



Pattern separation and pattern completion: Behaviorally separable processes?

Chi T. Ngo^{1,2} · Sebastian Michelmann³ · Ingrid R. Olson¹ · Nora S. Newcombe¹

© The Author(s) 2020

Abstract

Episodic memory capacity requires several processes, including mnemonic discrimination of similar experiences, termed *pattern separation*, and holistic retrieval of multidimensional experiences given a cue, termed *pattern completion*. Both computations seem to rely on the hippocampus proper, but they also seem to be instantiated by distinct hippocampal subfields. Thus, we investigated whether individual differences in behavioral expressions of pattern separation and pattern completion were correlated after accounting for general mnemonic ability. Young adult participants learned events comprised of a scene-animal-object triad. In the pattern separation task, we estimated mnemonic discrimination using lure classification for events that contained a similar lure element. In the pattern completion task, we estimated holistic recollection using dependency in retrieval success for different associations from the same event. Although overall accuracies for the two tasks correlated as expected, specific measures of individual variation in holistic retrieval and mnemonic discrimination did not correlate, suggesting that these two processes involve distinguishable properties of episodic memory.

Keywords Pattern separation · Pattern completion · Lure classification · Holistic recollection · Episodic memory

Introduction

Episodic memory binds together the diverse co-occurring elements that make up the specific events of our lives, forming distinctive and complex events that can guide ongoing behavior. This capacity requires many processes. Past experiences can share overlapping content, hampering the retrieval of a specific event against the backdrop of interfering memories. Thus, remembering a past event with high specificity is optimized by pattern separation processes, whereby similar experiences are assigned to non-overlapping neural codes in the service of preserving each event's distinctiveness. Further,

episodic memory retrieval unites various aspects of an event including where you were, specific people you met, and the objects that were encountered as an integrated unit (Tulving, 2002). Pattern completion processes enables this network of relations to be retrieved holistically, such that one constituent of an event can elicit the retrieval of other elements from the same event. Computational models posit that the hippocampus instantiates two crucial computations of *pattern separation* and *pattern completion* to support mnemonic discrimination and holistic recollection. Critically, hippocampal subfields differentially participate in these processes (Montaldi & Mayes, 2010; Neunuebel & Knierim, 2014; O'Reilly & McClelland, 1994; Rolls, 2016).

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13421-020-01072-y>) contains supplementary material, which is available to authorized users.

✉ Chi T. Ngo
ngo@mpib-berlin.mpg.de

¹ Department of Psychology, Temple University, Philadelphia, PA, USA

² Center for Lifespan Psychology, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany

³ Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA

Mnemonic discrimination via pattern separation

Accurate episodic memory requires remembering details with high specificity, so that they can be mnemonically discriminated from other similar memories to circumvent catastrophic interference. Pattern separation aids mnemonic discrimination by reducing the degree of representational similarity among overlapping experiences. The granule cells in the dentate gyrus (DG) assign distinct representations to similar inputs via sparse coding, thereby amplifying small differences from the entorhinal cortex. The pattern-separated signals are then

projected onto CA3 via the mossy fiber pathway. This mechanism helps minimize interference to support new learning (Marr, 1971; McClelland, McNaughton, & O'Reilly, 1995; Norman & O'Reilly, 2003).

In humans, one paradigm designed to index the behavioral outcome of pattern separation is the Mnemonic Similarity Task (MST). In this paradigm, participants first view a series of object images. At test, participants are shown targets (identical to the objects seen at encoding), lures (similar exemplars of the objects seen at encoding), and foils (new and dissimilar from those seen at encoding). The theoretically important part of this task is that it requires high-resolution mnemonic representations of studied objects in order to reject the highly-similar lure items. The ability to make this discrimination may relate to the DG/CA3 signals of the hippocampus, such that large responses were evoked by small changes in perceptual input (Lacy, Yassa, Stark, Muftuler, and Stark, 2011; Reagh & Yassa, 2014; reviewed in Yassa & Stark, 2011). Evidence from a study using ultra-high-resolution functional magnetic resonance imaging (fMRI) and multivariate pattern analysis further revealed that DG signals represent similar scenes in a less overlapping fashion compared to other hippocampal subfields and medial temporal regions (Berron et al., 2016). Causal evidence for a role of the DG in pattern separation comes from a case study of an individual with selective DG damage. This individual was impaired at making fine-grained target-lure discriminations (Baker et al., 2016), corroborating findings in rodents (Gilbert, Kesner, & Lee, 2001).

Holistic recollection via pattern completion

Retrieving a complex and multimodal episode also necessitates pattern completion, which occurs when a partial cue reactivates a complete event representation (Marr, 1971; McClelland, McNaughton, & O'Reilly, 1995; Norman & O'Reilly, 2003). Pattern completion is thought to rely in part on hippocampal CA3 network, which receives pattern-separated inputs from the DG as well as direct inputs from the entorhinal cortex via the perforant pathway, bypassing the DG. The CA3 pyramidal cells project onto themselves via the recurrent collaterals. This mechanism is thought to associate elements of an event to a shared representation in CA3, thereby enabling the recapitulation of the conjunctive representation of an entire event from a partial cue (Guzowski, Knierim, & Moser, 2004; Rolls, 2016). An important property of pattern completion is that a whole memory can be initiated from any part of the memory (Rolls, 2016).

One paradigm to test pattern completion in humans relies on the conceptualization that pattern completion enables the recovery of the entire event based on a partial cue so that all elements within an event can be retrieved successfully (or not). For instance, if retrieving a place successfully reminds us of the person we met there, it would also likely evoke our

memory of what objects we encountered in the same event. Findings from multiple studies have demonstrated that young adults indeed retrieve events in a holistic fashion (Bisby, Horner, Bush, & Burgess, 2018; Horner & Burgess, 2013, 2014; Ngo, Horner, Newcombe, & Olson, 2019). In the multi-element paradigm, participants first learned unique events, each comprised of a scene-person-object triad (e.g., kitchen-Obama-hammer). Later, participants performed a cued recognition task that included every possible cue-test element combination of each event (e.g., cue: kitchen, test: Obama), allowing for the examination of retrieval dependency between different associations of the same event. Evidence for significant retrieval dependency suggests that complex episodic memories may be retrieved as integrated units.

Findings on the neural bases of retrieval dependency point to the involvement of the hippocampus (Horner et al., 2015), and specifically the CA3 subfield (Grande et al., 2019). In a multi-element event task variant, participants learned scene-person-object triads (e.g., kitchen-Obama-hammer) in three separate trials of pairwise associations (e.g., kitchen-Obama; Obama-hammer; kitchen-hammer) along the encoding phase. After having seen the first two pairs of an event (e.g., kitchen-Obama, Obama-hammer), hippocampal activity during encoding of the final pair (e.g., kitchen-hammer) was associated with memory performance on other pairs of the same event. Furthermore, during cued recognition of pairwise associations (e.g., cue A; test B), neocortical activity corresponding to all event elements was reinstated, including the element that was incidental to a given trial (e.g., element C). Importantly, the extent of neocortical reinstatement of non-target elements correlated with hippocampal activity at retrieval, consistent with the presence of pattern completion (Horner et al., 2015; reviewed in Horner & Doeller, 2017). Crucially, this reinstatement signal is more robust within the CA3 compared to the dentate gyrus (Grande et al., 2019). Together, these findings support the idea that the retrieval of multiple elements of an event given the same cue is mutually contingent, and that this computation may, in part, rely on hippocampal pattern completion.

Behavioral relation between pattern separation and completion

Pattern separation and completion are distinct computations that may support complementary aspects of episodic capacity: laying down high-resolution memory traces and holistic retrieval of an integrated unit, respectively. If the relevant computations support different episodic features of a memory, do their individual differences in these mnemonic capacities track one another? The answer to this question hinges upon the conceptual link between the behavioral consequences of pattern separation and completion, but thus far, the interpretation of this link has been mixed.

In some investigations, the behavioral proxies of pattern separation and pattern completion are thought to occupy two ends of a lure discrimination performance scale (Yassa & Stark, 2011). The inverse operational definitions for pattern separation and completion may have resulted in the problematic interpretation that an improvement in one process necessarily comes with the impairment in the other (for a discussion of this criticism, see Hunsaker & Kesner, 2013). For example, pattern separation failure measured by false alarms in lures trials on the MST is sometimes interpreted as a gravitation of the system towards pattern completion (e.g., Holden & Gilbert, 2012; Kirwan & Stark, 2007; Lacy et al., 2011; Reagh & Yassa, 2014; Yassa et al., 2010). Although pattern completion likely underlies false alarm response in lure trials (Wynn, Ryan, & Buchsbaum, 2020), such an approach in measuring pattern completion behaviorally may pertain to specific instances in which the memory judgment reflects a tradeoff between the two processes.

It is likely that in some cases, pattern separation and completion play complementary roles in supporting episodic memory accuracy by establishing distinctive memory traces and holistic retrieval of complex experiences. The measurement dependency between pattern separation and pattern completion in paradigms that use similar lures, by design, does not usually permit independent indices of these processes (but see Ally, Hussey, Ko, & Molitor, 2013). Thus, investigating the behavioral relation between pattern separation and pattern completion requires an operational definition of pattern completion that is independent of pattern separation failure hallmarks (e.g., Ally, Hussey, Ko, & Molitor, 2013; Rollins & Cloude, 2018; Vieweg, Stangl, Howard, & Wolbers, 2015).

Current study

This study tested whether individual differences in holistic recollection were related to remembering the specific details of past events. In the same group of participants, we administered two tasks that shared the same encoding format. Participants learned a series of multi-element “events,” each consisting of a scene-animal-object triad. To operationalize pattern separation performance, we measured lure classification performance using a confusion matrix. We computed participants’ success in identifying the test events that included a lure element, i.e., lure classification. Importantly, a hallmark of pattern separation failure is confusing similar events with one another. Thus, we further corrected lure classification performance for the specific confusion patterns of misidentifying lure events as old events when measuring pattern separation. For pattern completion, we measured dependency – the degree to which the retrieval success of different associations from the same event is mutually contingent (all accurate or inaccurate).

We hypothesized that the behavioral expression of pattern separation through lure classification and pattern completion through holistic retrieval would not correlate with one another, as they were posited to rely on distinct neural computations. In addition, we tested whether general episodic memory performance that is nonspecific to pattern separation and completion would correlate between the two tasks. If this relation between gross accuracy of the two tasks exists, but the correlation between holistic retrieval and lure classification is not detected, this result would substantiate the idea that individual differences on pattern separation and pattern completion are unrelated. It is important to note, for naming heuristic purposes, pattern separation and pattern completion tasks here refer to the dependent variables of interest from each task, without the assumption that these processes are the only one at play in each task.

Methods

Participants

A total of 84 (42 female) undergraduate students ($M_{age} = 20.07$ years; $SD = 2.06$, range = 18–27) from Temple University participated for partial course credit. All participants gave informed consent and reported to have normal or corrected-to-normal vision. This experiment was completed in accordance with and approved by the Institutional Review Board committee at Temple University. This sample size was determined based on a power analysis to ensure there was sufficient power (0.80) to detect significant Pearson correlations with a medium effect size (Cohen’s $d = 0.3$) using G*power (see pre-registration).

Memory task

Materials

Pattern separation task We sampled 390 cartoon images of distinct scenes (e.g., living room), animals (e.g., panda), and object (e.g., comb) from Google image search engine. Every two images were pairs of perceptually similar exemplars (e.g., two perceptually similar living rooms). An independent group of 31 young adults rated the level of similarity of each exemplar pair in Qualtrics. We used the ranking results to assign scene-animal-object triad membership to minimize potential differences in difficulty levels, i.e., uneven levels of similarity, among different counterbalanced task versions across participants (see [Online Supplementary Material 1.1](#)). 360 images were used to create two sets of 60 events (sets A and B) such that event 1 in set A is made up of element exemplars of event 1 in set B. Four different task versions were created from the two sets of events to counterbalance the frequency with which

each event would appear in the study list or in the test list as target or lure events, across participants. The other 30 stimuli were used to create ten foil events (events containing elements that only appear at the test phase), which were fixed across all counterbalanced task versions.

Pattern completion task We sampled 72 cartoon images of distinct scenes, animals, and objects (24 per category) from Google image search engine to construct 24 scene-animal-object triads. The event assignment of the elements was randomized, with the exception that items with pre-experimental associations (e.g., books and library) were not assigned to the same event. The selected stimuli in this task did not overlap with those used in the pattern separation task.

Procedure

All participants were administered the pattern separation and pattern completion tasks in a counterbalanced order. Verbal instructions for both tasks were voice recorded.

Pattern separation task The task procedure comprised of two encoding test blocks with non-overlapping stimuli between the two blocks. Each block consisted of 30 encoding and 35 test trials, which led to a total of 60 encoding and 70 test trials. At encoding, participants were instructed that they would see many different events, each consisting of a scene-animal-object triad, and that they should pay close attention to all of the different elements altogether in each event (see Figure 1A). A practice phase preceded the encoding phase in order to acquaint the participants with the task (see [Online Supplementary Material 1.2](#) for details). At encoding, participants viewed the events sequentially (5 s each; 0.5-s intertrial interval (ITI)).

The test phase followed the encoding phase and consisted of 35 trials. Fifteen trials were targets – events that contained all three identical elements as studied events. Five trials were scene lures – events that contained identical animals and objects, but the *scenes* were similar exemplars of those seen at encoding. Five trials were *animal* lures – events that contained the identical scenes and objects, but the animals were similar exemplars of those seen at encoding. Five trials were object lures – events that contained identical scenes and animals, but the *objects* were similar exemplars of those seen at encoding. Five trials were foils – events that contained all novel and dissimilar elements from all studied events (see Figure 1B). Participants were instructed to categorize five test event types with five memory judgments (“identical,” “similar scene,” “similar animal,” “similar object,” and “completely new”) by pressing the color-coded keys on the keyboard. We provided a “cheat sheet” with each response and its corresponding color key to ensure that memory failures were not due to misremembering response-key mapping. There were no

missing responses as the response time was unrestricted. The memory task took approximately 25 min.

Pattern completion task At encoding, participants were told that they would see many different scene-animal-object triads called “events” and that they should pay close attention to all of the different elements including the scene, animal, and object altogether in each event. An example phase preceded the encoding phase in order to acquaint the participants with the task (see [Online Supplementary Material 1.2](#) for details). After the example phase, the encoding phase commenced, in which participants viewed 24 events sequentially (5 s each; 0.5-s ITI).

Immediately after the encoding phase, participants performed a self-paced four-alternative-forced-choice task that consisted of 144 trials. We tested participants on every possible cue-test combination of each studied event (e.g., cue: scene; test: animal), resulting in six test trials per event, which totaled 144 test trials. On each trial, a cue and four options were presented simultaneously on the screen (see Fig. 2). Among four options, one was a target (the correct item) because it belonged to the same event as the cue. The three lures were same-category elements from different events. The positions of the correct answer were counterbalanced across the entire test phase. Across all 24 events, any given two test trials that had overlapping cue items (e.g., A_B^1 and A_C^1) or tested items (e.g., B_A^1 and C_A^1) only shared one foil item (out of three) with respect to their event membership. For example, for the A_B test trial of event 1, the foils included the B elements from events 2, 3, and 4, whereas for the A_C trial of event 1, the foils included the C elements from events 3, 5, and 7 (one B and one C foil both from event 3). Furthermore, all items served as foils an equal number of times across all 144 test trials. Participants were asked to press the four keys on the number keypad marked with colored stickers that correspond to the four options presented on the right side of the screen. There were no missing responses as the response time was unrestricted. This task took approximately 30 min.

Verbal intelligence

Prior to the memory tasks, all participants were administered the 45-item American National Adult Reading Test (AMNART; Grober & Sliwinski, 1991) – an American version of the National Adult Reading Test (Nelson, 1982). This test measures the ability to read aloud irregular words. Pronunciation errors were tallied and AMNART-estimated verbal IQ scores were calculated using Grober and Sliwinski’s formula, which accounts for years of education. The AMNART was included as a covariate variable for all statistical analyses.

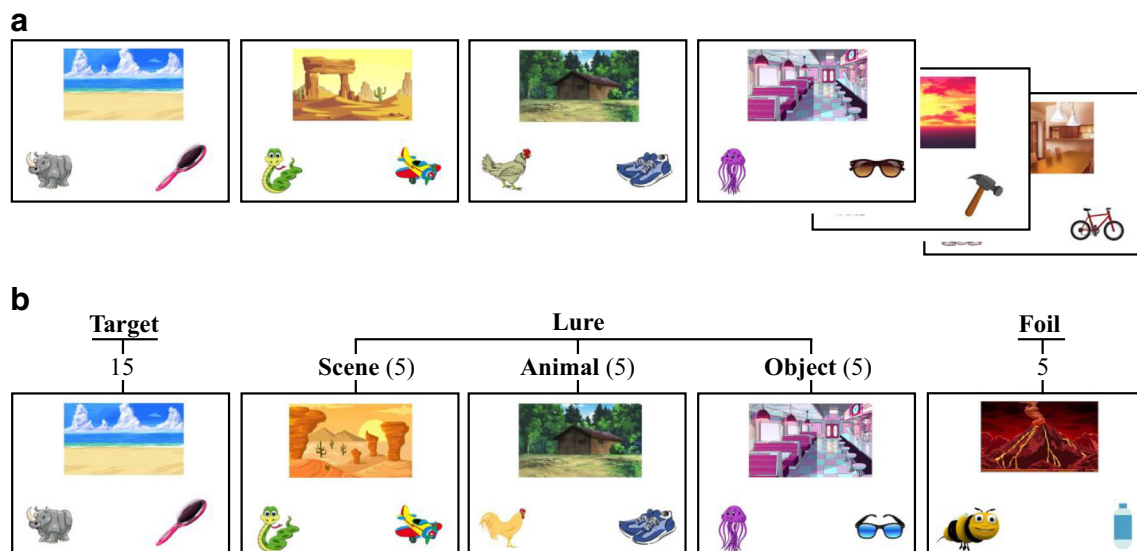


Fig. 1 (A) A schematic depiction of the multi-element event task procedure for each of the two blocks. In each block, participants first learned 30 events, each comprised of a scene, an animal, and an object. (B) The test

list consisted of 35 events including 15 target, five scene lures, five animal lures, five object lures, and five foil trials. Note that the stimuli between the two blocks were non-overlapping

Analytical approach

Pattern separation

The behavioral expression of pattern separation is characterized by two crucial elements: (1) The detection accuracy and response precision in classify a similar event as similar; and (2) The ability to avoid confusing similar events as old events (reviewed in Yassa & Stark, 2011). To best capture these features of participants' responses, we computed the confusion matrix for each participant.

Confusion matrix First, we constructed a 5×5 confusion matrix for every participant, with the rows of the matrix representing the classes of the test items (e.g., target, scene lure), and the columns representing a participant's response frequency. A value in each cell denotes the number of instances in which a participant identified the test items as a given class. The utility of using a confusion matrix is that it effectively handles the multiclass problem. It provides classification accuracy for each class (diagonal cells), and specifies inter-class confusion, i.e., the frequency with which each class is mistaken for each of the other (off-diagonal cells, see Fig. 3). Importantly, from these confusion matrices, we can derive indices of precision and sensitivity for each class (Powers, 2011). Precision of class A is the proportion of correctly identified trials out of the total number of trials a participant labeled the test items as class A (row), therefore capturing the response precision for a given class by a participant. Sensitivity of class A is the proportion of correctly identified trials out of the total number of trials in class A (column), capturing detection performance (see Fig. 4).

Next, we calculated an F1 score combining both precision and sensitivity for each class, which used a harmonic mean of the two indices – a more conservative mean measure as it punishes the extreme values (Powers, 2011). The F1-scores range from 0–1, where 0 indicates no discrimination between a given class from other classes, and 1 indicates a perfect classification for a given class without missing an item from a class or misclassifying items from other classes as a given class.

$$F1 \text{ score} = 2 * \text{Precision} * \text{Sensitivity} / (\text{Precision} + \text{Sensitivity})$$

Confusing lures as old items is the hallmark of pattern separation failure (reviewed in Yassa & Stark, 2011). Thus, specific kinds of inter-class confusion are crucial to succinctly characterize the underlying processes of pattern separation. In order to approximate the specific ability to pattern separate similar experiences (2), we computed the frequency at which lures were misclassified as targets. To do this, we derived a combined measure by correcting F1 scores for those misclassifications, i.e., we subtracted the frequency with which lures were misclassified as targets from the F1 scores of lures, given that mistaking a similar but different item as an old item is a hallmark of pattern separation failure (reviewed in Yassa & Stark, 2011).

- Scene Lure classification = Scene Lure F1 score – P("Target" | Scene Lure)
- Animal Lure classification = Animal Lure F1 score – P("Target" | Animal Lure)
- Object Lure classification = Object Lure F1 score – P("Target" | Object Lure)

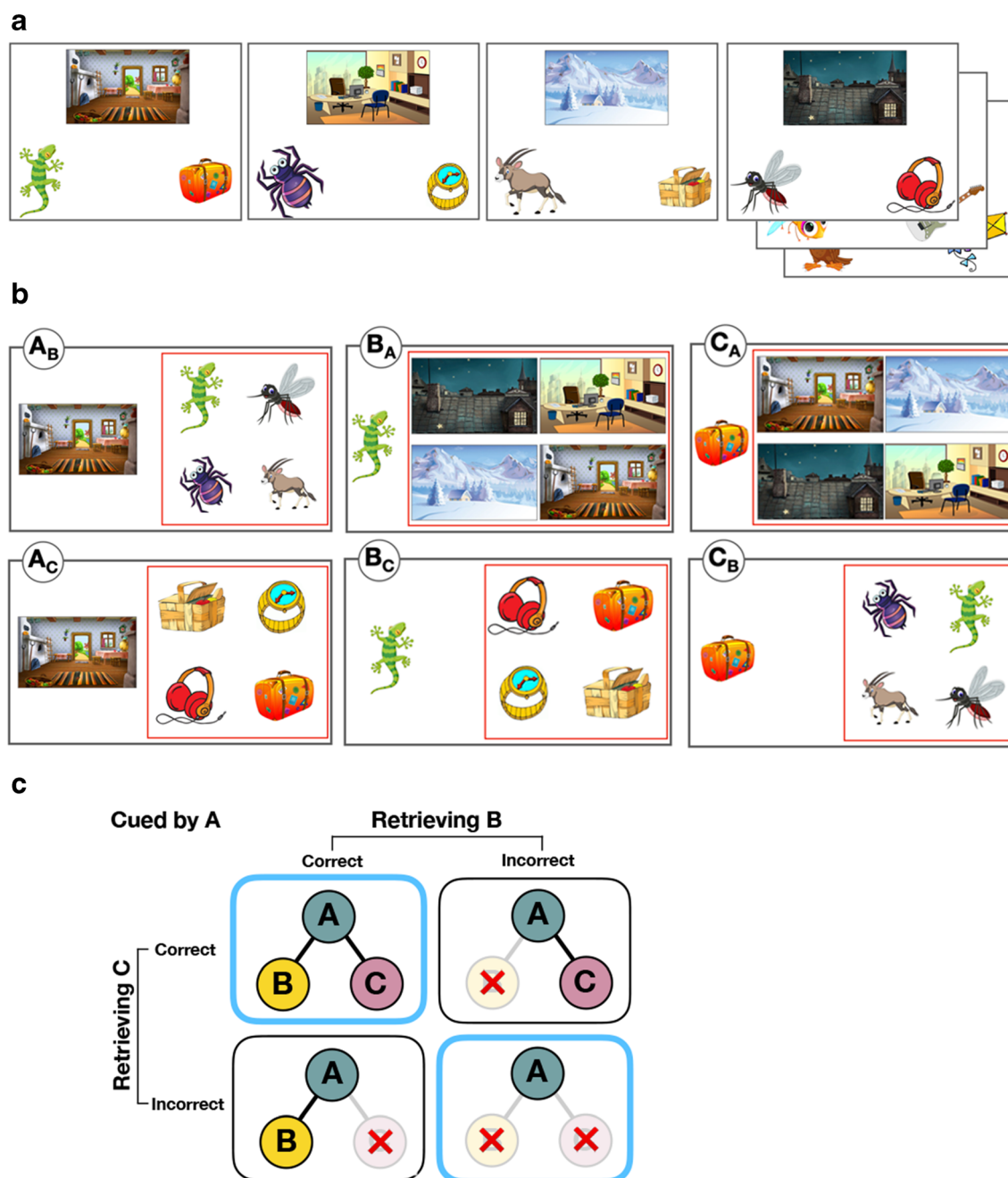


Fig. 2 (A) Examples of the scene-animal-object. (B) Examples of six retrieval types per event in the test phase. Each element of a studied event took turns serving as the cue (item presented on the left side of the screen) and the tested element (one of the four options presented

inside the red box). (C) A schematic depiction of how the proportion of joint retrieval for $A_B A_C$ pairs was computed for each participant by concatenating the proportion of events in the blue outlined boxes out of the total number of events

We took a similar approach with targets and foils. For targets, we corrected for the interclass confusion in which that an entirely old event was misclassified as entirely new.¹ We computed a corrected F1 score for targets by subtracting the frequency in which foils were misclassified as targets from the target F1 score. For foils, we corrected for the interclass confusion in which an entirely new event was mistaken as entirely old. We computed a corrected F1 score for foils by subtracting the frequency with which foils were misclassified as targets from the foil F1 score.

- Target classification = Target F1 score – P(“Foil” | Target)
- Foil classification = Foil F1 score – P(“Target” | Foil)

Pattern completion

Estimating retrieval dependency The retrieval dependency between retrieval successes for different associations within the same event was computed using the same methods as in

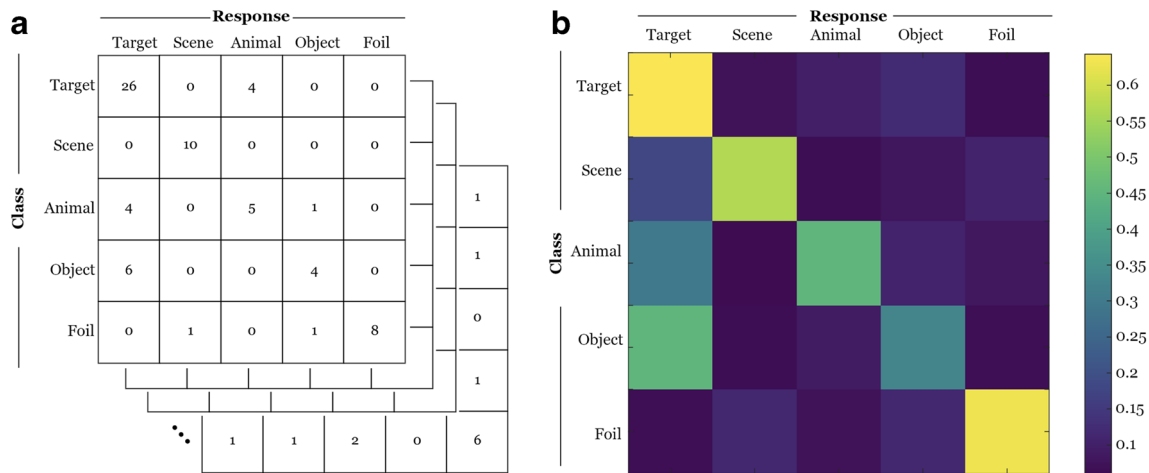


Fig. 3 (A) A 5×5 confusion matrix was constructed for each participant to illustrate the frequency of response (column) to each class of test item (row). (B) A group-level confusion matrix with classification accuracy

(diagonal cells) and errors (off-diagonal cells). Color intensity illustrates the classification frequency in each cell

previous studies (Horner & Burgess, 2013, 2014; Horner et al., 2015; Bisby et al., 2018). Six 2×2 contingency tables for the data and the predicted independent model were

computed for each participant based on their retrieval accuracy for each pairwise association in order to assess dependency between retrieving two elements when cued by the remaining

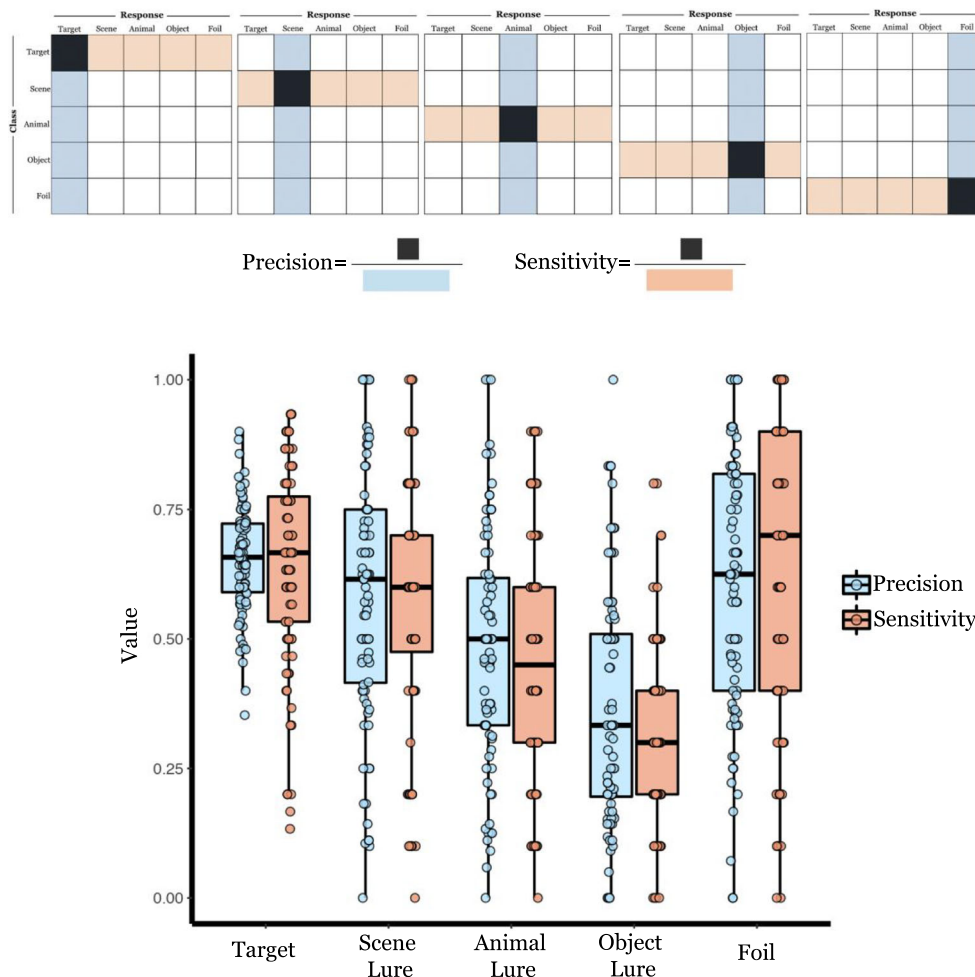


Fig. 4 A schematic depiction of the formulas derived from the confusion matrices (A) and the distribution of precision and sensitivity for each class of test item (B) for the pattern separation task

common element within an event ($A_B A_C$; i.e., cue with A and retrieve B, and cue with A and retrieve C), and the dependency between retrieving a common item when cued by the other two elements within an event ($B_A C_A$; i.e., cue with B and retrieve A, and cue with C and retrieve A). Each 2×2 contingency table for the data, for every participant, shows the proportion of events that fall within the four categories: both A_B and A_C are correct or incorrect, A_B correct and A_C incorrect, and vice versa. To measure retrieval dependency, we computed the proportion of joint retrieval for the data – defined as the proportion of events in which both associations were either correctly or incorrectly retrieved (cells 1, 1 and 2, 2 of each contingency table; see Fig. 2C). We then averaged this measure across six contingency tables (three tables for the $A_B A_C$ analysis, for each element-type, and three tables for the $B_A C_A$ analysis, for each element-type) for each participant.

The independent model estimates the degree of statistical dependency if retrieval success for different associations within the same event is independent from one another, given a participant's overall accuracy. If retrievals of event pairs are independent, the probability of the successful retrieval for both A_B and A_C is equal to $P_{AB} * P_{AC}$, where P_{AB} is the probability of retrieving B when cued by A across all events, and similarly for P_{AC} (see Table 1 for full details). The independent model serves as a predicted baseline for which we compare the proportion of joint retrieval in the data. Given that the proportion of joint retrieval in the data scales with accuracy, the main index of retrieval dependency was the difference between the joint retrieval in the data and independent model for each participant – referred to as dependency. If dependency is significantly greater than zero, this provides evidence for significant retrieval dependency (Horner & Burgess, 2014).

Statistical analyses

All planned statistical analyses were performed using JASP.

Data availability The stimuli in the memory tasks (<https://osf.io/pjy28/>) and second-level data (<https://osf.io/5k8jx/>) are publicly available through the Open Science Framework.

Table 1 Contingency table for the predicted independent model for proportion of correct and incorrect cued recognition over the total number of events for elements B and C when cued by A

Cued by A		Retrieving B	
		Correct	Incorrect
Retrieving C	Correct	$\sum_{i=1}^N P_{AB} P_{AC}$	$\sum_{i=1}^N P_{AC} (1 - P_{AB})$
	Incorrect	$\sum_{i=1}^N P_{AB} (1 - P_{AC})$	$\sum_{i=1}^N (1 - P_{AB}) (1 - P_{AC})$

Results

Pattern separation task

Overall accuracy

Overall accuracy was defined as the proportion of correct responses out of 70 total test trials. On average, young adults' overall accuracy was 0.56 ($SE = 0.02$), which was significantly above chance (20% given that they were five response choices), $t(83) = 20.98$, $p < .001$, 95% CI for mean difference [0.32, 0.39], $d = 2.29$. There were no sex differences in overall accuracy, $t(82) = -0.62$, $p = .54$, 95% CI for mean difference [-0.09, 0.05], $d = -0.13$, $BF_{01} = 3.72$. Overall accuracy did not significantly differ between participants who completed the pattern separation task first or second, $t(82) = -0.41$, $p = .69$, 95% CI for mean difference [-0.08, 0.05], $d = -0.09$, $BF_{01} = 4.09$. Accuracy on this task and verbal intelligence measured by AMNART showed a significant positive correlation, $r(82) = 0.25$, $p = .02$.

Frequency of each participant's response to each class item was calculated in a confusion matrix (see Fig. 3). Precision and sensitivity rates were then calculated for each class of the test item for each participant (see Fig. 4). Finally, we corrected F1 score (harmonic mean of precision and sensitivity) for each class to account for specific inter-class confusion to better characterize discriminability (see Fig. 5).

There was a main effect of test item types, $F(4, 83) = 75.56$, $p < .001$, $partial \eta^2 = 0.48$. Post hoc comparison with Bonferroni correction revealed that foil ($M = 0.53$, $SE = 0.04$) did not differ from target classification ($M = 0.45$, $SE = 0.04$), $t = -2.32$, $p = 0.23$, 95% CI for mean difference [-0.18, 0.02], $d = -0.25$, but was greater than scene lure ($M = 0.38$, $SE = 0.04$), $t = -3.76$, $p = .003$, 95% CI for mean difference [-0.26, -0.04], $d = -0.41$, animal lure ($M = 0.15$, $SE = 0.04$), $t = -9.08$, $p < .001$, 95% CI for mean difference [-0.50, -0.26], $d = -0.99$, and object lure ($M = -0.12$, $SE = 0.04$), $t = -13.28$, $p < .001$, 95% CI for mean difference [-0.79, -0.51], $d = -1.45$. Target classification did not differ from scene, $t = 1.79$, $p = 0.77$, 95% CI for mean difference [-0.04, 0.18], $d = 0.20$, but was greater than animal, $t = 6.64$, $p < 0.001$, 95% CI for mean difference [0.17, 0.43], $d = 0.73$, and objects, $t = 12.62$, $p < .001$, 95% CI for mean difference [0.44, 0.69], $d = 1.38$.

Lure classification differed across the stimulus categories. Participants were better at classifying scene lures than animal lures, $t = 5.62$, $p < .001$, 95% CI for mean difference [0.11, 0.35], $d = 0.61$, and object lures, $t = 11.77$, $p < .001$, 95% CI for mean difference [0.38, 0.62], $d = 1.28$. Animal lures were classified better than object lures, $t = 5.49$, $p < .001$, 95% CI for mean difference [0.13, 0.40], $d = 0.60$. It is worth noting that the similarity rating is higher for object lure than for scene lure (see Online Supplementary Material 1.1). Thus, differences in lure classification between scenes and objects may

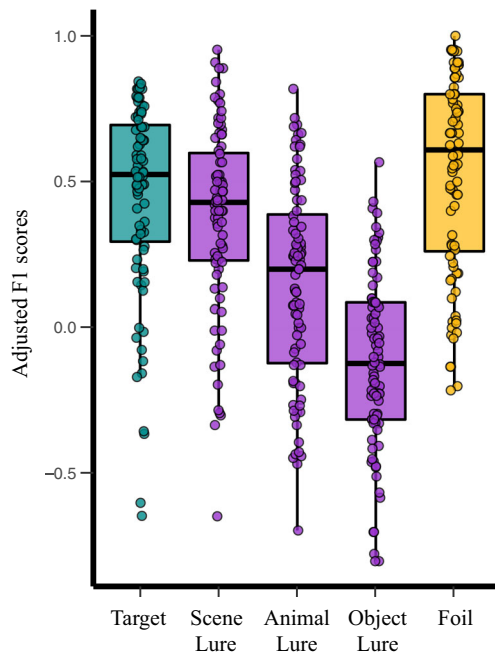


Fig. 5 The distribution of adjusted F1 score for each class of test item. The F1 score is the harmonic mean between the sensitivity and precision indices for a given class of test item derived from each participant's confusion matrix. F1 scores were adjusted for specific misclassification patterns (e.g., misclassify lures as targets, see text)

be due to the uneven perceptual similarity of the exemplar pairs between these two categories.

Pattern completion task

Overall accuracy

Overall accuracy is defined as the proportion of target selection across 144 test trials. Young adults performed significantly greater than chance level (0.25), $M = .69$, $SE = 0.02$, $t(83) = 18.21$, $p < .001$, 95% CI for mean difference [0.39, 0.49], $d = 1.99$ (see Fig. 6A). There were no sex differences in overall accuracy, $t(82) = 0.44$, $p = .66$, 95% CI for mean difference [-0.08, 0.12], $d = 0.10$, $BF_{01} = 4.04$. Overall accuracy did not significantly differ between participants who completed the pattern completion task first or second, $t(82) = 1.14$, $p = .26$, $d = 0.25$, $BF_{01} = 2.48$. Accuracy and verbal intelligence measured by AMNART showed a trend towards a positive correlation, $r(82) = 0.18$, $p = .10$.

Retrieval dependency

To test for evidence of holistic recollection, paired-samples t -tests were conducted to examine whether dependency in the data exceeded the independent retrieval model. As expected, we found significant dependency, such that dependency in the data ($M = 0.85$, $SE = 0.02$) exceeded the independent retrieval model ($M = 0.72$, $SE = 0.02$), $t(83) = 12.35$, $p < .001$, 95% CI

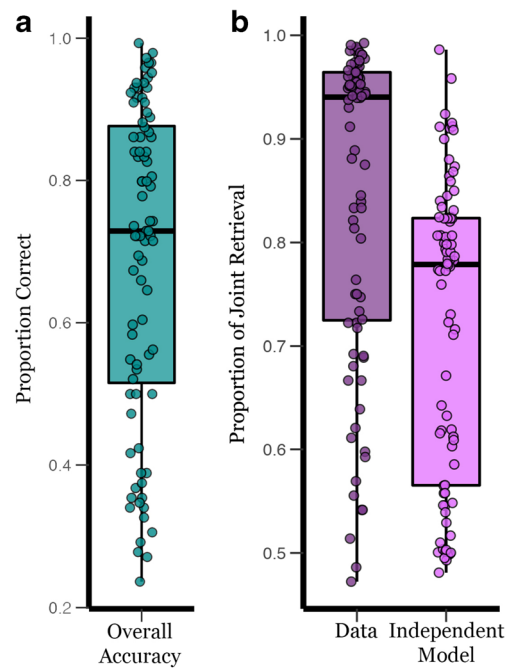


Fig. 6 The distribution of the overall accuracy (A) and the proportion of joint retrieval of the data and independent model (B) of the pattern completion task. Overall accuracy in the pattern completion task is the proportion of correct trials across 144 test trials. The proportion of joint retrieval in the data and the independent model was calculated from the contingency table for each participant

for mean difference [0.11, 0.15], $d = 1.35$ (see Fig. 6B). These results replicated previous studies showing significant retrieval dependency in young adults using different stimuli (e.g., Bisby et al., 2018; Horner & Burgess, 2013, 2014; Ngo et al., 2019). Together these studies demonstrate that memories for multi-element events may be represented as integrated episodic units.

We also tested whether different stimulus categories yielded uneven retrieval dependency. It has been suggested that scenes may play a robust cuing role in episodic memory (Robin, 2018). Thus, it is possible that scenes may yield stronger dependency compared to other element categories. To test this possibility, we calculated dependency scores for associations that shared the same cue ($A_B A_C$) or tested item ($B_A C_A$) separately for each category. A one-way ANOVA showed a nonsignificant effect of category on dependency, $F(2, 166) = 1.21$, $p = 0.30$, $\text{partial } \eta^2 = 0.01$, suggesting that retrieval dependency did not differ among the three categories of elements.

Pattern completion and pattern separation correlations

Dependency and lure classification correlations

To probe the behavioral relation between pattern separation and completion, we tested whether individual differences in

lure classification and dependency co-varied with one another. We averaged the adjusted F1 scores from the three lure categories to yield one overall lure classification index. Dependency was a difference between the proportion of joint retrieval in the data and the independent model of retrieval. A partial correlation between lure classification and dependency, with AMNART as a covariate, revealed a nonsignificant correlation, $r(81) = .12$, $p = .28$, 95% CI [-0.09, 0.32], $BF_{01} = 4.05$ (see Fig. 7A).

We further tested the possibility that pattern separation and completion only show related performances at the category-specific level, we correlated lure classification and dependency within each category separately (scene, animal, object; see pre-registration). Lure classification and category-specific dependency did not correlate for scene, $r(81) = .14$, $p = 0.20$, 95% CI [-0.07, 0.35], $BF_{01} = 3.22$, animal $r(81) = .08$, $p = 0.45$, 95% CI [-0.13, 0.29], $BF_{01} = 5.56$, or object, $r(81) = -.002$, $p = .89$, 95% CI [-0.23, 0.20], $BF_{01} = 7.27$, after partialing out AMNART (see Fig. 7B).

These results suggest that lure classification and holistic recollection may be behaviorally separable processes such that individual differences in one facet of episodic capacity do not track those in the other.

Pairwise associative memory and target classification correlation

Next, we reasoned that indices of general episodic memory performance – nonspecific to retrieval dependency or lure classification – should correlate between the two tasks as they both tap general episodic capacity. For the pattern separation task, lure classification specifically indexes pattern separation, whereas *target classification* indexes the abilities to recognize old events. For the pattern completion task, dependency indexes holistic recollection, whereas *overall accuracy* indexes pairwise associative recognition, irrespective of event membership. Both of these processes are general memory performances that are not uniquely attributed to pattern separation or pattern completion. Therefore, we predicted that overall accuracy in the pattern completion task and target classification in the pattern separation task would relate to one another. We treated this analysis as a “control” correlation analysis to ensure that the nonsignificant correlation between lure classification and retrieval dependency was not due to potential superficial differences between the two tasks, or due to power issues (see pre-registration). Indeed, they were strongly correlated, $r(81) = .51$, $p < .001$, after controlling for AMNART (see Fig. 7A). Crucially, this correlation significantly exceeded the dependency versus lure classification correlation, $Z = 2.79$, $p < 0.01$.

In our pre-registration, we planned to assess pattern separation performance using the raw lure identification rate (i.e., identifying a test event that contained a similar scene to a

scene lure trial). However, raw lure identification does not adjust either for response bias or for the specific false alarm rate relevant to pattern separation (i.e., misclassifying a lure as a target). Thus, we employed a confusion matrix approach to estimate pattern separation. The correlational results using the pre-registered approach are reported in [Online Supplemental Material 2.2](#).

General discussion

The hippocampus supports episodic memory capacity by (1) preserving distinctive features of specific past experiences, and (2) by recapitulating constituent parts of an episodic memory altogether at retrieval. Pattern separation and completion are the underlying neurocomputationally distinct processes that work in concert to support different aspects of episodic memory capacity. This study tested whether individual differences in the behavioral indices of these processes would relate to each other using an individual differences approach. In a novel combination of behavioral tasks of pattern separation and completion, we show that retrieval dependency and lure classification indeed do not track each other. These findings demonstrate a degree of inter-independence of the two processes, suggesting that different hippocampal computations play complementary roles in supporting different facets of episodic memory. The nonsignificant correlation between holistic recollection and lure classification aligned with our predictions, and was supported by an analysis using Bayes factors, which suggests that the null hypothesis is four times more likely than the alternative hypothesis. Further, gross performances on the two tasks – nonspecific to pattern separation and completion – were significantly correlated. Importantly, the gross-performances correlation was greater than that of the dependency versus lure classification correlation. These results suggest that certain aspects of the two tasks tap general episodic capacity, and that our null finding was unlikely to be due to power issues.

Our findings shed light on the debate on the computational interplay between pattern separation and pattern completion (for a discussion, see Hunsaker & Kesner, 2013). The behavioral dissociation between lure discrimination and retrieving old events from partial cues aligns with neuropsychological findings in a patient with selective dentate gyrus damage (Baker et al., 2016). As expected, this patient was severely impaired at lure discrimination on the MST. In addition to the MST, the authors administered the Memory Image Completion (MIC) task, which evaluates recognition memory for scenes (old and new) presented at varying degrees of completeness during test. Interestingly, the patient was spared on recognizing old scenes at all levels of completeness, suggesting that abilities to complete a true old memory from partial cues was neither impaired or enhanced compared to healthy

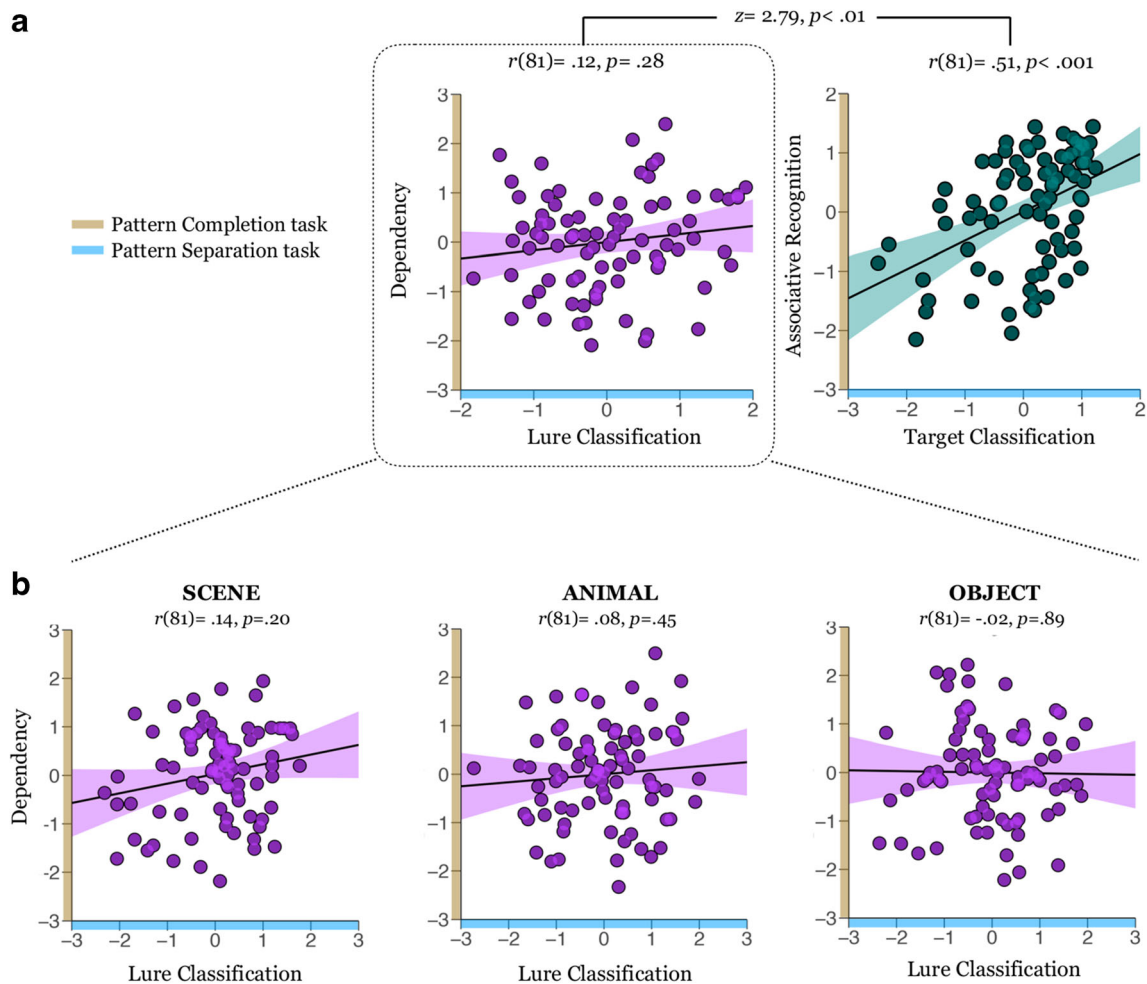


Fig. 7 Scatterplots of the standardized residuals illustrating the relation between dependency on the pattern completion task and lure classification on pattern separation task (**A, left**), between associative memory on the pattern completion task and the target classification on

the pattern separation task (**A, right**), and between dependency and lure classification for each stimulus category separately (**B**). All scatterplots depict the partial correlations after partialing out AMNART score

controls. However, he was impaired in correctly rejecting novel scenes at all levels of completeness compared to healthy controls, suggesting an increased tendency to pattern complete in the presence of noisy cues. Based on these findings, DG damage appears to be linked to the selective deficits in lure discrimination (i.e., poor lure rejection in both the MST and MIC), while sparing the ability to reconstruct old experience based on partial cues (i.e., intact target recognition in the MIC). This neuropsychological evidence that an impairment in lure discrimination does not necessarily come with an impairment (or increase) in pattern completing an old event converges with our current findings in healthy adults.

Our study employed separate paradigms from the pattern separation and completion literature that provide independent behavioral indices. We implemented critical methodological and analytical modifications to specifically probe their behavioral relation. Methodologically, we equated the encoding format of both tasks using multi-element events, departing from other MST-task variants, which typically test individual items

(e.g., as reviewed in Yassa & Stark, 2011). This design allowed us to examine the pattern separation-pattern completion relations at the category-agnostic and category-specific levels. Analytically, we leveraged computational principles from confusion matrices to extract one index of lure classification, accounting for hit rate, response precision, and false alarm.

It is important to note that there is variation in the behavioral paradigms designed to approximate pattern completion. Based on the definition of pattern completion as recovering an old experience from partial or degraded cues, some tasks have operationalized partial cues as fragments of learned scenes (Vieweg et al., 2015), whereas in the multi-element event task, partial cues are defined as elements within complex events (Horner & Burgess, 2013, 2014). Others have operationalized degraded cues as noisy cues through the use of similar lures that include both the original input and added novel features (e.g., perceptually similar object exemplars) (Yassa & Stark, 2011). The nature of the cues presented at test may impact the

degree to which pattern completion is engaged (O'Reilly & McClelland, 1994). Importantly, it has been suggested that the use of partial cues as opposed to noisy cues in behavioral paradigm may preferentially engage pattern completion (reviewed in Liu, Gould, Coulson, Ward, & Howard, 2016).

Further, when assessing pattern completion behaviorally, it may be important to consider the difference between a gravitation towards pattern completion in the presence of noisy cues and the capacity to pattern complete a true old event based on partial cues. Models of aging propose that over-activity of the CA3 auto-associative function leads to the over-expression of old information at the expense of discriminating new information (Wilson, Gallagher, Eichenbaum, & Tanila, 2006). Relative to young adults, older adults show a greater tendency to falsely recognizing lures as old scenes (Stark, Yassa, Lacy, & Stark, 2013; Vieweg et al., 2015; Yassa and Stark, 2011). At the same time, older adults are also less likely to pattern complete target items as old based on partial inputs (Vieweg et al., 2015). A reduction in the abilities to recover the learned experiences from partial cues may co-occur with an increase in a bias towards pattern completion old events when in the presence of noisy cues. However, the inferences of pattern separation and completion based on one behavioral continuum of lure discrimination performance may not make this distinction obvious.

Our work presented the initial evidence that capacities of remembering a past event in its totality and remembering individual elements with high resolution do not scale with one another in healthy young adults. These processes likely play complementary roles and contribute to distinct properties of complex episodic capacities. Interpretation of pattern completion using behavioral measures would also benefit from a systematic examination of whether tasks designed to assess pattern completion correlate with one another and how each may relate to pattern separation. Future research should tackle a full treatment of pattern separation and pattern completion dissociation by combining behavioral paradigms and neuroimaging methods.

Author Note We would like to thank Ying Lin, Elizabeth Eberts, Lizi Zhong, Richard Ho, Rebecca Adler, Emily Cerimele, and Nadhia Engle, for their help with stimuli development and/or with data collection. This work was supported in part by National Institutes of Health grants F31HD090872 to C.T.N, RO1 MH091113 to I.R.O., and R21 HD098509 to I.R.O. and N.S.N. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Health. The authors declare no competing or conflicting financial interests.

Open practices statement This study was formally pre-registered (<https://osf.io/b8c2h>). De-identified raw and second-level data and experimental materials have been made publicly available through the Open Science Framework (<https://osf.io/he3wn/>).

Author Contributions C. T. Ngo and N. S. Newcombe developed the research questions and the design of experiment. Data were collected by C. T. Ngo and others (see Author note). Data analyses and results interpretation were led by C. T. Ngo and S. Michelmann, under the supervision of N. S. Newcombe and I. R. Olson. C. T. Ngo drafted the manuscript, and all co-authors provided critical revisions. All authors approved the final version of the manuscript for submission.

Funding Information Open access funding provided by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ally, B. A., Hussey, E. P., Ko, P. C., & Molitor, R. J. (2013). Pattern separation and pattern completion in Alzheimer's disease: Evidence of rapid forgetting in amnesic mild cognitive impairment. *Hippocampus*, 23(12), 1246–1258.
- Baker, S., Vieweg, P., Gao, F., Gilboa, A., Wolbers, T., Black, S. E., & Rosenbaum, S. (2016). The human dentate gyrus plays a necessary role in discriminating new memories. *Current Biology*, 26, 2629–2634.
- Berron, D., Schutze, H., Maass, A., Cardenas-Blanco, A., Kuijff, H. J., Kumaran, D., & Düzel, E. (2016). Strong evidence for pattern separation in human dentate gyrus. *The Journal of Neuroscience*, 36(29), 7569–7579.
- Bisby, J. A., Horner, A. J., Bush, D., & Burgess, N. (2018). Negative emotional content disrupts the coherence of episodic memories. *Journal of Experimental Psychology: General*, 147(2), 243–256.
- Gilbert, P. E., Kesner, R. P., & Lee, I. (2001). Dissociating hippocampal subregions: double dissociation between dentate gyrus and CA1. *Hippocampus*, 11(6), 626–636.
- Grande, X., Berron, D., Horner, A. J., Bisby, J. A., Düzel, E., & Burgess, N. (2019). Holistic recollection via pattern completion involves hippocampal subfield CA3. *Journal of Neuroscience*, 39(41), 8100–8111.
- Grober, E., & Sliwinski, M. (1991). Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *Journal of Clinical and Experimental Neuropsychology*, 13(6), 933–949.
- Guzowski, J. F., Knierim, J. J., Moser, E. I. (2004). Ensemble dynamics of hippocampal regions CA3 and CA3. *Neuron*, 44(4), 581–584.
- Holden, H. M., & Gilbert, P. E. (2012). Less efficient pattern separation may contribute to age-related spatial memory deficits. *Frontiers in Aging Neuroscience*, 29, 4–9.
- Horner, A. J., & Burgess, N. (2013). The associative structure of memory for multi-element events. *Journal of Experimental Psychology: General*, 142(4), 1370–1383.
- Horner, A. J., & Burgess, N. (2014). Pattern completion in multielement event engrams. *Current Biology*, 24, 988–992.

- Horner, A. J., Bisby, J. A., Bush, D., Lin, W. J., & Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nature Communications*, 6, 7462.
- Horner, A. J., & Doeller, C. F. (2017). Plasticity of hippocampal memories in humans. *Current Opinion in Neurobiology*, 43, 102–109.
- Hunsaker, M. R., & Kesner, R. P. (2013). The operation of pattern separation and pattern completion processes associated with different attributes or domains of memory. *Neuroscience and Biobehavioral Reviews*, 37, 36–58.
- Kirwan, C. B., & Stark, C. E. L. (2007). Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. *Learning & Memory*, 14(9), 625–633.
- Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T., & Stark, C. E. (2011). Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learning & Memory*, 18(1), 15–18.
- Liu, K. Y., Gould, R. L., Coulson, M., Ward, E. V., & Howard, R. J. (2016). Tests of pattern separation and pattern completion in humans: A systematic review. *Hippocampus*, 26(6), 705–717.
- Marr, D. (1971). Simple memory: A theory for archicortex. *Philosophical transactions of the Royal Society of London*, 262(841), 23–81.
- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why are there complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3), 419–457.
- Montaldi, D., & Mayes, A. R. (2010). The role of recollection and familiarity in the functional differentiation of the medial temporal lobes. *Hippocampus*, 20(11), 1291–1314.
- Nelson, H. E. (1982). *National Adult Reading Test (NART): For the assessment of premorbid intelligence in patients with dementia: Test manual*. Windsor: NFER-Nelson.
- Neunuebel, J. P., & Knierim, J. J. (2014). CA3 retrieves coherent representations from degraded input: Direct evidence for CA3 pattern completion and dentate gyrus pattern separation. *Neuron*, 81(2), 416–427.
- Ngo, C. T., Horner, A. J., Newcombe, N. S., & Olson, I. R. (2019). Development of holistic episodic recollection. *Psychological Science*.
- Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychological Review*, 110(4), 611–646.
- O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus*, 4(6), 661–682.
- Powers, D. M. (2011). Evaluations: from precision, recall and F-measure to ROC, informedness, markedness, and correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63.
- Reagh, Z. M., & Yassa, M. A. (2014). Object and spatial mnemonic interference differentially engage lateral and medial entorhinal cortex in humans. *Proceedings of the National Academy of Sciences*, 111(40), E4264–E4273.
- Robin, J. (2018). Spatial scaffold effects in event memory and imagination. *Wiley Interdisciplinary Reviews: Cognitive Science*, 9(4), e1462.
- Rolls, E. T. (2016). Pattern separation, completion, and categorization in the hippocampus and neocortex. *Neurobiology of Learning and Memory*, 129, 4–28.
- Rollins, L., & Cloude, E. B. (2018). Development of mnemonic discrimination during childhood. *Learning & Memory*, 25, 294–297.
- Stark, S. M., Yassa, M. A., Lacy, J. W., & Stark, C. E. L. (2013). A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia*, 51(12), 2442–2449.
- Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, 53, 1–25.
- Vieweg, P., Stangl, M., Howard, L. R., & Wolbers, T. (2015). Changes in pattern completion — a key mechanism to explain age-related recognition memory deficits. *Cortex*, 64, 343–351.
- Yassa, M. A., & Stark, C. E. (2011). Pattern separation in the hippocampus. *Trends in Neuroscience*, 34(10), 515–525.
- Wilson, I. A., Gallagher, M., Eichenbaum, H., & Tanila, H. (2006). Neurocognitive aging: prior memories hinder new hippocampal encoding. *Trends in Neuroscience*, 20(12), 662–670.
- Wynn, J. S., Ryan, J. D., & Buchsbaum, B. R. (2020). Eye movements support behavioral pattern completion. *Proceedings of the National Academy of Sciences*, 117, 11, 6246–6254.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.