

Measuring memory integration: A metric tapping memory representation rather than inference

Wangjing Yu (wy2345@columbia.edu)

Department of Psychology, Columbia University, New York, NY 10027, United States

Margaret L. Schlichting (schlichting@psych.utoronto.ca)*

Department of Psychology, University of Toronto, Toronto, ON M5S 3G3, Canada

Katherine D. Duncan (duncan@psych.utoronto.ca)*

Department of Psychology, University of Toronto, Toronto, ON M5S 3G3, Canada

* equal contributions

Abstract

Our ability to link related events could be supported either by connecting their representations in memory, or by storing them separately but integrating their content when later drawing inferences. Here, we adapted classic memory contingency analyses to develop and validate an integration index designed to tap stored representations. We conducted three pre-registered experiments adopting this metric. We found positive recall dependency for associations experienced both within the same and across different events. Compared to a conventional inference test, we found that recall dependency was more sensitive to a manipulation of memory integration. Leveraging recall dependency to investigate individual differences revealed that better memory for contextual detail was associated with faster inference judgments, consistent with high-fidelity representations of related memories—but only for people who tended to store memories separately. Our approach, thus, provides an important tool to illuminate how related events are represented in memory.

Keywords: cued recall; episodic memory; memory integration; memory interference; pattern separation

Introduction

The elements which are shared across related events can be used to adaptively bridge across these learning experiences. Precisely how these related experiences are represented in memory, though, remains an open question. Two prominent representational schemes have been proposed: integration and pattern separation. According to the integration view, while encoding new information, overlap between current and prior events increases the likelihood that related memories are reactivated. This reactivation leads to the formation of an integrated memory representation linking the experiences (Shohamy & Wagner, 2008; Schlichting et al., 2014). According to the separation view, related experiences are stored as discrete memories, which allows us to retain the idiosyncratic details of each. However, these memories can be flexibly recombined at retrieval to guide inferences (Kumaran & McClelland, 2012; Banino et al., 2016). As both the encoding- and retrieval-based mechanisms could support judgments about relationships that span experiences, we need an approach to arbitrate between these possibilities.

One common test of memory integration comes from the

associative inference paradigm. In this paradigm, participants learn two sets of overlapping pairs (AB, AC). They are then asked to infer which items (B and C) were indirectly related to each other (Schlichting et al., 2015; Zeithamova & Preston, 2010). Making such judgments has been argued to benefit from integrating both pairs into one memory during encoding (Schlichting & Preston, 2015). This mechanism has gained support from neuroimaging evidence, with greater hippocampal activation (Schlichting et al., 2014) and reactivation of B learning during AC encoding (Zeithamova et al., 2012) being associated with subsequent inferences. However, correct inferences on this test can also be reached via the pattern separation mechanism. In this case, participants would make a choice by logically recombining information from independent memories when faced with an inference decision (Banino et al., 2016; Kumaran & McClelland, 2012; Zeithamova & Preston, 2010). In support of a retrieval mechanism, it has been shown that people more readily draw inferences across events when they have strong contextual memory for each—a relationship uniquely predicted by a computational model implementing the pattern separation mechanism (Banino et al., 2016).

Given either integrative encoding or recombination at retrieval can support equivalent performance on the inference test, it remains ambiguous under which conditions each solution is adopted. Here, we aim to resolve this ambiguity by adapting a classic analytic procedure developed to assess inter-memory dependency in the modified modified free recall (MMFR) task (Barnes & Underwood, 1959). In a MMFR, rather than explicitly drawing inferences, participants are simply presented with a cue (A) that linked two studied associations (AB, AC) and asked to recall both associated items (B and C). The contingency between recalling the items is then used to infer the memories' independence or dependence (DaPolito, 1966; Kahana, 2000). For example, if the conditional probability of recalling one of the two indirectly related items given the other was also correctly recalled is approximately equal to the unconditional probability of recalling that item (i.e., $P(C|B) \approx P(C)$), the two associations are inferred to be stored and retrieved independently. Alternatively, inequalities should arise if the associations either interfere with each other

(i.e., negative contingency, $P(C|B) < P(C)$) or facilitate each other (i.e., positive dependency, $P(C|B) > P(C)$). Importantly, linking B and C via integrative encoding is expected to result in facilitation (Wahlheim & Jacoby, 2013), while employing pattern separation is expected to result in independence. Early applications of this analysis to MMFR demonstrated that recall of B and C were surprisingly *independent* (DaPolito, 1966; Wichawut & Martin, 1971; Delprato, 1972), and only a few studies have shown positive contingencies (Tulving & Watkins, 1974). However, the interpretation of these early results was clouded by the fact that recall data were pooled across participants prior to estimating contingencies, such that individual differences in memory abilities would inflate contingency estimates (Hintzman, 1972). This approach then fell out of practice as its capacity to characterize memory-to-memory relationships remained in question.

However, the approach has been recently revived by studies that estimate contingencies within participants to avoid these issues (Caplan et al., 2014; Burton et al., 2017; Horner & Burgess, 2013, 2014). These newer studies all used contingency tables—which tabulate the frequency of recalling only B, only C, both B and C, or neither—rather than conditional probabilities to estimate memory dependency. Some studies (Caplan et al., 2014; Burton et al., 2017) used these tables to calculate Yule’s Q (Yule, 1912), a rank correlation between the recall of associations. They demonstrated that while recall contingencies reflect different encoding strategies, they tend to suggest independent or interfering memories for related events. However, Yule’s Q’s high sensitivity to zeroes in contingency tables becomes all the more problematic when estimated for each participant. To overcome this limitation, Horner & Burgess (2013, 2014) developed a modeling approach that generates personalized estimates of coherent retrieval rates (both B and C, or neither) that would be expected from independent memories given each participant’s recall ability. This model is then used to detect systematic deviations from independence. This approach has shown positive dependencies when all items are studied simultaneously or when all possible pairs are separately trained, but independence when only a subset of all possible pairs are explicitly learned—as in the associative inference paradigm. Thus, despite the revival of contingency analyses in recent years, none appear to reflect the integration exhibited in neural measures.

To address this discrepancy, we first compared different estimates of retrieval contingencies in a MMFR test when three items were all studied together within the same event (experiment 1). We found that a conditional probability-based index (our cued recall integration index) was highly sensitive to the expected positive contingencies (Horner and Burgess, 2013, 2014). In experiment 2, we then compared our integration index to a conventional associative inference test to demonstrate its superior sensitivity to a manipulation of memory integration. Lastly, in experiment 3, we turned to the substantial individual differences in memory integration (Schlichting et al., 2015; Zeithamova et al., 2012; Shohamy & Wagner, 2008) by placing participants along a continuum

of memory schemes—from integration (positive contingency), separation (minimal contingency), to interference (negative contingency)—based on our index. We revealed that these individual differences were associated with the varying relationships between detailed episodic memory and inference. All three experiments were pre-registered, including their sample size and exclusion criteria (Experiment 1: <https://osf.io/xje3f>; Experiment 2: <https://osf.io/6swqk>; Experiment 3: <https://osf.io/xfp2n>).

Cued Recall Integration Index

We defined our cued recall integration index as the mean of the proactive facilitation index (FI_p) and the retroactive facilitation index (FI_r), such that dependencies in both directions ($B \rightarrow C, C \rightarrow B$) are considered.

$$FI_p = \gamma_p (P(C|B) - P(C)) \quad FI_r = \gamma_r (P(B|C) - P(B))$$

$$\text{Cued recall integration index} = (FI_p + FI_r) / 2$$

To validate the index, we first conducted simulations which independently varied the probability of recalling one of the associations (ranging from 0.1 to 0.8) and the contingency between recalling both (ranging from -1 to 1). We computed the FI and examined to what extent FI recovered our predefined recall contingency at different performance levels using linear regressions that predicted the true contingency with FI. In line with prior research (Brady et al., 2012), our FI underestimated underlying contingencies when performance was high. Therefore, we derived proactive ($\gamma_p = 1.07 + 5.62 \times P(B)^c$) and retroactive ($\gamma_r = 1.07 + 5.62 \times P(C)^c$) correction factors which restored unified linear relationships between FI and the underlying memory contingency, regardless of overall memory performance.

Experiment 1

To assess the validity of our memory integration index and to compare it with other established measures, we examined whether it could replicate the dependency between retrieving multiple associations learned within the same event (Horner & Burgess, 2013, 2014).

Participants

Twenty-five undergraduate students (17 females, mean age=19.24 years) were recruited. Data from 5 participants were excluded due to low recall (<4 recalls of either B or C on the triad cued recall test). All reported results only included data from the remaining 20 participants (13 females, mean age=19.33).

Procedures

The experiment consisted of an object triad study task and a triad cued recall test. During the object triad study task, participants viewed 18 triads (ABC) of object images, which were presented simultaneously for 7s. They were asked to memorize the triads by creating visual or verbal stories. After a 15-minute delay filled with an unrelated distractor task, participants were tested on their memory for all 18 studied object triads in a self-paced triad cued recall test. They were

shown one of the objects in each triad on top of the screen (A) and were asked to recall (type in) the names of as many associated objects as possible (B, C, or both).

Results and Discussion

We first confirmed that the cued recall integration index was normally distributed (measure of sample skewness with $|g_1| > 1$ interpreted as skewed; $g_1 = -0.665$). We then confirmed that integration index was significantly above zero, demonstrating reliable positive dependency in recall of B and C (one-sample t-test, $t(19) = 10.8$, $p < 0.001$, 95% CI [0.333, 0.493], Cohen's $d = 2.415$).

To compare our metric to established approaches, we created two 2×2 contingency tables per participant. We first mapped each participant's frequencies of successes and failures in recalling B and C to calculate Yule's Q. Yule's Q was negatively skewed ($g_1 = -1.964$); however, it also revealed a significantly positive contingency (one-sample Wilcoxon signed rank test, $Z = 3.877$, $p < 0.001$, 95% CI [0.750, 0.889]). We then created a contingency table for the independent model developed by Horner & Burgess (2013, 2014) and calculated Dependency (difference between the dependency for observed data and independent model; Ngo et al., 2019). Dependency was not skewed ($g_1 = -0.890$) and also significantly exceeded zero (one-sample t-test, $t(19) = 10.431$, $p < 0.001$, 95% CI [0.139, 0.209], Cohen's $d = 2.333$).

In sum, the results of experiment 1 replicated a significant dependency between recalling within-event associations across three measures of memory contingencies. Our normally distributed integration index was validated and achieved an equivalent effect size of Horner & Burgess's modeling approach (2013, 2014). We, thus, used our computationally simpler metric in experiments 2 and 3.

Experiment 2

Having validated our metric's sensitivity to well-established contingencies in within-event memories, we applied it to measuring contingencies between indirectly associated memories. In experiment 2, we aimed to assess how sensitive our metric was to a manipulation of memory integration: presenting some related pairs on the same and others on different task-irrelevant scene backgrounds. Although prior work has shown that increased similarity between indirectly related items (B and C) in paired-associate lists facilitates new learning (Morgan & Underwood, 1950; Barnes & Underwood, 1959), this experiment is the first to our knowledge to test whether increasing cross-event similarity via a task-irrelevant aspect of the experience affects memory integration. We predicted that increasing similarity across events would lead to a stronger reactivation of the first pair memory while participants encode the second pair, increasing the likelihood of two events becoming integrated in memory.

Participants

Seventy-four undergraduate students (45 females, mean age = 19.14) were recruited. Data from 29 participants were excluded for poor learning (<20% of pairs recalled in either

direct pair recall test). Reported results included data from the remaining 48 participants (27 females, mean age = 18.61).

Procedures

This experiment consisted of an overlapping pair learning phase and a memory integration test phase (Fig. 1A). The learning phase consisted of two object pair study tasks and two direct pair recall tests. In each study task, participants were shown 48 pairs of object images, each displayed on top of a scene picture. Participants were instructed to memorize the pairs by creating a visual or verbal story. For each object pair in the first study task (AB), there was a corresponding pair in the second study task (AC) that had one object (A) in common. The three objects form an ABC triad. In each trial, the scene picture was presented alone first for 1s, then object A was presented on the left of the scene background for 1s before object B or C appeared on the right. The pair of objects was then presented together for 2.5s. Across the two study tasks, half of the triad pairs were studied on the same scene backgrounds (i.e., high similarity condition) while the other half were studied on different backgrounds (i.e., low similarity condition). Immediately following each study task, participants performed a self-paced direct pair recall test in which they were shown one of the objects in each pair (A) and were asked to recall (type) the name of the object it was paired with in the preceding study task (B for the first test; C for the second test). All pairs in each study task were tested.

After a 15-minute delay, we tested participants' memory integration. First, in the triad cued recall test, participants were shown the object A and were asked to recall as many associated objects as possible (B, C, or both). Twenty-four triads (12 in each similarity condition)—half of the total triads—were tested. These tested triads formed the cued condition, while the untested 24 triads formed the non-cued condition. We compared the cuing conditions to examine whether performing cued recall impacted the effect of our similarity manipulation on the following indirect pair test, in which all 48 triads were tested. In the indirect pair test, participants first were told that objects from an indirect pair (BC) did not appear together in the study task but were both paired with the same object (i.e., B and C were associated via A). Participants were then shown an object image on the top of the screen (B) and were asked to choose which among three options on the bottom formed its indirect pair (C). Foils were familiar objects from the same learning condition.

Results and Discussion

Comparing cued recall integration indices across similarity conditions revealed more evidence of positive memory contingency in the high than low similarity condition (paired t-test, $t(44) = 2.165$, $p = 0.036$, 95% CI [0.008, 0.211]), consistent with greater integration when events were more similar. Separate one-sample t-tests showed that the index in the high ($t(46) = 2.102$, $p = 0.041$, 95% CI [0.003, 0.137]) but not low ($t(45) = -0.608$, $p = 0.546$, 95% CI [-0.102, 0.055]) similarity condition was significantly different from 0 (Fig. 1B), suggesting positive dependency between recall of B and

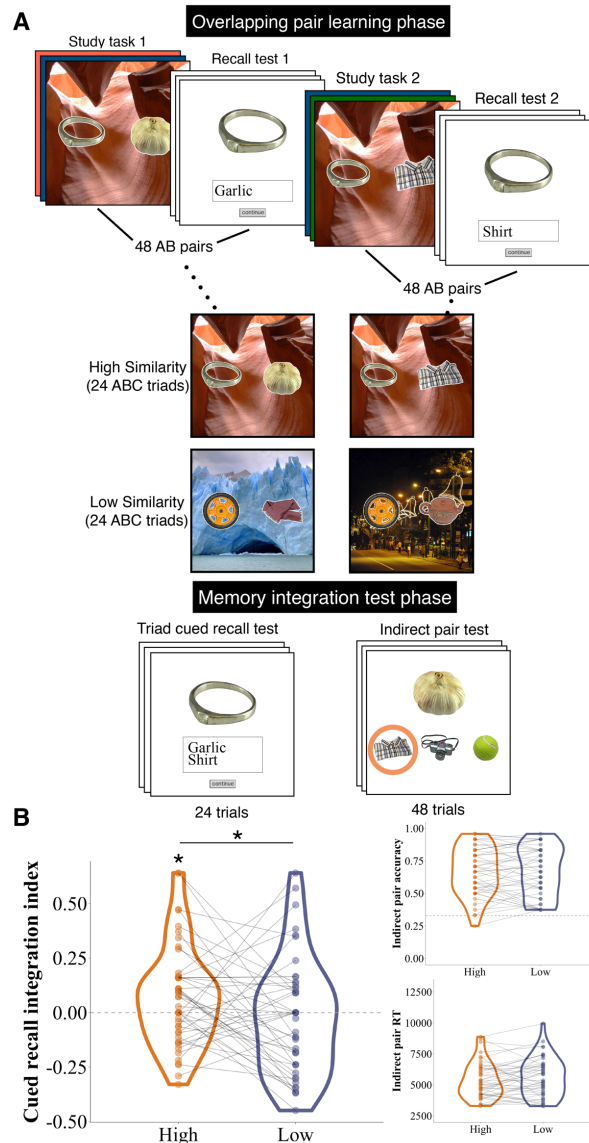


Figure 1: (A) Experiment 2 design. (B) Experiment 2 results. Left: cued recall integration index. Right: accuracy (top) and RT (bottom; outlier removed) on the indirect pair test. Each line represents one participant.

C when cross-event similarity was high.

We then assessed whether the similarity manipulation also influenced performance on the indirect pair test. Paired t-tests showed that accuracy across the two similarity conditions did not differ ($t(47)=0.24$, $p=0.81$, 95% CI [-0.040, 0.032]), but there was a trending effect of cross-event similarity on response times (RT) of the correct trials with marginally faster RTs in the high similarity condition, ($t(47)=1.73$, $p=0.089$, 95% CI [-53, 716]). However, this effect was driven by an outlier in the low similarity condition. After excluding the outlier, the difference was no longer significant ($t(46)=1.429$, $p=0.16$, 95% CI [-105, 618]; Fig. 1B).

Next, because only half of the triads tested in the indirect pair test were included in cued recall, we were able to

examine whether the effect of cross-event similarity on indirect pair test performance differed across cued and non-cued conditions. Repeated measures ANOVAs revealed no significant interaction between cross-event similarity and previous cuing on indirect pair test accuracy or RT, suggesting that the apparent lack of sensitivity of these inference measures was not driven by prior cued recall testing. Unsurprisingly, we found higher accuracy and faster RT in the cued vs. non-cued condition (accuracy: $F(1, 47)=5.608$, $p=0.022$; RT: $F(1, 47)=4.604$, $p=0.037$).

These results showed that our integration index captured the interdependency between memories of related events with higher similarity—not detectable in the indirect pair test—highlighting our metric as a more sensitive measure of memory integration.

Experiment 3

Having shown that our metric responded as expected to a manipulation of memory integration, we next asked whether individual differences in the natural tendency towards (or away from) integration as measured by our index impacts memory for the idiosyncratic details of experiences. We hypothesized that individuals with increased behavioral flexibility afforded by integrated memories might show a cost to memory for both perceptual and contextual details of the underlying memories, whereas those who store memories independently would be protected from such detail loss.

Participants

Sixty-three undergraduate students (42 females, mean age=18.6) participated in this study. Data from 15 participants were excluded for poor learning (<20% of pairs recalled in either of the direct pair recall tests). Reported analyses included data from the remaining 48 participants (33 females, mean age=18.46).

Procedures

This experiment consisted of a pair learning phase and a memory test phase (Fig. 2A). The pair learning phase was similar to the overlapping pair learning phase of experiment 2 with several modifications. Firstly, to get an unbiased measure of the individual differences in the encoding strategy, integration was not explicitly manipulated; all pairs were presented on white backgrounds. Secondly, to more accurately test memories for perceptual details that were acquired during initial learning, we used words instead of images as cues in the intervening direct pair recall tests. Object names were presented on the screen along with images during the study tasks to reduce name ambiguity. Thirdly, to account for individual differences in the general ability to remember details and isolate effects specific to associations that shared elements, we included 16 non-overlapping control pairs (XY) in addition to 32 overlapping AB/AC pairs.

Following a 15-minute delay, participants' memory for perceptual and contextual details of all studied objects was tested. In each trial, participants were first presented with two

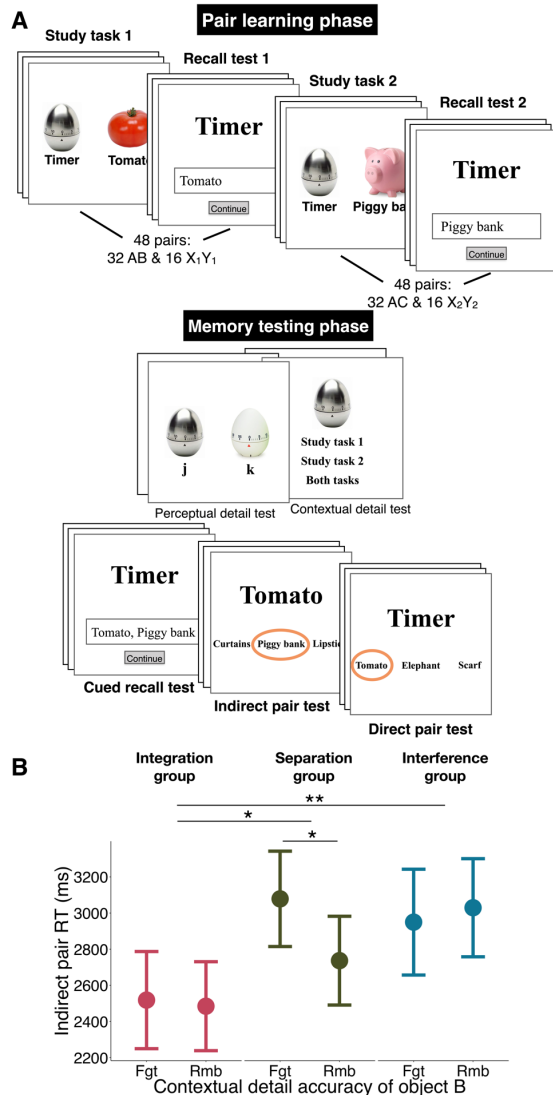


Figure 2: (A) Experiment 3 design. (B) Experiment 3 results. Predicted RT on trials of indirect pair test when object B contextual information of that triad was forgotten (Fgt) vs. remembered (Rmb).

perceptually similar objects (e.g., two apples) and were asked to choose the object they saw during the learning phase within 3s. Following each perceptual detail question, participants then indicated the study task(s) in which they studied the object (first, second, or both) within 2.5s in a contextual detail question. Participants then performed a self-paced cued recall test in which they were shown the name of either the linking object (A) or an object from a control pair (X) and were asked to recall all of the associated objects they could (B and C for ABC triads, Y for XY pairs). All ABC triads and XY pairs were tested. Following the cued recall test, participants performed the indirect pair test in which all BC pairs from the ABC triads were tested (5s response window). Participants then performed a direct pair test that used the same alternative forced-choice assessment to test all 96 directly studied associations (32 AB, 32 AC and 32 XY

pairs; 4s response window).

Results and Discussion

We found no evidence for integration or interference across all participants as measured using the integration index (one-sample t-test, $t(47)=-0.313$, $p=0.756$, 95% CI [-0.056, 0.041]), though variability in the indices was observed ($SD=0.168$). This variability, however, was not related to memory for perceptual or contextual details of overlapping pair items across participants, after adjusting for detailed memory of control pairs (multiple regression models, all $p>0.05$). In other words, contrary to our original hypothesis, the negative association between memory integration and memory for idiosyncratic details was not found across participants.

We next examined how inference performance was related to memory for details within participants. This was a conceptual replication and extension of Banino et al.'s finding that inference performance measured in the indirect pair test was positively related to memory for details (2016). Here, we separately asked whether memory for each detail type was associated with subsequent accuracy/RT on the indirect pair test, on an object by object basis. We used four sets of mixed-effect models to predict trial-wise indirect pair test accuracy/RT with perceptual/contextual detail accuracy for the corresponding triads. These analyses were restricted to triads for which participants correctly recognized both AB and AC associations during the direct pair test to focus on mechanisms specifically involved in making inferences. We compared each full model including perceptual/contextual detail accuracy for A, B, C items with a base model including only an intercept term using likelihood-ratio tests. All fixed effects (including the intercept) were included as random effects, grouped by participant.

The full model predicting RT on the indirect pair test with contextual detail accuracy better fit responses as compared to the corresponding base model ($\chi^2(6)=13.281$, $p=0.039$). This gained explanatory power was likely driven by better memory for B and C learning context being associated within-participant with faster memory inferences across the sample (object B: $b=-142$, $p=0.043$, 95% CI [-279, -5]); object C: $b=-216$, $p=0.013$, 95% CI [-382, -53]). This finding suggests that rather than trading off with memory for detail, the behavioral flexibility of memory inference was most accessible for more contextually rich memories. All models related to memory for perceptual details and models predicting indirect pair test accuracy with contextual memory did not show differences between the full and base models.

Given the positive relationship between inference performance and strong memory for individual events has been used to support the pattern separation account in prior work (Banino et al., 2016), we hypothesized that this relationship would be largest in those individuals most inclined to store memories independently as measured using our metric derived from the cued recall test. In an exploratory analysis, we divided participants into three equally sized groups based on their integration indices—the integration group with high index values (median=0.161, $SD=0.092$;

one-sample t-test vs. 0, $t(15)=7.358$, $p<0.001$), separation group with index values close to zero (median=0, $SD=0.043$; $t(15)=0.57$, $p=0.577$) and interference group with low index values (median=-0.181, $SD=0.077$; $t(15)=-10.295$, $p<0.001$).

We then demonstrated that adding the group designation and its interaction with contextual memory for all item types (A, B, C) better explained indirect pair inference RT than the simpler model that did not account for individual differences using model comparison ($\chi^2(8)=17.807$, $p=0.023$). The new model showed that the integration group made the fastest inferences (vs. separation group: $b=-380$, $p=0.030$, 95% CI [-706, -55]; vs. interference group: $b=-500$, $p=0.007$, 95% CI [-840, -164]); the RT difference between separation and interference groups was not significant ($p=0.496$). Further, consistent with our expectations, the separation group showed a stronger relationship between object B contextual memory and inference RT than did either the integration ($b=-336$, $p=0.037$, 95% CI [-646, -17]), or interference group ($b=-460$, $p=0.011$, 95% CI [-806, -106]) (Fig. 2B). The relationship between contextual memory for the other items (A & C) and inference RT was not statistically different across the 3 groups. Moreover, separate mixed models for each group revealed that only the separation group showed faster inference for triads with superior object B contextual memory ($b=-368$, $p=0.002$, 95% CI [-599, -130]). These findings align with the idea that object B memory is most likely to be modified in these paradigms, with potential reinstatement during AC encoding being an opportunity for memory modification (Zeithamova et al., 2012; Gershman et al., 2013; Hupbach et al., 2007)—here, emphasizing that such modification might take the form of enhancement.

Taken together, these results suggest that there are prominent individual differences in how people store memories of related events. We found that on average people were able to more efficiently draw memory inferences about indirectly related content when those memories contained stronger contextual information. This is broadly consistent with the idea that related experiences are stored in separate memory traces that, if rich enough, can be recombined and used flexibly at retrieval (Banino et al., 2016). However, building on prior work, the incorporation of our integration index—designed to tap the manner in which related memories are encoded and stored—adds a crucial caveat to this interpretation. Richly representing individual episodes in memory only facilitated inferences in those who tended to store memories with related content independently (one-third of our sample), whereas participants inclined toward integration or interference did not exhibit such effects. Nevertheless, people with integrated memory representation exhibited the biggest behavioral benefit among three groups.

General Discussion

Related experiences can be represented in various fashions in memory (Duncan & Schlichting, 2018), but the nature of these representations is ambiguous in simple measurements of memory performance. To address this issue, we adapted classic memory contingency analyses into a new metric.

Experiment 1 demonstrated that our integration index was sensitive to dependencies in recalling multiple elements experienced within an event, replicating findings from previous work (Horner & Burgess, 2013, 2014). In experiment 2, we used this index to assess memory integration across events. We found that our metric was more sensitive to different levels of memory integration than are more commonly used inference test. Importantly, these results provided novel evidence for positive memory contingencies when only a subset of—not all—possible associations across events were directly experienced. Experiment 3 adopted the metric to uncover how individual differences in memory integration are associated with memory for details. We found that preservation of idiosyncratic contextual details in memory facilitated the inference judgment across related events—but only in individuals who tended to store related memories separately, while overall fastest inference judgments were observed in individuals inclined towards integration. Taken together, our metric provided valuable insights into the properties of different memory schemes.

Several other approaches have been used to measure the integration of episodic elements within individual events (Horner & Burgess, 2013, 2014) and those related across events (Horner & Burgess, 2014; Burton et al., 2017; Preston et al., 2004). Importantly, our integration index achieved an equivalent effect size to the modeling approach of Horner & Burgess (2013, 2014). Experiment 1 also demonstrated that exhaustive testing which requires multiple test trials per triad adopted in prior work (Horner & Burgess, 2013, 2014) is not necessary to reliably estimate Dependency, opening the door to significantly simpler paradigms. Further, our metric exhibits sensitivity superior to the inference test, more precisely tapping memory representations. Lastly, extending on prior work using Yule's Q (Burton et al., 2017), we show that individual differences in memory contingencies are meaningfully related to other aspects of memory performance, not just subjective reports of encoding strategies.

It is worth noting that our metric has some limitations. In particular, our approach aggregates across memories within participants to estimate their personalized memory scheme. Therefore, someone who shows positive contingencies for some memories and negative for others would on average be placed around the separation region of our continuum—a shortcoming shared by all memory contingency analyses. Our results, however, show that this limitation can be overcome in practice: aggregating across trials can still reveal how manipulations bias the tendency to adopt particular memory schemes (experiment 2); and individual differences in this metric nevertheless can show meaningful relationships with other aspects of memory (experiment 3).

In sum, our integration index is a useful behavioral measure for assessing inter-memory relationships that could be applied across a wide range of studies. It presents an important tool that contributes to the emerging understanding of how related information is represented in memory.

Acknowledgements

We thank Jaden E. Murray for assistance with data collection and everyone in the Budding Minds Lab and the Duncan Memory Lab for insightful discussion. This work was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN-04933-2018 (MLS), RGPIN-05582-2016 (KDD)), Canada Foundation for Innovation (JELF, Project #34479 (MLS), 36876 (KDD)), University of Toronto startup funds (MLS), Canadian Institutes of Health Research (CIHR, PJT-159492).

Reference

- Banino, A., Koster, R., Hassabis, D., & Kumaran, D. (2016). Retrieval-Based Model Accounts for Striking Profile of Episodic Memory and Generalization. *Scientific Reports*, 6(August), 1–15.
- Barnes, J. M., & Underwood, B. J. (1959). “Fate” of first-list associations in transfer theory. *Journal of Experimental Psychology*, 58(2), 97–105.
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2013). Real-world objects are not represented as bound units: independent forgetting of different object details from visual memory. *Journal of Experimental Psychology: General*, 142(3), 791.
- Burton, R. L., Lek, I., & Caplan, J. B. (2017). Associative independence revisited: competition between conflicting associations can be resolved or even reversed in one trial. *Quarterly Journal of Experimental Psychology*, 70(4), 832–857.
- Caplan, J. B., Rehani, M., & Andrews, J. C. (2014). Associations compete directly in memory. *The Quarterly Journal of Experimental Psychology*, 67(5), 955–978.
- DaPolito, F. J. (1967). Proactive effects with independent retrieval of competing responses. *Dissertation Abstracts International*, 27(7-B), 2522–2523.
- Delprato, D. J. (1972). Pair-specific effects in retroactive inhibition. *Journal of Verbal Learning and Verbal Behavior*, 11(5), 566–572.
- Duncan, K. D., & Schlichting, M. L. (2018). Hippocampal representations as a function of time, subregion, and brain state. *Neurobiology of Learning and Memory*, 153, 40–56.
- Gershman, S. J., Schapiro, A. C., Hupbach, A., & Norman, K. A. (2013). Neural context reinstatement predicts memory misattribution. *The Journal of Neuroscience*, 33(20), 8590–8595.
- Hintzman, D. L. (1972). On testing the independence of associations. *Psychological Review*, 79(3), 261–264.
- 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(May).
- 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(9), 988–992.
- Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning & Memory*, 14, 47–53.
- Kahana, M. J. (2000). Contingency analyses of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (p. 59–72). Oxford University Press.
- Kumaran, D., & McClelland, J. L. (2012). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, 119(3), 573–616.
- Morgan, R. L., & Underwood, B. J. (1950). Proactive inhibition as a function of response similarity. *Journal of Experimental Psychology*, 40(5), 592–603.
- Ngo, C. T., Horner, A. J., Newcombe, N. S., & Olson, I. R. (2019). Development of holistic episodic recollection. *Psychological science*, 30(12), 1696–1706.
- Preston, A. R., Shrager, Y., Dudukovic, N. M., & Gabrieli, J. D. E. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus*, 14(2), 148–152.
- Schlichting, M. L., Zeithamova, D., & Preston, A. R. (2014). CA subfield contributions to memory integration and inference. *Hippocampus*, 24(12), 1248–1260.
- Schlichting, Margaret 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, 1–10.
- Schlichting, Margaret L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, 1, 1–8.
- Shohamy, D., & Wagner, A. D. (2008). Integrating Memories in the Human Brain: Hippocampal-Midbrain Encoding of Overlapping Events. *Neuron*, 60(2), 378–389.
- Tulving, E., & Watkins, M. J. (1974). On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning and Verbal Behavior*, 13(2), 181–193.
- Wahlheim, C. N., & Jacoby, L. L. (2013). Remembering change: The critical role of recursive reminders in proactive effects of memory. *Memory & Cognition*, 41(1), 1–15.
- Wichawut, C., & Martin, E. (1971). Independence of A-B and A-C associations in retroaction. *Journal of Verbal Learning and Verbal Behavior*, 10(3), 316–321.
- Yule, G. U. (1912). On the Methods of Measuring Association Between Two Attributes. *Journal of the Royal Statistical Society*.
- Zeithamova, D., & Preston, A. R. (2010). Flexible Memories: Differential Roles for Medial Temporal Lobe and Prefrontal Cortex in Cross-Episode Binding. *Journal of Neuroscience*, 30(44), 14676–14684.
- Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron*, 75(1), 168–179.