The Role of Attention in Three-way binding in Episodic Memory

Hyungwook Yim (yim.31@osu.edu)
Department of Psychology & Center for Cognitive Science
The Ohio State University
209C Ohio Stadium East, 1961 Tuttle Park Place
Columbus, OH 43210, USA

Simon J. Dennis (dennis.210@osu.edu)
Department of Psychology & Center for Cognitive Science
The Ohio State University
200E Lazenby Hall, 1827 Neil Avenue
Columbus, OH 43210 USA

Vladimir M. Sloutsky (sloutsky.1@osu.edu)
Department of Psychology & Center for Cognitive Science
The Ohio State University
208C Ohio Stadium East, 1961 Tuttle Park Place
Columbus, OH 43210, USA

Abstract

The current study examines the role of attention in forming complex binding structures in episodic memory. Previous research (Yim, Dennis, & Sloutsky, 2011a, 2011b) indicated that three-way bindings can be formed within an explicit memory task. Here we attempt to reduce explicit attending by presenting participants with a variant of statistical learning. The paradigm was modified to accommodate the constraints in two list learning paradigms ABABr, and ABCD. Only the ABABr paradigm requires a three-way binding in order to perform above chance. Evidence of learning was derived from learning curves, accuracies and reaction time at test. Results show evidence of learning for the ABCD condition but not for the ABABr condition. This finding indicates that whereas the ABCD structure can be learned implicitly, learning of ABABr lists depends on attention and requires explicit learning.

Keywords: episodic memory, attention, three-way binding, context use, statistical learning

Introduction

One of the memory tasks that we confront daily is finding where we parked our car at work. Even though the only thing that one need remember seems to be the parking spot, one's car, and the link between the two, we often fail. One of the reasons that we are sometimes unable to retrieve this memory is because our previous memories of parking the car interfere with today's memory. Therefore in order to correctly retrieve today's event, we also need to remember the link between the context (i.e. today) and the event (i.e. parking the car in a specific spot).

As depicted in the above example, one's success in retrieving an episode depends on what was remembered, and how it was structured when it was remembered. Depending on the episode that one has to remember, there could be differences in the number of components that have to be remembered, and the structure that has to be formed. The process by which one forms a memory representation containing multiple components is called *binding* (Cohen & Eichenbaum, 1993; Schacter & Tulving, 1994). Along with the ability to store the bounded memory components, it has been argued that control processes such as memory strategies and metacognitive operations also play a central role in forming episodic memory (Ghetti & Lee, 2011). However, how binding occurs and the nature of the structure it forms are still not clear.

Yim, Dennis, & Sloutsky (2011a, 2011b) proposed that a key determinant of performance during episodic memory formation is simultaneously attending to the components that should be remembered (e.g. events, context). This attentional mechanism was argued to be used during encoding and/or retrieval especially when the binding structure is complex. To support the argument, they presented a list learning paradigm to different age groups (i.e. 4-year olds, 7-year olds, and adults). The participants had to remember two lists each consisting of six pairs of pictures. The two lists were separated by a retention interval and the pairs were presented one at a time. At the end of studying the two lists a cued recall test was given with a context cue and an item cue (see Figure 1-(a) left slide for an example of a trial which has a context cue and two items). The main manipulation was the structure of the lists which varied the complexity of their required binding structure.

In an ABABr condition two lists had an identical picture set with the only difference between the two lists being how they were paired (Porter & Duncan, 1953; Postman, 1964, see Figure 1-(b) left). For example, in the first list there would be two pairs, [apple]-[dog] and [chair]-[car], and in the second list there would be two different pairs, [apple]-[car] and [chair]-[dog], which is a re-arrangement of the first two pairs. It has been logically illustrated by Humphreys, Bain, & Pike (1989) that to correctly answer a given cued recall test (e.g. what was paired with apple in list1?) one must have formed, at the minimum, a *three-way binding* structure that includes the two items [apple] and

[dog], and the context [list1] together. Suppose that one formed a simpler binding structure such as a two-way binding structure (e.g. [apple] – [dog] or [apple] – [car]). In this case when cued with "apple" and "list1" in a cued recall test (e.g. what was paired with apple in list1?), the [apple] will not only elicit [dog] in list1 but also [car] in list2, which would make one's recall ambiguous. However, if one formed a three-way binding structure, [apple] and [list1] will act as a compound cue, and will elicit the correct answer [dog].

In an ABAC condition (see Figure 1-(b) middle), the cues were identical between the two lists whereas the targets were different. In this condition the minimal binding structure that is required to remember the episode correctly are known to be *two two-way binding* structures (Barnes & Underwood, 1959; Postman, 1962). For example, one should have formed at least the binding between the two items (e.g. [apple] – [dog]) and the binding between an item and the context ([list1] – [dog]) to correctly answer at test (e.g. what was paired with apple in list1?). If one had only formed a single two-way binding structure in this case, it would be hard to recall the correct answer since [apple] has multiple two-way bindings, which are [dog] and [rat]. However, if one had formed two two-way binding structures, the item to context binding (e.g. [list1] – [dog]) would restrict the multiple bindings and lead to a correct response, [dog].

Finally, in an ABCD condition two lists contained different items (see Figure 1-(b) right). Therefore a single two-way binding structure between the two items would be sufficient at test without considering the context (provided the items are selected so that there are no preexisting bindings between them).

A multinomial process tree (MPT) model (see Batchelder & Riefer, 1999 for a review) was used to quantify the relative contributions of the cue-target, context-target and the three-way binding structures. Differences among age groups were mainly restricted to the three-way binding structure. In particular, the improved ability to form three-way bindings differentiated the 7 year olds and adults, suggesting that development of the critical mechanisms is extended, perhaps through the teenage years. In a follow up study, Yim and colleagues increased the saliency of the context cue for 4 year olds (e.g. visiting Elmo's house instead of visiting a green house, see Figure 1-(a) right slide). By increasing the saliency of the context cue, 4 year olds increased the ability to use the context information (list). If the formation of episodic memory (or three-way binding structure) only relied on having the representational capacity to bind memory components, the manipulation of stimulus attention (or saliency) should not affect one's performance. However, the results suggest that simultaneous attention to components is required for successful binding. It was also argued that since children have low attentional control (Zelazo, Carlson, & Kesek, 2008), the developmental mechanism underlying episodic binding would be relying on attentional control during encoding and/or retrieval.

From the above results, it seems that attention plays an important role in binding, especially for the three-way binding structure. The goal of this research is to test this hypothesis directly. To ensure low or no attention, a statistical learning paradigm (Saffran, Aslin, & Newport, 1996) was used by modifying the task used in Yim, Dennis &, Sloutsky (2011). In the modified task, participants saw a sequence of cartoon characters one at a time and their task was to distinguish whether the character was a male or a female while their accuracy and reaction time were measured. The sequence had a specific pattern resembling the ABCD condition and the ABABr condition that were used in the previous study. However, instead of presenting the two items and context together, the items and context were presented sequentially using pictures of cartoon characters (see Figure 1-(c)). Therefore, there was always a picture representing the list context followed by two pictures which represented the items. The main prediction was that if two items and context are bound together during learning, the triplet would be segmented as in the original statistical learning paradigm. Therefore, it would be possible to predict the third picture after seeing the first two pictures in the triplet (Turk-Browne, Simon, & Sederberg, in press). As a result, faster reaction time or higher accuracy at the third item in a triplet would be an indicator of learning the triplet.

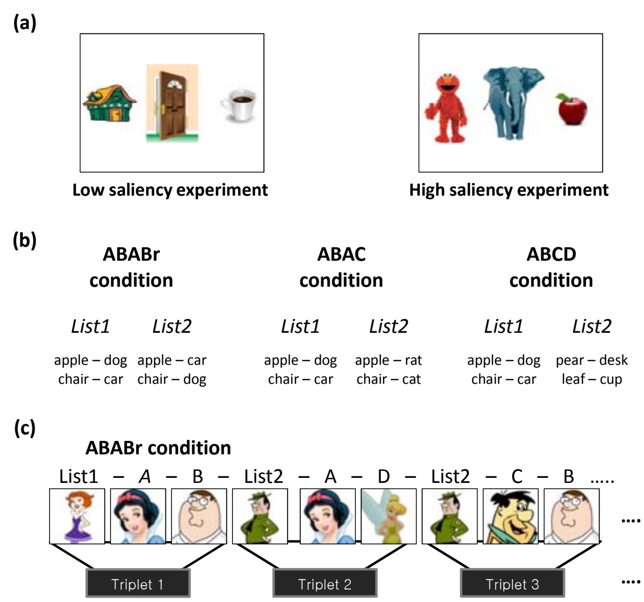

Figure 1: An illustration of experiment used in Yim, Dennis, & Sloutsky (2011a, 2011b) and its modification. (a) Stimuli used in the original experiment. The 'green house' and 'Elmo' represents the context cue whereas the 'door'/'apple' represents the item cue, and the 'cup'/'elephant' represents the target items (b) An example of the structure of the list in three conditions. (c) A modification of the original list learning paradigm (ABABr condition) into a statistical learning paradigm.

Experiment

Method

Participants Eighty undergraduate students at The Ohio State University participated for course credit (44 females, *M* = 19.11 years, *SD* = 1.38 years). Ten additional participates were excluded from the sample due to pressing the wrong key (N=5), chance level accuracy (N=4), and not understanding the instructions (N=1).

Stimuli The stimuli were thirty six pictures of cartoon characters where half were male and the other half female. Post-experimental questions indicated that the sex of each character was easily determined by the participants. Each picture was presented on a white squared patch (10.55cm × 10.55 cm), which was centered on a black background of a 17inch computer monitor. For every participant, the pictures were randomly assigned to each experimental condition following the constraint of each condition.

The practice phase preceded the experiment and had five male and five female pictures where the order of presentation was randomized. The pictures in the practice were not used in the experiment. All experimental conditions including the baseline condition consist of triplets (i.e. context and two items). In each experimental condition, there were four unique triplets in the learning phase and an additional four unique triplets at test (see Figure 2-(a), (b)). The triplets in the learning phase followed the structure of the two list learning paradigms (i.e. ABABr, and ABCD) but were presented one at a time following the order of [context]-[item1]-[item2]. At test, participants were presented with triplets consisting of the learned elements, such that some of the triplets conformed to the learned statistics (i.e., Congruent) and some did not (Incongruent). Since the current paradigm examines the existence of a binding structure via the predictability of the third item based on the first two items, it is important to manipulate the third item at test while controlling the first two items. For example, in the ABABr condition if the first triplet (X1-A-B) was learned, which would mean that a three-way binding structure among the context (X1) and the two items (A and B) is formed, after seeing a sequence of X1 and A, it would be able to predict that the next item would be B. Therefore, response time or/and accuracy of item B would be faster or higher. On the other hand, a corresponding incongruent triplet (i.e. X1-A-D) would results in a slower response time or lower accuracy since the sequence violates the learned statistics and would interfere with the learned prediction "B".

In each learning phase, there were ten repetitions for each unique triplet. In a repetition there were four unique triplets which result in a total of 120 (10 (repetition) × 4 (number of triplets) × 3 (pictures in a triplet)) trial per learning phase. In addition to the learning phase, the test phase consisted of 8 triplets, which had 4 congruent triplets identical to the learning phase and 4 incongruent triplets. Therefore the test phase had 24 trials (8 (triplets) × 3 (trials per triplet)) Therefore, there was a total of 144 trials for each condition.

There was also a baseline condition, which was designed to have no predictability of the third picture based on the first two pictures. Therefore, the triplets were all possible combinations using two pictures at each position (see Figure 2-(c)). The goal of the baseline condition was to measure the latency decrease that was due to task familiarity rather than to statistical learning. The total number of trials for the baseline condition was same as the experimental conditions. However, since there were eight unique triplets in the baseline condition, the whole triplet was repeated five times instead of ten. Finally, all conditions had the same number of male and female picture in each position of the triplet, and the frequency of the last picture in each triplet was equated for each condition.

(a) ABABr

learning	*test* *congruent*	*test* *incongruent*
X1-A-B	X1-A-B	X1-A-D
X1-C-D	X1-C-D	X1-C-B
X2-A-D	X2-A-D	X2-A-B
X2-C-B	X2-C-B	X2-C-D

(b) ABCD

learning	*test* *congruent*	*test* *incongruent*
X3-G-H	X3-G-H	X3-G-K
X3-M-H	X3-M-H	X3-M-K
X4-J-K	X4-J-K	X4-J-H
X4-N-K	X4-N-K	X4-N-H

(c) Baseline

learning	*test* *congruent*
X5-O-P	X5-O-P
X5-Q-P	X5-Q-R
X5-O-R	X6-Q-P
X5-Q-R	X6-O-R
X6-O-P	
X6-Q-P	
X6-O-R	
X6-Q-R	

Figure 2: The structure of stimuli in each condition. Every letter represents a picture that was presented sequentially where X denotes a context while other letters denote items.

Procedure The experiment consisted of three blocks in addition to the practice block, where each block was assigned to a specific condition (i.e. the ABABr condition, ABCD condition and, the baseline condition). Each block consisted of a learning phase and a test phase, and the transition between learning and test was unbeknownst to the participants while there was a break between blocks. Each participant experienced all four conditions, where the order of the conditions was randomized and the participants did not know the identity of each condition. The procedure of each condition was identical except for the stimuli and the sequence of each stimuli being presented as explained in the stimuli section. Participants were told that they would see cartoon characters and their job was to distinguish whether it was a male or a female. The pictures were presented on the screen until the participants respond, and the next picture was presented after a 750msec ISI. The participants were not informed that there was a triplet structure in each condition. The test phase consist of 4 old triplets from the learning phase and 4 new triplets that where incongruent to the triplets in the learning phase. The order of the triplets was randomized.

After the experiment, participants were asked whether they saw a pattern in the sequence and whether there were any cartoon characters for which it was difficult to determine their sex.

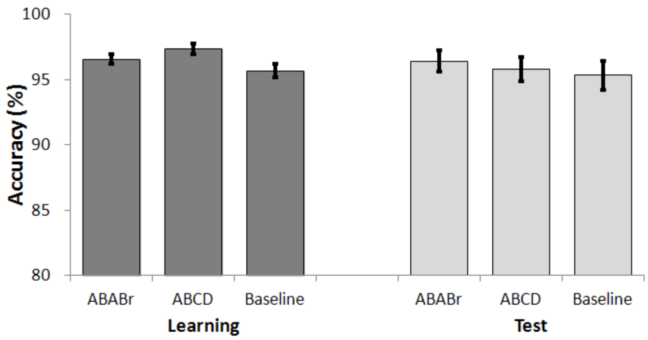

Figure 3: Accuracy for each condition at learning and test. Error bars indicate +/- one standard error.

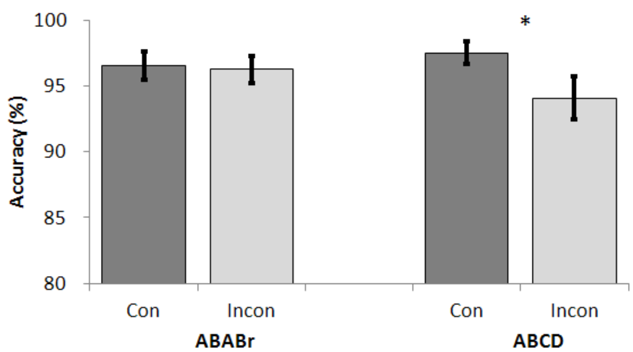

Figure 4: Comparing accuracy for congruent (Con) and incongruent (Incon) triplets at each condition at test. Error bars indicate +/- one standard error. * indicates $p < .05$

Results

The results from the post-experimental questions show that all pictures were distinguishable. Moreover, there were no participants who reported finding a pattern in the sequences, thus confirming that learning was indeed implicit.

For the analysis, only the reaction time and accuracy of the third position in the triplet was used. The overall accuracy was 96.48% (*SD* = 2.71%), with 96.64% (*SD* = 2.69%) for learning and 95.66% (*SD* = 5.03%) for test. A 2 × 3 (Phase: Learning vs. Test by Condition: ABABr, ABCD, and baseline) within-subjects ANOVA was conducted on accuracy showed no effect on Phase (p = .247), nor Condition (p = .220), and no interaction (p = .565) (see Figure 3). A 2 × 2 (Congruency × Condition) within-subjects ANOVA conducted on the accuracy at test only showed a marginal main effect for Congruency ($F(1, 79) = 2.95$, $p = .09$, $\eta_p^2 = .036$), whereas there was no significant main effect for Condition (p = .625) nor significant interaction (p = .159) (see Figure 4). Conducting a paired t-test between the congruent and incongruent triplets in each condition showed a significant difference in the ABCD condition ($t(79) = 1.95$, $p < .05$, $d = .30$) but not in the ABABr condition (p = .81).

Before analyzing the reaction time data (RT), values below 250msec and above 2500msec were excluded as outliers. The excluded data was .23% of the total learning data, and .08% for the test data. Also the median value was used for each subject's RT for each repetition in a condition.

The overall mean RT for learning was 577msec (*SD* = 203msec), 589msec (*SD* = 190msec) for the ABABr condition, 573msec (*SD* = 189msec) for the ABCD condition, and 551msec (*SD* = 80msec) for the Baseline condition (see Figure 5-(a)).To analyze learning during each condition, the asymptote of each learning curve was calculated. The asymptotic point was chosen by examining the last four repetitions for the experimental conditions and the last two repetitions for the baseline condition (cf. note that experimental conditions of four unique triplets and the baseline condition has eight unique triplets). The median value of each subject's RT was calculated among the asymptotic points and was analyzed using one way within-subjects ANOVA. Results show that there was a significant difference among conditions ($F(1.95, 154.30) = 3.49$, $p < .05$, $\eta_p^2 = .042$). Conducting a pair-wise comparison with Bonferroni adjustments showed that the ABABr condition was significantly different from the ABCD condition ($p < .05$), and not different from the Baseline condition (p = 1.00). Difference between the Baseline condition and the ABAC condition was also marginally significant ($p < .10$) (see Figure 6).

The mean RT for the ABABr condition at test was 572msec (*SD* = 121msec) for the congruent trials, and 563msec (*SD* = 150msec) for the incongruent trials. For the ABCD condition it was 543msec (*SD* = 118msec) for the congruent trials, and 577msec (*SD* = 162msec) for the incongruent trials. The congruent trials for the Baseline condition was 546msec (*SD* = 86msec). The analysis of test data suggests that for the ABCD conditions the congruent triplets elicited faster response than the incongruent triplets, which was not the case for the ABABr condition (see Figure 5-(b)). A 2 × 2 (Condition × Congruency) within-subjects ANOVA conducted on the test data showed no main effect for Condition ($p < .518$) nor Congruency ($p < .201$), but a significant interaction ($F(1, 79) = 7.63$, $p < .005$, $\eta_p^2 = .088$). Further comparison between the congruent and incongruent triplets within each condition was conducted by a paired t-test. Results showed a significant difference in the ABCD condition between congruent and incongruent triplets ($t(79) = 3.24$, $p < .005$, $d = .24$). However there was no significant difference between the congruent and incongruent triplets in the ABABr condition (p = .534).

In sum, the learning data showed that the ABCD condition had a significant difference in the amount of learning compared to the baseline condition, whereas the ABABr did not. Results from the test data indicated that only the ABCD condition, but not the ABABr condition, exhibited evidence of learning.

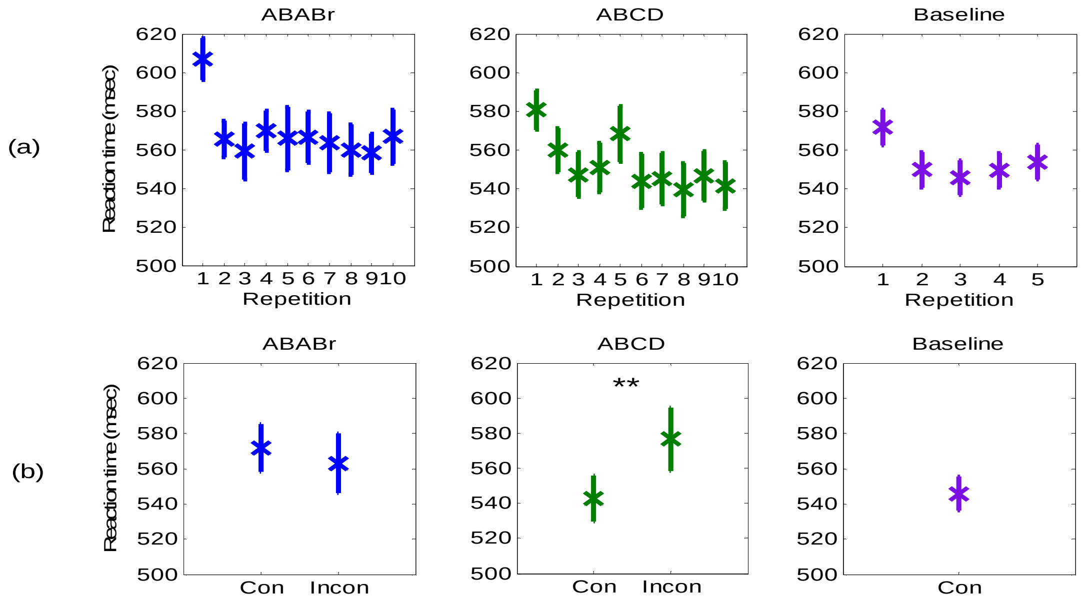

Figure 5: RT at each condition during (a) learning at each repetition, (b) test at congruent (Con) and incongruent (Incon) trials. Error bars indicate +/- one standard error. ** indicates $p < .005$.

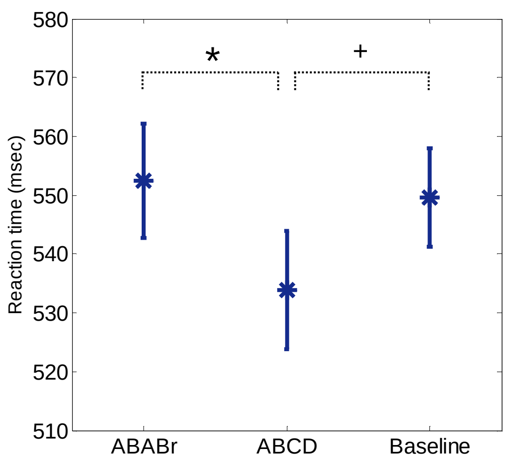

Figure 6: Comparing the asymptotic point for each condition during learning. Error bars indicate +/- one standard error. * indicates $p < .05$, + indicates $p = .10$

General Discussion

The current study examined the possibility of forming a complex binding structure without attention by presenting participants with an implicit statistical learning task. Results showed that there was evidence of learning only in the ABCD condition but not in the ABABr condition. These findings suggest that whereas forming a two-way binding structure does not require attention, a three-way binding structure does require attention.

The results could be viewed from the point of view of the sequential learning literature (e.g. Reber, 1993) where the conditional probability of the 1st item (context item) and the 2nd item (item1) in a triplet plays a role in predicting the 3rd item (item2). Based on the predictability of each condition, the ABABr in nature is the hardest condition to learn since the first two items conjointly should predict the third item. This argument is also consistent within the memory literature where a three-way binding structure is required to correctly recall in an ABABr condition - the two cures are conjointly bound with the target item (Humphreys, Bain, & Pike, 1989). On the other hand, the ABCD condition only requires association between the 2nd and 3rd item, which is also consistent with the memory literature – two-way binding structure.

However, predictability might not be the only factor that plays a role in the ABABr condition. The baseline condition has no predictability since it consist of all possible combinations using binary values at each position of a triplet. Thus, there should be no advantage from predicting the 3rd item. However, comparing the ABABr condition and baseline condition at test and learning shows no difference between the two conditions. Since there is evidence that 2nd order sequences are learnable (Stadler & Frensch, 1998), learning the ABABr condition should gain from predictability. Therefore, it is possible that the ABABr condition confronts additional interference which could not be alleviated without attention (cf. however, it is arguable that the ABABr condition requires more learning trials.)

The argued binding mechanism has some similarities compared to the binding mechanisms in other fields. In visual perception, binding concerns with object recognition (Treisman, 1996). To recognize an object and differentiate is from others, one should properly bind the properties (e.g. color, shape, location, etc.) to the correct object. Known that different properties are processed in different areas of the brain (Hubel & Wiesel, 2004), correct binding could not be done without attention (Treisman & Schmidt, 1982). Relational reasoning also requires a binding mechanism

(Hummel & Holyoak, 2003), where one has to bind correct fillers (e.g. *cat, tiger*) to correct roles (e.g. *B* is **bigger than** *A*) to form a relational representation (e.g. a *tiger* is **bigger than** a *cat*), and to further generalize or infer the representation.

The similarities among these bindings are that incoming components have to be bound together for further processing, and that incorrect binding will produce an erroneous response. However, the level of binding seems to be quite different even though they could be on the continuum. The components in the binding process of visual perception are features that are within the visual object. Moreover, due to dedicated feature detectors the binding process during visual perception would require less elaborated attention than in episodic memory. On the other hand, binding in relational reasoning requires more elaborate attention than episodic memory since the binding structure (i.e. role) is not a simple association but has a specific structure (e.g. ~ is bigger than). There is no evidence that the binding process has the same underlying mechanism across these domains. However, based on the similarities at the computational level, it is possible that there could be common mechanisms that involve the attentional system.

Finally, the fact that the participants failed to learn in the ABABr condition without attentional mechanism raises a question about the nature of the ABAC condition. Both ABAC and ABABr conditions require a complex binding structure to correctly remember the episode, which is a three-way binding and two two-way binding respectively. In forming a three-way binding structure, one should attend to all three elements that should be bound together. Therefore, attention is critical mostly at encoding but also at retrieval. However, it is possible that the two two-way binding structures are formed as two independent two-way bindings and integrated at retrieval. Thus, attention will mostly be required at retrieval. To address these issues, future studies should utilize methods that could measure attention during encoding and retrieval separately.

References

Barnes, J. M., & Underwood, B. J. (1959). "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology*, *58*(2), 97-105.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychological Bulletin & Review*, *6*(1), 57-86.

Cohen, N. J., & Eichenbaum, H. (1993). Memory, amnesia, and the hippocampal system. Cambridge, MA: MIT Press.

Ghetti, S., & Lee, J. (2011). Children's episodic memory. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*, 365-373.

Hubel, D. H., & Wiesel, T. N. (2004). *Brain and Visual Perception: The Story of a 25-Year Collaboration*. New York: Oxford University Press.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review*, *110*, 220-264.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different Ways to Cue a Coherent Memory System. *Psychological Review*, *96*(2), 208-233.

Porter, L. W., & Duncan, C. P. (1953). Negative Transfer in Verbal Learning. *Journal of Experimental Psychology*, *46*(1), 61-64.

Postman, L. (1962). Transfer of training as a function of experimental paradigm and degree of first-list learning. *Journal of Verbal Learning and Verbal Behavior*, *1*, 109-118.

Postman, L. (1964). Studies of Learning to Learn II. Changes in Transfer as a Function of Practice. *Journal of Verbal Learning and Verbal Behavior*, *3*, 437-447.

Reber, A. S. (1993). *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. New York: Oxford University Press.

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

Schacter, D. L., & Tulving, E. (1994). *Memory Systems 1994* (1st ed.). Cambridge, MA: The MIT Press.

Stadler, M. A., & Frensch, P. A. (1998). *Handbook of Implicit Learning*. Tousand Oaks, CA: Sage Publications, Inc.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, *6*, 171-178.

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*, 107-141.

Turk-Browne, N. B., Sederberg, P. B., & Simon, M. G. (in press). Scene representations in parahippocampal cortex depend on temporal context. *Journal of Neuroscience*.

Yim, H., Dennis, S., & Sloutsky, V. (2011a). The Development of Context Use and Three-way Binding in Episodic Memory. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3000-3005). Austin, TX: Cognitive Science Society.

Yim, H., Dennis, S., & Sloutsky, V. (2011b). The Development of Context Use and Three-way Binding in Episodic Memory. *Poster Presented at the Seventh Biennial Meeting of the Cognitive Development Society (CDS).* Philadelphia, PA.

Zelazo, P. D., Carlson, S. M., & Kesek, A. (2008). The Development of Executive Function in Childhood. In C. A. Nelson & M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience*. Cambridge, MA: MIT Press.