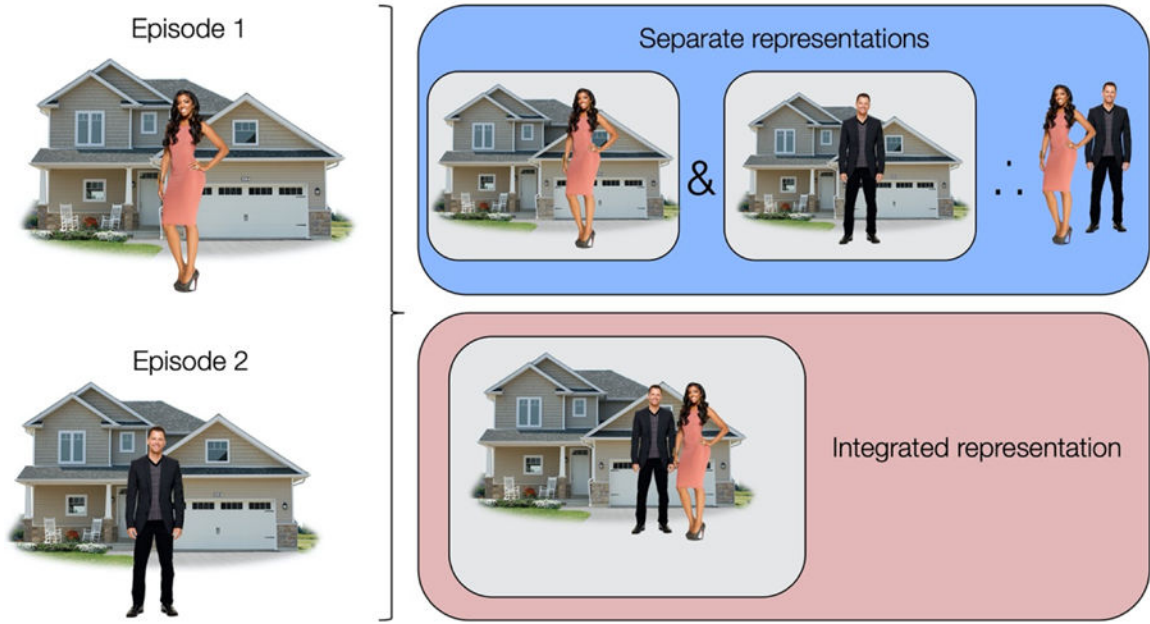


Figure 4.

Two episodes with overlapping elements can be encoded as two separate representations (blue) or linked together into an integrated representation (red). New information (here, the man and the woman are a couple) can be inferred by retrieval and joint consideration of separate specific memories (blue). Alternatively, integrated memories represent the inferred relationship directly (red).

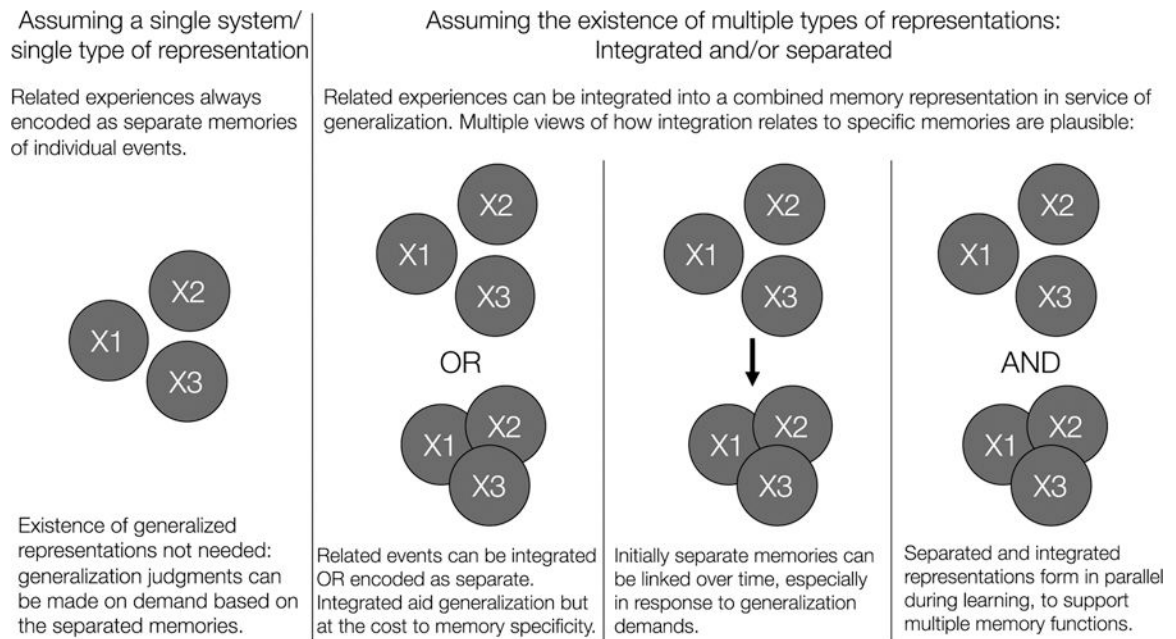


Figure 5.

Multiple views of how related experiences (X1, X2, X3) may be represented in memory to support generalization across experiences. From left to right: Generalization may rely on the same representations as memory specificity—separated memories of individual events. Alternatively, related events may be encoded into an integrated representation that span experiences. Integrated memories may form *instead of* separated memories, *from* initially separated memories, or *alongside* separated memories.