Title: Multi-step planning can be improved across the human lifespan with individualized memory interventions

Sharon M Noh[1]*, Keiland W Cooper[2], Craig EL Stark[2,3], Aaron M Bornstein[1,3]*

1. Department of Cognitive Sciences, The University of California, Irvine, Irvine, CA, 92697
2. Department of Neurobiology and Behavior, The University of California, Irvine, Irvine, CA, 92697
3. Center for the Neurobiology of Learning and Memory, The University of California, Irvine, Irvine, CA, 92697

   * To whom correspondence should be addressed: nohsm@uci.edu, aaron.bornstein@uci.edu

Keywords: learning, memory, reasoning, judgment, aging, planning, inference

Acknowledgements

The authors wish to thank Mohit Nadkarni for extensive assistance in data collection, Tyson Kerr for programming assistance, and Judith Fan for helpful discussions. Funding was provided by NIH R21AG072673 and R01NS119468 (to AMB), R01AG066683 and P30AG066519 (to CELS), F32AG072836 (to SMN), and NSF DGE-1839285 (to KWC).

Abstract

Memory and judgment are both affected by age-related cognitive decline. We test the hypothesis that these declines stem from a common mechanism. We adapted a latent learning task that measures one's ability to learn and make judgments about associations among interconnected stimuli and evaluated performance across the lifespan. Younger adults generally showed evidence of successful latent learning and inference, and variability in judgment performance was explained by mnemonic discrimination ability. In a second experiment, we examined if performance can be improved for those with lower memory ability by separating overlapping object pairs across time, in an attempt to reduce memory interference during learning. We found that mnemonic discrimination ability interacted with training condition to improve judgment performance. This interaction was also captured using an artificial neural network simulation that varied training sequence and network capacity to predict judgment performance. These findings suggest that age-related declines in complex judgments can be improved via targeted training interventions.

Significance Statement

Humans across the lifespan are often required to make complex decisions for themselves and others, yet research shows that older adults are impaired at various forms of decision-making. We aimed to understand how this impairment arises by examining learning and reasoning behaviors that underlie goal-directed decisions. We found that age-related impairments in memory precision can explain deficits in the learning and reasoning processes involved in goal-directed planning, and further found that memory-based training interventions can improve these learning and reasoning behaviors for those with weaker memory abilities. These results suggest that deficits in older adult decision-making may result from age-related declines in memory precision and introduce promising avenues for interventions that can preserve learning and reasoning behaviors underlie goal directed decisions across the lifespan.

Introduction

Humans across the lifespan are often asked to make decisions with lasting impacts for themselves and loved ones. Recent trends have shown that increasing numbers of older adults are holding positions of power at the highest levels of government and the top companies in the country, which makes the decision-making abilities of these top executives particularly important for their constituents or employees[1–4]. Yet with age also comes a decline in several cognitive abilities, including forms of decision-making[5–7]. In particular, older adults have been shown to have impairments in making judgments that require inferring multi-step associations, such as in forward planning[8–10]. However, the mechanism by which this capability declines with age is unclear, frustrating the search for interventions that can reverse or slow these effects. One challenge in understanding why older adults struggle with such judgments is that current behavioral assays do not distinguish whether declines in performance can be attributed to the judgment process itself or to the learning of associations needed to make such judgments, for instance in model-based planning[11,12].

In many real-world settings, humans often need to infer structure from memories of disparate experiences to make flexible, goal-directed plans. For example, let's imagine that a student (Alice) is making travel plans to present at a conference held out of town. She books a 4PM flight on Friday and plans to call a Lyft at noon to ensure she will get to the airport with enough time to spare. Even if this is her first conference, she can engage in this type of multi-step planning by stitching together information gained from various related experiences to accomplish her goal (namely, getting to the conference): having experienced many flight delays on domestic airlines, she may choose to book a flight that gets in the night before her presentation, rather than the morning of. She may use her knowledge of the local traffic patterns to anticipate the need to leave an extra hour of commute time on Friday afternoons. Her decision to take a Lyft to the airport rather than her own car may have been influenced by prior experiences in which she had driven to the airport and struggled to find parking because the lot was full.

Recent work in younger adults has shown that this type of goal-directed decision-making may indeed depend on one's ability to infer structure from their environment. Rmus and colleagues[13] developed a task in which participants learned to infer associative structure among a collection of image pairs arranged along a latent "graph"[13]. Participants were able to implicitly acquire these complex associative structures through a randomized sequence of exposures to individual edges of the graph (*latent structure learning*), revealed via their use of this knowledge to reason about shortest paths on the graph (*multi-step associative inference*). Furthermore, the degree to which participants' behavior reflected the underlying graph corresponded to increased use of model-based planning on a separate decision-making task. A major strength of the task developed by Rmus et al.[13] is its ability to sensitively measure the extent of latent structure learning and multi-step inference by testing judgments between nodes that vary in the number of intermediate associations between them. One open question, however, is *how* structure learning occurs, and what type of knowledge representations might support the formation of these latent structures.

When trying to understand the mechanism by which individuals successfully acquire latent structures from experience, it seems reasonable to assume that memory would play a role in the learning process. Recent work has suggested that model-based planning involves hippocampal contributions and effective memory search[14–16]. In the task by Rmus et al., structural inference likely depends on accurate memory representation of the individual associations that make up the graph[13]. More specifically, memory *precision* may be important when trying to form knowledge representations necessary for efficient and flexible model-based planning. For example, remembering what traffic is like on Friday afternoons specifically, as opposed to Monday afternoons or Friday evenings, would help Alice make more accurate predictions for how to plan her trip. In episodic memory, high memory precision can be achieved via *pattern separation*–a neural computation in which competing information is represented as distinct, non-overlapping neural patterns[17,18]. By reducing overlap in neural representations, pattern separation protects against interference between competing memories and in doing so supports mnemonic discrimination of similar events[19,20]. Pattern separation is well documented to show age-related decline[17,21–23], and this decline in pattern separation ability could explain why older adults are highly susceptible to memory interference and related memory failures[24,25]. Thus, one goal of the present study was to determine whether the deficits seen in multi-step planning in older adults are a result of failures in the learning of latent structures due to increased memory interference and age-related declines in memory precision.

If older adults' increased susceptibility to memory interference is indeed hindering their ability to adequately learn the latent structures necessary to engage in multistep planning, then training interventions designed to improve memory encoding may improve structure learning and multistep associative inference. In support of this idea, work in episodic memory has shown that manipulating learning sequences or separating the presentation of associations with overlapping elements across time can reduce memory interference and bias the formation of distinct neural representations to support associative inference via different learning strategies[26–28].

What is the nature of the representations that support structural inference?

Research in memory has suggested that different types of knowledge may be supported by different kinds of neural representations[19,29,30]. Studies that have examined associative inference have identified two types of neural representations that can support memory-based inference[26,27]. In typical associative inference paradigms, participants are told to memorize associations with overlapping elements. In the standard version of the task, participants study A-B pairs, then later encounter B-C pairs. Participants are then tested on their knowledge of the indirect association between A and C, despite never having been presented with an A-C pair. If participants encoded A-B and B-C as separate episodes during learning via pattern separation, they may form localist (orthogonalized) representations of each pair[27,31] and successful A-C inference would occur through effortful retrieval and recombination of individual memories at time of test (e.g., "A was paired with B, and B was paired with C, so A is associated with C"). While localist representations lead to high memory precision and are resistant to memory interference, these representations are less efficient for making inferences, as they require effort and could pose a working memory burden as one engages in memory search to retrieve the relevant information and recombine the information as necessary to make inference judgments[26,27,32].

Another possibility is that learners may encode A-B and, upon encountering B-C, reactivate their memory of A-B and update this existing representation with the newly encountered information (C). By reactivating existing knowledge during encoding, connections can be formed between existing memories and new incoming information[26,32–34]. This process of memory integration could form more distributed representations in which neurons represent multiple entities such as A-B-C. In this scenario, learners can more rapidly make inferences between A and C, as they may have learned to represent A, B, and C within the same distributed network. While distributed representations support rapid generalization and may be ideal for inferring structure from experience, these representations are prone to interference and false memories[27,35,36], which may imply that forming these representations can be difficult and may not be ideal for those who are already susceptible to memory interference due to age-related cognitive decline. If age-related failures in inference are due to age-related failures in memory precision, we expect that a learning manipulation that can reduce memory interference should have the potential to improve latent structure learning and subsequent judgments.

In the present study, we examined the degree to which age-related decline in memory precision influences performance on a multi-step associative inference task that relies on latent learning, which we refer to as the *graph task*[13]. In the first experiment, participants studied a series of object pairs presented randomly during the study phase (Fig. 1A, *graph learning*). Unbeknownst to the participants, the objects were drawn from an underlying graph structure (Fig. 1B) made up of 12 nodes (objects) and 16 edges (distinct pairs). On a later test, participants were asked to judge the relative distances between two object pairs to examine the degree to which participants were able to mentally navigate their learned associations and use them to make pairwise shortest-path judgments (Fig. 2A). These types of judgments are a fundamental computation for high-level planning[37]. We expected to see deficits in performance on the structural inference task as a function of increasing age, but more specifically due to low memory precision. We hypothesized that performance on the judgment task should depend on the precision of memory representations participants form during the graph learning phase, so participants also completed a separate Mnemonic Similarity Task (Fig. 3, MST) to get an independent measure of individual differences in memory ability (specifically mnemonic discrimination). The MST was selected for its ability to tax pattern separation processes, and its sensitivity in tracking age-related declines in memory precision through its measure of mnemonic discrimination[38]. In a follow-up experiment, we examined whether judgments can be improved for older adults and/or those with lower memory precision by manipulating the presentation sequence of object pairs shown in the study phase to reduce memory interference during learning. We predicted that by separating overlapping object pairs across time during learning (blocked training), those with low memory precision may show improvements in structural inference and subsequent model-based planning (Fig. 1C). To affirm that this relationship depended on memory precision specifically, we simulated performance of the task using variants of artificial neural network models that differed only in their internal representational capacity and compared the model outputs to our behavioral findings.

Method

Experiment 1 and 2 were identical in procedure other than the presentation order that was used during learning. As such, we have combined the method across both experiments and specified the differences across experiments accordingly.

Participants. 219 total participants (112 female, ages 19-84, mean(sd) age=55.7(14.2)) were recruited online (via Amazon Mechanical Turk). 113 participants completed Experiment 1 (59 female, ages 22-84, mean(sd) age=56.7(13.8)), and 106 additional participants were recruited for Experiment 2 (53 female, ages 19-79, mean(sd) age=54.8(14.7)). Participants received monetary compensation for participating in the study. Participants had to complete a brief tutorial and screener of the task to be eligible for the full experiment. The screener consisted of a simple rotation detection task during practice trials of the learning phase to ensure participants were paying attention to each trial. Individuals had 2 opportunities to complete 10 practice trials with a minimum 70% accuracy to continue on to the full experiment.

Procedure

<u>Structural Inference (“Graph”) Task</u>

Study Phase. Participants performed an associative learning task in which they were asked to view and remember a series of object pairs (Fig. 1A). Each pair was presented on the screen for 1 second in a randomized order, and participants had to make a rotation judgment on each trial to ensure they were paying attention throughout the task. Participants are told that they have to memorize the object pairs for a later test, but are not given any additional information about the relationship between object pairs. Unbeknownst to participants, the object pairs are part of an underlying graph structure based on the associations encountered during learning (Fig. 1B). In Experiment 1, each of the 16 unique object pairs were repeated 44 times (for a total of 704 trials) and presented in a randomized sequence during the learning phase (Fig. 1C, “intermixed sequence”). In Experiment 2, object pairs were grouped into 4 mini “blocks,” with each mini-block containing 4 unique object pairs. The 4 pairs presented in each mini-block did not have any overlapping nodes–rather, overlapping nodes were separated and assigned to different mini-blocks. The goal of this manipulation was to improve learning of the graph structure by reducing the potential for memory interference during encoding by separating confusable (overlapping) edges across time.

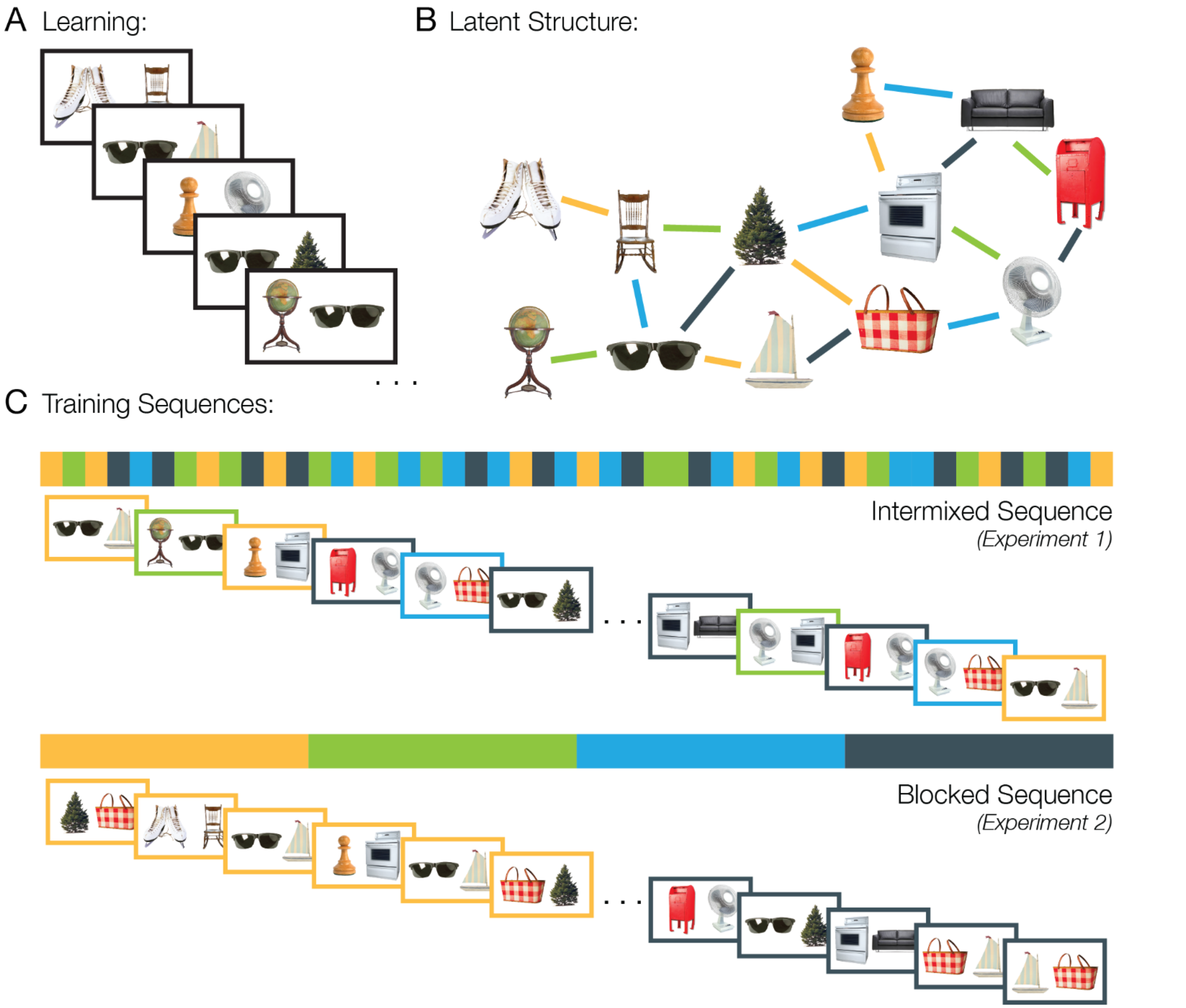


Figure 1. Structural Inference Task learning phase. Participants learn individual edges (A) drawn from a latent structure made up of several overlapping edges (B). In Experiment 1, edges are drawn randomly and presented in an intermixed fashion (C, intermixed), whereas in Experiment 2, overlapping edges are separated in time across different mini-blocks (C, blocked).

Following the study phase, participants completed two different tests to measure how well participants can *use* their knowledge of the graph structure (Fig. 2A, *Judgment Test*), as well as how well they *learned* the graph structure (Fig. 2B, *Graph Reconstruction Test)*.

Judgment Test. After completing the learning phase, participants performed a relative distance judgment task. This test is used to measure how well participants are able to use the information they learned during the study phase to make structural inference-based judgments. During this judgment task, participants were asked to determine the relative distance between three objects taken from the study phase. Participants were asked to indicate which object was closest to the central object (reference node) based on the indirect relationships they learned about the objects in the previous learning phase. Participants used the left and right keyboard buttons to indicate their choice on each trial and completed 204 trials with up to 10 seconds to respond to each trial. Each unique object was selected as a central reference node 17 times, and the two options used on each trial were randomly selected with the constraint that neither object was directly paired with the central node during the study phase, and that the two objects had a different relative distance from the central node. Judgment phase trials varied in difficulty based on the degree to which the choice options differed in associative distance from the reference node (Fig. 2A). The difference in relative distance between each choice option and the reference node ranged from 1 (most difficult) to 3 (easiest). Accuracy was computed within each difficulty bin (distance difference 1, 2, and 3).

Because data collected via online methods are unsupervised and therefore tend to be noisy, data from the judgment test were screened for outliers and filtered prior to analysis. In the judgment phase, data were combined across all conditions (age and training condition) and screened using reaction time measures, as reaction times were not used for any of our analyses of interest. The median reaction time for judgment phase trials was 1368 ms. Judgment trials were screened via a 2-step process. First, we set a reasonable range of reaction time data to consider: trials with reaction times less than 400 ms were excluded, as reaction times faster than this cutoff were likely unrealistic (the lowest average reaction time in the judgment phase measured across all participants was 463.2 ms). Trials with reaction times greater than 3000 ms were also excluded, as these reaction times are unrealistically long and could suggest significant connectivity or lag issues, and/or lack of focus or attention on the task from the participant. Once this general range was set [400 ms, 3000 ms], we retained trials within 2 times the interquartile range ([Q3-1.5*IQR, Q1+1.5*IQR]) of the remaining reaction times [494ms, 2226ms]. This restricted range still covered the interquartile range of the original, unfiltered reaction time data (median=1368ms, IQR=[951, 2062]). This focused range allows us to have enough variability to examine differences in strategy that may influence judgment performance, while excluding trials in which participants may not be properly focused or engaged in the task. These exclusion criteria and methods are similar to other methods that have used reaction time measures to filter out noise in data collected through online methods[39].

Data from the judgment phase were analyzed using linear mixed-effects models in R's nlme package[40]. Difficulty (distDiff=1, 2, or 3, with 1 being the most difficult), Age, LDI, and Sequence (blocked vs. interleaved) were specified as fixed effects and participants were specified as random effects in the models.

Graph Reconstruction Test. After the judgment test, participants completed a graph reconstruction task to test knowledge of the graph structure they learned during the study phase. During the graph reconstruction task, participants were shown all 12 objects used during the learning phase and were asked to arrange the objects on a "canvas" (indicated by a large white rectangle on the computer screen) by using their mouse to click and drag the objects onto the canvas. Participants were asked to link the objects that were directly paired together during the study phase. Objects can be linked or unlinked by clicking two objects on the canvas, one after the other. Linked objects were connected with a straight line between them. All objects had to be placed on the canvas, and each object had to be connected to at least one other object before participants could submit their final graph and complete the task.

We computed several metrics using participants' graph reconstruction. If participants connected two objects that were indeed paired together during the study phase, the drawn edges were counted as "hits" (Fig. 5A). When participants connected two objects that were not directly paired together during the study phase, the drawn edges were counted as "false alarms" (Fig. 5B). We also examined the total number of edges that were drawn by each participant (Fig. 5C). Finally, we computed two different "accuracy" metrics to assess how accurately participants reconstructed what they learned: 1) accuracy (Fig. 5D) was computed for each participant by taking the proportion of correctly connected pairs ("hits"/16) and subtracting the proportion of incorrectly connected pairs ("false alarms"/50) and 2) a relative measure of accuracy (Fig. 5E) was computed for each participant using the following equation: $(hits - 0.32 * false\ alarms)/(total\ edges\ drawn)$. This scaled accuracy metric takes into account the expected number of hits and false alarms given the total number of edges drawn by each participant. For instance, if a participant constructed their graph with only 12 edges, but those 12 edges were all "true" edges, then this metric would assign an accuracy of 1 for that participant. Thus, the scaled accuracy measure ranges from 1 to -1: a value of 1 would mean that the participant's graph consists of only true edges, whereas a value of -1 indicates a participant's graph consists entirely of false edges. A value of 0 would be the expected value if participants were drawing arbitrary edges. Four participants were excluded from the analysis for having drawn over 25 false edges, which is more than half the number of total possible false edges on the graph and well outside of the interquartile range of false alarms in the unfiltered data (median=5, IQR=[3, 7]).

Data from the graph reconstruction phase were analyzed in R. Using a linear regression approach (lm function in R), we separately analyzed how the number of hits, false alarms, total edges drawn, and our two measures of reconstruction accuracy differed as a function of mnemonic discrimination ability (LDI) and training Sequence (blocked vs. intermixed training). For each analysis, the full model was specified as follows: DV~LDI+Sequence+LDI*Sequence. In the absence of a significant interaction between LDI and training sequence, the interaction term was removed to test for any main effects of LDI and training sequence[41].

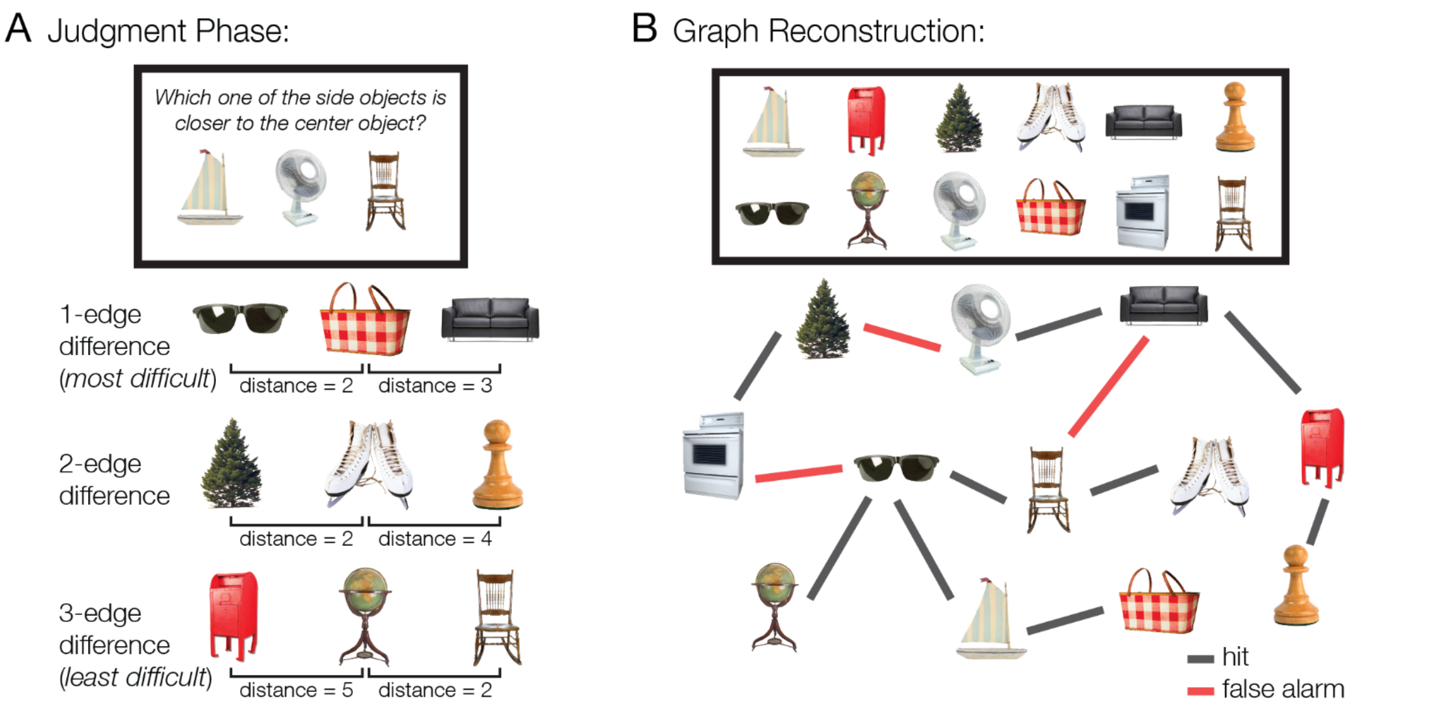


Figure 2. Structural Inference Task test phases. (A) After the learning phase, structural inference-based judgments were assessed for each participant. Participants were presented with 3 objects and asked whether the object on the left or right was closer to the center object, based on the associations they learned in the previous study phase. Judgment phase trials varied in difficulty based on the difference between choice options. The most difficult trials were ones in which options differed by an associative distance of 1 (distance 2 vs. 3, 3 vs. 4, or 4 vs. 5), whereas the easiest trials were ones in which choice options differed by an associative distance of 5 (distance 2 vs. 5). (B) After the judgment phase, participants were asked to reconstruct the graph to the best of their knowledge by placing all studied objects on a "canvas" on their screen and connecting objects only if they had been directly paired together during the study phase. Correctly drawn connections were classified as "hits," whereas incorrectly drawn connections were classified as "false

alarms." Participants were required to place all object on the canvas, and each object had to be connected to at least one other object to complete this phase.

Mnemonic Similarity Task (MST)

Participants also completed the Mnemonic Similarity Task (Fig. 3). This task was used as an independent measure for assessing individual differences in memory ability and how mnemonic discrimination ability might explain individual differences in performance on the graph task.

Incidental Encoding Phase. During the encoding phase, participants viewed a series of object images and completed a cover task which consisted of making indoor/outdoor judgments for each object. None of the images used in the MST overlapped with any of the images used in the graph task, and participants were not informed that they would have to memorize the object images.

Mnemonic Discrimination Test. On a later surprise discrimination test, participants are shown object images that are identical (old), similar (lures), or completely different (foils) relative to the studied items from the encoding phase and are told to indicate whether the object shown is "old", "similar", or "new" relative to what they saw before. A Lure Discrimination Index (LDI) was computed for each participant using the following metric: p('sim'|lure) - p('sim'|foil). Higher LDIs correspond to the ability to more accurately categorize lures, rather than foils, as being "similar" to previously studied items, which is indicative of better mnemonic discrimination and memory encoding precision[38].

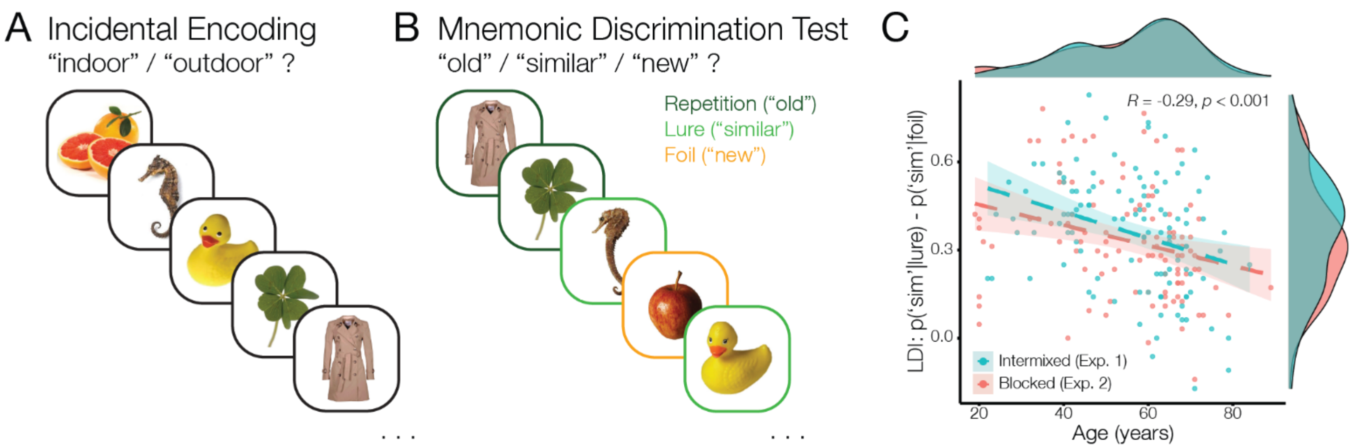


Figure 3. The Mnemonic Discrimination Task. (A) Participants view a sequence of objects during an incidental encoding phase in which participants are asked to classify each object as an indoor or outdoor object. (B) In a surprise discrimination test, participants view a series of objects and are asked to determine whether each object is an "old," "new," or "similar" item relative to what was shown during the encoding phase. (C) Relationship between chronological age and LDI across participants in both experiments. Shaded bands indicate 95% confidence intervals around the best fit regression line for each experiment (dashed lines).

Artificial Neural Network Model (ANN)

To further examine the behavior of participants under different simulated conditions, we constructed a neural network model which implements a minimal-assumption approach to latent structure learning. To avoid confusion between terminology between the graph and the neural network model, node/edge will be used when referring to the graph, and unit/weight will be used when referring to the neural network. The network's input represented the presented stimulus of one node of the pair and was trained to learn the mapping between the pairs' complementary vertex. To simulate variance in memory encoding ability, the number of units in the second layer of the network was varied. Specifically, we constructed a fully connected five-layer feedforward neural network implemented in PyTorch (Version 1.13.1[42]), with 12 units in the input and output layers (one for each of the 12 items in the experiments) and 3 units in the hidden layer. To model varying degrees of pattern separation, the width of the second and fourth layers were varied at incremental intervals at a fourth the size of the input vector, with either 6, 9, 12, 15, or 18 units. At smaller unit widths, encoding incoming information requires more many-to-one mappings, reflecting low memory capacity and increased potential for memory interference. The model's activation function at unit *i* of layer *l* was defined as:

$$h_l(i) = f(w^T_{il} h_{l-1})$$

where *f* is a pointwise function, $\mathbf{w}_{il}$ is a vector of learnable parameters, $\mathbf{h}_{l-1}$ represents the output of the previous layer, and *T* is the matrix transposition operation. The input layer is referred to as h_0.

For each of the two conditions, the networks were exposed to the same image pair stimuli as participants, *<x, y>*. Each of the 12 images were one-hot encoded as an input unit corresponding to each image of the pair, set to a value of 1, and with all other inputs set to 0. Similarly, the output unit corresponding to the other paired image was set to a value of 1, and all other outputs were set to 0.

The model was trained to minimize the difference between each node in the pair: given one node of the pair, the network was trained to pattern-complete its complement. To accomplish this, we estimated the parameters $\mathbf{w}_{il}$, for all *i* and *l*, using stochastic gradient descent with the Adam optimizer[43] to minimize the cosine similarity between the input and output. We used a rectified nonlinear unit for the pointwise function, f, for all layers except the output layer, where we used a sigmoid function. To account for the variance in the stochastic weight initialization and optimize training time, 7 models were trained for each condition with weights randomly initialized from a uniform distribution between –0.5 and +0.5. Models were trained for 25 epochs, with each epoch containing 704 trials. In the blocked condition, 4 blocks of 176 trials (totaling 704 trials) each were used.

To assess the ability of the model to infer the latently learned graph, a relative distance judgment task was performed, which required the model to judge the relative distance between three randomly-selected nodes, unconstrained by edge relationship. Since nodes which share edges will be presented more frequently to the model, we hypothesized that the model would represent nodes with shorter graph path lengths as being more similar. Conversely, we hypothesized that nodes with longer path lengths would be represented dissimilarly. To indicate which of the two nodes in the judgment task was closer to the reference node, we used the minimum cosine distance between the model's internal representation (the third network layer) for each of the three presented nodes.

Results

Experiment 1: Better memory encoding ability improves structural inference-based judgments

Judgment Test. To examine the role of age-related memory decline on structural inference-based judgments, we conducted a linear mixed effects regression predicting accuracy on the judgment test (Fig. 4A). We tested whether trial difficulty (measured by relative distance difference between choice options) and chronological age (in years) were significant predictors of judgment test accuracy: judgment accuracy~agedistDiffage*distDiff, random=~1|subject (adjusted R^2=0.23, AIC=-152.55). There was a main effect of age (β_{age}=-0.030, SE=0.012, t(110)=-2.57, p=0.011), with older adults generally performing worse than younger adults on structural inference judgments. There was also a main effect of difficulty ($\beta_{distDiff}$=0.032, SE=0.009, t(216)=3.50, p < 0.001): participants performed near chance for the most difficult trials (distDiff=1), and performance steadily improved with decreasing trial difficulty (i.e., increasing relative distance between choice options). This result confirmed that trials did indeed increase in difficulty as a function of the associative difference between choice options, and we were able to observe the full range of performance (i.e., chance performance in distDiff=1 to perfect accuracy in distDiff=3) across participants in this task. There was also a significant interaction between trial difficulty and age ($\beta_{age*distDiff}$=-0.022, SE=0.009, t(216)=-2.41, p=0.017) such that older adults performed worse than younger adults on structural inference judgments, but this age-related deficit was most apparent in easier trials. Specifically, there was no difference in performance as a function of age on the most difficult trials (distDiff=1), as judgment accuracy on those trials was no different from chance regardless of age. However, younger adults showed evidence of better memory-based inference as trials got easier (i.e., with increasing relative distance difference) and the degree to which younger adults outperformed older adults increased for lower-difficulty trials.

While we predicted that older adults would perform worse than younger adults on structural inference judgments, we had a more specific hypothesis that judgment accuracy would decline as a function of decreasing memory ability, as structural inference judgments are likely dependent on the precision of memory representations formed during the graph learning phase. To test the possibility that memory precision could drive the age-related decline seen in structural inference performance, we took a multiple regression approach and added mnemonic discrimination ability scores, measured separately with the MST, into our linear mixed effects model. LDI was added as a fixed effect into the model, with higher LDIs corresponding to better memory precision (i.e., greater mnemonic discrimination ability). The full model was specified as: judgment accuracy~age+distDiff+LDI+age*distDiff+age*LDI+distDiff*LDI+age*distDiff*LDI, random=~1|subject (adjusted R^2=0.26, AIC=-133.98). This model revealed that LDI, rather than chronological age, was a better predictor of judgment performance. When LDI was considered in the model, the age*distDiff interaction was no longer significant ($\beta_{age*distDiff}$=-0.012, SE=0.010, t(214)=-1.24, p=0.218), and neither was the main effect of age (β_{age}=-0.021, SE=0.012, t(108)=-1.72, p=0.088). Instead, there was a significant main effect of memory ability (β_{LDI}=0.031, SE=0.013, t(108)=2.42, p=0.017) such that better memory precision predicted better performance on structural inference-based judgments. There was also a significant interaction between trial difficulty and memory ability ($\beta_{LDI*distDiff}$=0.033, SE=0.010, t(214)=3.34, p=0.001) due to the fact that those with better memory precision showed better structural inference judgments, but the performance advantage only emerged in easier trials. There was no significant interaction between chronological age and LDI, and the three-way interaction between age, LDI, and difficulty was also not significant.

Because LDI appeared to be a better predictor of structural inference-based judgments than chronological age, we conducted another linear mixed-effects model using LDI, rather than chronological age, in predicting judgment accuracy (Fig. 4B). We tested whether trial difficulty and mnemonic discrimination ability (LDI) were significant predictors of judgment test accuracy: judgment accuracy~LDI+distDiff+LDI*distDiff, random=~1|subject (adjusted R^2=0.26, AIC=-164.19). There was a main effect of memory ability (β_{LDI}=0.035, SE=0.012, t(110)=2.99, p=0.003) such that greater mnemonic discrimination ability predicted better structural inference judgments. There was also a main effect of difficulty ($\beta_{distDiff}$=0.032, SE=0.009, t(216)=3.50, $p < 0.001$) with judgment performance improving for easier trials (i.e., greater relative distance between choice options). There was also a significant interaction between trial difficulty and memory ability ($\beta_{LDI*distDiff}$=0.034, SE=0.009, t(216)=3.84, $p < 0.001$) such that better memory ability predicted better performance on structural inference judgments, but this pattern emerged for lower-difficulty trials.

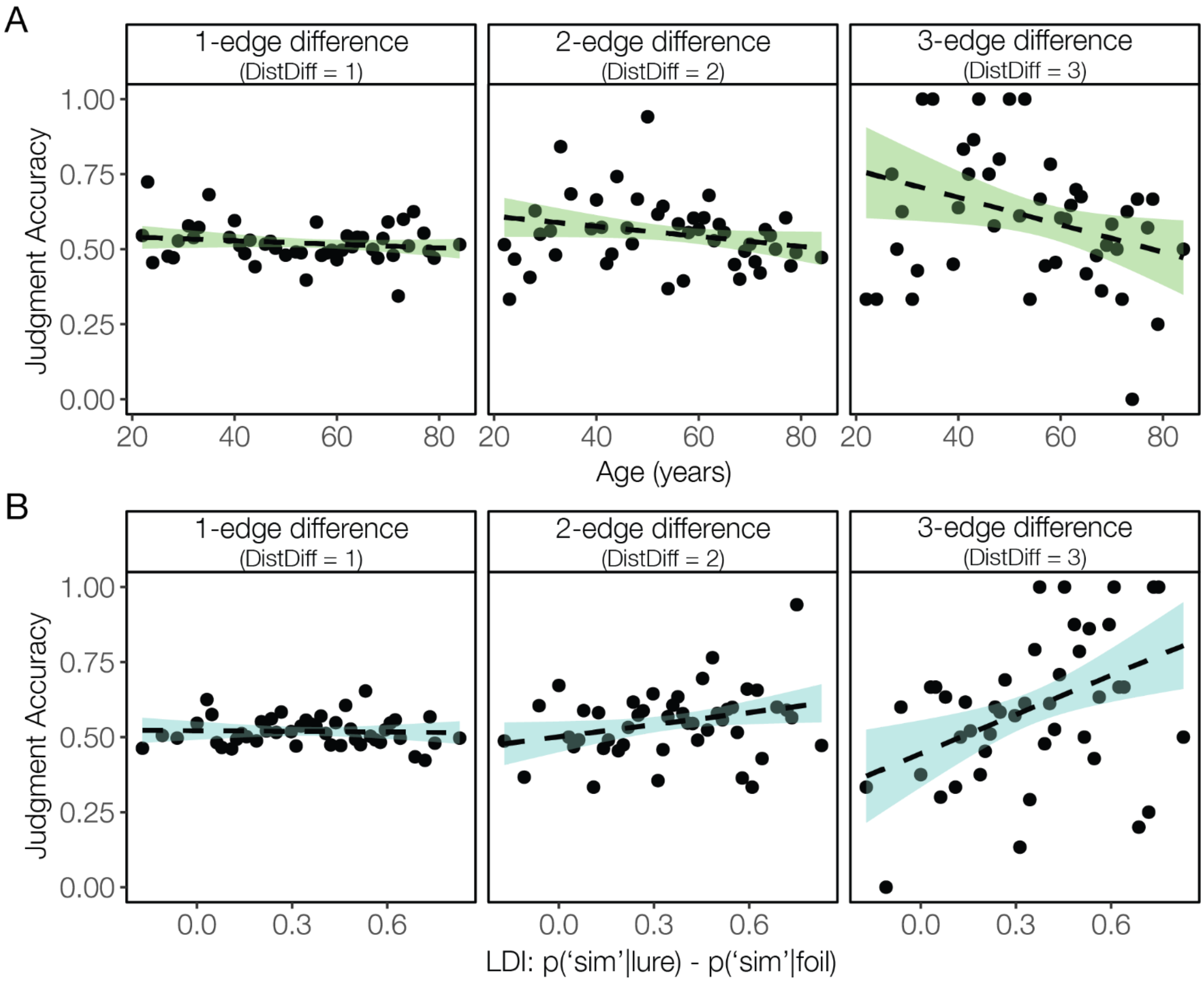


Figure 4. Behavioral results from Experiment 1. (A) Judgment accuracy as a function of chronological age and trial difficulty (distance difference between choice options). The green band indicates 95% confidence intervals around the best fit regression line. (B) Judgment accuracy as a function of mnemonic discrimination ability (LDI) and trial difficulty (distance difference between choice options). The blue band indicates 95% confidence intervals around the best fit regression line.

The findings of Experiment 1 support the idea that some individuals can perform this multi-step inference task by extracting latent structure across multiple presentations of paired object associations. Further, we show that the ability to reliably extract this structure is critically dependent on memory ability to precisely represent the individual edge pairs, with the greatest difference observed when asked to compare across larger distance differences between choice options. However, one concern was that individuals with weaker memory abilities (low LDI) did not show any evidence of latent structure learning as measured by the judgment task (performance was at chance across all levels of difficulty). We hypothesized that this could be due to the increased susceptibility to memory interference in individuals with weak mnemonic discrimination ability. One strategy to overcome such interference is to train on non-overlapping subsets of the training set one at a time. Thus, in Experiment 2, we introduced this training sequence (blocked training) which could help improve judgment performance for participants with lower LDIs by reducing the potential for memory interference during the learning phase by spacing out overlapping information across time.

Experiment 2: Memory-based inference can be improved via individualized training

To examine whether different training conditions may be beneficial for improving memory-based inference as a function of memory ability, we conducted a linear mixed effects regression predicting accuracy on the judgment test (Fig. 5). We combined data from 113 participants in Experiment 1 (intermixed learning sequence) with the 106 participants collected in Experiment 2 (blocked learning sequence) to examine whether training schedule interacted with mnemonic discrimination ability (LDI) to affect structural inference-based judgments. We specified trial difficulty (distDiff), mnemonic discrimination ability (LDI), and training sequence (blocked vs. intermixed study) as fixed factors and assessed whether these variables were significant predictors of judgment test accuracy. The full model was specified as: judgment accuracy~LDI+distDiff+Sequence+LDI*distDiff+Sequence*LDI+Sequence*distDiff+LDI*Sequence*distDiff, random=~1|subject (adjusted R^2=0.21, AIC=-304.55). This analysis revealed a 3-way interaction between mnemonic discrimination ability, difficulty, and learning sequence ($\beta_{LDI*distDiff*sequence}$=0.054, SE=0.013, t(422)=4.10, p < 0.001). Importantly, there was an interaction between learning sequence and mnemonic discrimination ability ($\beta_{LDI*sequence}$=0.038, SE=0.016, t(214)=2.26, p=0.025) such that structural inference-based judgments were optimized by different training conditions (blocked vs. intermixed learning sequences) depending on memory ability (high vs. low mnemonic discrimination ability). Specifically, individuals with low memory ability benefitted from blocked training, whereas those with high memory ability benefitted from intermixed training. This crossover interaction was apparent only in the easiest trials (distDiff=3), as blocked training was not sufficient to overcome memory deficits to improve judgment performance above chance for more difficult structural inference-based judgments (distDiff < 3). Nonetheless, these results demonstrate that reducing memory interference during the learning of latent structures (via blocked training) can improve structural inference for those with weaker memory abilities such that they are able to make more accurate inference judgments.

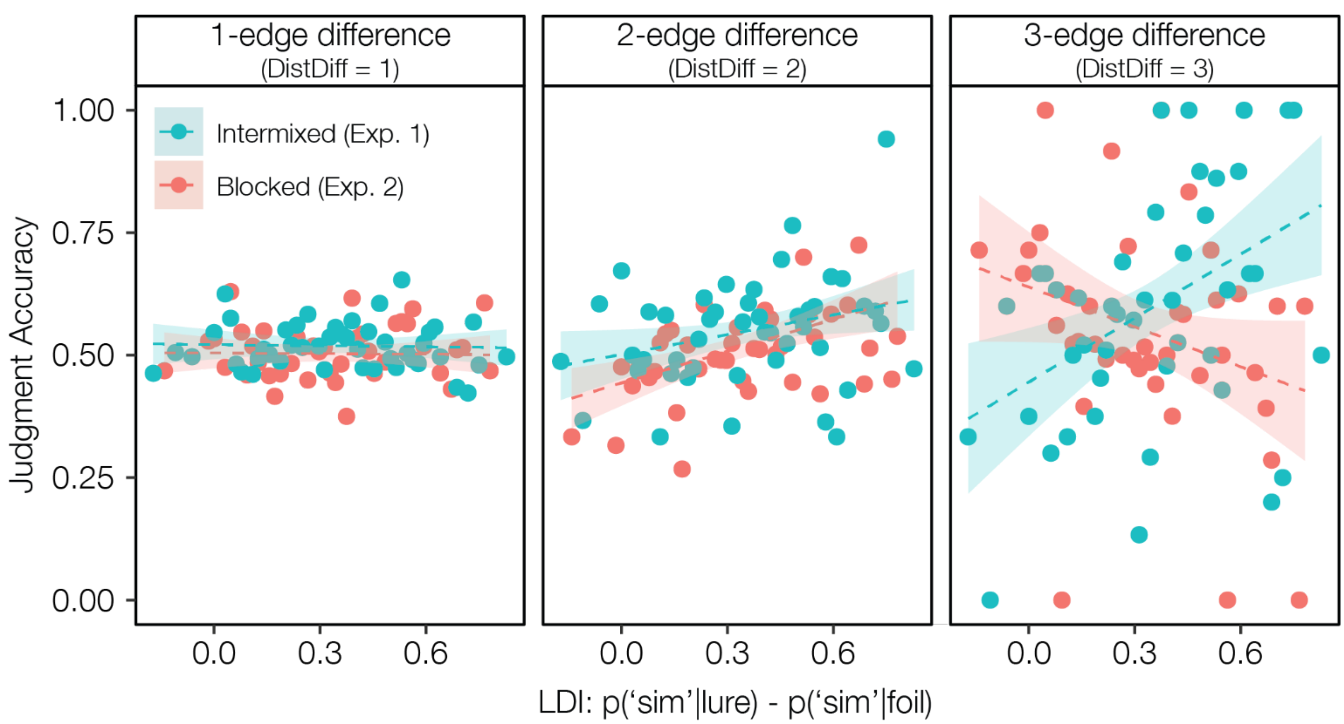


Figure 5. Behavioral results combining data from Experiment 1 (blue) and Experiment 2 (red) showing judgment accuracy as a function of trial difficulty (distance difference between choice options), mnemonic discrimination ability (LDI), and

training condition (blocked vs. intermixed learning sequence). Shaded bands indicate 95% confidence intervals around the best fit regression line for each training condition (dashed lines).

Artificial neural network captures relationship between training sequence and memory capacity in distance-dependent judgment performance.

Across our experiments, we observed an interaction between training sequence and memory encoding ability (measured using LDI on the Mnemonic Similarity Task) in participants' behavior on the judgment task. But is LDI really capturing pattern separation ability? To test this idea, we built a simple ANN to see if we could reproduce our behavioral results solely by manipulating the network's "memory capacity."

The results of Experiment 1 and Experiment 2 suggest that performance in the graph task depends on one's ability to form precise representations (i.e., of each graph node) that simultaneously reflect the latent, lower-dimensional structure in the input (i.e., the graph's edges). To investigate this hypothesis more directly, we implemented a five-layer feedforward neural network (Fig. 6), trained to associate paired items, and examined how performance of the network was modulated by representation precision. First, instances of the network were trained using the same method as humans who performed the intermixed version of the task, using 704 randomized presentations of the edge pairs (see *Methods*, subsection *Artificial Neural Network Model* for details). Then, we had the network perform the judgment task, by measuring the distance between the network's internal representations of the three presented items. Consistent with the hypothesis that such a network expresses a minimal mechanism for performing multi-step inference, we found that the pattern of distance-dependent judgment performance matched that of Rmus et al (2022) (aged 18-27, mean=22; Fig. 4a in that paper) and the young/high-LDI participants in Experiment 1 (Fig. 4). To test the hypothesis that inter-individual variability in performance in this task depends on the precision of internal representations (as indexed by the LDI), we further modified the network to have variable capacity in the second and fourth hidden layers (Fig. 6a). Memory precision can depend on the density of neurons in memory encoding regions, as denser neural populations can allow for more precise memory representations by providing the ability to store information in distinct, non-overlapping neural populations. Loosely inspired by pattern separation in the dentate gyrus [17], the large hidden layers in our network model corresponds to more distinct representations of each node (high LDI), while a smaller hidden layers imply greater chance of "merging" representations before the latent structure can be extracted by the third hidden layer (low LDI). Consistent with this hypothesis, we found that judgment task performance tracked the size of the second and fourth layers, with the greatest distinction observed on distDiff=3 trials, which compared large distances (those spanning multiple edges) to smaller ones (which required precise, orthogonalized representations of individual directly-observed edge pairs; Fig. 6b). We conducted a linear mixed effects regression predicting accuracy on the judgment test as a function of difficulty (distDiff: 1/2/3), training sequence (Sequence: blocked/intermixed) and layer width (Width: 6/9/12/15/18) of the model (Fig. 6). The full model was specified as: judgment accuracy~Width+distDiff+Sequence+Width*distDiff+Sequence*Width+Sequence*distDiff+Width*Sequence*distDiff, random=~1|model_ID (adjusted R^2=0.40, AIC=1560.7). This analysis revealed a 3-way interaction between layer width, difficulty, and training sequence ($\beta_{width*distDiff*sequence}$=3.662, SE=1.20, t(136)=3.06, p=0.003). Consistent with our behavioral results, there was an interaction between training sequence and the models' layer width ($\beta_{width*sequence}$=4.78, SE=1.85, t(66)=2.58, p=0.012) such that structural inference-based judgments were optimized by different training conditions (blocked vs. intermixed learning sequences) depending on the neural network model's capacity (larger vs. smaller hidden layer width).

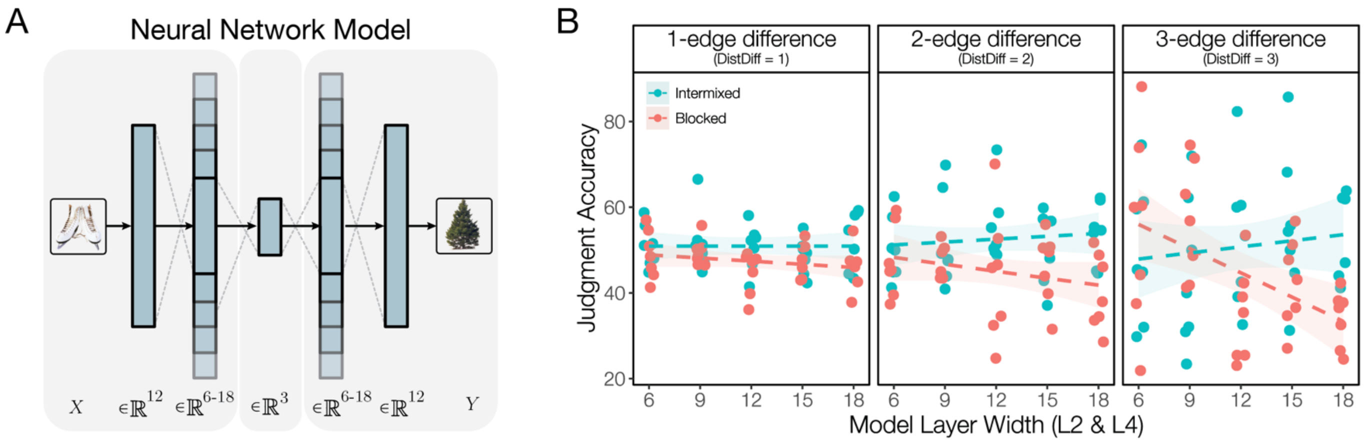


Figure 6. (A) Schematic of the artificial neural network. To model memory capacities' effect on judgment accuracy, the width of the first and third hidden layer were varied from 6 units (low-LDI; reflecting the hypothesized diminished capacity for separating inputs) to 18 units (high-LDI; allowing for sparse, highly separated representations of each input object). (B) Model judgment task results across varying layer widths and training regimens, matching the pattern observed in human participants with varying memory precision.

Graph Reconstruction. To get a sense of the memory representations that participants may have formed during learning, we analyzed data from the graph reconstruction phase. Using a linear regression approach, we separately analyzed how the number of hits, false alarms, and total edges drawn differed as a function of mnemonic discrimination ability (LDI) and training sequence (blocked vs. intermixed). First, we tested whether mnemonic discrimination ability (measured by LDI) and training sequence were significant predictors of the total number of correct edges (hits) drawn in participants' graph reconstructions (Fig. 7A). This overall regression (hits~LDI+Sequence) was statistically significant (adjusted R^2=0.03, F(2, 212)=4.09, p=0.018). There was a main effect of memory encoding ability (β_{LDI}=0.65, SE=0.23, t(212)=2.81, p=0.005) such that those with better memory ability had a greater number of hits on their graph. There was no effect of training sequence on the number of hits on participants' graphs. We also examined whether memory ability and training sequence predicted the total number of incorrect edges (false alarms) drawn in participants' graph reconstructions with the following model (Fig. 7B). There was no significant interaction between LDI and training sequence, so we assessed the following model for main effects: false_alarms~LDI+Sequence (adjusted R^2=0.03, F(2, 212)=4.07, p=0.018). We observed a main effect of training sequence ($\beta_{sequence}$=1.53, SE=0.55, t(212)=2.77, p=0.006) such that participants in the intermixed condition drew more false edges on their graphs. There was no main effect of memory ability on the number of false edges drawn during graph reconstruction. Finally, we examined the effects of memory ability and training sequence on the total number of edges participants drew on their graph (Fig. 7C). Again, there was no significant interaction between LDI and training sequence, so the following model was examined for main effects: total_edges~LDI+Sequence (adjusted R^2=0.05, F(2, 212)=6.38, p=0.002). This analysis revealed 2 main effects: there was a main effect of memory ability (β_{LDI}=0.79, SE=0.33, t(212)=2.36, p=0.019) such that those with higher LDIs were more likely to draw more edges on their graph. Additionally, there was a main effect of training sequence ($\beta_{sequence}$=1.71, SE=0.67, t(212)=2.55, p=0.011) such that participants in the intermixed condition drew more edges on their graph.

We also tried to assess graph reconstruction accuracy using two different accuracy metrics. The first metric ("accuracy") used the proportion of correct edges and subtracted the proportion of false alarms for each graph. We used a linear regression to examine how LDI and training sequence influenced graph accuracy (Fig. 7D). The interaction between LDI and training sequence was not significant, so the model was specified as accuracy~LDI+Sequence (adjusted R^2=0.02, $F(2, 212)$=3.12, p=0.046). This analysis revealed a main effect of memory ability such that those with higher LDIs showed better graph reconstruction accuracy (β_{LDI}=0.04, SE=0.01, $t(212)$=2.49, p=0.014). We also computed a relative accuracy measure scaled to each participants' graph (Fig 7E). The relative accuracy measure takes into account the total number of edges drawn for each graph and scales accuracy according to the expected number of hits and false alarms given the total edges drawn by each participant. We added this measure as a way to assess relative graph accuracy across conditions by taking into account the fact that those with poor memory abilities are likely to draw fewer edges in their graph overall. The interaction between LDI and training sequence was not significant, so the model was specified as relative_accuracy~LDI+Sequence (adjusted R^2=0.02, $F(2, 212)$=3.03, p=0.050). This analysis revealed two trending effects: a trending effect of memory ability such that those with higher LDIs showed better graph reconstruction accuracy (β_{LDI}=0.04, SE=0.02, $t(212)$=1.72, p=0.088), and a trending effect of training sequence such that those in the blocked condition showed better graph reconstruction accuracy ($\beta_{sequence}$=-0.09, SE=0.05, $t(212)$=-1.86, p=0.065).

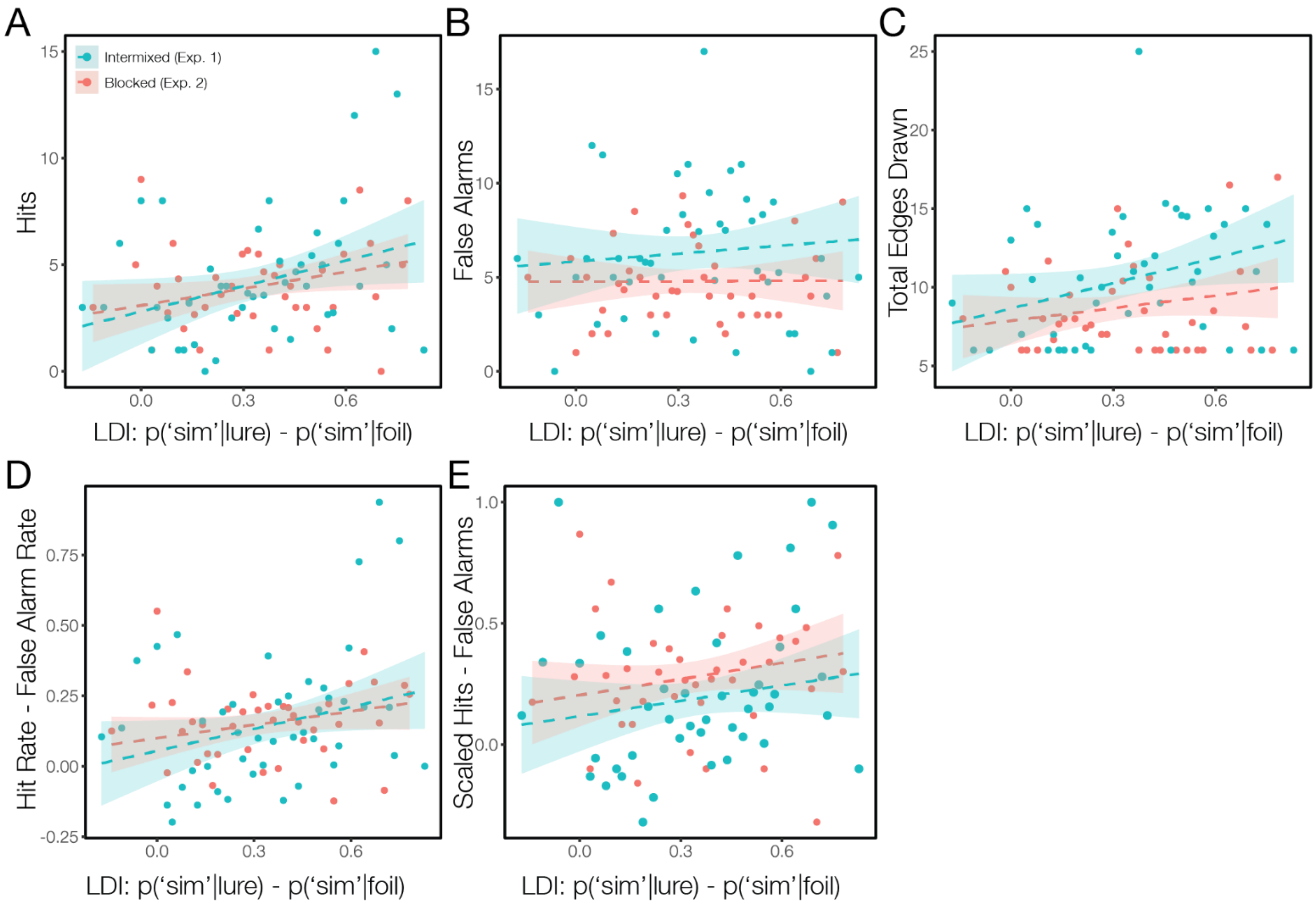


Figure 7. Behavioral results combining data from Experiment 1 (blue) and Experiment 2 (red) for the graph reconstruction phase. Shaded bands indicate 95% confidence intervals around the best fit regression line for each training condition (dashed lines).

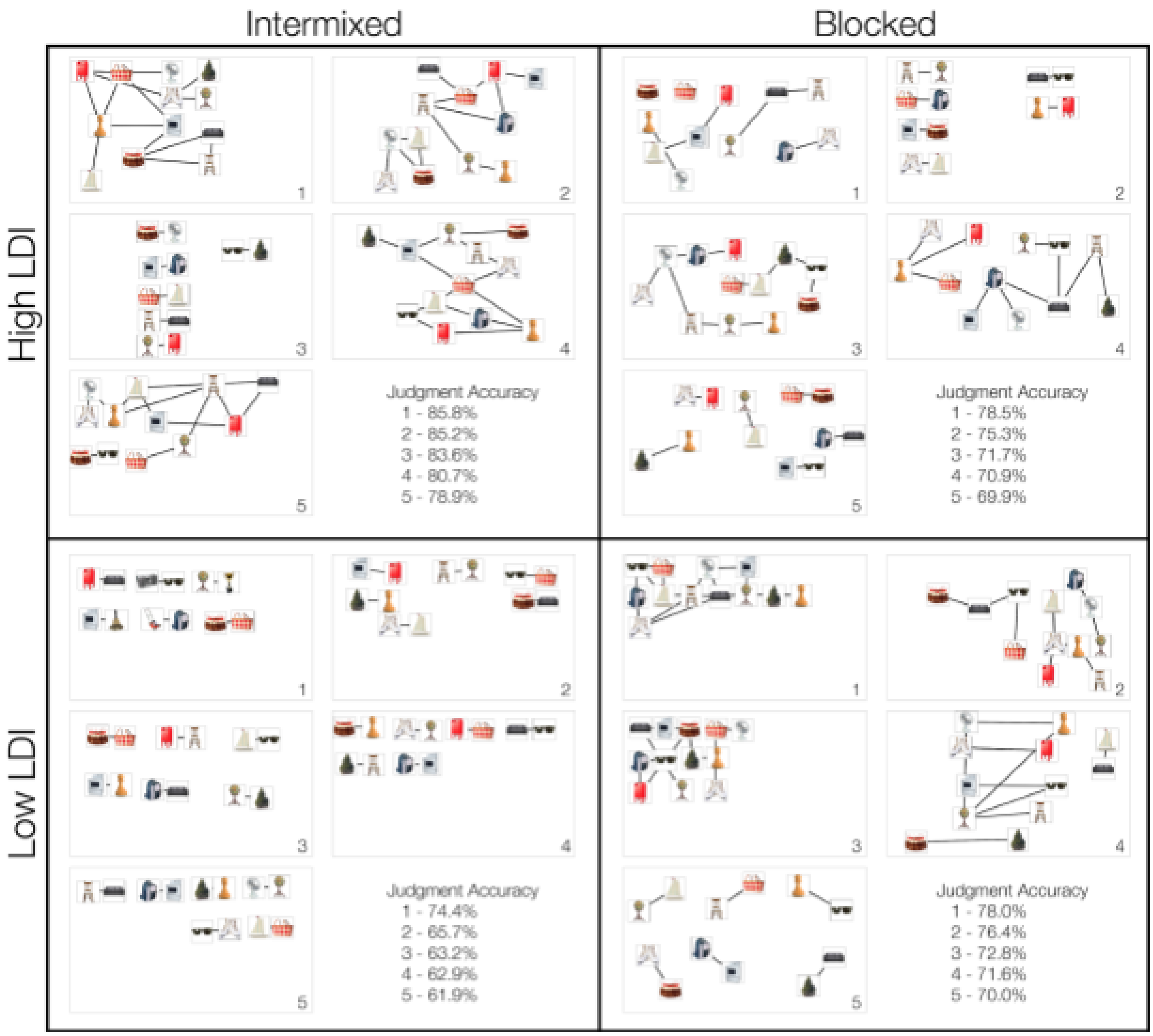
Discussion

In the present study, we examined whether multi-step inference judgments that depend on latent structure learning are affected by age-related cognitive decline. We found evidence in support of this hypothesis (Experiment 1), and further that these deficits were attributable to individual differences in mnemonic discrimination ability. We captured these effects in a computational model and discovered that the model performed categorically differently when presented with the same training samples in either structured or randomized order. In a follow-up experiment in an additional set of human subjects, we found that structural inference can be improved with individualized training interventions based on individual differences in memory ability: intermixing overlapping associations during learning improves judgment performance for those with high memory precision, but spacing out overlapping associations into separate learning blocks improves judgment performance for those with low memory precision. Taken together, our results suggest that the way in which individuals encode and organize incoming information may interact with their memory abilities to form representations that are more or less effective for making accurate inferences. Thus, configuring learning to form memory representations in a way that is congruent with one's cognitive abilities may be critical for improving judgment performance.

In Experiment 1, evidence of latent structure learning was only seen in those with high memory (specifically mnemonic discrimination) ability. While those with better memory ability (high LDIs) were able to perform structural inference-based judgments with reasonable levels of accuracy (especially as trials got easier), those with low memory ability showed no evidence of latent structure learning, as judgment performance was at chance for these individuals. This was further supported by the fact that those with better memory ability scores demonstrated better graph reconstruction across several metrics (number of correct hits, graph reconstruction accuracy, and relative graph reconstruction accuracy). However, those with better memory ability also tended to have higher false alarms and drew more edges in their graphs overall. There are two possible explanations for this pattern: the first is that the act of making inference judgments during the judgment phase (which preceded the reconstruction phase) created false memories, as participants misremember inferred information as having been directly observed[35]. However, this explanation does not fully explain our results, since all participants (including those with low memory ability) completed the graph reconstruction at the end, and any increase in false memories due to order effects should be equally likely in all conditions. Another possible explanation is that those with high memory ability are more likely to form distributed representations during the learning phase[27,31]. According to parallel distributed processing models such as C-HORSE, distributed representations better support inference and flexible reasoning (relative to localist representations), but have a tendency to produce more false alarms[27]. Our data are indeed in line with this theory: those with higher LDIs showed better structural inference judgments, but they also had higher false alarms.

In Experiment 2, we introduced a blocked training sequence during latent structure learning in hopes that separating overlapping edges across time would reduce memory interference and improve structural inference for those with low memory abilities. Indeed, we found that blocked training improved structural inference in those with low mnemonic discrimination ability, however, we found that blocked training impaired judgment performance for those with high mnemonic discrimination ability. Thus, our results suggest that the benefits of blocked or intermixed training sequences depend on individual differences in memory ability. To better understand the types of representations that drive successful inference judgments as a function of memory ability and training condition, we visually examined participant's data from the graph reconstruction phase (Supplemental Fig. 1). Specifically, we were curious to see if there were visual differences in memory representations for those who demonstrate high overall judgment accuracy (computed by averaging across the mean judgment accuracy for distance difference 1, 2, and 3 bins) across our conditions. To do this, we did a median split by LDI and examined the graphs constructed by individuals with the 5 best judgment accuracy scores within each condition (high-intermixed, high-blocked, low intermixed, low-blocked). By observing the graph reconstructions produced by those with the best judgment accuracy within each condition, we can infer the types of representations within each condition that appear to be driving successful judgment performance. This observation revealed several insights: for those with high memory encoding abilities (Supp. Fig. 1, High LDI), intermixed training led to more distributed representations–participants who had high judgment accuracy generally tended to create highly interconnected graphs (Supp. Fig. 1, High LDI; Intermixed). In contrast, blocked training seemed to create more localist representations for those with high memory ability, as participants with the highest judgment accuracy within this group appeared to form graphs that were more “broken up” relative to the intermixed group (Supp. Fig. 1, High LDI; Blocked).

Interestingly, this pattern was reversed in those with low memory encoding abilities (Supp. Fig. 1, Low LDI). In the intermixed condition, those who made the best inference judgments were relying on memory representations consisting of just a few individual object pairs (Supp. Fig. 1, Low LDI; Intermixed). While individuals with poor memory abilities produced highly localized graphs in general, it appears that blocked training seemed to encourage participants to form more interconnections in their graphs relative to those in the intermixed group (Supp. Fig. 1, Low LDI; Blocked). While these are just visual inspections of the data, it seems reasonable that those with high memory encoding abilities prefer to form more distributed memory representations, whereas those with low memory abilities are unable to form distributed memory representations and instead are more likely to form and depend on localist representations to support learning. If memory encoding ability biases the type of memory representations that individuals prefer to form, it is possible that the training sequences interact with memory encoding abilities such that training conditions that support the type of representations that individuals prefer to form will lead to improved performance. Thus learning can be improved when training manipulations are congruent with individual learning biases.
This framework may also help explain seemingly contradictory findings in the associative memory literature that has found that intermixed sequences appear to encourage the formation of distributed representations according to some studies[27] but more distinct, localist representations in others[26] and vice versa (blocking led to more distributed representations in Schlichting et al., whereas blocking led to more localist representations in Zhou et al). Specifically, the blocked condition in Zhou et al. did not reduce the potential for memory interference, as blocked trials were mixed in with intermixed trials all within the same learning phase. Thus the task used in Zhou et al. may have been better suited for those with high LDI, and low LDI participants may not have reached performance levels that were high enough to meaningfully contribute to the overall effects found in their task.



Supplementary Figure 1. Sample graph reconstructions from participants with the overall best judgment accuracy within each condition. LDI groups were created using a median split.

Recent work using neural network models has shown that distributed representations are more likely to form when information is presented in an intermixed (rather than blocked) sequence[27]. Neural network models that employ distributed representations show that blocked learning can cause memory interference as new information is encountered during learning[27,36]. This may explain our performance differences as a function of training sequence for those with high memory encoding abilities. If it is the case that individuals with high mnemonic discrimination abilities have a tendency to form and use distributed representations, we would predict that these individuals would benefit from intermixed training, as intermixing supports the formation of distributed representations[27,36]. Similarly, neural network models would predict that blocked learning could overwrite knowledge of previously encountered information and lead to catastrophic interference[27,36,44]. Our behavioral data within those with high memory encoding ability support this pattern: the high-LDI intermixed group show the best inference performance, but they also have the highest false alarms during reconstruction, consistent with a reliance on distributed representations. Similarly, the high-LDI blocked group demonstrates the most accurate memory of individual pairs, as evidenced by their low false alarm rates (Fig. 7B) and high relative accuracy on graph reconstruction (Fig. 7E). However, this increase in memory precision appears to come at the cost of impaired judgment performance (Fig. 5) relative to the high-LDI intermixed group. This impairment in inference judgments seems consistent with other findings in associative memory that show that while information presented close together in time tends to support integrative encoding[28,45,46], separating information with overlapping features across time may hurt inference performance if inference depends on bridging relevant overlapping information encountered across a larger time delay[27,28]. Visual inspection of the graph reconstruction data suggests that those who are more successful in structural inference judgments in the high-LDI blocked group are relying on graph structures that are less cohesive than the high-LDI intermixed group (Supp. Fig. 1), which further supports the idea that blocking may create disruptions in an otherwise distributed network to form more localized representations which makes inference across those disjointed representations more challenging.

For those with low memory encoding ability (such as older adults with age-related cognitive decline), the susceptibility to memory interference may have catastrophic effects on learning. If individuals are already prone to memory interference due to weaker memory abilities, an intermixed learning schedule may create learning conditions in which it is too difficult for any successful encoding to occur. Learning 16 associative pairs (edges) that consist of many overlapping elements is an extremely challenging memory problem, especially because of the high potential for memory interference due to shared elements across several different pairs. In this type of scenario, the key to successful learning for someone with weaker memory abilities may be to selectively focus on a few pairs to encode so that *some* successful learning can occur. Indeed, this is a learning strategy often employed by older adults–for instance, older adults learn to selectively attend to high valued information so they could strategically allocate cognitive resources to successfully remember highly valuable information over others[47,48]. When examining the graph reconstruction data from those that show the best judgment performance within the low-LDI group within the intermixed training condition, we see evidence of this as well: all 5 of the top performers created graph reconstructions that consisted of just 6 pairs–the bare minimum required to complete the graph reconstruction task (Supp. Fig. 1). This may suggest that successful learners in the low-LDI intermixed group are relying on highly localized representations of a small subset of studied associations to make their inference judgments. Since successful inference on our graph task is likely facilitated by a distributed representation, the challenge within the low-LDI group might be to find a way to overcome their memory deficits by reducing the crippling memory interference burden that they face during learning[49–51]. Thus, our blocked manipulation within the low-LDI group may have improved learning within this population by alleviating the most challenging obstacle to successful learning, which is their high susceptibility to memory interference during encoding. If it is the case that the memory interference burden is the primary concern that needs to be addressed for any learning to occur, then blocking training would improve learning by separating overlapping elements across time. By splitting up the learning phase into smaller “chunks” of information, we may be giving poor learners the chance to first strengthen their representations of a subset of the graph (ex: A-B) before encountering overlapping information that may compete and interfere with this knowledge. Once a relatively stable representation is formed through repetition and practice, introducing new overlapping information (ex: B-C) may allow for integrative encoding of the newly encountered information (-C) with existing prior knowledge through the formation of overlapping memory traces (ex: A-B-C[26]. Thus, for low-LDI individuals, the blocked training manipulation gives learners an opportunity to separate and scaffold learning in a way that could allow them to overcome their memory deficits by encouraging them to build more integrated representations. Our results provide support for this explanation: when looking at individuals with low memory encoding ability, blocked training led to improved performance in structural inference-based judgments (Fig. 5) and also led to fewer false alarms in graph reconstruction (Fig. 7B), which may be indicative of improved memory performance via reduced memory interference during learning. Additionally, when comparing the visual graph reconstructions produced by the best judgment performers in the low-LDI blocked vs. intermixed conditions, it appears that blocked training promotes the formation of more integrated representations. This is consistent with work in associative memory that has found that separating the learning of associative pairs with overlapping elements into separate phases (i.e., blocked training) encouraged integrated neural representations of overlapping pairs, whereas intermixing associative pairs with overlapping elements within the same learning phase encouraged more distinct, pattern separated representations between pairs with overlapping elements[26].

Collectively, our results show that successful structural inference is best achieved by those with strong memory encoding abilities, but that structural inference and related inference-based judgments can be improved with individualized training interventions. Our results therefore suggest that age-related failures in model-based planning and decision-making may be attributable to memory failures related to forming adequate latent structures necessary to support multi-step planning and inference. Future research should more directly examine how disruptions in latent structure learning might cause specific impairments in multi-step planning, and subsequently model-based decision making. The present study was somewhat limited in that it used a general measure of memory encoding ability (LDI) and did not directly measure how failure to remember individual associations within participants may have caused specific impairments in judgment accuracy. Future studies could include more direct and specific tests of memory for individual associations after latent structure learning, which would allow us to more precisely measure how memory affects judgment accuracy for trials in which the intermediate associations were more or less correctly recalled. Future studies could also examine the formation of memory representations underlying structural inference more directly, perhaps using neural data. Given that our findings suggest that successful judgment performance could be achieved through seemingly different memory representations, it would also be important to test the flexibility and limitations of different types of memory representations in how well they can support different types of multi-step inference and goal-directed planning. Despite these limitations, we believe the present study provides valuable insight into how training conditions interact with individual differences in memory encoding ability to form different memory representations that can be used to make multi-step inference decisions. These results offer the promise of developing individualized learning interventions that can help mitigate the effects of age-related cognitive decline on decision-making across the lifespan.

References

1. Akhtar, A. Corporate America is seeing a spike in the age of CEOs being hired - and yes, they're overwhelmingly white men. *Business Insider* https://www.businessinsider.in/strategy/news/corporate-america-is-seeing-a-spike-in-the-age-of-ceos-being-hired-and-yes-theyre-overwhelmingly-white-men/articleshow/71856494.cms (2019).
2. Schnoor, A. The Problem of the Aging CEO. *Medium: The Intelligence of Everything* https://medium.com/the-intelligence-of-everything/the-problem-of-the-aging-ceo-e09e8d21bb65 (2020).
3. Hall, M. & Hickey, W. This one chart shows an unprecedented recent spike in age for Congress, posing a huge problem for American democracy. *Insider* businessinsider.com/congress-oldest-history-gerontocracy-lawmakers-2022-9 (2022).
4. Ingraham, C. Yes, Congress is getting older – but so are the rest of us. *The Washington Post* https://www.washingtonpost.com/news/wonk/wp/2014/04/01/yes-congress-is-getting-older-but-so-are-the-rest-of-us/ (2014).
5. Craik, F. I. M. & Salthouse, T. A. *The Handbook of Aging and Cognition: Third Edition*. (Psychology Press, 2011).
6. Hess, T. M., Strough, J. & Löckenhoff, C. *Aging and Decision Making: Empirical and Applied Perspectives*. (Academic Press, 2015).
7. Raz, N. Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. *The handbook of aging and cognition., 2nd ed.* **2**, 1–90 (2000).
8. Drummond, N. & Niv, Y. Model-based decision making and model-free learning. *Curr. Biol.* **30**, R860–R865 (2020).
9. Eppinger, B., Walter, M., Heekeren, H. R. & Li, S.-C. Of goals and habits: age-related and individual differences in goal-directed decision-making. *Front. Neurosci.* **7**, 253 (2013).
10. Eppinger, B. & Bruckner, R. Chapter 4 - Towards a Mechanistic Understanding of Age-Related Changes in Learning and Decision Making: A Neuro-Computational Approach. in *Aging and Decision Making* (eds. Hess, T. M., Strough, J. & Löckenhoff, C. E.) 61–77 (Academic Press, 2015).
11. Konovalov, A. & Krajbich, I. Mouse tracking reveals structure knowledge in the absence of model-based choice. *Nat. Commun.* **11**, 1893 (2020).
12. Silva, C. F. da, da Silva, C. F. & Hare, T. A. Humans primarily use model-based inference in the two-stage task. *Nature Human Behaviour* vol. 4 1053–1066 Preprint at https://doi.org/10.1038/s41562-020-0905-y (2020).
13. Rmus, M., Ritz, H., Hunter, L. E., Bornstein, A. M. & Shenhav, A. Humans can navigate complex graph structures acquired during latent learning. *Cognition* **225**, 105103 (2022).
14. Bornstein, A. M. & Daw, N. D. Cortical and hippocampal correlates of deliberation during model-based decisions for rewards in humans. *PLoS Comput. Biol.* **9**, e1003387 (2013).
15. Doll, B. B., Shohamy, D. & Daw, N. D. Multiple memory systems as substrates for multiple decision systems. *Neurobiol. Learn. Mem.* **117**, 4–13 (2015).
16. Vikbladh, O. M. *et al.* Hippocampal Contributions to Model-Based Planning and Spatial Memory. *Neuron* **102**, 683–693.e4 (2019).
17. Yassa, M. A. & Stark, C. E. L. Pattern separation in the hippocampus. *Trends Neurosci.* **34**, 515–525 (2011).
18. Bakker, A., Kirwan, C. B., Miller, M. & Stark, C. E. L. Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science* **319**, 1640–1642 (2008).
19. Poppenk, J., Evensmoen, H. R., Moscovitch, M. & Nadel, L. Long-axis specialization of the human hippocampus. *Trends Cogn. Sci.* **17**, 230–240 (2013).
20. Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T. & Stark, C. E. L. Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learn. Mem.* **18**, 15–18 (2011).
21. Toner, C. K., Pirogovsky, E., Kirwan, C. B. & Gilbert, P. E. Visual object pattern separation deficits in nondemented older adults. *Learn. Mem.* **16**, 338–342 (2009).
22. Burke, S. N., Wallace, J. L., Nematollahi, S., Uprety, A. R. & Barnes, C. A. Pattern separation deficits may contribute to age-associated recognition impairments. *Behav. Neurosci.* **124**, 559–573 (2010).

23. Stark, S. M., Yassa, M. A., Lacy, J. W. & Stark, C. E. L. A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia* **51**, 2442–2449 (2013).
24. Wilson, I. A., Gallagher, M., Eichenbaum, H. & Tanila, H. Neurocognitive aging: prior memories hinder new hippocampal encoding. *Trends Neurosci.* **29**, 662–670 (2006).
25. Campbell, K. L., Hasher, L. & Thomas, R. C. Hyper-binding: a unique age effect. *Psychol. Sci.* **21**, 399–405 (2010).
26. Schlichting, M. L., Mumford, J. A. & Preston, A. R. Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nat. Commun.* **6**, 8151 (2015).
27. Zhou, Z., Singh, D., Tandoc, M. C. & Schapiro, A. C. Distributed representations for human inference. *bioRxiv* 2021.07.29.454337 (2021) doi:10.1101/2021.07.29.454337.
28. Zeithamova, D. & Preston, A. R. Temporal Proximity Promotes Integration of Overlapping Events. *J. Cogn. Neurosci.* **29**, 1311–1323 (2017).
29. Eichenbaum, H. The hippocampus and declarative memory: cognitive mechanisms and neural codes. *Behav. Brain Res.* **127**, 199–207 (2001).
30. Eichenbaum, H. Memory: Organization and Control. *Annu. Rev. Psychol.* **68**, 19–45 (2017).
31. Kumaran, D. & McClelland, J. L. Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychol. Rev.* **119**, 573–616 (2012).
32. Zeithamova, D., Schlichting, M. L. & Preston, A. R. The hippocampus and inferential reasoning: building memories to navigate future decisions. *Front. Hum. Neurosci.* **6**, 70 (2012).
33. Schlichting, M. L. & Preston, A. R. Memory integration: neural mechanisms and implications for behavior. *Curr Opin Behav Sci* **1**, 1–8 (2015).
34. Morton, N. W., Sherrill, K. R. & Preston, A. R. Memory integration constructs maps of space, time, and concepts. *Curr Opin Behav Sci* **17**, 161–168 (2017).
35. Bowman, C. R., de Araujo Sanchez, M.-A., Hou, W., Rubin, S. & Zeithamova, D. Generalization and False Memory in an Acquired Equivalence Paradigm: The Influence of Physical Resemblance Across Related Episodes. *Front. Psychol.* **12**, 669481 (2021).
36. McCloskey, M. & Cohen, N. J. Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem. in *Psychology of Learning and Motivation* (ed. Bower, G. H.) vol. 24 109–165 (Academic Press, 1989).
37. Huang, Z., Liu, F. & Su, H. Mapping state space using landmarks for universal goal reaching. *arXiv [cs.LG]* (2019).
38. Stark, S. M., Kirwan, C. B. & Stark, C. E. L. Mnemonic Similarity Task: A Tool for Assessing Hippocampal Integrity. *Trends Cogn. Sci.* **23**, 938–951 (2019).
39. Ratcliff, R. & Hendrickson, A. T. Do data from mechanical Turk subjects replicate accuracy, response time, and diffusion modeling results? *Behav. Res. Methods* **53**, 2302–2325 (2021).
40. Pinheiro, J. *et al.* Package ‘nlme’. *Linear and nonlinear mixed effects models, version* **3**, (2017).
41. Grace-Martin, K. Testing and dropping interaction terms in regression and ANOVA models. *The analysis factor* https://www.theanalysisfactor.com/testing-and-dropping-interaction-terms/ (2011).
42. Paszke, A. *et al.* PyTorch: An imperative style, high-performance deep learning library. *arXiv [cs.LG]* (2019).
43. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. *arXiv [cs.LG]* (2014).
44. McClelland, J. L. & Rumelhart, D. E. Distributed memory and the representation of general and specific information. in *Connectionist psychology: A text with readings* 75–106 (Psychology Press, 2020).
45. Pudhiyidath, A. *et al.* Representations of temporal community structure in hippocampus and precuneus predict inductive reasoning decisions. *Journal of Cognitive Neuroscience* **34**, 1736–1760 (2022).
46. Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B. & Botvinick, M. M. Neural representations of events arise from temporal community structure. *Nat. Neurosci.* **16**, 486–492 (2013).
47. Castel, A. D. The Adaptive and Strategic Use of Memory By Older Adults: Evaluative Processing and Value-Directed Remembering. in *Psychology of Learning and Motivation* (eds. Benjamin, A. S. & Ross, B. H.) vol. 48 225–270 (Academic Press, 2007).

48. Castel, A. D., Benjamin, A. S., Craik, F. I. M. & Watkins, M. J. The effects of aging on selectivity and control in short-term recall. *Mem. Cognit.* **30**, 1078–1085 (2002).
49. Shimamura, A. P. & Jurica, P. J. Memory interference effects and aging: Findings from a test of frontal lobe function. *Neuropsychology* **8**, 408–412 (1994).
50. Pettigrew, C. & Martin, R. C. Cognitive declines in healthy aging: evidence from multiple aspects of interference resolution. *Psychol. Aging* **29**, 187–204 (2014).
51. Kirwan, C. B. & Stark, C. E. L. Overcoming interference: an fMRI investigation of pattern separation in the medial temporal lobe. *Learn. Mem.* **14**, 625–633 (2007).