

An investigation of memory specificity and generalization in young children and adults

Presented by Elizabeth Hipskind

In partial fulfillment of the requirements for graduation with the Dean's Scholars Honors Degree in the Department of Neuroscience

4/19/2020

Date

Alison Preston
Supervising Professor

4/29/2020

Date

George Pollak
Honors Advisor in Neuroscience

Abstract

Optimal behavior in familiar and novel contexts depends on retrieval and consideration of past experiences. In adults, hippocampus supports retrieval of prior memories based on partially overlapping cues (Mack & Preston, 2016). Given that the hippocampus develops through childhood and adolescence (Keresztes et al., 2017), in the present research we investigated developmental differences in flexible memory retrieval during new experiences. Four-year-olds ($N=15$) and adults ($N=20$) learned a series of common object-novel shape associations. Following learning, participants were cued with a shape and tasked with retrieving the target object associate. On half of the trials, participants were cued with an identical shape from learning. On the remaining trials, participants were cued with a similar but non-identical shape morph, enabling examination of whether participants can flexibly generalize across similar but non-identical experiences to retrieve related memories. Accuracy and response times were measured for adults, and accuracy was measured for children. Both adults and children demonstrated reliable retrieval when cued with similar yet non-identical shapes. Whereas adults showed slower and less accurate retrieval for the non-identical versus identical cues, children showed no differences in retrieval as a function of cue similarity. These findings have important implications for our understanding of how mnemonic specificity and generalization interact across development. In particular, our findings suggest that mnemonic generalization in early childhood is a consequence of less detailed memory representation. Conversely, the more mature form of generalization evidenced in adulthood is accomplished through dual processing of the commonalities and specific differences between similar yet non-identical experiences.

Introduction

In our everyday lives, we not only face the challenge of forming specific memories of individual events, but also generalizing across similar events to extract general knowledge. No two experiences are identical, but some may be highly similar. Whereas generalization may guide behavior in novel, highly similar situations, specificity prevents confusion between distinct events. For example, a student attending their first lecture in college might be assigned a seat near the front of the lecture hall where they note how easy it is to hear the professor clearly. When the student arrives at their next lecture, they may recognize the common layout between the different lecture halls. In this way, generalizing across the similar yet distinct experiences supports inference and decision making, allowing the student to choose an optimal seat toward the front to facilitate hearing. However, it is also important to encode the unique elements of those experiences, such as the specific seat chosen in each setting. Failure to do so will lead to later confusion between the distinct lecture halls and may result in the individual sitting in the wrong seat. As this example demonstrates, mature memory behaviors rely on recognition of both the unique and common elements of highly related experiences. Consistent with this idea, recent work suggests that adults simultaneously code the specific and generalizable features of overlapping memories in the hippocampus (Schlichting, Mumford, & Preston, 2015). However, much less is known about whether children are sensitive to similarities and differences between overlapping experiences. Memory specificity and generalization are supported by the hippocampus (Schlichting et al., 2015; Keresztes et al., 2017, Zeithamova & Preston, 2010), which develops throughout childhood and adolescence (Schlichting, Guarino, Schapiro, Turk-Brown, & Preston, 2017). Therefore, investigating the development of these dual mnemonic processes in children provides an opportunity to study the building blocks of mature memory systems as they develop.

The ability to retrieve memories from partially overlapping cues is known as representational flexibility (Eichenbaum, 1997), which is central to encoding the similarities and

differences across related experiences. Hippocampal pattern completion mechanisms support reactivation of prior, related memories during new experience (Gershman, Schapiro, Hupbach, Norman, 2013; Zeithamova, Dominick, & Preston, 2012). Enhanced reactivation of related memories during novel experience has been linked to superior inference behavior in adults (Zeithamova et al., 2012). Along a similar vein, enhanced reactivation of prior memories during new experiences has also been linked to superior neural differentiation between related episodes (Kim, Norman, Turk-Browne, 2017). In particular, pattern separation is a process whereby similar inputs are orthogonalized to produce distinct memory representations (McClelland, McNaughton, & O'Reilly, 1995; Norman & O'Reilly, 2003; Stark, Yassa, Lacy, & Stark, 2013), which may be actively engaged when new experiences differ from memory-based expectations (Preston & Eichenbaum, 2013; Van Kesteren, Ruiters, Fernández, & Henson, 2012). That is, theoretical accounts posit that complex memory-guided behaviors are initiated through retrieval-mediated pattern completion, with resolution of the perceived conflict between newly and previously learned information resulting in dual memory representations that code the similarities and differences among individual events (Schlichting et al., 2015). Support for this view comes from the finding that during encoding of a separate yet related event, activity in hippocampal subfield CA1 is predictive of success on subsequent inferential judgments, a region that is thought to play a role in novelty detection (e.g., Larkin, Lykken, Tye, Wickelgren, & Frank, 2014; Schlichting, Zeithamova, & Preston, 2014). The proposal is that CA1 serves as a comparator that triggers a cascade of subsequent processes required to resolve, link, and generalize across the events. Furthermore, Varga and Bauer (2017) showed that detection of a difference between newly and previously learned information triggered subsequent mnemonic processes involved in representing the commonalities among them. Together, these findings suggest that the ability to represent the similarities and differences between overlapping experiences fundamentally relies on successful retrieval of prior memories, coupled with further attention to the unique and overlapping features.

The present research was an examination of developmental differences in the extent to which young children and adults successfully retrieve memories for previous events during similar but nonidentical new experiences. Understanding memory retrieval during development is important because a developmental analysis can be used as a theoretical tool to understand the building blocks that support the fully functioning memory system. Developmentally, a bias toward enhanced memory specificity may protect children from potential harm caused by responding to stimuli that differ from those already encountered (Barr & Brito, 2014). Indeed, research on young infants supports the idea that specificity may precede generalization in development. For example, in operant conditioning paradigms, infants learn to associate their kicking with the movement of a crib mobile that is tied to their foot, and consistently produce this action when they are shown the same mobile repeatedly. However, when more than one novel object is substituted on the mobile, or if markings on the mobile are more than 25% larger or smaller, 3-month-olds significantly reduce their kicking behavior (Hayne, Greco, Earley, Griesler, & Rovee-Collier, 1986; Gerhardstein, Adler, & Rovee-Collier, 2000). This finding suggests that the kicking action associated with the prior event (i.e., crib mobile) is only retrieved when new experience identically matches the previous experience. Memories are thus highly specific early in life.

Representational flexibility, which supports memory retrieval based on incomplete or even novel cues, has been shown to develop through early childhood (Barr & Brito, 2014). For example, Allen, Nurmsoo, and Freeman (2016) use a paradigm where an experimenter states they are going to draw a particular object (e.g., a balloon) then creates an ambiguous line drawing. They then ask if the drawing could be interpreted as something else (e.g., a lollipop) in the presence of distractor items. When the distractor items are perceptually distinct (e.g., a snake), both 4- and 6-year-old children say that yes, the drawing could be interpreted differently. However, when the distractor items are perceptually similar (e.g., a lollipop), only 6-year-olds

accept multiple interpretations of the ambiguous drawing. The 6-year-olds are able to overcome the interference of perceptually similar stimuli to flexibly recognize objects. Four-year-olds cannot overcome interference from similar stimuli and are instead more rigid in their representations in the presence of similar yet nonidentical cues, even though they are flexible when the distractor is distinct, suggesting that representational flexibility increases with age. (Allen et al., 2016). A developmental representational flexibility hypothesis supports an increase in the ability to generalize across stimuli with age. Furthermore, recent work examining the ability of young children to link separate yet related information across learning episodes indicates that this generalization behavior improves substantially between four and six years of age (Bauer, King, Larkina, Varga, & White, 2012; Bauer & San Souci, 2010; Bauer, Varga, King, Nolen, & White, 2015) and continues through adolescence (Schlichting et al., 2017). In addition to behavioral work which points to protracted development of mnemonic generalization, it is noteworthy that the posterior hippocampus, which supports the encoding of specific details, develops earlier than the anterior hippocampus, which supports flexible encoding and retrieval (Schlichting et al., 2017).

Contrary to the theoretical proposal that memory representation proceeds from specific to general, recent work has argued that children form general but not specific memories (Keresztes, Ngo, Lindenberger, Werkle-Bergner, & Newcombe, 2018; Ngo, Newcombe, & Olsen, 2018). According to this proposal, a bias toward enhanced generalization may facilitate the development of general semantic knowledge which may be more critical than discriminating fine details between similar episodes. To test this claim, the authors used a mnemonic similarity task adapted from adult research, where participants see a series of common objects and later judge whether a set of identical pictures (targets), highly similar but nonidentical pictures (lures), and completely new pictures (distracters) are the same, similar, or new. Developmental improvements in performance were observed between 4-year-olds, 6-year-olds, and adults, such that the ability to correctly judge similar lure items as “similar” increased with age (Ngo et

al., 2018). Moreover, only adults showed a reliable ability to successfully discriminate between similar but non-identical items, suggesting that memory specificity is not evidenced until adulthood. This pattern of results has thus been taken as support for the theoretical view that memory is over-general early in life and gradually develops specificity.

The conclusion that memory develops from general to specific conflicts with work demonstrating that children form detailed memory representations early in life. Inconsistent patterns of specificity and generalization across development may result from the use of different paradigms, and the use of novel or familiar materials in particular. Robust prior exposure to individual events before encoding related events has been shown to promote memory generalization in young adults (Schlichting et al., 2015). Therefore, the use of common stimuli in the mnemonic discrimination task may have biased children toward enhanced generalization behavior. Because children had extensive prior experience with the encoded exemplars (e.g., rubber ducks, bicycles, etc.), it may have been more difficult to encode the specific details of each studied exemplar in the experimental setting. That is to say, encoding of the specific elements of each exemplar required that it be effectively separated from all prior experiences with that type of stimulus—a task that may be particularly challenging for younger children. Additionally, this task requires monitoring of retrieved memories to decide whether items are the “same”, “similar”, or “new”, which relies on non-mnemonic decision processes that develop late into the school-age years (Ghetti, Lyons, Lazzarin, & Cornoldi, 2008). Hence, children may be able to recall the previous item, but unable to hold it in their mind, compare it to the picture in front of them, then make a decision about its similarity leading to an apparent bias toward disproportionate over-generalization early in life.

Due to the discrepancy between the research in infants that suggests memory develops from specific to generalizable and the research in preschool-age children indicating that memory becomes more specific across developmental time, in the present work we investigated the development of specificity and generalization in preschool-age children with a task more akin to

the infant paradigms. In order to reduce the effects of prior knowledge, we employed novel stimuli. Here, we asked if children and adults generalize across similar, but not identical, novel stimuli after a single, experimentally-controlled exposure to that previous item. Participants learned associative pairs consisting of a novel 3D object and a common object (see Figure 1A). After learning the unique pairs, participants were presented with a similar but non-identical 3D object, referred to as a morph (see Figure 1B), and asked to retrieve the paired common object. Through employing a direct test of retrieval via pattern completion, the present design allowed us to isolate how 4-year-old children treat related but different experiences during the first overlapping encounter, providing insight into the basic mechanisms in place to support overlapping learning in the absence of extensive prior knowledge. As discussed above, accumulating neuroscientific evidence suggests that the mature memory system forms both specific and general representations for overlapping events (Schlichting et al., 2015). Evidence of both specificity and generalization would be supported if individuals reliably retrieve the paired item when cued with the non-identical morphed shape (i.e., generalization) but show lower accuracy or slower reaction time as compared to retrieving paired items when cued with the originally learned 3D shape (i.e., specificity). Another possibility is that participants show robust specificity in the absence of generalization, which would be evidenced by unreliable retrieval when cued with a shape morph. That is, a retrieval deficit in the lure condition would indicate a failure of pattern completion, implying the existence of a highly specific memory for the originally experienced item. Finally, participants may show generalization in the absence of specificity, as evidenced by equivalent retrieval when cued with either the shape morph or the original item, indicating that they are more sensitive to the commonalities between novel stimuli.

One of the problems when studying generalization is that it may be the result of either perceptual confusion, where two similar stimuli cannot be discriminated, or acquired equivalence, where two similar stimuli are discriminated but treated the same (Barr & Brito, 2014). Perceptual confusion may indicate a lack of encoding specificity. Previous studies have

used a lack of specificity when comparing similar exemplars of common objects to conclude that children's memory develops from general to specific (Ngo et al., 2018). However, they do not attempt to disentangle perceptual confusion from acquired equivalence. In the present study, we employ a working memory perceptual discrimination task in order to separate perceptual confusion from acquired equivalence. This allows us to control for trials in the lure condition for which participants were unable to discriminate between the target and lure, thereby ruling out the possibility that over-general memory retrieval is due to an inability to perceptually discriminate between the original and morphed cue items.

Method

Participants

Participants were 15 4-year-olds (8 females) and 20 adults between 18-23 years (15 females). One additional child participant took part in the study but was excluded from analysis due to failure to complete the primary retrieval task. Moreover, five 4-year-old participants were excluded from the main retrieval analyses due to incomplete or unusable perceptual discrimination data, as we were not able to rule out the possibility that their mnemonic retrieval was influenced by failures in encoding specificity. Children were recruited through an existing pool of volunteer parents who had expressed prior interest in participating in child development research. Adults were recruited through undergraduate psychology courses at a public university. Children were compensated with \$10 to acknowledge their participation. Adults received partial course credit for participation. The protocol and procedures were reviewed and approved by the university Institutional Review Board. Informed consent was obtained from parents and adult participants and assent was obtained from children prior to the start of the study.

Experimental Design

The primary aim of the present research was to clarify whether, when confronted with similar but non-identical experiences, individuals show evidence for memory

specificity, memory generalization, or both. Unlike previous paradigms which have focused on memory specificity through behavioral “old” or “new” judgments, here we assessed behavioral evidence for specificity through a paired associate inference task. As depicted in Figure 1A, individuals learned a series of object-object associations. Following learning, participants were asked to retrieve the target object when cued with the paired object (see Figure 1B). The critical manipulation pertained to how memory retrieval was cued. On some trials, participants were cued with the exact object learned previously. However, on other trials, participants were cued with a similar yet not identical object (i.e., a morph), experimentally creating the conundrum experienced in everyday life—whether to process and represent the overlapping similarities, the unique differences, or both. By testing children and adults, the design therefore enabled examination of each of these potential outcomes at different points in development.

Stimuli

The stimuli consisted of 14 pairs of objects. As depicted in Figure 1A, each pair consisted of a common object (e.g., chair, hammer, bicycle) and a novel 3D object. In order to maximize learning of the association between the common and novel objects, the set of possible common objects was chosen based on the likelihood of 4-year-olds being familiar with them and prioritized objects with a functional use. As discussed above, previous reports of diminished memory specificity in 4-year-olds may have been influenced by the use of highly familiar stimuli which are difficult to discriminate (Ngo et al., 2018). To address this concern, we developed novel 3D objects. As depicted in Figure 1B (right), two versions of the novel 3D objects were generated, whereby one key component of the original object was morphed into a distinguishable, separate shape. Through changing one key feature (e.g., grey pointy versus rounded feature) while holding the remaining features constant (e.g., teal ribbon), these newly designed “shape morph” objects enabled examination of whether individuals show specificity for the original target shapes when cued with a morphed object. Images of novel 3D objects were

created using Blender, an open source 3D software. Novel objects were designed to be realistically plausible, such that they seemed like they could exist, while still remaining distinct from real-world objects to ensure novelty. Each novel object consisted of approximately two or three features and incorporated a variety of colors so as to be interesting to children.

Procedure

Stimulus Selection. All tasks were completed on laptops using MATLAB. The first part of the study consisted of a “Know” game in which participants were shown common objects individually and asked to label them. Stimuli were only included in the subsequent tasks if the participant was able to name the object, as a lack of familiarity may have interfered with the ability to learn the association between the common object and the novel 3D object. In an effort to maximize the similarity of the stimulus sets between participants, all participants received the same order of objects. Participants had an unlimited amount of time to name each object. The task terminated when participants labeled 14 objects successfully.

Learning Phase. Following stimulus selection, participants completed the paired associative learning task in which they learned 14 pairs of objects (Figure 1A). To provide sufficient time for encoding, participants viewed the pair of objects for 8 seconds at a time. The object on the left was a 3D novel object, as described above, and the object on the right was a common object that the participant was familiar with. Two counterbalanced sets of novel 3D objects were created, which balanced factors such as size, salience, and color across the 14 possible objects. Each set of 3D objects was tested equally in the target and morph condition within each age group. These novel 3D objects were randomly paired with common objects to create a unique set of stimulus pairs for each participant.

To facilitate learning, participants were instructed to come up with a story that used the two pictures *together*. Prior to beginning the learning task, participants completed six practice items to ensure that they understood how to form associations between the objects. Participants first completed two untimed encoding practice trials in which participants were shown object

pairs and asked to come up with a story about how they could go together. If the participants could not come up with a story, they were provided with an example. Once the participant was able to verbalize how the two objects might go together, the same procedure was repeated for four additional timed examples that mimicked the 8-second time limit implemented in the main task. Some participants required extra time to think of a story on the practice items or were given an example of how to use the objects together, but all participants were able to imagine the items together after corrective feedback. Participants were informed that there would be a memory test following the learning phase but were not provided with further details so as not to promote a particular strategy during encoding.

As depicted in Figure 1A, following the 8-second encoding interval, participants were asked to describe their imaginative story to the experimenter. In order to control for learning in later analyses, participants were scored on whether they were able to successfully put the two objects together for each trial. Success was defined as the participant producing a story that used both items and related the items to each other. Participants could begin describing the story during the 8-second encoding interval or after the trial and were given enough time to finish their description before moving on to the next trial. The order of pairs was pseudo-randomized, with the constraint that no more than three pairs from the same condition (target versus shape morph) appeared in a row.

Test Phase. In order to assess retrieval success based on identical (i.e., target) and nonidentical (i.e., morph) cues, participants completed a three-alternative forced-choice task that was given immediately after learning. Memory for each of the pairs learned in the previous task was tested, with seven associative pairs tested in each condition. In the target condition (Figure 1B; left), participants were cued with the same 3D novel object that they learned previously, while in the morph condition (Figure 1B; right), participants were cued with a similar, but not identical, version of the 3D novel object. Participants were given a choice of three objects, one of which was paired with the novel object during learning. The two distractor items

were common objects that were previously paired with a different novel object during learning. Participants were instructed to select the object on the bottom that went with the one on top during the story game (learning) and were not informed that some of the novel objects were slightly different from what was originally viewed.

Test items were presented in a pseudo-random order, with the constraint that no more than three trials of the same condition appeared in a row. In order to minimize the effects of a systematic strategy, the correct answer was located in each position (left, middle, or right) an approximately equal number of times across the test phase. Each common object appeared three times, once as the correct answer and twice as a distractor. Moreover, to mitigate interference across test trials, correct objects and distractor objects never appeared in back-to-back trials. This task was self-paced, and participants had as much time to select an answer as they needed. Participants were instructed that they should make their best guess if they did not know the answer. Adults responded by pressing the corresponding button on the keyboard, and response time was collected when they made their answer choice. Based on piloting work, allowing younger children to make button presses increased errors, as some children tried to press the button as fast as they could. As such, 4-year-olds responded either verbally or by pointing to their answer choice, and the experimenter recorded their response.

Perceptual Discrimination. The last task was a perceptual discrimination task to ensure that participants could visually discriminate the features of the shape morphs that changed between study and test. As depicted in Figure 1C, participants saw the target novel object from the learning task for 4 seconds followed by a one-second visual mask. Following the mask, participants saw either the identical target image or its shape morph (Figure 1C). Participants were instructed to tell us if the second object was exactly the same or different from the first one. In order to use this perceptual discrimination task as a control for the associative retrieval task, objects in the perceptual discrimination task were tested in the same experimental condition (identical target versus morph). Thus, objects in the target condition in the associative

retrieval task should have elicited a “same” response in the discrimination task. On the other hand, objects in the morph condition in the associative retrieval task should have elicited a “different” response. Importantly, matching the trial conditions across the mnemonic retrieval and perceptual discrimination tasks allowed us to control for trials in which participants were not able to discriminate the changed stimulus morph and therefore would not be expected to show mnemonic specificity for the change.

In pilot work, adults looked for any possible change and consequently false negatives (responding “different” to two pictures that were the same) were common. To circumvent this issue, participants completed practiced trials in which they were shown examples of correct same and different judgments. Participants were instructed that when the objects were different, it was because part of the shape had changed, though the size and color were always the same. Participants were given two practice trials, one morph (i.e., different) and one target (i.e., same). If the participant answered incorrectly, the instructions were repeated, and the trial was given again. Participants could repeat a trial up to two times if necessary. If the participant was unable to answer either trial correctly after repetition of the instructions twice, the session was ended ($N=2$ 4-year-olds). Once again, 4-year-olds responded verbally, while adults responded using a keyboard. Participants had as much time as necessary to respond. The experimenter manually started each trial to ensure that participants were attending to the screen before presentation of the next trial. In order to increase engagement and motivation in 4-year-olds at this latter half of the testing session, visual and audio feedback was provided after responding. This task was divided into two runs, with eight trials in the first run and six trials in the second run. The order of trials was pseudo-randomized, with the constraint that no more than three trials of the same condition appeared in a row, and each run had an equal number of “same” and “different” trials.

Results

Associative encoding

To ensure the associative pairs were initially encoded, we first examined imagination performance during the learning phase. Four-year-old children successfully imagined the two items together on 83% of the trials on average (*Range* = 14-100%), which was reliably above chance (50%), $t(14)=4.97$, $p<.001$. Although one 4-year-old imagined only 14% of the pairs together, the remaining child participants were successful on 42% or more of the trials. Adults successfully imagined the stimuli together on an average of 97% of trials (*Range* 71-100%), which significantly differed from chance, $t(19)=29.33$, $p<.001$. Thus, both age groups exhibited highly reliable encoding of the novel associative pairs, suggesting that participants were able to form associative memories during learning.

We next examined imagination success as a function of age group and later retrieval condition. As depicted in Figure 2A, a 2x2 mixed ANOVA with a between-subjects factor of age group and a within-subjects factor of condition (target versus morph) revealed a main effect of age, ($F(1,33) = 5.42$, $p = .03$, partial $\eta^2 = .14$), such that adults exhibited higher imagination success than children. However, there was no significant effect of condition ($F(1,33)=1.02$, $p=.37$, partial $\eta^2=.03$) nor an age group x condition interaction, ($F(1,33) = .48$, $p = .50$, partial $\eta^2 = .01$), indicating that imagination success was equivalent regardless of whether the stimulus pairs were assigned to the target or morph condition. The ability to associate the novel stimuli during learning thus did not differ by later retrieval condition, indicating that there were no systematic differences between the target and morph stimuli groups at learning. However, to control for age-related differences in encoding success, subsequent analyses of associative retrieval include only trials in which participants exhibited successful associative encoding during learning.

Perceptual discrimination

If participants were unable to perceptually distinguish between the target and morph stimuli during encoding, then it is unlikely that there would be a mnemonic effect at retrieval. Thus, to ensure that participants were able to perceive the differences between the target and morph stimuli, we further examined performance on the perceptual discrimination task. Overall, 4-year-olds correctly discriminated target and morphed shapes on 68% of the total trials, which was significantly above chance (50%) based on a one-sample t-test, $t(10)=3.99$, $p=.003$. Adults correctly discriminated target and morphed shapes on 96% of all trials, which was also significantly above chance, $t(19)=35.14$, $p<.001$.

We next examined perceptual discrimination success as a function of age group and later retrieval condition. As revealed in Figure 2B, a 2x2 mixed ANOVA with a between-subjects factor of age group and a within-subjects factor of condition (target versus morph) showed a main effect of age, $F(1,29)=55.75$, $p<.001$, partial $\eta^2 = .66$, such that adults had higher overall accuracy than children. However, there was no main effect of condition, $F(1,29)=1.59$, $p=.22$, partial $\eta^2 = .05$ nor an age x condition interaction, $F(1,29)=1.59$, $p=.22$, partial $\eta^2 = .05$, indicating that perceptual discrimination success did not vary between the target and morph conditions. Although both age groups were able to reliably perceive the differences between the target and morph conditions, our main associative retrieval analyses controlled for age-related differences in sensitivity to perceptual details of the stimuli by including only target and morph trials for which participants were able to successfully discriminate.

Associative retrieval

The primary aim of the present research was to test whether associative retrieval differs when cued with identical features of previous experience as compared to similar but non-identical features, as well as whether this retrieval effect differs across age. To isolate retrieval processes, as opposed to differences in encoding of the original pairs and perception of the experimentally manipulated shapes, these analyses only include trials where the participants

both successfully imagined the pair of objects together at learning and were able to successfully discriminate between targets and morphs in the subsequent perceptual control task. We first examined whether participants exhibited reliable associative retrieval, regardless of experimental condition. Four-year-olds retrieved 53% of the total paired associates, which significantly differed from chance (33%) based on a one sample t-test, $t(9)=3.41$, $p=.008$. Adults retrieved 96% of the paired associates, which was significantly above chance, $t(19)=26.73$, $p<.001$. Thus, both young children and adults reliably retrieved the previously learned associative pairs.

Given that both children and adults exhibited reliable retrieval of the original associative pairs, we next addressed whether retrieval varied as a function of cue specificity and age. As depicted in Figure 3, a 2 x 2 mixed ANOVA examining forced-choice accuracy with a between-subjects factor of age group and a within-subjects factor of condition (target versus morph) revealed a main effect of age, $F(1,28)=61.27$, $p<.001$, partial $\eta^2 = .69$, such that retrieval was more robust in adults relative to children. Although we did not observe an effect of condition, $F(1,29)=.95$, $p=.34$, partial $\eta^2 = .03$, the condition x age interaction was marginally significant, $F(1,29)=3.94$, $p=.06$, partial $\eta^2 = .123$. Follow-up paired sample t-tests were conducted separately for each age group to test whether there were differences in retrieval between the target versus morph conditions. For the 4-year-olds, there was no significant difference in forced-choice accuracy between the target and morph conditions, $t(9)=-1.08$, $p=.31$. For adults, the t-test revealed a significant difference in the target versus morph condition, $t(19)=2.31$, $p=.03$, such that retrieval was more robust in the target compared to the morph condition (see Figure 3).

In light of the effect of condition on retrieval success in adults, we also explored whether there were differences in response speed on trials in which adults successfully retrieved the paired associate to provide an additional measure of sensitivity to the manipulation of the retrieval cue. Notably, three adults were excluded from this analysis because their average

response times were greater than 2 standard deviations above the mean. A paired sample t-test on the remaining sample revealed a significant difference in response time in the target and morph conditions, $t(16)=-2.66$, $p=.02$, such that adults were slower to choose the paired associate when cued with a morph compared to when cued with a target (Figure 4). Together, this pattern of results suggests that while adults were sensitive to the specificity of the retrieval cues, as evidenced by a retrieval deficit and retrieval speed cost when cued with a morphed object that differed from the originally experienced object, children showed no evidence of memory specificity for the originally encoded associative elements.

Controlling for effects of trial count and sample size on age-related retrieval differences

While the associative retrieval analyses reported above controlled for age-related encoding differences, we further examined whether there were significant differences in the number of trials submitted to the target and morph conditions, which may have impacted the retrieval patterns reported. As reflected in Table 1, fewer overall trials went into the analyses for the 4-year-olds compared to the adults, due to their reduced success on both the imagination and perceptual discrimination tasks. Critically, however, paired sample t-tests revealed that the number of trials entered into the target and morph conditions was not significantly different for children, $t(9)=.43$, $p=.68$ or adults, $t(19)=.27$, $p=.79$. There is thus no reason to believe that the null condition effect in 4-year-olds is driven by differentially reduced power for one condition or the other.

Although the primary analyses isolated mnemonic differences in memory retrieval by controlling for encoding and perception, which did not systematically affect trial counts between conditions, a third of the child sample was excluded due to failure to complete the perceptual discrimination task. Furthermore, exclusion of 35% of the overall trials for the included 4-year-olds (see Table 1) still contributes to overall reductions in analytical power. To address this concern, we also performed analyses that only controlled for imagination success, so that all participants were included while still ensuring that participants had learned the associative pairs.

Retrieval performance for the target and morph cues is depicted in Figure 5, separately for children and adults. Consistent with the previous analyses, paired sample t-tests revealed that there was no significant difference in accuracy between the target and morph conditions for the 4-year-olds, $t(14)=.17$, $p=.87$, while the effect of condition on retrieval success was replicated in the adults, $t(19)=2.36$, $p=.03$.

Association between retrieval success and perceptual discrimination

Finally, in light of the finding that perceptual discrimination success was strikingly lower in children as compared to adults (Figure 2B), we further explored whether individual differences in perceptual encoding were related to retrieval success. That is, despite controlling for associative encoding through the imagination task, the ability to encode the finer details of the novel 3D shapes may still vary across individuals within an age group, which should predict later retrieval ability. Consistent with this idea, we observed a positive correlation between overall retrieval success (controlling for imagination success) and perceptual discrimination in both children, $r(10) = .72$, $p=.01$ (Figure 6A) and adults, $r(19) = .62$, with $p=.004$ (Figure 6B). For both age groups, as perceptual discrimination ability increased, so did retrieval accuracy, suggesting a relationship between encoding specificity and mnemonic retrieval success.

Discussion

Consistent with previous research, we find evidence in adults for dual signatures of memory specificity and generalization (Schlichting et al., 2015). Specificity is suggested by the increase in response time in the morph condition compared to the target condition, as well as a decrease in accuracy in the morph condition. However, the overall accuracy is still high in the morph condition, indicating that adults are able to reliably generalize across the morph and target exemplars to retrieve the paired associate. This pattern of results suggests that adults are able to recognize the differences between similar episodes while still drawing on the similarities to retrieve the elements associated with the original event. On the other hand, we do not find

evidence of memory specificity in 4-year-olds. As expected, overall performance was less robust in children relative to adults. However, there were no significant differences in accuracy between the target and morph conditions, indicating that they were not sensitive to changes in the retrieval cue, which suggests less memory specificity. Importantly, this null effect in children was evident after controlling for potential differences at encoding, indicating that this is a mnemonic difference, not a perceptual difference.

The present data are consistent with the idea that memory develops from general to specific. Unlike Ngo and colleagues (2018), we find evidence of dual memory signatures in adults, in that they show evidence of both memory specificity and generalization across similar stimuli. While Ngo used a specificity paradigm that relied on explicit memory monitoring, we provide converging lines of evidence for a lack of memory specificity in 4-year-olds using an incidental generalization paradigm. The present research replicates and extends previous work. Robust prior experience has been shown to promote generalization across learning episodes in adults (Schlichting et al., 2015), so the use of common objects by Ngo may have promoted general encoding of the stimuli, rather than encoding the specific details, making it harder to discriminate targets and highly similar lures. However, we similarly find diminished memory specificity in 4-year-olds for novel materials, ruling out the potentially confounding influence of prior knowledge. Additionally, by implicitly probing cued retrieval rather than requiring explicit memory monitoring, we eliminated the influence of potential non-mnemonic processes that may have accounted for apparent developmental difference in memory between young children and adults. Our research thus corroborates the theoretical proposal that memory is less specific earlier in development.

Developmental differences in representational flexibility may explain the different pattern of results between adults and 4-year-olds. Representational flexibility supports retrieval based on partially overlapping cues. When adults are cued with the morph, a partially overlapping cue, they retrieve the originally-learned pair. Children also retrieve the originally-learned pair, given

that they perform above chance, but without the increase in response time and lower accuracy in the morph condition observed in adults, indicating that there are differences in how adults and children process the partially-overlapping cue. The accuracy deficit seen in adults in the morph condition suggests that they have high specificity for the originally-learned pair, such that sometimes they fail to reactivate the original associate when current experience deviates from prior experience in some way. One explanation for the increase in response time is therefore based on novelty detection. It has been suggested that memory integration may be triggered by associative novelty signals (Schlichting et al., 2014; Shohamy & Wagner, 2008). Under this account, adults reactivate the original pair based on the morph. However, because part of the novel object has changed, the neural representation of the morph deviates from that of the target, prompting novelty signaling. This causes additional encoding to extract the commonalities and differences between the related experiences. In the target condition, a lack of novelty signaling in the target condition leads to a faster response time, because the retrieval cue is identical to part of the original experience, precluding the need for additional time to resolve the conflict between the target and morph representations.

Unlike adults, in 4-year-olds, the lack of a significant difference in accuracy between the target and morph conditions suggests that they retrieve indiscriminately. As suggested by Lukowski and Bauer (2014), this flexibility in recall may be the result of forgetting the specific details of the originally-learned experience. A lack of specificity may bypass novelty signaling, such that the original experience is not retrieved with enough detail to conflict with the nonidentical morph. Therefore, retrieval in 4-year-olds in the morph condition would not tax representational flexibility, as the retrieved representation is not detailed enough to require flexibility to process the commonalities and differences with the morph. The differences in processing between adults and 4-year-olds therefore suggests that they are not engaging in the same type of generalization. Whereas the adults recognize and overcome the specific

differences between non-identical stimuli to generalize across them, 4-year-olds appear to generalize as a consequence of diminished representational specificity for past experience.

Interestingly, work in older adults demonstrates the same pattern of results with age-related decline in memory specificity. In one study (Vieweg, Stangl, Howard, & Wolbers, 2015), participants viewed five line drawings of scenes (e.g., kitchen, library). At test, they were shown the previously learned scenes and novel scenes with varying degrees of mask obscuring the picture. Participants were asked to identify the scene (e.g., “kitchen”) or select “none of these” if it was novel. Relative to young adults, older adults were more likely to identify novel scenes as one of the previously-learned scenes. This indicates an increase in pattern completion to nonidentical stimuli, similar to our observations that 4-year-olds indiscriminately retrieve to morphs. It is possible that age-related changes in the hippocampal CA3 region, which supports formation of specific memories, may drive declines in memory specificity with age. Consistent with this interpretation, when participants view objects with varying levels of mnemonic similarity, Yassa and colleagues (2011) found that older adults required much larger stimulus changes to exhibit separation-like BOLD responses compared to young adults. This suggests that age-related differences in the CA3 region may influence indiscriminate retrieval in the face of identical and highly similarity cues, which is primarily driven by diminished representation specificity for learned items.

It may seem contradictory that memory appears to develop from general to specific, given the work in infants that suggests even infants display specificity in some paradigms. Yet it is important to emphasize that some of these tasks that rely on hippocampally-independent non-declarative memory systems (Thompson & Steinmetz, 2009), such as the operant conditioning kicking mobile paradigm (e.g., Hayne et al., 1986, Gerhardstein et al., 2000), so behavioral changes likely follow a different developmental trajectory. Moreover, additional evidence from deferred and elicited imitation paradigms documenting memory specificity for temporal order in infancy shows that general constraints on the order in which sequences can feasibly occur

facilitates memory for those action (see Lukowski & Bauer, 2014 for review). In other words, prior knowledge and built-in environmental constraints may promote encoding specificity and retrieval of those memories. Indeed, memory specificity for arbitrarily ordered items is not evidenced until later infancy. Even once specificity for arbitrary items appears, infants are limited in the number of items they can encode and reliably retrieve, suggesting continued room for improvement into early childhood. Consistent with the idea that differences in task demands may influence patterns of memory specificity, when infants show generalization of temporal sequences with non-identical props, subsequent discrimination of target items from distracter props is taken as evidence of generalization based on specificity. However, it is important to note that memory for specific features is typically tested in forced-choice format in which infants can readily compare targets and distractors to one another. In contrast, when provided with similar but non-identical props (or shape morphs) and tested for retrieval of the associated actions (or objects), individuals must compare the present cues to their stored memory trace. If the mnemonic trace is not highly detailed, infants will not show a behavioral difference between morphs and targets. Hence, it is highly likely that infants do not demonstrate specificity in these more demanding and naturalistic retrieval conditions, consistent with the pattern of indiscriminate retrieval performance shown in 4-year-olds here.

The correlation between perceptual discrimination and successful retrieval further suggests a more domain general role of encoding specificity, such that encoding may constitute an important rate-limiting factor for retrieval sensitivity. Although we controlled for trials where the differences between the identical and nonidentical shapes were not perceived, it is not necessarily the case that details stored in short-term memory are were consolidated into long-term memory or able to be retrieved from long-term memory. Advances in multivariate neuroimaging methods have recently enabled measurement of specificity of neural representations during initial encoding in children (Fandakova, Leckey, Driver, Bunge, & Ghetti, 2019). Therefore, future research should measure and compare the specificity of neural

representations during encoding to those retrieved in response to a cue, thereby enabling examination of how encoding and retrieval processes contribute to the protracted developmental time course of associative retrieval success and improvements in memory specificity with age.

Finally, although a lack of a difference in accuracy in 4-year-olds between the target and morph conditions suggests a lack of specificity, it is important to note that response time was not collected for this age group due to task constraints. It is thus possible that additional implicit measures, such as eye tracking, may reveal some sensitivity to the differences between similar stimuli. For instance, Koski and colleagues (2013) used eye tracking to assess relational memory in 4-year-olds. After learning a series of face-scene pairs, eye movements were recorded during a three-alternative forced-choice retrieval test. They found that for trials in which children chose the correct associate, they fixated on the correct item longer than the distractors. This suggests that eye tracking may provide an additional index of memory, prior to overt behavioral responses. Likewise, Molitor and colleagues (2014) compared eye movements during encoding and later presentation of objects and similar lures in adults. Participants were asked to label objects as “old,” “similar,” or “new” as they viewed a series of pictures. They found that for trials where the lure was incorrectly labelled as “old,” there were fewer fixations during encoding of the initial object, suggesting that eye tracking may be sensitive to differences in encoding. Based on the use of eye tracking to isolate encoding and retrieval processes, future research should implement this technique to determine if children evidence implicit sensitivity to the change shape morphs.

In conclusion, both children and adults were able to form memories of novel stimuli and reliably retrieve these memories in response to both identical and slightly altered cues. The data suggest that children are able to retrieve equally in the target and morph conditions because of a deficit in memory specificity, such that they are not sensitive to small changes in the retrieval cue. On the other hand, we find evidence that adults exhibit both memory specificity, such that

they recognize the differences between the target and morph exemplars, and generalization, as they still exhibit a high degree of success in retrieving the paired associate in response to the altered cue. One potential explanation is that because adults have high specificity, the morphs initiate novelty detection processes. Pattern separation would encode the differences between the target and morph, but also slow down response time compared to the target condition, where pattern separation would not occur. This would create two distinct memory representations, and interference between these may result in the observed accuracy deficit in the morph condition. The present research is consistent with the proposal that memory specificity increases with age. Unlike previous studies, our task design uses novel stimuli to control for the effects of prior knowledge, and also attempts to control for both encoding and perceptual differences. While both children and adults showed evidence of generalization, these may be different forms of generalization, where in children it is mediated by a lack of specificity, while adults notice and overcome the differences between nonidentical experiences. Additional work is needed to assess when this more complex mnemonic representation forms and should use neuroimaging methods to directly measure the contents of the retrieved representations.

References

- Allen, M. L., Nurmsoo, E., & Freeman, N. (2016). Young children show representational flexibility when interpreting drawings. *Cognition*, 147, 21–28.
- Barr, R., & Brito, N. (2014). From Specificity to Flexibility: Early Developmental Changes in Memory Generalization. In *The Wiley Handbook on the Development of Children's Memory* (pp. 453–479).
- Bauer, P. J., King, J. E., Larkina, M., Varga, N. L., & White, E. A. (2012). Characters and clues: Factors affecting children's extension of knowledge through integration of separate episodes. *Journal of Experimental Child Psychology*, 111, 681–694.
- Bauer, P.J., & San Souci, P. (2010). Going beyond the facts: young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology*, 107(4), 452-65.
- Bauer, P.J., Varga, N.L., King, J.E., Nolen, A.M., White, E.A. (2015). Semantic Elaboration through Integration: Hints Both Facilitate and Inform the Process. *Journal of Cognition and Development*, 16(2), 351-369.
- Eichenbaum, H. (1997). Declarative Memory: Insights from Cognitive Neurobiology. *Annual Review of Psychology*, 48, 547–572.
- Fandakova, Y., Leckey, S., Driver, C. C., Bunge, S. A., & Ghetti, S. (2019). Neural specificity of scene representations is related to memory performance in childhood. *NeuroImage*, 199, 105–113.
- Gerhardstein, P., Adler, S. A., & Rovee-Collier, C. (2000). A dissociation in infants' memory for stimulus size: Evidence for the early development of multiple memory systems. *Developmental Psychobiology*, 36, 123-135.
- Gershman, S.J., Schapiro, A.C., Hupbach, A., Norman, K.A. (2013). Neural context reinstatement predicts memory misattribution. *Journal of Neuroscience*, 33(20), 8590-8595.
- Ghetti, S., Lyons, K.E., Lazzarin, F., & Cornoldi, C. (2008). The development of metamemory monitoring during retrieval: The case of memory strength and memory absence. *Journal of Experimental Child Psychology*, 99(3), 157-181.
- Hayne, H., Greco, C., Earley, L. A., Griesler, P. C., & Rovee-Collier, C. (1986). Ontogeny of early event memory: II. Encoding and retrieval by 2- and 3-month-olds. *Infant Behavior and Development*, 9, 441–460.
- Keresztes, A., Bender, A. R., Bodammer, N. C., Lindenberger, U., Shing, Y. L., & Werkle-Bergner, M. (2017). Hippocampal maturity promotes memory distinctiveness in childhood and adolescence. *PNAS*, 114(34), 9212–9217.

- Keresztes, A., Ngo, C.T., Lindenberger, U., Werkle-Bergner, M., Newcombe, N.S. (2018). Hippocampal Maturation Drives Memory from Generalization to Specificity. *Trends in Cognitive Science*, 22(8), 676-686.
- Kim, G., Norman, K. A., & Turk-Browne, N. B. (2017). Neural differentiation of incorrectly predicted memories. *Journal of Neuroscience*, 37(8) 2022-2031.
- Koski, J., Olson, I. R., & Newcombe, N. S. (2013). Tracking the eyes to see what children remember. *Memory*, 21(3), 396–407.
- Larkin, M.C., Lykken, C., Tye, L.D., Wickelgren, J.G., Frank, L.M. (2014). Hippocampal output area CA1 broadcasts a generalized novelty signal during an object-place recognition task. *Hippocampus*, 24(7), 773-83.
- Lukowski, A. F., & Bauer, P. J. (2014). Long-term memory in infancy and early childhood. In P. J. Bauer & R. Fivush (Eds.), *The Wiley handbook on the development of children's memory* (p. 230–254). Wiley Blackwell.
- Mack, M.L., & Preston, A.R. (2016). Decisions about the past are guided by reinstatement of specific memories in the hippocampus and perirhinal cortex. *NeuroImage*, 127, 144-157.
- 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 Reviews*, 102(3), 419–457.
- Molitor, R. J., Ko, P. C., Hussey, E. P., & Ally, B. A. (2014). Memory-related eye movements challenge behavioral measures of pattern completion and pattern separation. *Hippocampus*, 24(6), 666–672.
- Ngo, C. T., Newcombe, N. S., & Olson, I. R. (2018). The ontogeny of relational memory and pattern separation. *Developmental Science*, 21:e12556, 1–11.
- Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, 110(4), 611–646.
- Preston, A.R., & Eichenbaum, H. (2013). Interplay of the hippocampus and prefrontal cortex in memory. *Current Biology*, 23(17), R764-R773.
- Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal structure predicts statistical learning and associative inference abilities during development. *Journal of Cognitive Neuroscience*, 29(1), 37–51.

- Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, 6(8151), 1–10.
- Schlichting, M.L., Zeithamova, D. & Preston, A.R. (2014). CA1 contributions to memory integration and inference. *Hippocampus*, 24(10), 1248-1260.
- Shohamy, D., & Wagner, A. D. (2008). Integrating Memories in the Human Brain: Hippocampal-Midbrain Encoding of Overlapping Events. *Neuron*, 60, 378–389.
- Stark, S. M., Yassa, M. A., Lacy, J. W., & Stark, C. E. L. (2013). A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia*, 51, 2442–2449.
- Thompson, R.F., & Steinmetz, J.E. (2009). The role of the cerebellum in classical conditioning of discrete behavioral responses. *Neuroscience*, 162(3), 732-55.
- Van Kesteren, M.T., Ruiters, D.J., Fernández, G., & Henson, R.N. (2012). How schema and novelty augment memory formation. *Trends in Neuroscience*, 35(4), 211-219.
- Varga, N. L., & Bauer, P. J. (2017). Young adults self-derive and retain new factual knowledge through memory integration. *Memory and Cognition*, 45, 1014–1027.
- Vieweg, P., Stangl, M., Howard, L. R., & Wolbers, T. (2015). Changes in pattern completion - A key mechanism to explain age-related recognition memory deficits? *Cortex*, 64, 343–351.
- Yassa, M.A., Mattfeld, A.T., Stark, S.M., Stark, C.E. (2011). Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *PNAS*, 108(21), 8873-78.
- 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.
- 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.

Table 1. Mean number and range of trials included in retrieval analyses after controlling for encoding and perception, separately for children and adults and each experimental condition.

	Child		Adult	
	<i>M (SD)</i>	<i>Range</i>	<i>M (SD)</i>	<i>Range</i>
Target	4.70 (1.81)	2.00-7.00	6.50 (.89)	4.00-7.00
Morph	4.40 (1.35)	3.00-7.00	6.55 (.67)	5.00-7.00
All	9.10 (2.13)	5.00-7.00	13.05 (1.36)	9.00-14.00

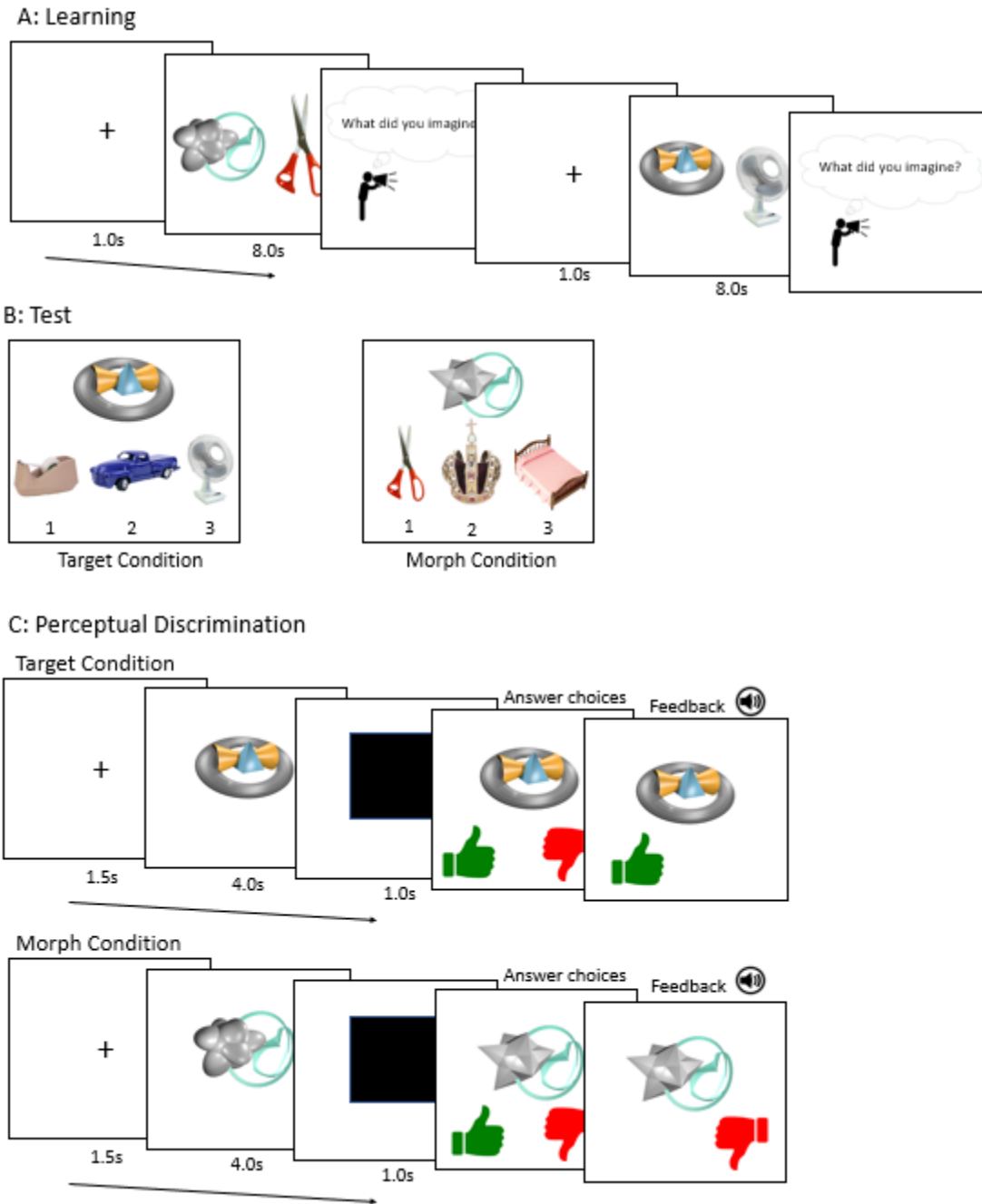
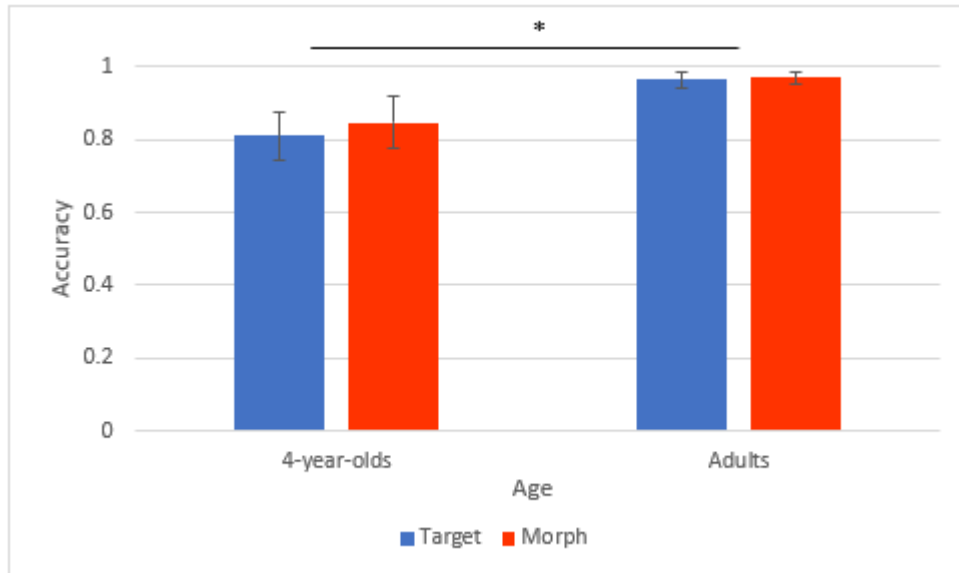


Figure 1. Schematic of each task phase. A: Associative learning phase. Participants were presented with a series of pairs consisting of a novel object and a common object and asked to form an association between them through imagining how they might go together. B: Associative retrieval phase. Target and shape morphs served as cues to test for differences in memory retrieval as a function of whether participants attempted to retrieve the associate through an identical or nonidentical item. The participant selected the common object paired associate from among three choices. C: Perceptual discrimination task. Participants viewed a target novel object and were asked to judge whether the subsequent object was the same (i.e., a target) or different (i.e., a shape morph) as compared to the first image. Audio and visual feedback were provided after each response.

Panel A: Imagination Success



Panel B: Perceptual Discrimination Success

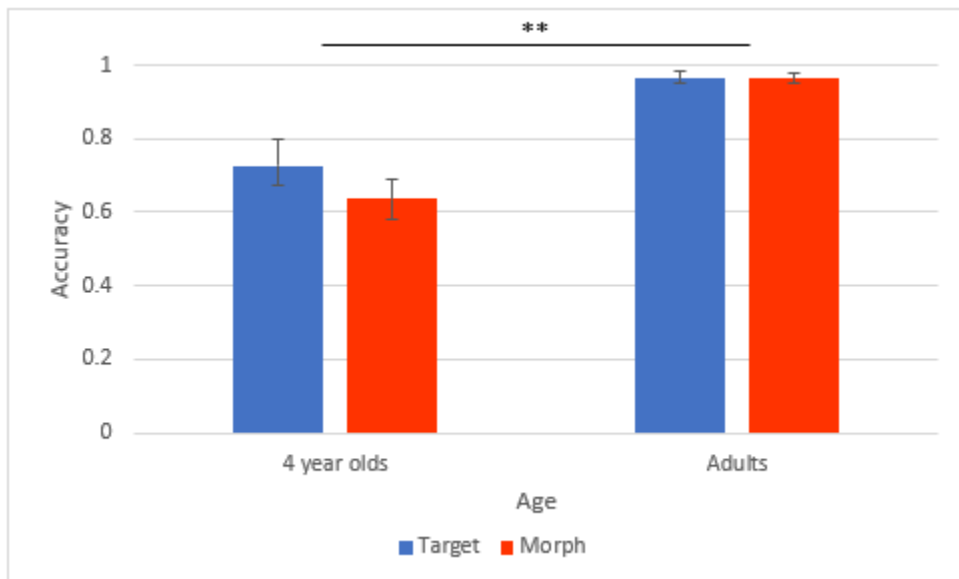


Figure 2. Panel A: Imagination success. Adults were more successful at imagining the pairs of objects together than 4-year-olds. Within each age group, imagination success did not differ between the target and morph conditions. Panel B: Perceptual discrimination accuracy. Adults were more successful at discriminating between target and morph objects. Perceptual discrimination success did not differ between target and morph conditions for either age group. Error bars reflect one standard error of the mean. (* represents $p < .05$, ** represents $p < .001$)

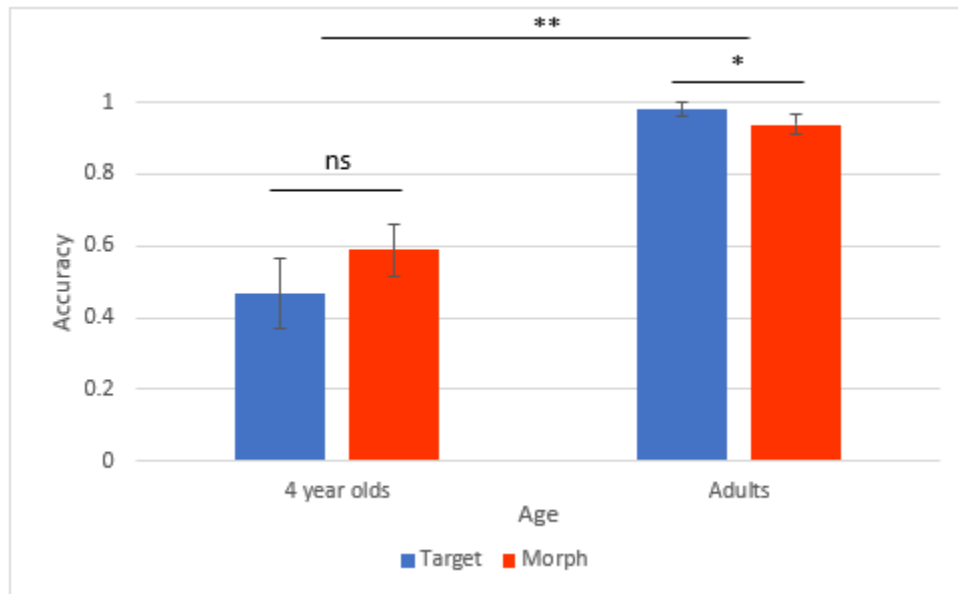


Figure 3. Accuracy in the three-alternative forced-choice task when controlling for encoding and perceptual success. Four-year-olds had lower overall retrieval success than adults. For the 4-year-olds, there was no significant difference in accuracy between the target and morph conditions. Adults were more accurate in the target condition than in the morph condition. (represents $p < .05$, ** represents $p < .001$)*

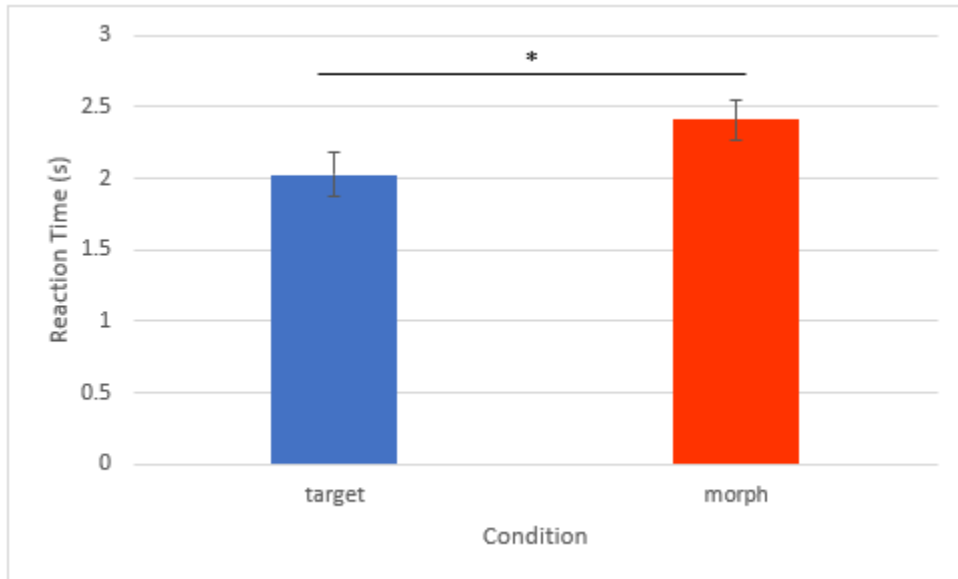


Figure 4. Adult response times in the three-alternative forced-choice test for the target versus morph conditions. Adults responded significantly faster in the target condition than in the morph condition. (represents $p < .05$)*

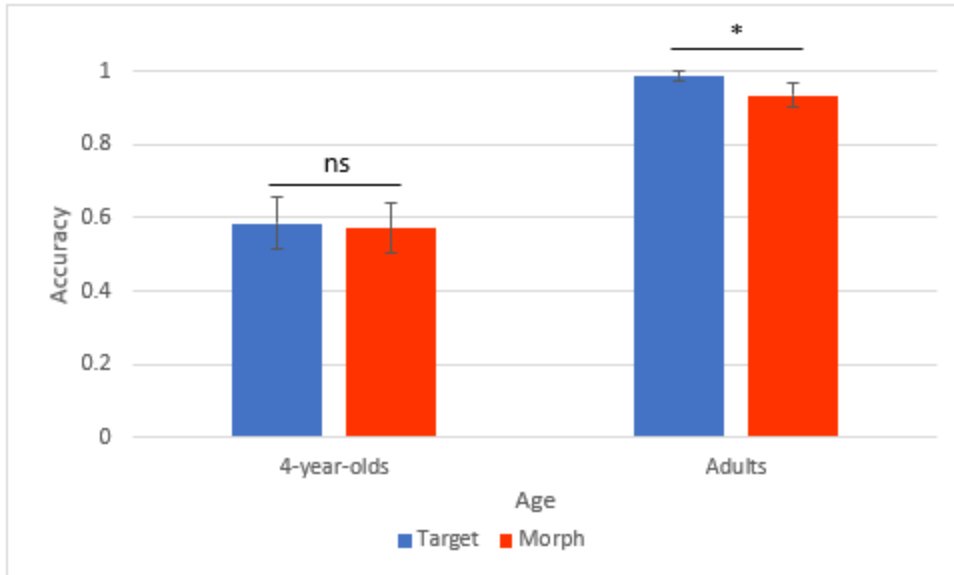


Figure 5. Accuracy in the 3-alternative forced-choice task when controlling only for encoding imagination success. There was no significant difference in accuracy in 4-year-olds between the target and morph conditions. Adults were more accurate in the target condition than the morph condition. (* represents $p < .05$)

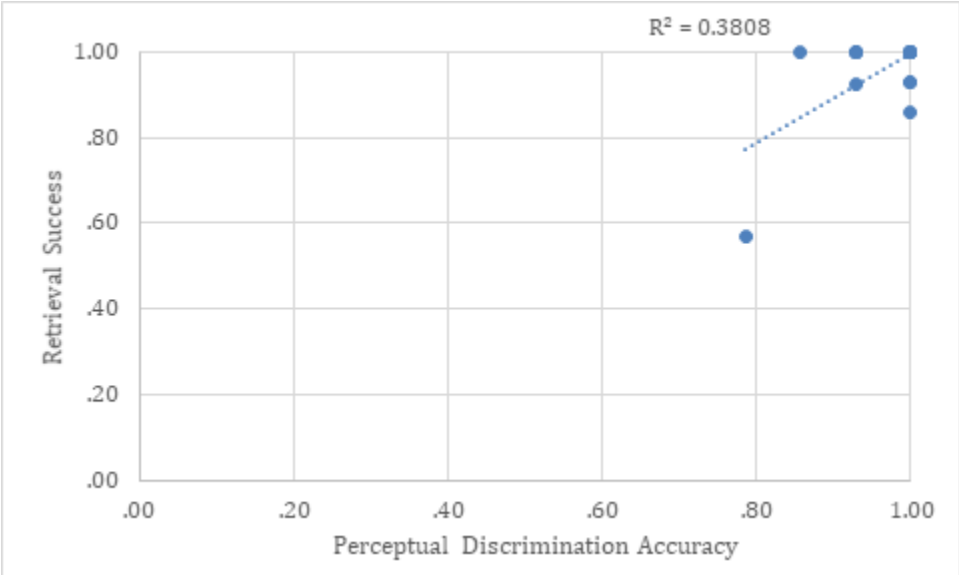
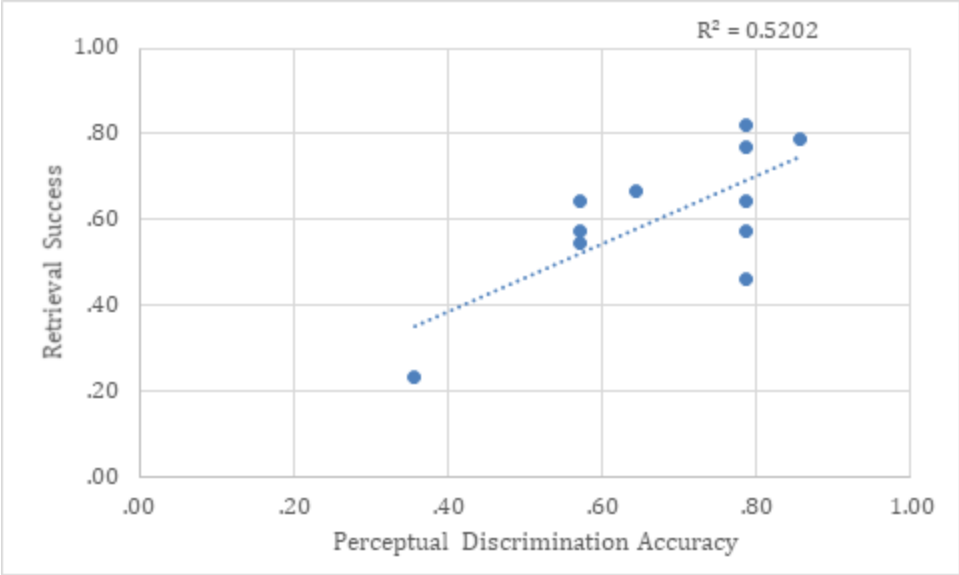


Figure 6. Top: Correlation in 4-year-olds between overall accuracy in the three-alternative forced-choice test and the perceptual discrimination task. Bottom: Correlation in adults between overall accuracy in the three-alternative forced-choice test and the perceptual discrimination task. Only trials with successful imagination were used to calculate the test accuracy.