

1
2
3 **Multi-step inference across the human lifespan can be improved with individualized**
4 **memory interventions**
5
6
7
8
9

10 Sharon M. Noh, PhD¹,, Keiland W. Cooper, PhD²,, Shuheng Guo, BS¹, Dale Zhou, PhD^{1,2},,
11
12 Craig E. L. Stark, PhD^{2,3},, Aaron M. Bornstein, PhD^{1,3,*},

- 13
14
15
16
17 1. Department of Cognitive Sciences, The University of California, Irvine, Irvine,
18 California, United States
19
20 2. Department of Neurobiology and Behavior, The University of California, Irvine, Irvine,
21 California, United States
22
23 3. Center for the Neurobiology of Learning and Memory, The University of California,
24 Irvine, Irvine, California, United States
25
26
27
28
29
30

31
32
33 *Address correspondence to: Aaron M. Bornstein, PhD. E-mail: aaron.bornstein@uci.edu
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57

Abstract

Objectives. Effective goal-directed decision-making relies on memory and planning—processes that are known to decline with age. We tested the hypothesis that these declines stem from a common mechanism by focusing on *mnemonic discrimination*, a measure of memory precision that shows unique vulnerability to age-related decline.

Methods. We used a latent learning task that measures the ability to learn and make judgments about associations among interconnected stimuli, assessing performance across the adult lifespan. This task allows us to measure multi-step inference judgments that reflect how individuals organize relational structure, previously shown to capture the internal model-construction processes that support model-based planning. In Experiment 1, we examined relationships between judgment performance and memory precision. In Experiment 2, we tested whether a “blocked” learning schedule designed to reduce memory interference by separating overlapping objects could improve performance for individuals with weaker memory abilities. Across the lifespan, both young and older adults showed evidence of successful latent learning and inference, but variability in judgment performance was explained by mnemonic discrimination ability. In Experiment 2, mnemonic discrimination interacted with training condition: intermixed training benefited those with high memory precision, whereas blocked training benefited those with low memory precision. We also implemented artificial neural network simulations, which reproduced these qualitative patterns.

Discussion. These findings suggest that age-related declines in complex judgments stem from declines in mnemonic discrimination and demonstrate that individualized, memory-based training interventions can improve learning and reasoning processes that support goal-directed

1
2
3 planning, offering a promising approach to preserving decision-making abilities across the
4
5 lifespan.
6

7
8 **Keywords:** Individual differences, Latent learning, Memory precision, Training intervention,
9
10 Goal-directed planning
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Downloaded from <https://academic.oup.com/psychsocgerontology/advance-article/doi/10.1093/geronb/gbag038/8524505> by Barbara Scarpine user on 18 March 2026

Introduction

Humans across the lifespan must often make decisions with lasting consequences for themselves and others. Increasingly, older adults occupy positions of power at the highest levels of government and major corporations, making their decision-making abilities particularly consequential for constituents and employees (Akhtar, 2019; Hall & Hickey, 2022; Ingraham, 2014; Schnoor, 2020). Yet aging is associated with declines in several cognitive abilities, including decision-making (Craik & Salthouse, 2011; Hess et al., 2015; Raz, 2000). In particular, older adults show impairments in complex decision-making requiring multi-step associations and judgments, such as forward planning (Drummond & Niv, 2020; Eppinger & Bruckner, 2015; Eppinger et al., 2013). However, the mechanism by which multi-step inference and planning decline with age is unclear, limiting efforts to develop effective interventions to reverse age-related cognitive decline. One challenge is that existing behavioral assays do not distinguish whether age-related declines arise from the judgment process itself or from difficulties learning the associations needed for such judgments, as in model-based planning (Feher da Silva & Hare, 2020; Konovalov & Krajbich, 2020).

Many real-world decisions require inferring structure from disparate experiences to form flexible, goal-directed plans. For instance, consider a student, Alice, preparing to present at an out-of-town conference. She books a 4 PM Friday flight and plans to call a taxi at noon to allow ample time for travel. Even if this is her first conference, she can engage in multi-step planning by assembling relevant knowledge from related experiences: after many delayed flights, she opts to arrive the night before her presentation (as opposed to the morning of); she anticipates heavier Friday afternoon traffic; and she chooses a taxi over driving due to past struggles finding airport parking.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

Recent work in younger adults suggests that such goal-directed planning depends on inferring structure from the environment. Rmus et al., (2022) developed a task that measures how participants learn the associative structure among image pairs (edges) arranged along a latent “graph”. Participants implicitly learned complex associative structures (*latent structure learning*) from randomized exposures to individual graph edges. This knowledge was assessed via shortest path judgments requiring multi-step inference. Importantly, performance on the judgment task predicted greater use of model-based planning on a goal-directed decision-making task. A major strength of the task developed by Rmus et al. (2022) is its sensitivity to both the degree of latent structure learning and the ability to measure performance across varying associative distances. One open question, however, is *how* structured learning occurs, and what type of knowledge representations support multi-step inference judgments.

29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

One reasonable assumption is that memory may influence how individuals learn latent structures from experience, as model-based planning involves hippocampal contributions and effective memory search (Bornstein & Daw, 2013; Doll et al., 2015; Vikbladh et al., 2019). In the task by Rmus et al., an accurate representation of individual associations is critical for learning the graph. More specifically, memory *precision* may be especially important for forming representations that enable efficient planning. For example, remembering Friday afternoon traffic patterns (rather than Monday afternoon or Friday evening) allows Alice to make more accurate predictions for her trip. In episodic memory, high precision can be achieved through pattern separation, in which competing information is encoded as distinct neural patterns (Bakker et al., 2008; Yassa & Stark, 2011). Pattern separation reduces or resolves memory interference, enabling discrimination of similar events (Lacy et al., 2011; Poppenk et al., 2013), but this ability declines with age (Burke et al., 2010; Stark et al., 2013; Toner et al., 2009; Yassa

1
2
3 & Stark, 2011). Declines in pattern separation may explain older adults' increased susceptibility
4 to interference and related memory failures (Campbell et al., 2010; Wilson et al., 2006). We
5
6 therefore asked whether age-related deficits in multi-step planning stem from impaired latent
7
8 structure learning due to reduced precision and greater interference.
9
10

11
12
13 If older adults' increased susceptibility to memory interference undermines their ability
14 to learn the structures necessary for multistep planning, training interventions aimed at bolstering
15 encoding may improve multi-step inference judgments needed for complex decision-making
16
17 tasks. In support of this idea, work on episodic memory has shown that manipulating learning
18
19 sequences or temporally separating overlapping associations can reduce memory interference
20
21 and bias formation of distinct neural representations to support associative inference
22
23
24 (Zeithamova & Preston, 2017; Zhou et al., 2023).
25
26
27

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

29
30
31 Research suggests that associative inference can be supported by at least two kinds of neural
32
33 representations (H. Eichenbaum, 2001; Howard Eichenbaum, 2017; Poppenk et al., 2013;
34
35 Schlichting et al., 2015; Zhou et al., 2023). In standard associative inference tasks, participants
36
37 learn A–B pairs, then B–C pairs, and are later tested on their knowledge of the indirect
38
39 association between A and C, despite never having been presented with an A–C pair. If
40
41 participants encode A–B and B–C as separate episodes during learning via pattern separation,
42
43 they may form localist (orthogonalized) representations of each pair (Kumaran & McClelland,
44
45 2012; Zhou et al., 2023) and A–C inference requires effortful retrieval and recombination at test
46
47 (e.g., “A paired with B, and B paired with C, so A is associated with C”). While localist
48
49 representations yield high precision and resist interference, they may be less efficient for making
50
51
52
53
54
55
56
57
58
59
60

1
2
3 inferences due to the additional effort required for information retrieval and recombination.
4
5
6 (Zeithamova et al., 2012; Zhou et al., 2023).
7

8 Alternatively, encountering B–C may reactivate A–B, allowing integration of new (-C)
9
10 information into an updated A–B–C representation (Morton et al., 2017; Schlichting & Preston,
11
12 2015; Zeithamova et al., 2012). These distributed representations support rapid inference but are
13
14 prone to interference and false memories (Bowman et al., 2021; McCloskey & Cohen, 1989;
15
16 Zhou et al., 2023), making them potentially challenging for individuals such as older adults who
17
18 are already vulnerable to interference. If age-related declines in memory precision cause
19
20 inference failures, then a learning method that reduces memory interference should improve
21
22 latent structure learning and judgments.
23
24

25 26 **Current Study**

27
28 In the present study, we tested whether age-related declines in memory precision affect
29
30 performance on a multi-step associative inference task (the “graph” task; (Rmus et al., 2022).
31
32 Unlike standard associative inference tasks that use independent triads to test inference (e.g.,
33
34 AB-BC → AC; (Carpenter et al., 2021; Schlichting et al., 2015; Zhou et al., 2023), our task
35
36 requires participants to integrate many overlapping associations into a coherent cognitive graph
37
38 and to estimate relative distances within that graph. These representational operations have been
39
40 shown to predict the degree to which individuals rely on model-based planning strategies (Rmus
41
42 et al., 2022; Yoo et al., 2024). In Experiment 1, participants studied randomly presented object
43
44 pairs drawn from an underlying graph of 12 nodes (objects) and 16 edges (object pairs). At test,
45
46 they judged relative distances between object pairs (considered to be a fundamental computation
47
48 for high-level planning; (Huang et al., 2019)) assessing their ability to mentally navigate
49
50 associations and make shortest-path judgments. While the task does not involve overt goal-
51
52
53
54
55
56
57
58
59
60

1
2
3 directed action or planning, the multi-step inference judgments directly probe the internal model-
4 construction processes that support goal-directed planning. Accordingly, we interpret
5 performance on our inference judgment test as indexing the fidelity of internal relational models
6 that underlie overt planning behavior. We expected performance deficits with age, specifically
7 driven by poor memory precision, so participants also completed the Mnemonic Discrimination
8 Task (MDT) to independently measure *mnemonic discrimination*, a sensitive behavioral index of
9 pattern separation ability that tracks age-related decline (Stark et al., 2019) and predicts decision-
10 making beyond chronological age (Noh et al., 2023). In Experiment 2, we tested the prediction
11 that separating overlapping edge pairs during learning (blocked training) would reduce working
12 memory load and improve multi-step inference judgment performance for older adults or those
13 with low memory precision (Schlichting et al., 2015). To confirm the relationship between
14 training condition and memory precision, we simulated task performance using variants of
15 artificial neural network models that differed only in their internal representational capacity and
16 compared the model outputs to our behavioral findings.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

35 **Method**

36
37
38 Experiments 1 and 2 were identical except for the learning-phase presentation order
39 (intermixed vs. blocked). Data were collected concurrently, with participants randomly assigned
40 to the intermixed condition (Experiment 1) or blocked condition (Experiment 2). Although the
41 two training conditions were implemented under a single overarching protocol with random
42 assignment of participants to conditions, we present the work as two experiments to reflect their
43 distinct conceptual aims. Experiment 1's task closely followed Rmus et al. (2022) but included a
44 lifespan sample with older adults, was administered online, and incorporated an additional
45 measure of memory specificity. Thus, Experiment 1 uses the original intermixed schedule (Rmus
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

et al., 2022) to replicate prior findings and to test whether age or mnemonic discrimination better explains multi-step inference performance across the adult lifespan. Experiment 2 implements a temporally separated schedule to test the hypothesis that reducing interference would preferentially benefit low-precision learners. Anticipating that older adults might find the original task difficult, we implemented the blocked sequence in Experiment 2, based on prior work in associative memory suggesting that temporally separating overlapping associations might reduce cognitive load and improve performance in populations with lower memory capacity (Schlichting et al., 2015). Note that the “blocked” manipulation used here (i.e., separating overlapping edges across time over the course of training) is consistent with similar manipulations of “blocking” used in episodic memory (Beukers et al., 2024; Schlichting et al., 2015; Zhou et al., 2023) and is somewhat distinct from how “blocking” is used in the category learning literature (Brunmair & Richter, 2019; Kang & Pashler, 2012; Noh et al., 2024, 2016). Methods are described jointly below, with procedural differences between the two experiments noted explicitly.

Participants

Sample size was based on prior work with similar designs ($N = 81$; Rmus et al., 2022), with a ~25% increase to offset higher noise in our online data (relative to the supervised in-person data collected by Rmus et al., 2022). We aimed for ~100 participants per experiment post-exclusion. A total of 219 participants (112 female, 107 male; ages 19-84, mean[*sd*] age = 55.7[14.2]) were recruited online (see Supplementary Material; *Online Data Collection*): 113 in Experiment 1 (59 female, 54 male, ages 22-84, mean[*sd*] age = 56.7[13.8]), and 106 in Experiment 2 (53 female / 53 male, ages 19-79, mean[*sd*] age = 54.8[14.7]). Participants received monetary compensation and completed a tutorial plus a rotation-detection screener (10 practice trials; $\geq 70\%$ accuracy

required within two attempts to ensure attention to task instructions) to be eligible to participate in the full task. All protocols were classified as an “Exempt Online Survey” by the Institutional Review Board of the University of California, Irvine.

Procedure

Mnemonic Discrimination Task (MDT)

Participants completed the MST (Figure 1) as an independent measure of memory ability to assess whether mnemonic discrimination explained individual differences in graph task performance.

Encoding Phase. Participants viewed object images and made indoor/outdoor judgments (cover task). No MST images overlapped with those used in the graph task, and participants were unaware that their memory would later be tested.

Mnemonic Discrimination Test

On a surprise discrimination test, participants saw object images that were identical (old), similar (lures), or novel (foils) relative to objects shown during encoding, and judged each object as “old”, “similar”, or “new”. A Lure Discrimination Index (LDI) was calculated for each participant: $p(\text{sim}|\text{lure}) - p(\text{sim}|\text{foil})$. Higher LDI values indicate better ability to classify lures as “similar” relative to foils, reflecting greater mnemonic discrimination and better memory encoding precision (Stark et al., 2019).

Structural Inference (“Graph”) Task

Study Phase. Participants viewed and were told to memorize a series of object pairs (Figure 2 Panel A), each presented for 1s in random order. Participants provided a rotation judgment on each trial to ensure attention. No information about the underlying structure was given, though pairs were drawn from a hidden 12-node, 16-edge graph (Figure 2 Panel B). In

1
2
3 Experiment 1, all 16 unique edge pairs were repeated 44 times (704 trials total) in a random
4
5 sequence (Figure 2 Panel C, “intermixed”). Experiment 2 grouped object pairs into 4 mini-
6
7 blocks, each with 4 unique, non-overlapping pairs (Figure 2 Panel C, “blocked”). This separation
8
9 of potentially confusing (overlapping) edges across time aimed to reduce memory interference
10
11 and improve learning of the graph structure.
12
13

14
15 Following the study phase, participants completed two different tests to measure how
16
17 well participants can *use* their knowledge of the graph structure (Figure 3 Panel A, *Judgment*
18
19 *Test*), as well as how well they *learned* the graph structure (Figure 3 Panel B, *Graph*
20
21 *Reconstruction Test*).
22
23

24
25 *Judgment Test.* After completing the learning phase, participants completed a relative
26
27 distance judgment task. This task measured participants’ ability to use learned information for
28
29 structural inference. Participants were asked to judge which of two objects (left vs. right) was
30
31 closer to a central object, based on indirect relationships learned previously (i.e., graph distance).
32
33 Responses were made using keyboard buttons, across 204 trials with up to 10 seconds per trial.
34
35 Each unique object served as a central node 17 times. Choice options were randomly selected
36
37 with the following constraints: 1) neither option was directly paired with the central node during
38
39 the study, and 2) the shortest path length between the central node and the two options
40
41 (associative distance) was not equal. Trial difficulty varied based on the difference in associative
42
43 distance between choice options and the reference node (ranging from 1 to 3). Accuracy was
44
45 calculated within each difficulty bin.
46
47
48

49
50 Because online data are unsupervised and noisier, judgment test data were screened for
51
52 outliers prior to analysis using a two-step reaction time (RT)–based procedure, as RTs were not
53
54 analyzed directly (see Supplementary Material; *RT-based Filters* for additional details). Filtered
55
56
57
58
59
60

1
2
3 judgment-phase data were analyzed using linear mixed-effects models (R nlme package;
4 (Pinheiro et al., 2017), with difficulty (distDiff = 1, 2, 3; 1= most difficult), age, LDI, and
5
6 sequence (blocked vs. intermixed) specified as fixed effects, and participant as a random effect.
7
8
9

10 *Graph Reconstruction Test.* Following the judgment test, participants completed a graph
11 reconstruction task to assess explicit knowledge of the learned structure. All 12 study-phase
12 objects were displayed above a blank canvas, and participants arranged them by clicking and
13 dragging each object onto the canvas. Participants were asked to place and link objects that were
14 directly paired during the study by selecting two objects sequentially; linked items were
15 connected with a straight line. All objects had to be placed on the canvas and linked to at least
16 one other object before submission.
17
18
19
20
21
22
23
24
25

26 We used a Signal Detection Theory approach to analyzing the reconstruction data
27 (Carterette, 1967). We scored reconstructions as follows: “hits” for correctly linked study pairs,
28 “false alarms” for incorrectly linked pairs, and the total number of edges drawn. Accuracy was
29 computed in two ways: 1) sensitivity (d'), defined as “hit rate - false alarm rate”, and 2)
30 precision. Across the 12 nodes, there are a total of 66 possible edges that could be drawn if every
31 object pair were connected. Of these, 16 edges are true edges in the underlying graph structure
32 (“hits”), and 50 represent non-edges (“false alarms”). Thus, the hit rate is calculated as hits/16,
33 and the false-alarm rate is calculated as false alarms/50, and d' is computed as hit rate minus
34 false-alarm rate. We also examined an additional measure of accuracy (precision), taken from
35 machine learning (Blair, 1979), to complement d' . While d' separately considers hits and false
36 alarms relative to the total possible true and false edges, precision instead evaluates the
37 proportion of drawn edges that were correct (hits / total edges drawn). Precision is useful for
38 comparing reconstruction accuracy across individuals who may differ in memory capacity. For
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Results

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

Judgment Test

We used a linear mixed-effects model to examine whether trial difficulty (relative distance difference between choice options) and chronological age predicted judgment accuracy (Figure 4): judgment accuracy \sim age + distDiff + age*distDiff, random = \sim 1|subject (adjusted $R^2 = 0.23$, AIC = -152.55). There was a main effect of age ($\beta_{\text{age}} = -0.030$, SE = 0.012, $t(110) = -2.57$, $p = 0.011$), with older adults performing worse than younger adults. There was also a main effect of difficulty ($\beta_{\text{distDiff}} = 0.032$, SE = 0.009, $t(216) = 3.50$, $p < 0.001$), with performance improving as relative distance increased. The age \times difficulty interaction was significant ($\beta_{\text{age*distDiff}} = -0.022$, SE = 0.009, $t(216) = -2.41$, $p = 0.017$) such that younger adults showed greater performance gains as trials got easier (i.e., greater relative distance between options).

Adding mnemonic discrimination ability (LDI) to the model revealed it was a stronger predictor of judgment accuracy than age: judgment accuracy \sim age + distDiff + LDI + age*distDiff + age*LDI + distDiff*LDI + age*distDiff*LDI, random = \sim 1|subject (adjusted $R^2 = 0.26$, AIC = -133.98). When both age and LDI were included, the main effect of age ($\beta_{\text{age}} = -0.021$, SE = 0.012, $t(108) = -1.72$, $p = 0.088$) and its interaction with difficulty ($\beta_{\text{age*distDiff}} = -0.012$, SE = 0.010, $t(214) = -1.24$, $p = 0.218$) were no longer significant. Instead, higher LDI predicted better performance ($\beta_{\text{LDI}} = 0.031$, SE = 0.013, $t(108) = 2.42$, $p = 0.017$), with a significant LDI \times difficulty interaction ($\beta_{\text{LDI*distDiff}} = 0.033$, SE = 0.010, $t(214) = 3.34$, $p = 0.001$) indicating that memory precision benefits emerged primarily in easier trials. Age \times LDI and the three-way interaction were non-significant. Given LDI's predictive strength, we re-ran the model

excluding age (Figure 5, intermixed condition; adjusted $R^2 = 0.26$, $AIC = -164.19$). This model confirmed significant main effects of LDI ($\beta_{LDI} = 0.035$, $SE = 0.012$, $t(110) = 2.99$, $p = 0.003$) and difficulty ($\beta_{distDiff} = 0.032$, $SE = 0.009$, $t(216) = 3.50$, $p < 0.001$) as well as a robust LDI \times difficulty interaction ($\beta_{LDI*distDiff} = 0.034$, $SE = 0.009$, $t(216) = 3.84$, $p < 0.001$), with higher LDI predicting greater gains in easier trials.

These results suggest that structural inference performance depends on the precision of encoded associations, with the largest differences observed in trials requiring comparisons across larger associative distances (i.e., easier trials). Notably, participants with low memory abilities performed at chance across all difficulty levels, suggesting minimal or no latent structure learning. We hypothesized that this may reflect greater susceptibility to memory interference, which might be mitigated by training on non-overlapping subsets one block at a time. Thus, Experiment 2 introduced the “blocked” training intervention to reduce interference for low-LDI participants during learning and potentially improve subsequent judgment performance.

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

To test whether training conditions modulate structural inference as a function of memory ability, we combined data from Experiment 1 ($n = 113$, intermixed sequence) and Experiment 2 ($n = 106$, blocked sequence) and fit a linear mixed-effects model (Figure 5): judgment accuracy $\sim LDI + distDiff + Sequence + LDI*distDiff + Sequence*LDI + Sequence*distDiff + LDI*Sequence*distDiff$, random = $\sim 1|subject$ (adjusted $R^2 = 0.21$, $AIC = -304.55$). There was a significant 3-way interaction between mnemonic discrimination ability, difficulty, and learning sequence ($\beta_{LDI*distDiff*sequence} = 0.054$, $SE = 0.013$, $t(422) = 4.10$, $p < 0.001$), indicating that the relationship between mnemonic discrimination ability (LDI) and performance differed by training sequence and trial difficulty. Importantly, there was an interaction between learning

1
2
3 sequence and mnemonic discrimination ability ($\beta_{\text{LDI}*\text{sequence}} = 0.038$, $\text{SE} = 0.016$, $t(214) = 2.26$, p
4 $= 0.025$): participants with low LDI benefitted from blocked training, whereas those with high
5 LDI performed better with intermixed training. This crossover interaction was evident only in the
6 easiest trials ($\text{distDiff} = 3$). Still, reducing memory interference via blocked training improved
7 structural inference for individuals with weaker memory ability, enabling more accurate
8 inference judgments when associative distances were large.

9
10
11
12
13
14
15
16
17 *Artificial neural network captures relationship between training sequence and memory*
18
19
20 *capacity in distance-dependent judgment performance*

21
22 Given the observed interaction between training sequence and mnemonic discrimination
23 ability (LDI) in Experiments 1 and 2, we tested whether such effects could be reproduced by
24 manipulating a neural network's representational capacity (Figure 6 Panel A). To test this, we
25 implemented a five-layer feedforward autoencoder, which jointly optimizes encoding and
26 decoding of sequential inputs (Blanco Malerba et al., 2024; Chandak et al., 2024), trained to
27 associate paired items (edges) from the graph task (704 presentations; see Supplementary
28 Material: *Artificial Neural Network Model*).

29
30
31
32
33
34
35
36
37
38 To model variability in memory precision, we manipulated the number of units in the
39 second and fourth layers (L2/L4; Figure 6 Panel A). Larger widths corresponded to more
40 distinct, nonoverlapping representations of each node (analogous to higher LDI), whereas
41 smaller widths simulated representational merging before latent structure extraction (analogous
42 to lower LDI). Consistent with our behavioral results, there was a 3-way interaction between
43 layer width, difficulty, and training sequence ($\beta_{\text{width}*\text{distDiff}*\text{sequence}} = 0.0032$, $\text{SE} = 0.0006$, $t(23996)$
44 $= 5.58$, $p < 0.001$) (Figure 6 Panel B). Additionally, there was an interaction between training
45 sequence and the models' layer width ($\beta_{\text{width}*\text{sequence}} = 0.0147$, $\text{SE} = 0.0008$, $t(11996) = 18.37$, $p <$
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

0.001) 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) in a similar pattern to those seen in the human data (blocked > intermixed for low LDI, intermixed > blocked for high LDI). These interactions show that lower-capacity models benefited more from blocked training, while higher-capacity models benefited more from intermixed training—precisely the crossover pattern observed between memory precision and sequence in our human participants.

Graph Reconstruction. To get a sense of the memory representations that participants may have formed during learning, we analyzed hits (correct edges), false alarms (incorrect edges), total edges drawn, and two accuracy metrics (sensitivity and precision) from the graph reconstruction phase as a function of mnemonic discrimination ability (LDI) and training sequence (blocked vs. intermixed). Each of these five reconstruction metrics (hits, false alarms, total edges, sensitivity, and precision) was used as the outcome measure with LDI and sequence as independent predictors (outcome ~ LDI + Sequence + LDI*Sequence). Across all five regressions, the LDI × Sequence interaction was not significant, so interaction terms were removed and models were refit to assess main effects; the results reported here are from those main-effects models. For hits, the linear regression ($hits \sim LDI + Sequence$) was significant, adjusted $R^2 = .03$, $F(2, 212) = 4.09$, $p = .018$, with a main effect of LDI, $\beta = 0.65$, $SE = 0.23$, $t(212) = 2.81$, $p = .005$, indicating that participants with better memory ability produced more correct edges; there was no effect of training sequence (Figure 7 Panel A). For false alarms, the model ($false\ alarms \sim LDI + Sequence$) was significant, adjusted $R^2 = .03$, $F(2, 212) = 4.07$, $p = .018$, with a main effect of training sequence, $\beta = 1.53$, $SE = 0.55$, $t(212) = 2.77$, $p = .006$, such that participants in the intermixed condition produced more false edges; LDI was not significant

(Figure 7 Panel B). For total edges, the model ($total\ edges \sim LDI + Sequence$) was significant, adjusted $R^2 = .05$, $F(2, 212) = 6.38$, $p = .002$, with main effects of LDI, $\beta = 0.79$, $SE = 0.33$, $t(212) = 2.36$, $p = .019$, and Sequence, $\beta = 1.71$, $SE = 0.67$, $t(212) = 2.55$, $p = .011$, indicating that higher-LDI participants and those in the intermixed condition drew more edges overall (Figure 7 Panel C). We also assessed reconstruction accuracy using two metrics: sensitivity (proportion of hits - proportion of false alarms) and precision (number of hits divided by the total number of drawn edges). For sensitivity, the model ($sensitivity \sim LDI + Sequence$) was significant, adjusted $R^2 = .02$, $F(2, 212) = 3.12$, $p = .046$, revealing a main effect of LDI, $\beta = 0.04$, $SE = 0.01$, $t(212) = 2.49$, $p = .014$, with higher-LDI participants showing better reconstruction accuracy (Figure 7 Panel D). For precision, which adjusts for the tendency of low-memory participants to draw fewer edges, the model ($precision \sim LDI + Sequence$) was marginal, adjusted $R^2 = .02$, $F(2, 212) = 3.03$, $p = .050$, with trending effects of LDI, $\beta = 0.04$, $SE = 0.02$, $t(212) = 1.72$, $p = .088$, and Sequence, $\beta = -0.09$, $SE = 0.05$, $t(212) = -1.86$, $p = .065$, suggesting higher precision for participants with better memory ability and for those trained in the blocked condition (Figure 7 Panel E).

Discussion

The present study examined whether multi-step inference judgments requiring latent structure learning are affected by age-related cognitive decline. In Experiment 1, we found that older adults performed worse on structural inference judgments, but these deficits were better explained by individual differences in mnemonic discrimination ability (LDI) than by chronological age. Experiment 2 demonstrated that structural inference can be improved through individualized training: intermixing overlapping associations benefited those with high LDI, while blocked training benefited those with low LDI. These behavioral findings were reinforced

1
2
3 by artificial neural network model simulations that varied in memory capacity to examine multi-
4
5 step inference performance under blocked and interleaved training conditions. Although the
6
7 models are a simplified system, manipulating representational capacity in these networks
8
9 successfully reproduced the sequence x memory precision interaction observed in human
10
11 participants, thereby providing converging evidence for our hypotheses.
12
13

14
15 Collectively, our findings suggest that the way individuals encode and organize
16
17 information interacts with memory ability to shape representations that vary in their effectiveness
18
19 for supporting accurate inference. Matching training to an individual's representational
20
21 tendencies may be critical for optimizing judgment performance. Although our task did not
22
23 require participants to execute planned sequences of actions, the multi-step inference judgments
24
25 it elicited engage the representational mechanisms thought to support model-based planning
26
27 (Rmus et al., 2022; Yoo et al., 2024). Specifically, the judgment test requires estimating which of
28
29 two stimuli is "closer" within a learned associative graph, effectively querying the internal
30
31 relational model used to chain together multiple steps of the structure. Prior work using similar
32
33 graph-learning tasks has demonstrated that the integrity of these relational models predicts
34
35 individuals' reliance on model-based strategies (Rmus et al., 2022). Thus, our findings offer
36
37 insight into how age- and memory-related variability in representational fidelity may constrain
38
39 the computations that underlie goal-directed planning.
40
41
42
43

44
45 In Experiment 1, evidence of latent structure learning emerged only in participants with
46
47 high mnemonic discrimination ability. High-LDI individuals achieved above-chance inference
48
49 performance, especially on easier trials, whereas low-LDI participants performed at chance. The
50
51 lack of learning observed in low-LDI participants who performed our task may explain why
52
53 certain populations such as older adults and individuals with mild cognitive impairments may
54
55
56
57
58
59
60

1
2
3 exhibit deficits in goal-directed planning (Eppinger & Bruckner, 2015; Samanez-Larkin, 2013),
4
5 as prior work shows that these populations are particularly vulnerable to deficits in memory
6
7 precision (Yassa & Stark, 2011; Yassa et al., 2010). High-LDI participants also produced more
8
9 accurate graph reconstructions across multiple metrics (number of correct edges, reconstruction
10
11 accuracy measured using sensitivity and precision), but they also generated more false alarms
12
13 and total edges. One explanation is that making inference judgments before reconstruction
14
15 induced false memories by misattributing inferred information as directly experienced (Bowman
16
17 et al., 2021). However, this order effect should have applied equally to all participants, making it
18
19 unlikely to explain the selective increase in false alarms for high-LDI individuals. A more
20
21 plausible account is that high-LDI individuals tend to form distributed representations during
22
23 learning (Kumaran & McClelland, 2012; Zhou et al., 2023). According to parallel distributed
24
25 processing models such as C-HORSE, distributed representations support flexible inference but
26
27 are prone to false alarms (Zhou et al., 2023). Our data fit this prediction: high-LDI participants
28
29 outperformed others on inference judgments but also produced more false edges.
30
31
32
33
34

35
36 In Experiment 2, we tested whether separating overlapping pairs in time (blocked
37
38 training) would reduce memory interference and improve inference for low-LDI individuals.
39
40 Indeed, blocked training improved structural inference for low-LDI participants but impaired
41
42 performance for high-LDI participants, producing a crossover interaction. Visual inspection of
43
44 graph reconstructions of the best performers in each condition (Supplementary Figure 1)
45
46 suggests why: high-LDI participants in the intermixed condition tended to create highly
47
48 interconnected, distributed graphs, whereas blocked training led them to form more fragmented,
49
50 localist graphs. In contrast, low-LDI participants generally produced localized graphs, but
51
52 blocked training encouraged more interconnections relative to intermixing. These patterns
53
54
55
56
57
58
59
60

1
2
3 suggest that memory encoding ability may bias representational style—distributed for high-LDI,
4 localist for low-LDI—and that training sequences aligning with these preferences yield better
5 performance. This hints at the idea that blocked vs. interleaved learning sequences can be
6 leveraged to improve multi-step inference and other forms of inferential reasoning as a function
7 of one’s memory capacity.
8
9

10
11 This framework may help reconcile conflicting findings in associative memory research.
12 Some studies report that intermixing promotes distributed representations (Zhou et al., 2023),
13 whereas others find it encourages more localist coding (Schlichting et al., 2015) with the reverse
14 pattern for blocked learning. A key methodological difference is that Zhou et al.’s “blocked”
15 condition still included intermixed trials within the same learning phase, likely negating
16 interference-reduction benefits. This design may have disproportionately favored high-LDI
17 participants while preventing low-LDI learners from reaching performance levels sufficient to
18 influence overall effects. Supporting this, Zhou et al. excluded many low performers, yet
19 performance remained low relative to Schlichting et al.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

36 Recent neural network modeling shows that distributed representations are more likely to
37 emerge when information is presented in an intermixed sequence (Zhou et al., 2023). Models
38 with distributed coding predict that blocked learning increases memory interference as new
39 information is acquired (McCloskey & Cohen, 1989; Zhou et al., 2023). Our behavioral data
40 support this: the high-LDI intermixed group showed the best inference performance but also the
41 highest false alarms during reconstruction—consistent with distributed coding—while the high-
42 LDI blocked group exhibited low false alarms (Figure 7 Panel B) and high graph reconstruction
43 precision (Figure 7 Panel E), suggesting more localist coding. However, this gain in precision
44 came at the cost of impaired judgment performance (Figure 5). This tradeoff parallels evidence
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 that temporally contiguous presentation facilitates integrative encoding (Pudhiyidath et al., 2022;
4 Schapiro et al., 2013; Zeithamova & Preston, 2017; Zhou et al., 2023), whereas separating
5
6 overlapping information can disrupt inference if it requires linking across longer temporal gaps
7
8 (Zeithamova & Preston, 2017; Zhou et al., 2023). Visual inspection supports this: high-LDI
9
10 blocked learners' graphs were less cohesive than those in the high-LDI intermixed group
11
12 (Supplementary Figure 1), consistent with blocking fragmenting an otherwise distributed
13
14 network into more localized subgraphs, making cross-representation inference harder.
15
16
17
18

19 For low-LDI learners—such as older adults with age-related decline—susceptibility to
20
21 memory interference can severely impair learning and memory encoding in an intermixed
22
23 schedule. With 16 overlapping pairs, interference may be so severe that learners can only focus
24
25 on memorizing a small subset of pairs—similar to older adults' strategy of selectively encoding
26
27 high-value items (Castel, 2007; Castel et al., 2002). Indeed, the top performers in the low-LDI
28
29 intermixed condition reconstructed only the minimum six edges required to complete the task,
30
31 suggesting encoding of a few localized subsets. Since our graph task likely benefits from
32
33 distributed representations, the main challenge for low-LDI learners is overcoming interference
34
35 during encoding (Kirwan & Stark, 2007; Pettigrew & Martin, 2014; Shimamura & Jurica, 1994).
36
37 Blocked training may mitigate this by spacing overlapping pairs across time, allowing learners to
38
39 consolidate subsets (e.g., A–B) before encountering related pairs (e.g., B–C). Previous studies in
40
41 episodic memory have also suggested that stabilizing A-B representations prior to introducing
42
43 overlapping B-C episodes can prevent memory interference and promote integration of the
44
45 related episodes (Noh et al., 2026; Schlichting et al., 2015). Once initial representations are
46
47 stabilized through repetition, new overlapping pairs can be integrated into existing knowledge
48
49 (A–B–C) by pattern completing the overlapping memory traces (Schlichting et al., 2015).
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Consistent with this account, low-LDI participants in the blocked condition showed better
4 inference judgments (Figure 5) and fewer false alarms (Figure 7 Panel B), indicating reduced
5 interference. This pattern was also replicated with the ANN model simulations, showing that
6 smaller-capacity models performed better under blocked (vs. intermixed) training conditions.
7
8 Visual reconstructions suggest blocked training promoted more integrated networks in low-LDI
9 learners than intermixing, consistent with prior work showing that blocking overlapping pairs
10 fosters integrated representations, while intermixing promotes pattern-separated representations
11 (Schlichting et al., 2015).
12
13
14
15
16
17
18
19
20
21

22 Collectively, these results indicate that structural inference depends on both memory
23 encoding ability and the alignment of training structure with representational tendencies. For
24 high-LDI learners, intermixing supports distributed networks that facilitate inference; for low-
25 LDI learners, blocking reduces interference and fosters integration. These findings suggest that
26 some age-related deficits in model-based planning and decision-making may stem from failures
27 to form adequate latent structures. Because multi-step inference over a cognitive graph provides
28 a substrate for evaluating relational distances and future action paths, these findings have direct
29 implications for models of goal-directed decision making. The present task therefore contributes
30 to understanding how aging- and memory-related representational constraints shape the cognitive
31 computations underlying planning, even in the absence of overt behavioral choice. Future work
32 should test how disruptions in latent structure learning impair multi-step planning (Harhen &
33 Bornstein, 2023; Yoo et al., 2024), directly measure memory for individual associations, and use
34 neuroimaging to track the formation of representations during learning. Given that successful
35 inference can emerge from different representational formats, future studies should also examine
36 their flexibility and limitations across different inference and planning contexts. Despite
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 limitations, our results offer a mechanistic basis for individualized learning interventions to
4
5 mitigate cognitive decline effects on decision-making across the lifespan.
6
7

8 **Funding**

9
10 This work was supported by the National Institute of Aging at the National Institutes of Health
11
12 (R21AG072673 and R01AG088306 to AMB, R01AG066683 and P30AG066519 to CELS, and
13
14 F32AG072836 to SMN); and the National Science Foundation (NSF DGE-1839285 to KWC).
15
16

17 **Conflict of Interest**

18
19 The authors declare no competing financial interests.
20
21

22 **Data Availability**

23
24 All de-identified data and the final versions of all R scripts used to generate each data figure, as
25
26 well as Pytorch model code, are freely available at the UCI CCNL GitHub repository
27
28 (<https://github.com/uciccnl>). This study was not preregistered.
29
30

31 **Author Contributions**

32
33 SMN and KWC served as co-first authors for this study.
34

35 **SMN:** conceptualization, methodology, software, validation, formal analysis, investigation, data
36
37 curation, writing – original draft, writing – review & editing, visualization, supervision, project
38
39 administration, funding acquisition
40
41

42 **KWC:** conceptualization, methodology, software, validation, formal analysis, investigation, data
43
44 curation, writing – review & editing, visualization, supervision
45
46

47 **SG:** methodology, software, validation, formal analysis, data curation, writing – review &
48
49 editing, visualization
50
51

52 **DZ:** software, supervision
53
54
55
56
57
58
59
60

1
2
3
4 **CELS:** conceptualization, methodology, resources, writing – review & editing, supervision,
5
6 funding acquisition

7
8 **AMB:** conceptualization, methodology, validation, resources, writing – review & editing,
9
10 supervision, project administration, funding acquisition

11 12 13 **References**

- 14
15 Akhtar, A. (2019, November 1). *Corporate America is seeing a spike in the age of CEOs being*
16
17 *hired - and yes, they're overwhelmingly white men*. Business Insider.
18
19 [https://www.businessinsider.in/strategy/news/corporate-america-is-seeing-a-spike-in-the-](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)
20
21 [age-of-ceos-being-hired-and-yes-theyre-overwhelmingly-white-](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)
22
23 [men/articleshow/71856494.cms](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)
24
25
- 26 Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. L. (2008). Pattern separation in the human
27
28 hippocampal CA3 and dentate gyrus. *Science*, *319*(5870), 1640–1642.
29
30 <https://doi.org/10.1126/science.1152882>
31
32
- 33 Beukers, A. O., Collin, S. H. P., Kempner, R. P., Franklin, N. T., Gershman, S. J., & Norman, K.
34
35 A. (2024). Blocked training facilitates learning of multiple schemas. *Communications*
36
37 *Psychology*, *2*(1), 28. <https://doi.org/10.1038/s44271-024-00079-4>
38
39
- 40 Blair, D. C. (1979). Information Retrieval, 2nd ed. C.J. Van Rijsbergen. London: Butterworths;
41
42 1979: 208 pp. Price: \$32.50. *Journal of the American Society for Information Science*.
43
44 *American Society for Information Science*, *30*(6), 374–375.
45
46 <https://doi.org/10.1002/asi.4630300621>
47
48
- 49 Blanco Malerba, S., Micheli, A., Woodford, M., & Azeredo da Silveira, R. (2024). Jointly
50
51 efficient encoding and decoding in neural populations. *PLoS Computational Biology*,
52
53 *20*(7), e1012240. <https://doi.org/10.1371/journal.pcbi.1012240>
54
55
56
57
58
59
60

- 1
2
3 Bornstein, A. M., & Daw, N. D. (2013). Cortical and hippocampal correlates of deliberation
4 during model-based decisions for rewards in humans. *PLoS Computational Biology*,
5 9(12), e1003387. <https://doi.org/10.1371/journal.pcbi.1003387>
6
7
8
9
10 Bowman, C. R., de Araujo Sanchez, M.-A., Hou, W., Rubin, S., & Zeithamova, D. (2021).
11 Generalization and False Memory in an Acquired Equivalence Paradigm: The Influence
12 of Physical Resemblance Across Related Episodes. *Frontiers in Psychology*, 12, 669481.
13 <https://doi.org/10.3389/fpsyg.2021.669481>
14
15
16
17
18
19 Brunmair, M., & Richter, T. (2019). Similarity matters: A meta-analysis of interleaved learning
20 and its moderators. *Psychological Bulletin*, 145(11), 1029–1052.
21 <https://doi.org/10.1037/bul0000209>
22
23
24
25
26
27 Burke, S. N., Wallace, J. L., Nematollahi, S., Uprety, A. R., & Barnes, C. A. (2010). Pattern
28 separation deficits may contribute to age-associated recognition impairments. *Behavioral*
29 *Neuroscience*, 124(5), 559–573. <https://doi.org/10.1037/a0020893>
30
31
32
33
34 Campbell, K. L., Hasher, L., & Thomas, R. C. (2010). Hyper-binding: a unique age effect.
35 *Psychological Science*, 21(3), 399–405. <https://doi.org/10.1177/0956797609359910>
36
37
38
39 Carpenter, A. C., Thakral, P. P., Preston, A. R., & Schacter, D. L. (2021). Reinstatement of item-
40 specific contextual details during retrieval supports recombination-related false
41 memories. *NeuroImage*, 236(118033), 118033.
42 <https://doi.org/10.1016/j.neuroimage.2021.118033>
43
44
45
46
47 Carterette, E. C. (1967). Toward a more rigorous psychophysics: Signal detection theory and
48 psychophysics . David M. green and John A. swets. Wiley, New York, 1966. 467 pp.,
49 illus. \$12.95. *Science*, 156(3775), 632–633. <https://doi.org/10.1126/science.156.3775.632>
50
51
52
53
54 Castel, A. D. (2007). The Adaptive and Strategic Use of Memory By Older Adults: Evaluative
55
56
57
58
59
60

- 1
2
3 Processing and Value-Directed Remembering. In A. S. Benjamin & B. H. Ross (Eds.),
4 *Psychology of Learning and Motivation* (Vol. 48, pp. 225–270). Academic Press.
5
6 [https://doi.org/10.1016/S0079-7421\(07\)48006-9](https://doi.org/10.1016/S0079-7421(07)48006-9)
7
8
9
10 Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on
11 selectivity and control in short-term recall. *Memory & Cognition*, *30*(7), 1078–1085.
12
13 <https://doi.org/10.3758/bf03194325>
14
15
16
17 Chandak, S., Shah, P., Borkar, V. S., & Dodhia, P. (2024). Reinforcement learning in non-
18 Markovian environments. *Systems & Control Letters*, *185*(105751), 105751.
19
20 <https://doi.org/10.1016/j.sysconle.2024.105751>
21
22
23
24 Craik, F. I. M., & Salthouse, T. A. (2011). *The Handbook of Aging and Cognition: Third Edition*.
25 Psychology Press. <https://play.google.com/store/books/details?id=YeJ4AgAAQBAJ>
26
27
28 Doll, B. B., Shohamy, D., & Daw, N. D. (2015). Multiple memory systems as substrates for
29 multiple decision systems. *Neurobiology of Learning and Memory*, *117*, 4–13.
30
31 <https://doi.org/10.1016/j.nlm.2014.04.014>
32
33
34
35
36 Drummond, N., & Niv, Y. (2020). Model-based decision making and model-free learning.
37 *Current Biology: CB*, *30*(15), R860–R865. <https://doi.org/10.1016/j.cub.2020.06.051>
38
39
40 Eichenbaum, H. (2001). The hippocampus and declarative memory: cognitive mechanisms and
41 neural codes. *Behavioural Brain Research*, *127*(1–2), 199–207.
42
43 [https://doi.org/10.1016/s0166-4328\(01\)00365-5](https://doi.org/10.1016/s0166-4328(01)00365-5)
44
45
46
47 Eichenbaum, Howard. (2017). Memory: Organization and Control. *Annual Review of*
48 *Psychology*, *68*, 19–45. <https://doi.org/10.1146/annurev-psych-010416-044131>
49
50
51 Eppinger, B., & Bruckner, R. (2015). Chapter 4 - Towards a Mechanistic Understanding of Age-
52 Related Changes in Learning and Decision Making: A Neuro-Computational Approach.
53
54
55
56
57
58
59
60

- In T. M. Hess, J. Strough, & C. E. Löckenhoff (Eds.), *Aging and Decision Making* (pp. 61–77). Academic Press. <https://doi.org/10.1016/B978-0-12-417148-0.00004-2>
- Eppinger, B., Walter, M., Heekeren, H. R., & Li, S.-C. (2013). Of goals and habits: age-related and individual differences in goal-directed decision-making. *Frontiers in Neuroscience*, 7, 253. <https://doi.org/10.3389/fnins.2013.00253>
- Feher da Silva, C., & Hare, T. A. (2020). Humans primarily use model-based inference in the two-stage task. *Nature Human Behaviour*, 4(10), 1053–1066. <https://doi.org/10.1038/s41562-020-0905-y>
- Grace-Martin, K. (2011, April 2). *Testing and dropping interaction terms in regression and ANOVA models*. The Analysis Factor. <https://www.theanalysisfactor.com/testing-and-dropping-interaction-terms/>
- Hall, M., & Hickey, W. (2022, September 14). *This one chart shows an unprecedented recent spike in age for Congress, posing a huge problem for American democracy*. Business Insider. <https://www.businessinsider.com/congress-oldest-history-gerontocracy-lawmakers-2022-9>
- Harhen, N. C., & Bornstein, A. M. (2023). Overharvesting in human patch foraging reflects rational structure learning and adaptive planning. *Proceedings of the National Academy of Sciences of the United States of America*, 120(13), e2216524120. <https://doi.org/10.1073/pnas.2216524120>
- Hess, T. M., Strough, J., & Löckenhoff, C. (2015). *Aging and Decision Making: Empirical and Applied Perspectives*. Academic Press.
- Huang, Z., Liu, F., & Su, H. (2019). Mapping state space using landmarks for universal goal reaching. arXiv. <https://doi.org/10.48550/arXiv.1908.05451>

- 1
2
3
4 Ingraham, C. (2014, April 1). *Yes, Congress is getting older – but so are the rest of us*. The
5
6 Washington Post. [https://www.washingtonpost.com/news/wonk/wp/2014/04/01/yes-
9
10 congress-is-getting-older-but-so-are-the-rest-of-us/](https://www.washingtonpost.com/news/wonk/wp/2014/04/01/yes-
7
8 congress-is-getting-older-but-so-are-the-rest-of-us/)
- 11 Kang, S. H. K., & Pashler, H. (2012). Learning Painting Styles: Spacing is Advantageous when
12
13 it Promotes Discriminative Contrast. *Applied Cognitive Psychology*, *26*(1), 97–103.
14
15 <https://doi.org/10.1002/acp.1801>
- 16
17 Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. arXiv.
18
19 <https://doi.org/10.48550/arXiv.1412.6980>
- 20
21 Kirwan, C. B., & Stark, C. E. L. (2007). Overcoming interference: an fMRI investigation of
22
23 pattern separation in the medial temporal lobe. *Learning & Memory*, *14*(9), 625–633.
24
25 <https://doi.org/10.1101/lm.663507>
- 26
27
28 Konovalov, A., & Krajbich, I. (2020). Mouse tracking reveals structure knowledge in the
29
30 absence of model-based choice. *Nature Communications*, *11*(1), 1893.
31
32 <https://doi.org/10.1038/s41467-020-15696-w>
- 33
34
35 Kumaran, D., & McClelland, J. L. (2012). Generalization through the recurrent interaction of
36
37 episodic memories: a model of the hippocampal system. *Psychological Review*, *119*(3),
38
39 573–616. <https://doi.org/10.1037/a0028681>
- 40
41
42 Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T., & Stark, C. E. L. (2011). Distinct
43
44 pattern separation related transfer functions in human CA3/dentate and CA1 revealed
45
46 using high-resolution fMRI and variable mnemonic similarity. *Learning & Memory*,
47
48 *18*(1), 15–18. <https://doi.org/10.1101/lm.1971111>
- 49
50
51 McCloskey, M., & Cohen, N. J. (1989). Catastrophic Interference in Connectionist Networks:
52
53 The Sequential Learning Problem. In G. H. Bower (Ed.), *Psychology of Learning and*
54
55
56
57
58
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Motivation* (Vol. 24, pp. 109–165). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60536-8](https://doi.org/10.1016/S0079-7421(08)60536-8)
- Morton, N. W., Sherrill, K. R., & Preston, A. R. (2017). Memory integration constructs maps of space, time, and concepts. *Current Opinion in Behavioral Sciences*, *17*, 161–168. <https://doi.org/10.1016/j.cobeha.2017.08.007>
- Noh, S. M., Bjork, R. A., & Preston, A. R. (2024). General knowledge and detailed memory benefit from different training sequences. *Journal of Applied Research in Memory and Cognition*, *13*(3), 329–341. <https://doi.org/10.1037/mac0000193>
- Noh, S. M., Singla, U. K., Bennett, I. J., & Bornstein, A. M. (2023). Memory precision and age differentially predict the use of decision-making strategies across the lifespan. *Scientific Reports*, *13*(1), 17014. <https://doi.org/10.1038/s41598-023-44107-5>
- Noh, S. M., Yan, V. X., Bjork, R. A., & Maddox, W. T. (2016). Optimal sequencing during category learning: Testing a dual-learning systems perspective. *Cognition*, *155*, 23–29. <https://doi.org/10.1016/j.cognition.2016.06.007>
- Noh, S. M., Zhou, D., Cooper, K. W., Guo, S., Dinh, E. T., & Bornstein, A. M. (2026). Sparsity and memory constraints interact with training sequence to bias learning of associative maps. *Neuropsychologia*, *109396*, 109396. <https://doi.org/10.1016/j.neuropsychologia.2026.109396>
- Ogletree, A. M., & Katz, B. (2021). How Do Older Adults Recruited Using MTurk Differ From Those in a National Probability Sample? *International Journal of Aging & Human Development*, *93*(2), 700–721. <https://doi.org/10.1177/0091415020940197>
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M.,

- 1
2
3
4 Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., ... Chintala, S. (2019). PyTorch: An
5
6 imperative style, high-performance deep learning library. arXiv.
7
8 <https://doi.org/10.48550/arXiv.1912.01703>
9
- 10 Pettigrew, C., & Martin, R. C. (2014). Cognitive declines in healthy aging: evidence from
11
12 multiple aspects of interference resolution. *Psychology and Aging, 29*(2), 187–204.
13
14 <https://doi.org/10.1037/a0036085>
15
- 16
17 Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., & Maintainer,
18
19 R. (2017). Package ‘nlme.’ *Linear and Nonlinear Mixed Effects Models, Version, 3*(1).
20
21 <http://cran.rapporter.net/web/packages/nlme/nlme.pdf>
22
23
- 24 Poppenk, J., Evensmoen, H. R., Moscovitch, M., & Nadel, L. (2013). Long-axis specialization of
25
26 the human hippocampus. *Trends in Cognitive Sciences, 17*(5), 230–240.
27
28 <https://doi.org/10.1016/j.tics.2013.03.005>
29
30
- 31 Pudhiyidath, A., Morton, N. W., Duran, R. V., Schapiro, A. C., Momennejad, I., Hinojosa-
32
33 Rowland, D. M., Molitor, R. J., & Preston, A. R. (2022). Representations of temporal
34
35 community structure in hippocampus and precuneus predict inductive reasoning
36
37 decisions. *Journal of Cognitive Neuroscience, 34*(10), 1736–1760.
38
39 https://doi.org/10.1162/jocn_a_01864
40
41
- 42 Ratcliff, R., & Hendrickson, A. T. (2021). Do data from mechanical Turk subjects replicate
43
44 accuracy, response time, and diffusion modeling results? *Behavior Research Methods,*
45
46 *53*(6), 2302–2325. <https://doi.org/10.3758/s13428-021-01573-x>
47
48
- 49 Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of
50
51 structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *The*
52
53 *handbook of aging and cognition* (2nd ed., pp. 1–90). Lawrence Erlbaum Associates
54
55
56
57
58
59
60

Publishers.

- Rmus, M., Ritz, H., Hunter, L. E., Bornstein, A. M., & Shenhav, A. (2022). Humans can navigate complex graph structures acquired during latent learning. *Cognition*, *225*, 105103. <https://doi.org/10.1016/j.cognition.2022.105103>
- Samanez-Larkin, G. R. (2013). Financial Decision Making and the Aging Brain. *APS Observer*, *26*(5), 30–33. <https://www.ncbi.nlm.nih.gov/pubmed/23946614>
- Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nature Neuroscience*, *16*(4), 486–492. <https://doi.org/10.1038/nn.3331>
- Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, *6*, 8151. <https://doi.org/10.1038/ncomms9151>
- Schlichting, M. L., & Preston, A. R. (2015). Memory integration: neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, *1*, 1–8. <https://doi.org/10.1016/j.cobeha.2014.07.005>
- Schnoor, A. (2020, February 6). *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>
- Shimamura, A. P., & Jurica, P. J. (1994). Memory interference effects and aging: Findings from a test of frontal lobe function. *Neuropsychology*, *8*(3), 408–412. <https://doi.org/10.1037/0894-4105.8.3.408>
- Stark, S. M., Kirwan, C. B., & Stark, C. E. L. (2019). Mnemonic Similarity Task: A Tool for Assessing Hippocampal Integrity. *Trends in Cognitive Sciences*, *23*(11), 938–951.

- 1
2
3
4 <https://doi.org/10.1016/j.tics.2019.08.003>
5
6 Stark, S. M., Yassa, M. A., Lacy, J. W., & Stark, C. E. L. (2013). A task to assess behavioral
7
8 pattern separation (BPS) in humans: Data from healthy aging and mild cognitive
9
10 impairment. *Neuropsychologia*, *51*(12), 2442–2449.
11
12 <https://doi.org/10.1016/j.neuropsychologia.2012.12.014>
13
14
15 Toner, C. K., Pirogovsky, E., Kirwan, C. B., & Gilbert, P. E. (2009). Visual object pattern
16
17 separation deficits in nondemented older adults. *Learning & Memory*, *16*(5), 338–342.
18
19 <https://doi.org/10.1101/lm.1315109>
20
21
22 Turner, A. M., Engelsma, T., Taylor, J. O., Sharma, R. K., & Demiris, G. (2020). Recruiting
23
24 older adult participants through crowdsourcing platforms: Mechanical Turk versus
25
26 Prolific Academic. *AMIA Annual Symposium Proceedings, 2020*, 1230–1238.
27
28 <https://www.ncbi.nlm.nih.gov/pubmed/33936499>
29
30
31 Vikbladh, O. M., Meager, M. R., King, J., Blackmon, K., Devinsky, O., Shohamy, D., Burgess,
32
33 N., & Daw, N. D. (2019). Hippocampal Contributions to Model-Based Planning and
34
35 Spatial Memory. *Neuron*, *102*(3), 683–693.e4.
36
37 <https://doi.org/10.1016/j.neuron.2019.02.014>
38
39
40 Wilson, I. A., Gallagher, M., Eichenbaum, H., & Tanila, H. (2006). Neurocognitive aging: prior
41
42 memories hinder new hippocampal encoding. *Trends in Neurosciences*, *29*(12), 662–670.
43
44 <https://doi.org/10.1016/j.tins.2006.10.002>
45
46
47 Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in*
48
49 *Neurosciences*, *34*(10), 515–525. <https://doi.org/10.1016/j.tins.2011.06.006>
50
51
52 Yassa, M. A., Stark, S. M., Bakker, A., Albert, M. S., Gallagher, M., & Stark, C. E. L. (2010).
53
54 High-resolution structural and functional MRI of hippocampal CA3 and dentate gyrus in
55
56
57
58
59
60

1
2
3 patients with amnesic Mild Cognitive Impairment. *NeuroImage*, 51(3), 1242–1252.

4
5
6 <https://doi.org/10.1016/j.neuroimage.2010.03.040>

7
8 Yoo, J., Chrastil, E. R., & Bornstein, A. M. (2024). Cognitive graphs: Representational

9
10 substrates for planning. *Decision*, 11(4), 537–556. <https://doi.org/10.1037/dec0000249>

11
12 Zeithamova, D., & Preston, A. R. (2017). Temporal Proximity Promotes Integration of

13
14 Overlapping Events. *Journal of Cognitive Neuroscience*, 29(8), 1311–1323.

15
16
17 https://doi.org/10.1162/jocn_a_01116

18
19 Zeithamova, D., Schlichting, M. L., & Preston, A. R. (2012). The hippocampus and inferential

20
21 reasoning: building memories to navigate future decisions. *Frontiers in Human*

22
23
24 *Neuroscience*, 6, 70. <https://doi.org/10.3389/fnhum.2012.00070>

25
26 Zhou, Z., Singh, D., Tandoc, M. C., & Schapiro, A. C. (2023). Building integrated

27
28 representations through interleaved learning. *Journal of Experimental Psychology.*

29
30
31 *General*, 152(9), 2666–2684. <https://doi.org/10.1037/xge0001415>

Figures

Figure 1. 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).

Alt Text: 3-panel schematic of the Mnemonic Discrimination Task (MDT): incidental indoor/outdoor object encoding phase; old/new/similar discrimination phase. Scatterplot showing LDI versus age across both experiments, with regression lines and confidence bands.

Figure 2. 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).

Alt Text: Diagram of latent graph learning: individual edges sampled from an overlapping network. Schematic contrasts intermixed random presentation in Experiment 1 with mini-blocks separating overlapping edges in Experiment 2.

Figure 3. 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 3 (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 objects on the canvas, and each object had to be connected to at least one other object to complete this phase.

Alt Text: 2-panels: Judgment phase - select which option is closer to center; difficulty varies by distance. Reconstruction phase - place 12 objects on canvas with links scored as hits and false alarms.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

Figure 4. Behavioral results from Experiment 1.

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.

Alt Text: Experiment 1 plot of judgment accuracy by trial difficulty and age. Points and regression line show lower accuracy on harder trials, with 95% confidence interval bands around the fit.

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 5. Behavioral results combining data from Experiment 1 (blue) and Experiment 2 (red).

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).

Alt Text: Combined plot of judgment accuracy by trial difficulty, LDI, and training sequence. Fitted lines for intermixed and blocked conditions, colored by experiment, show how schedule relates to accuracy.

Figure 6. Neural network design and simulation results.

(A) Schematic of two example artificial neural networks (the smallest and largest capacity models). To model memory capacities' effect on judgment accuracy, the width of the second and fourth layers was varied from 6 units (low-LDI; reflecting the hypothesized diminished capacity for separating inputs) to 36 units (high-LDI; allowing for sparse, highly separated representations of each input object). (B) Model judgment task results across varying layer widths and training conditions, matching the pattern observed in human participants with varying memory precision. Shaded bands indicate 95% confidence intervals around the best fit regression line for each training condition (dashed lines), and error bars around individual points reflect 95% confidence intervals around each mean.

Alt Text: Panel A shows neural networks with varying hidden-layer widths as memory capacity. Panel B plots model judgment accuracy across widths and training sequences, with 95% confidence interval bands around means.

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).

Alt Text: Graph reconstruction results across experiments: performance by LDI and training sequence. Regression lines for intermixed and blocked conditions, colored by experiment, with 95% confidence interval bands around each fit.

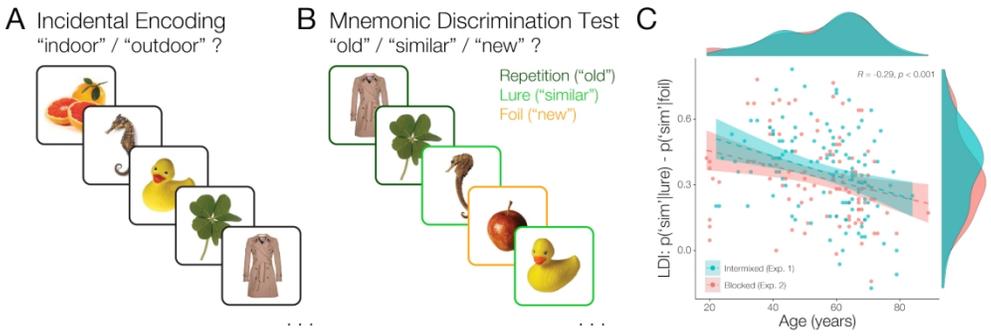


Figure 1

177x59mm (300 x 300 DPI)

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article/doi/10.1093/geronb/gbag038/8524505 by Barbara Scarpine user on 18 March 2026

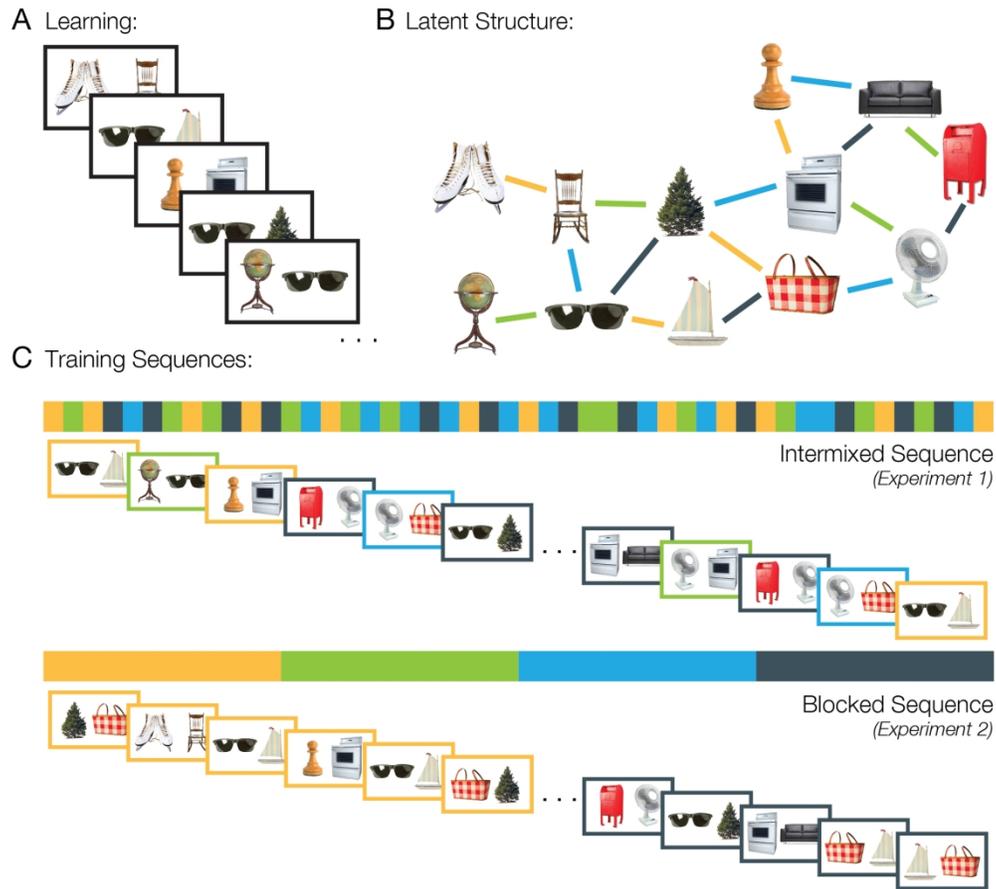


Figure 2

180x160mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

A Judgment Phase:



B Graph Reconstruction:

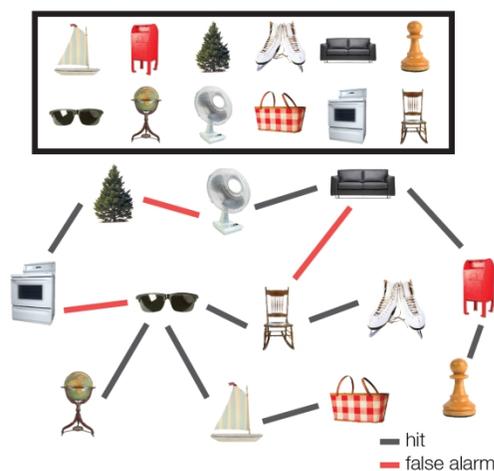


Figure 3

179x92mm (300 x 300 DPI)

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article/doi/10.1093/geronh/gbag038/8524505 by Barbara Scarpine user on 18 March 2026

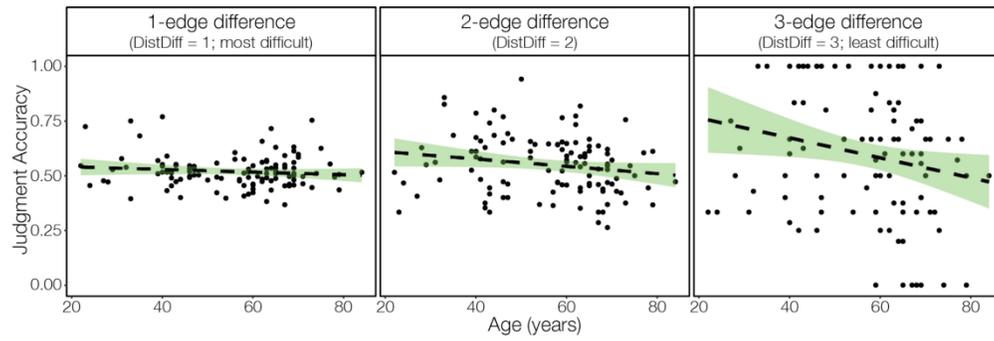


Figure 4

177x58mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

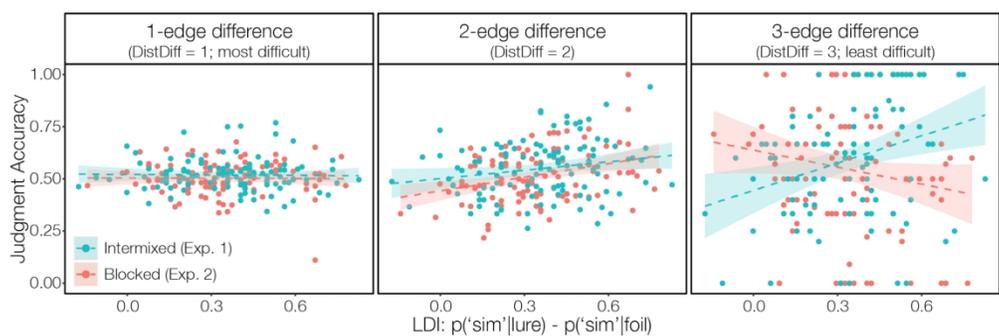


Figure 5

177x57mm (300 x 300 DPI)

Downloaded from <https://academic.oup.com/psychsocgerontology/advance-article/doi/10.1093/geronb/gbag038/8524505> by Barbara Scarpine user on 18 March 2026

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

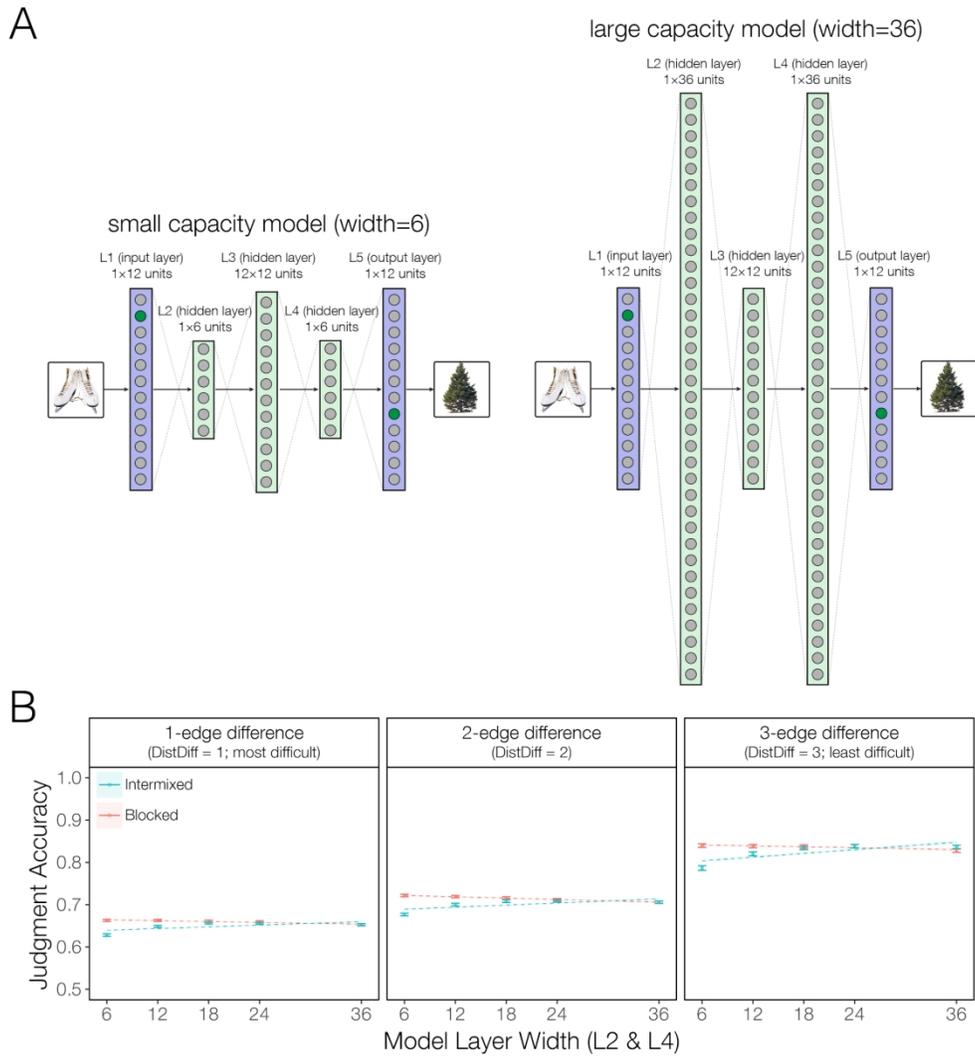


Figure 6

177x188mm (300 x 300 DPI)

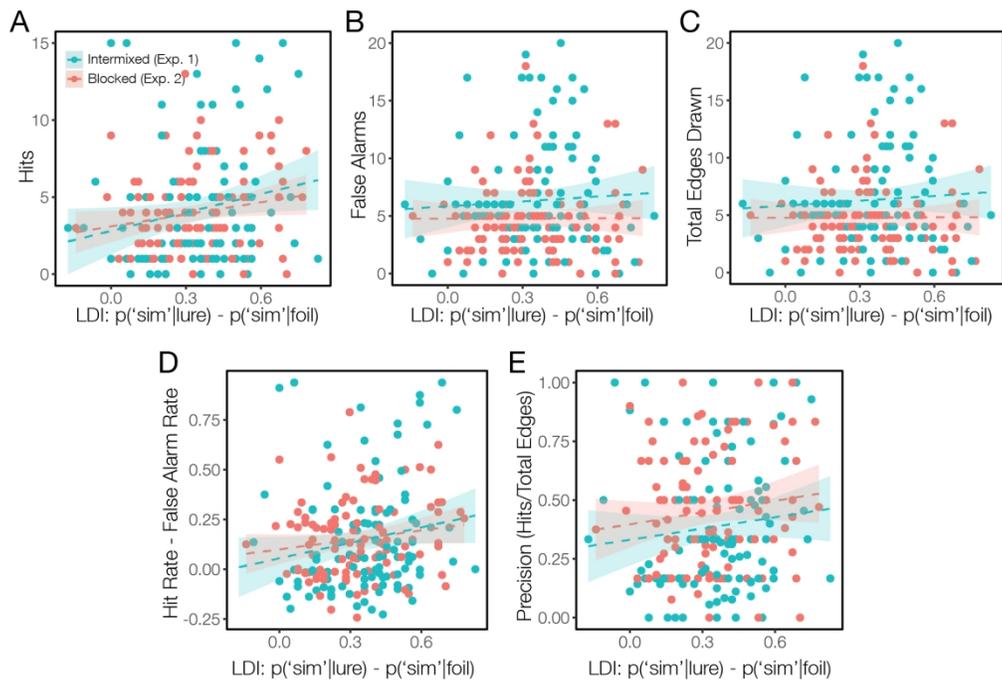


Figure 7

177x120mm (300 x 300 DPI)

Downloaded from <https://academic.oup.com/psychsocgerontology/advance-article/doi/10.1093/geronh/gbag038/8524505> by Barbara Scarpine user on 18 March 2026