Interleaving facilitates the rapid formation of distributed representations

Zhenglong Zhou (zzhou34@sas.upenn.edu)
Marlie Tandoc (tandoc@sas.upenn.edu)
Dhairyya Singh (dsin@sas.upenn.edu)
Anna Schapiro (aschapir@sas.upenn.edu)
Department of Psychology, University of Pennsylvania, Philadelphia, PA, 19104

Abstract

Distributed representations, in which information is encoded in overlapping populations of neuronal units, are essential to the remarkable success of artificial neural networks (ANNs) in many domains, and have been posited to be employed throughout the brain, especially in neocortex. A fundamental signature of ANNs employing distributed representations is that learning requires exposure to information in an interleaved order; exposure to new information in a blocked order tends to overwrite prior knowledge (i.e., 'catastrophic interference'). Because it is difficult to match human learning to the learning conditions of these networks, it is not known whether human learning exhibits these properties, which, if true, would implicate use of similar representations. To test this, we leveraged a recent proposal that parts of the hippocampus host distributed representations of the kind typically ascribed to neocortex, and adopted a hippocampally dependent task that contrasts the effects of interleaved versus blocked learning on a short timescale. Experiments 1a and 1b demonstrate that interleaved exposure facilitates the rapid perception of shared structure across items. Experiment 2 shows that only interleaved exposure permits useful inference when item associations need to be inferred based on statistical regularities. Together, these results demonstrate the power of interleaved learning and implicate the use of distributed representations in human rapid learning of structured information.

Keywords: associative inference; catastrophic interference; hippocampus; neural network models

Introduction

A fundamental dichotomy in the nature of neural representation is that between distributed and localist representations: entities can be processed using distributed, overlapping populations of neuronal units that reflect their shared structure, or localist representations that orthogonalize activity patterns despite any similarity across items. Distributed representations have been critical to the successes of artificial neural networks (ANNs), promoting efficient learning in many domains (LeCun et al., 2015). However, apart from this efficiency during learning, it is difficult to characterize the unique advantages of distributed over localist representations. Both kinds of representations can support generalization, for example: indirect item associations can be inferred from overlapping features encoded in distributed representations, or through recirculation of activity amongst localist representations (Kumaran & McClelland, 2012). In this work, we ask: What conditions promote the formation of distributed representations in humans, and what advantages do these representations confer for learning?

We explore these questions in humans through a behavioral lens, enabled by a fundamental behavioral signature of ANNs employing distributed representations: Learning useful distributed representations requires exposure to entities in an interleaved order (McClelland et al., 1995). Interleaving allows the learning algorithm to uncover the shared structure across the exposed information. If networks are exposed to one set of information entirely before a second related set, in a blocked fashion, the second set tends to overwrite knowledge of the first ('catastrophic interference'; McCloskey & Cohen, 1989).

There are two basic challenges to testing these ideas in humans (or other animals). First, neocortex is typically considered to be the part of the brain likely to employ distributed representations akin to those used by ANNs (e.g., Yamins et al., 2014). But neocortex generally learns complex novel information quite slowly, on the order of days at the fastest but sometimes months or years (McClelland et al., 1995). Testing the ideas in neocortex would thus involve difficult multisession experiments.

The second challenge is that it is difficult to separate trial-by-trial attentional effects, which are not at play in basic ANN learning, from the effects of interleaving versus blocking described above, which are fundamental to ANN learning. Indeed, there is a large literature on interleaving versus blocking in category learning, in which trial-by-trial attentional mechanisms play an important role (Carvalho & Goldstone, 2015). In order to test for ANN-like learning, we need to avoid these attentional effects.

We overcome the first challenge by leveraging our recent model proposing that a subfield of the hippocampus hosts distributed representations of the kind typically ascribed to neocortex, but that it can learn much more quickly – on the timescale of minutes to hours (Schapiro et al., 2017). The hippocampus is typically thought to use orthogonalized representations, but it is specifically subfields dentate gyrus (DG) and CA3 that employ these 'pattern-separated' representations, whereas CA1 appears to have more cortex-like properties. We can thus assess the use of overlapping representations in fast timescale learning by using a hippocampally dependent task.

To address the second challenge, we adopted a task with many separate simple associations to learn, such that attention to certain features in adjacent trials would neither benefit nor harm learning. The task is an associative inference task, known to depend on the hippocampus (Bunsey & Eichenbaum, 1996), in which participants learned to associate object pairs AB and BC and were later tested on the transitive AC relationship. During learning, related pairs could either be interleaved, appearing in alternating order amongst other pairs, or blocked, with ABs only appearing in the first half of

learning and BCs only in the second. Each participant learned some triads that appeared in interleaved order and others in blocked order.

There have been several proposals as to how different representations in the hippocampus might support AC inference in this task. One strategy, which we will call the distributed strategy, encodes A and C using overlapping populations of neurons, as a result of both being experienced with B (Schapiro et al., 2017). This representation supports an automatic, direct association between A and C at test, as these items have come to be represented in a directly overlapping way. We hypothesized that interleaved presentation of the direct pairs would afford use of this strategy. Another strategy, which we will call the localist strategy, proposed by Kumaran & McClelland (2012) in the REMERGE model, involves keeping memories of each pair separate during learning by encoding orthogonal conjunctive representations of AB and BC (as would be expected in subfields DG and CA3). Then, at retrieval, spreading activation across recurrent connections links from A to C via the separately encoded AB and BC. This strategy, at least as currently implemented, does not perform differently as a function of interleaved versus blocked exposure.

The distributed strategy promotes implicit, automatic association between A and C at test, whereas the localist strategy requires additional, and perhaps more explicit, processing. Our model claims that both of these strategies are available to the hippocampus, implemented in separate pathways (Schapiro et al., 2017). However, the model predicts that interleaving will afford use of the distributed strategy, as there will be more robust representational overlap between related AB and BC after interleaved than blocked learning. We therefore predict that participants should exhibit better performance in a rapid, implicit test of indirect associations in the interleaved condition than in the blocked condition. In a scenario where there is plenty of time for recurrent processing, there should be no difference. In Experiments 1a and 1b we tested these ideas in a paradigm in which direct associations were explicitly taught to participants. In Experiment 2, we embedded the direct associations in a continuous stream, a scenario where we predicted a qualitatively stronger difference between the more explicit and implicit strategies.

Experiments 1a-b

Does interleaved learning facilitate representations that support rapid AC associations? To address this question, in our first two experiments, participants explicitly learned AB and BC associations that were either interleaved or blocked (Fig. 1a). At test, participants completed two tasks. During the speeded recognition task, participants quickly judged whether two objects on the screen were shown as a pair during learning (Fig. 1b). We hypothesized that for objects that were never seen directly together but were associated transitively (AC), greater representational overlap would make it more difficult to correctly reject them as not paired together. Such

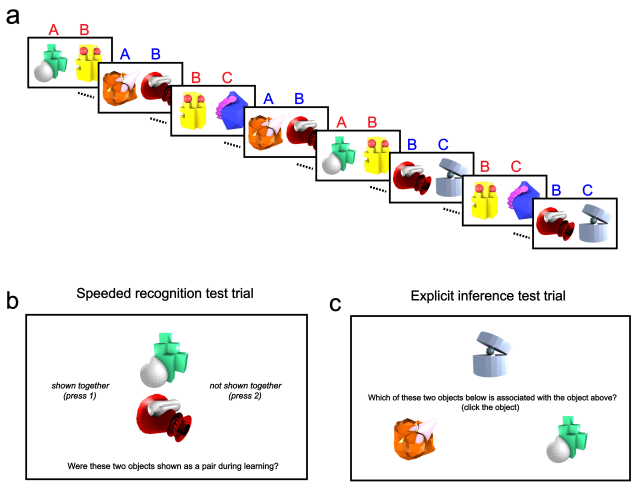

Figure 1: Exps 1a-b design. (**a**) Participants saw related ABC triads in either an interleaved (red) or a blocked (blue) order during learning. (**b**) Speeded recognition post-test. (**c**) Explicit inference post-test.

difficulty should be manifested in a slower response when indicating that they had not been seen as a pair, or a higher false alarm rate in identifying them as a presented pair. Thus, we predicted slower reaction time (RT) and higher false alarm rates for interleaved than blocked ACs, and than foil pairs that consisted of matched unrelated pairs.

Methods

Participants In Experiment 1a, we recruited 33 participants (mean age = 34.75, SD = 9.57), with 26 participants remaining after exclusions. Experiment 1b was a pre-registered study (https://osf.io/ag42z/) for which 83 participants were recruited, with 54 participants remaining after exclusions. For RT analyses in both experiments, we excluded participants who missed all trials of a condition of interest (e.g., participants who missed all interleaved AC trials were excluded from analyses of RT differences between interleaved and blocked AC trials). After looking at the collected data, we realized that a preregistered criterion to exclude subjects based on ability to indicate that target pairs were old would not exclude participants who indicated that all pairs were old, so we switched to using a criterion based on d-prime (d-prime for speeded recognition had to be higher than 1.5). We also dropped our age exclusion, as we found that behavior was matched across age groups. The results are qualitatively similar using the original exclusion plan. Participants were recruited through Amazon Mechanical Turk. The study protocol was approved by the University of Pennsylvania Institutional Review Board.

Design and procedure During learning, each participant was shown a sequence consisting of presentations of 12 pairs of novel visual objects randomly sampled from 36 artificial object images (Schlichting et al., 2015) for each subject. Each

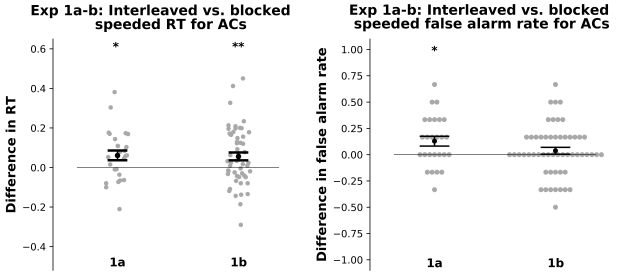

Figure 2: RTs and false alarm rates for interleaved — blocked AC trials during speeded recognition in Exps 1a and 1b. Each gray dot is an individual participant, with the mean across participants in black. Error bars represent +/- 1 SEM. *p<0.05; **p<0.01; ***p<0.001.

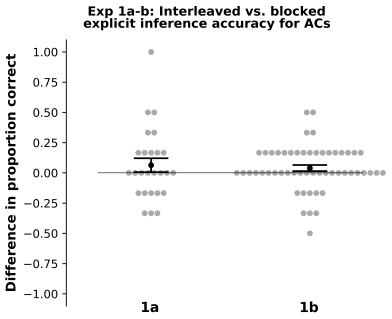

Figure 3: Accuracy differences between interleaved and blocked AC trials during explicit inference in Exps 1a and 1b.

"direct" pair AB was uniquely related to another direct pair BC through a linking item B. Each direct pair was shown 30 times. Among the six ABC triads shown to each participant, three were interleaved (e.g., $A_{in_3}B_{in_3}$ and $B_{in_3}C_{in_3}$ appear in alternation) and three were blocked (e.g., all presentations of $A_{bl_2}B_{bl_2}$ occur before the first presentation of $B_{bl_2}C_{bl_2}$), with the constraint that pairs that share an object were never shown back-to-back. Participants were instructed to remember the pairings of objects by creating quick narratives about how the two objects might interact. Each participant completed a total of 360 trials, with the two objects displayed horizontally side-by-side on the screen for 1 s. Across repeated presentations of an object pair AB, A appeared on the left side or on the right side of the screen randomly. Subsequent to each pair presentation, participants saw the question "on a scale of 1 (failed to visualize a story) to 5 (clearly visualized a story), how well were you able to visualize a story linking the objects?" and responded with a numerical rating by pressing the corresponding key within 7 s.

After learning, participants completed two tasks that probed their memory of learned object associations: a speeded recognition task (Fig. 1b) followed by an explicit inference task (Fig. 1c). During the speeded recognition task, on each trial, two objects were displayed for 1500 ms and participants were asked to respond within 3500 ms with a button press to indicate whether the two objects had been shown as a pair during the learning phase. The task had 24 trials with paired objects (e.g., A_1B_1, B_3C_3) which we refer to as direct trials, 12 trials with unrelated objects (e.g., A_1C_3) referred to as foil trials, and 12 trials with indirectly related objects (e.g., A_1C_1) which we refer to as AC trials. Each object pair appeared in two trials, with two different vertical positions (A above C or C above A), with the constraint that pairs could not be repeated on back-to-back trials. During the explicit inference task, on each trial, participants saw a cue object at the top of the screen (e.g., C_1) and were instructed to select which of two objects (e.g., A_1 and A_3) shown below the cue object was indirectly related to it within 7 s. The explicit inference task consisted of 12 trials whereby each AC association was tested twice, such that A_i served as the cue object in one trial and as the target object in the other. The target object in each trial was paired with a matched foil, such that a C_i target would be paired with a C_j foil (i.e., a C object from a different ABC triad).

Analysis Across all studies, our four main variables of interest were (a) RT differences between interleaved AC trials and foil trials, (b) RT differences between interleaved AC trials and blocked AC trials, (c) false alarm rate differences between interleaved AC trials and blocked AC trials during the speeded recognition task, and (d) accuracy differences between interleaved AC trials and blocked AC trials during the explicit inference task.

Across all experiments, RTs for trials on which participants responded correctly were log-transformed before averaging. For each variable of interest, we performed paired t-tests to test the significance of within-subject differences between two conditions of interest.

Experiment 1a Results

Analyses revealed significant (a) RT differences between speeded interleaved AC and foil trials ($t(25)=5.67$, $p<0.00001$; Fig 2), (b) RT differences between speeded interleaved AC and blocked AC trials ($t(25)=2.46$, $p=0.021$), and (c) false alarm rate differences between speeded interleaved AC and blocked AC trials ($t(25)=2.76$, $p=0.01$), and identified non-significant (d) within-subject accuracy differences between explicit interleaved AC and blocked AC trials ($t(25)=1.11$, $p=0.28$). Furthermore, false alarm rate was significantly higher for interleaved ACs than for matched foils ($t(25)=3.27$, $p=0.0031$), and was not significantly different between blocked AC and matched foil trials ($t(25)=1.09$, $p=0.29$).

Additionally, we observed that participants demonstrated lower accuracy for blocked AB pairs (mean = 0.83) than for blocked BC pairs (mean = 0.97) during the speeded recognition task ($t(25)=-3.93$, $p=0.00059$), suggesting some retroactive interference or other forgetting. Overall accuracy for interleaved direct (i.e., AB or BC) pairs (mean = 0.92) was not significantly different than accuracy on blocked direct pairs (mean = 0.90, $t(25)=0.99$, $p=0.33$). Explicit inference performance was above chance for both interleaved ACs (mean

= 0.85, SD = 0.14, t(25)=12.7, p<0.00001) and blocked ACs (mean = 0.79, SD = 0.27, t(25)=5.41, p=0.000013). Mean reaction time was not significantly different between interleaved and blocked trials in the explicit inference task (t(24)=0.46, p=0.65).

Experiment 1a Discussion

The results suggest, as predicted, that representations that support rapid association between A and C at test benefitted from interleaved exposure, as participants demonstrated slower responses to and higher false alarm rates for interleaved ACs than for blocked AC pairs and foil pairs during the speeded recognition. The demonstrated benefit of interleaved exposure for rapid recognition of AC did not translate into an advantage for interleaved ACs during the explicit inference task, consistent with our hypothesis that explicit AC inference can be solved using a strategy compatible with blocked presentation.

Experiment 1b Results

To increase confidence in the results from Experiment 1a, we preregistered and replicated the same study. Exp 1b replicated the main effects observed in 1a, including significantly slower responses to interleaved ACs than to blocked ACs (t(52)=2.82, p=0.011) and to foils (t(52)=4.72, p=0.000018) during speeded recognition, and non-significant accuracy (t(53)=1.60, p=0.1152) and RT (t(53)=-0.047, p=0.96) differences between interleaved AC trials and blocked AC trials during explicit inference. Explicit inference performance was above chance for both interleaved ACs (group = 0.85, SD = 0.18, t(53)=14.26, p<0.00001) and blocked ACs (mean = 0.81, SD = 0.18, t(53)=12.62, p<0.00001). We did not replicate the difference in false alarm rates between interleaved and blocked AC trials during speeded recognition (t(53)=1.14, p=0.26). False alarm rate was significantly higher for interleaved ACs than for matched foils (t(53)=4.56, p=0.000031), and for blocked AC than for matched foil trials (t(53)=3.50, p=0.00095).

Experiment 1b Discussion

Together, Exps 1a and 1b provide strong evidence that interleaving facilitates the formation of representations that support automatic, implicit association of indirectly related items.

Experiment 2

Exps 1a and 1b present evidence that interleaved learning facilitates representations that support rapid association of items based on shared features, yet failed to identify a difference between interleaved and blocked learning for inferring object associations in a more explicit scenario where time is abundant. We hypothesized that this is due to the availability of an alternative strategy in this case (i.e., AC association through recurrent computation as proposed in the REMERGE model; Kumaran & McClelland, 2012). This motivates us to ask the following question: Are there certain kinds of learning problems where interleaved exposure to information is required for successful behavior even in explicit settings? One domain where a strategy like that employed by REMERGE is likely to struggle is when the direct items to be associated are not clearly demarcated during encoding. For instance, in statistical learning paradigms, objects are presented one at a time in a continuous stream with embedded temporal regularities (Saffran, 1996). In such paradigms, an object will mostly appear with certain temporal associates but can also appear with other objects with lower frequency. A strategy like that employed by REMERGE, which quickly forms robust conjunctive representations of every observed temporal co-occurrence, would encode both reliable pairings of frequently co-occurring pairs, and unreliable, infrequently co-occuring ones. Inference through spreading activation using this strategy would activate both reliable and unreliable pairings. We therefore hypothesized that inference after blocked exposure would be difficult in a statistical learning version of our paradigm, even in the slow explicit inference test. Thus, we predicted this paradigm would reveal an advantage of inference over interleaved associations in the slow explicit inference test. In Experiment 2, we exposed participants to a sequence of objects in which object pairings were defined by the relative frequency of consecutive occurrence.

Methods

Participants In Experiment 2, 104 participants (mean age = 37.74, SD = 11.58) were recruited through Amazon Mechanical Turk, resulting in 43 participants after excluding participants who responded incorrectly to one third or more of the speeded recognition direct pair trials, or who missed more than half of the cover task responses during learning.

Design and procedure For each participant, a sequence of visual object pairs was generated following the same protocol as Exp 1a-b, except that each pair repeated 24 times. During learning, however, each participant saw objects presented one at a time in a continuous stream with no breaks, such that two objects from the same pair were shown consecutively followed by objects from a different pair. For each occurrence of an object pair AB (or BC), the order of the objects was randomized. Therefore, for an object pair AB, A is both the item that precedes and the item that follows B with the highest frequency among all objects (except C, which is equal). As a cover task, participants were instructed to quickly respond as to whether the current object appeared heavier than the previous object. Participants pressed one key if the current object seemed heavier than the preceding object, and a different key if not. Each object was displayed on the screen for 1 s followed by an inter-trial interval of 500 ms.

After learning, participants were informed that there were object pairs embedded within the sequence they saw. Participants then completed the speeded recognition task followed by the explicit inference task using the same procedure specified in Exps 1a-b.

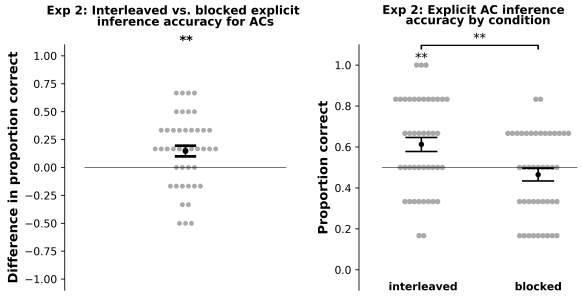

Figure 4: Explicit inference accuracy differences between interleaved and blocked AC trials (left) and explicit inference accuracy by condition (right) in Exp 2

Experiment 2 Results

As in Exp 1a-b, during the speeded recognition task, subjects demonstrated slower responses (t(37)=2.54, p=0.016) and more false alarms (t(42)=5.49, p<0.00001) to interleaved ACs than to blocked ACs, and showed marginally slower responses to interleaved ACs than to foils (t(37)=1.90, p=0.065). There was no significant difference between accuracy of interleaved and blocked direct trials (t(42)=0.81, p=0.42). False alarm rate was not significantly different between interleaved AC and matched foil trials (t(42)=1.22, p=0.23), or between blocked AC and matched foil trials (t(42)=0.18, p=0.86).

Unlike Exp 1a-b, however, we observed higher accuracy for interleaved AC pairs (mean = 0.61, SD = 0.23) than for blocked AC pairs (mean = 0.47, SD = 0.21) during the explicit inference task (t(42)=3.06, p=0.0039). Performance was above chance in the interleaved condition (t(42)=3.25, p=0.0023) but not in the blocked condition (t(42)=-1.115, p=0.27). There was no significant difference between RTs for interleaved and blocked trials (t(42)=-1.12, p=0.27).

Experiment 2 Discussion

These results suggest that the benefit of interleaved learning for rapid AC association is preserved in this setting where object pairings need to be inferred based on statistics of object occurrence, though responses to foils in the speeded recognition task were stronger than prior experiments. Unlike the previous experiments, participants' ability to explicitly infer associations between blocked AC pairs was impaired relative to interleaved AC pairs. Therefore, Exp 2 demonstrates that interleaved learning benefits even slow, explicit inference in a scenario where the basic object associations (i.e., AB, BC) need to be inferred based on statistical regularities.

General Discussion

A hallmark of successful ANNs is the processing of inputs using overlapping populations of neuronal units, as opposed to localist, orthogonalized representations. Distributed representations have been posited to be employed throughout the brain, especially in neocortex (Yamins et al., 2014). In ANNs, interleaved exposure to information is indispensable for building distributed representations that support adaptive behavior, with learning in a blocked setting tending to induce the catastrophic forgetting of previously learned information. Does interleaved learning similarly facilitate distributed representations that support generalization in the brain? What advantages do such representations confer for human learning?

We recently proposed that the CA1 subfield of the hippocampus hosts distributed representations of the kind typically ascribed to neocortex (Schapiro et al., 2017), which allows for the possibility of contrasting the effects of interleaved and blocked learning on a short timescale. To this end, we adopted a hippocampally dependent task that contrasts participants' ability to infer item associations after interleaved or blocked learning. We hypothesized that the benefit of interleaved learning for overlapping representations would be revealed during rapid association of items based on their common features, since inference through retrieval-based strategies that operate on pattern-separated representations require extra time and computation. In a scenario with plenty of time for this computation, we thus predicted no difference between conditions. Indeed, a previous study found no behavioral difference between interleaved versus blocked learning conditions when participants were asked to deliberately infer associations between related items (Schlichting et al., 2015). We additionally hypothesized that interleaved exposure would promote a more qualitative benefit when item associations need to be inferred based on statistical regularities, as pattern separated representations should be less sensitive to graded feature co-occurrence frequencies.

As predicted, we found no difference between interleaved and blocked ACs during the explicit inference task in Exps 1a and 1b, but a significant slowing in the response to interleaved ACs relative to blocked ACs and foils. In Exp 2, participants saw objects presented one-at-a-time with each object always co-occurring most frequently with the object from the same pair. Consistent with our prediction that this protocol would produce a qualitatively stronger benefit for interleaved learning, participants displayed better performance on interleaved ACs than blocked ACs during explicit inference. Together, the results implicate the rapid formation of overlapping representations of related items through interleaved learning, and demonstrate the particular advantage of this kind of representation in a statistical learning setting.

There is extensive prior work on the effects of interleaved and blocked exposure on learning categories of multi-dimensional stimuli. Studies have found differential effects of interleaved and blocked exposure driven by trial-by-trial attentional effects: interleaved exposure facilitates noticing differences between categories, whereas blocked exposure facilitates learning commonalities (Carvalho & Goldstone, 2015). Our present study differs from these category learning tasks in that attention to specific features in adjacent trials neither benefits nor harms learning, as there are no shared features across triads. Therefore, the observed benefits of interleaved

exposure in the present study are likely due to different underlying mechanisms.

One recent category learning paper found better performance under blocked than interleaved conditions, which the authors argued was supported by more 'factorized' representations that may be difficult to explain from an attentional account (Flesch et al., 2018). They interpret the results as evidence that neocortex may not be as susceptible to catastrophic interference as ANNs would predict. An alternative interpretation, consistent with the current framework, is that orthogonalized representations within areas DG and CA3 of the hippocampus learn factorized representations and thus can support behavior in scenarios where it is not advantageous to integrate across conditions, as was the case in that study.

In addition to the distributed and localist strategies considered above, another influential proposal for how the hippocampus may carry out associative inference is known as 'integrative encoding' (Shohamy & Wagner, 2008). Integrative encoding posits that after studying AB and upon study of BC, observing B triggers reinstatement of the AB memory through pattern completion mechanisms, and an overlapping representation of AB and BC is then encoded. Although this strategy, similar to our account, employs overlapping AC representations formed during encoding to support inference, it relies on the episodic encoding and pattern completion mechanisms of DG and CA3. It is not clear whether integrative encoding predicts an advantage for blocked or interleaved presentation: blocking may lead to strong AB memory that permits better reinstatement during BC learning, or temporal proximity between related AB and BC presentations during interleaving may help promote the formation of an overlapping AB-BC representation (Schlichting et al., 2015). Future work using computational models that implement integrative encoding could fruitfully explore these possibilities and their relationship to our predictions and findings.

Taken together, our results suggest that interleaved exposure promotes the formation of overlapping representations of associated entities. These representations support rapid, implicit inference, as well as graded sensitivity to co-occurrence statistics. These properties are suggestive of a distributed neural code and mirror the properties of ANNs using distributed representations. We thus take the findings as evidence that the brain, and in this case the hippocampus in particular, may benefit from distributed representations in a similar way to these models.

References

Bunsey, M., & Eichenbaum, H. (1996). Conservation of hippocampal memory function in rats and humans. *Nature*, *379(6562)*, 255-257.

Carvalho, P. F., & Goldstone, R. L. (2015). What you learn is more than what you see: what can sequencing effects tell us about inductive category learning? *Frontiers in Psychology*, *6*, 505.

Flesch, T., Balaguer, J., Dekker, R., Nili, H., & Summerfield, C. (2018). Comparing continual task learning in minds and machines. *Proceedings of the National Academy of Sciences*, *115(44)*, E10313-E10322.

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.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521(7553)*, 436-444.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102(3)*, 419.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. *Psychology of Learning and Motivation*, *24*, 109-165.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274(5294)*, 1926-1928.

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372(1711)*, 20160049.

Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, *6(1)*, 1-10.

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron*, *60(2)*, 378-389.

Yamins, D. L., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences*, *111(23)*, 8619-8624.