Developmental Science

Developmental Science 17:5 (2014), pp 743–756 DOI: 10.1111/desc.12162

PAPER

Two rooms, two representations? Episodic-like memory in toddlers and preschoolers

Nora S. Newcombe, Frances Balcomb, Katrina Ferrara, Melissa Hansen and Jessica Koski

Department of Psychology, Temple University, USA

Abstract

Episodic memory involves binding together what-where-when associations. In three experiments, we tested the development of memory for such contextual associations in a naturalistic setting. Children searched for toys in two rooms with two different experimenters; each room contained two identical sets of four containers, but arranged differently. A distinct toy was hidden in a distinct container in each room. In Experiment 1, which involved children between 15 and 26 months who were prompted with a very explicit cue (a part of the hidden toy), we found a marked shift in performance with age: while 15- to 20-month-olds concentrated their searches on the two containers that sometimes contained toys, they did not distinguish between them according to context, but 21–26-month-olds did. However, surprisingly, without toy cues, even the youngest children showed a fragile ability to disambiguate the two containers by room context. In Experiment 2, we tested 34- to 40-month-olds and 64- to 72-month-olds without toy cues. The 5-year-olds were nearly perfect, and the 3-year-olds showed a significant preference for the correct container given only the context. In Experiment 3, we filled in the age range, and also investigated the effects of the use of labels (i.e. names of experimenters and rooms) and of familiarization time, in groups of 34- to 40-month-olds, 42- to 48-month-olds, and 50- to 56-month-olds. Neither labels nor familiarization time had an effect. Across experiments, there was regular age-related improvement in context-based memory. Overall, the results suggest that children's episodic memory may undergo an early qualitative change, yet to be precisely characterized, and that continuing increments in the use of contextual cues occur throughout the preschool period.

A video abstract of this article can be viewed at https://www.youtube.com/watch?v=DkwEFw0UEz4&list=PLwxXcOKHPC0llAPVcJyW4EtzlA934A2Rz&index=1

Introduction

Episodic memory is a unique form of explicit memory, binding together aspects of context to represent highly specific information, often autobiographical in nature (e.g. Tulving, 1972). In humans, episodic memory is thought to involve context-specific binding (what-where-when). In addition, Tulving (1983) proposed the criterion of an awareness of one's own presence in the episodic memory (autonoetic consciousness). However, consciousness is hard to evaluate in very young children, who are largely non-verbal, and who do not show clear evidence of being aware of themselves as agents who exist across time (Povinelli, Landau & Perilloux, 1996; Povinelli, Landry, Theall, Clark & Castille, 1999). An alternative view of episodic memory emphasizes a 'minimalist approach' focusing on phenomenological experience without a consciousness criterion (Russell & Hanna, 2012).

Episodic memory should be distinguished from another kind of explicit memory, namely semantic memory, which does not require contextual binding. The investigation of the neural bases of episodic memory has focused on a circuit involving various cortical regions together with the hippocampus and its associated areas (Cohen & Eichenbaum, 1993; Wheeler, Stuss & Tulving, 1997). In contrast, although there is some controversy, considerable evidence suggests that semantic memory is not dependent on, or at least is less dependent on, the hippocampal formation, based in part on findings that children with hippocampal damage can develop relatively normal semantic memory while

Address for correspondence: Nora S. Newcombe, 318 Weiss Hall, Temple University, Philadelphia, PA 19122, USA; e-mail: newcombe@temple.edu

showing significant deficits in episodic memory (e.g. Bindschaedler, Peter-Favre, Maeder, Hirsbrunner & Clarke, 2011; Vargha-Khadem, Gadian, Watkins, Connelly, Van Paesschen & Mishkin, 1997).

Infants and toddlers clearly possess the ability to form explicit memories; the clearest evidence comes from studies using deferred imitation (e.g. Meltzoff, 1988; Barr, Dowden & Hayne, 1996). However, these memories may be semantic (concerning generalized knowledge not tied to a specific event) rather than episodic, as argued by Mandler (2002) and Newcombe, Lloyd and Ratliff (2007). Given the hippocampal substrate of episodic memory, it is an important fact that success in hippocampal-dependent spatial memory tasks such as place learning does not emerge until towards the second birthday (Newcombe, Huttenlocher, Drummey & Wiley, 1998; Ribordy, Jabès, Banta Lavenex & Lavenex, 2013; Sluzenski, Newcombe & Satlow, 2004). To the extent that spatial memory is an important component of episodic memory (Russell, Cheke, Clayton & Meltzoff, 2011), this observation suggests that episodic memory is not likely to emerge before the end of the second year of life at the earliest.

Such an age transition is supported by the fact that a period of dense infantile amnesia appears to lift at around the age of 2 years (e.g. Eacott & Crawley, 1998; for review see Newcombe *et al.*, 2007), and by studies suggesting an increase in episodic memory performance at the end of the second year. For example, Bauer and Lukowski (2010) observed memory for details of an event and long-term recollection among 20-month-olds, but not 16-month-olds. As another example, Bauer and Leventon (2013) found that 13-month-olds required multiple experiences to show elicited imitation over long delays, but that by 20 months, children could recall temporal order, even after a single exposure. Imaging research adds to these behavioral findings, showing a sharp increase in hippocampal volume during the first 2 years of human life (Utsunomiya, Takana, Okazaki & Mitsudome, 1999). More detailed anatomical work with non-human primates also supports a protracted timetable of hippocampal development but with an important inflection that would correspond to the second birthday in humans (Lavenex & Banta Lavenex, 2013).

Episodic memory is clearly not at an adult level when it first emerges. Memories that involve the binding of objects or facts to locations or sources seem to develop substantially across the preschool years (Drummey & Newcombe, 2002; Lloyd, Doydum & Newcombe, 2009; Sluzenski, Newcombe & Kovacs, 2006), as does the ability to recall events in detail (Hayne, Gross, McNamee, Fitzgibbon & Tustin, 2011) and to recall temporal order (Riggins, Miller, Bauer, Georgieff & Nelson, 2009). At the neural level, there are developmental changes in hippocampally associated brain regions throughout childhood, albeit at a slower pace than during the first 2 years of life (Gogtay, Nugent, Herman, Ordonez, Greenstein, Hayashi, Clasen, Toga, Giedd, Rapoport & Thompson, 2006; Lavenex & Banta Lavenex, 2013; Richmond & Nelson, 2007; Utsunomiya *et al.*, 1999). In fact, changes in episodic memory appear to continue well into the elementary school years (Bauer, Doydum, Pathman, Larkina, Güler & Burch, 2012; Ghetti & Angelini, 2008; Picard, Cousin, Guillery-Girard, Eustache & Piolino, 2012; Rhodes, Murphy & Hancock, 2011; Shing & Lindenberger, 2011; Townsend, Richmond, Vogel-Farley & Thomas, 2010), as does the development of the relevant neural substrates (Ghetti, DeMaster, Yonelinas & Bunge, 2010; Ghetti & Bunge, 2012; Güler &Thomas, 2013; Ofen, 2012).

The developmental origins of episodic memory are clearly difficult to probe with standard memory paradigms that rely on verbal report. However, it is possible to test whether young children can remember specific contextual associations, based on a one-time experience (an episode). Work with non-human animals (henceforth termed animals) provides a framework for using nonlinguistic experimental methods to explore contextual binding and episodic-like memory. We are aware of two prior efforts to adapt these non-verbal paradigms for use with young children.

Russell *et al.* (2011) looked at what-where-when binding in 3-, 4- and 5-year-olds, in a future-oriented task based on research with scrub jays (e.g. Clayton & Dickinson, 1998). Children were familiarized with the effects of hot and cold storage boxes on chocolate or cookies over short and long delay times, i.e. that the chocolate melts in the hot box if the delay is long. They were then asked either to predict what box they would choose if they had to leave for a short or a long time, or to decide what food they wanted to put in the hot box if they had to leave for a short or a long time. The 3-year-old children did very poorly, although performance improved with age, but even 5-year-olds did not do very well, perhaps because inhibiting the desire to get chocolate, melted or not, was challenging to preschoolers.

Hayne and Imuta (2011) took a different approach, studying what-where-when memory in 3- and 4-year-old children using a hide-and-seek paradigm 'designed with the scrub jay procedure in mind' (p. 318). Children selected three toys to hide in three different rooms in their own homes. After 5 minutes of book reading, the children were first asked verbally about the order in which they went to each room, what toy they hid in each room and exactly where in each room that toy was hidden. They were then asked to take the experimenter on a tour to retrieve the toys, giving nonverbal evidence of memory. Four-year-olds were better than 3-year-olds at verbal recall, as well as showing better behavioral recall for the order of rooms. The other behavioral measures were at ceiling.

This experiment is intriguing, but it invites follow up. First, conducting research in children's homes may be problematic; extremely familiar spatial contexts seem to support different patterns of performance from novel contexts, both in studies of spatial memory in infants (Feldman & Acredolo, 1979) and episodic memory (Hupbach, Gomez & Nadel, 2009). Second, once children enter a particular room, the hiding location is unique to that room (e.g. Big Bird is under the purple couch). Thus, there is no need to bind together particular cues and particular contexts in a contingent relational fashion. Third, it would be desirable to be able to test children as young as 18 months, in order to assess whether there is a discontinuity in binding towards the end of the second year.

One specific paradigm developed for use with animals provided the point of departure for the current work. Eacott and Norman (2004) capitalized on rats' natural instinct to explore novel items by testing whether rats could detect changes in the overall configuration of objects in context, when the contexts and objects themselves were equally familiar. In this task, rats were familiarized with two objects (A and B) in two enclosures that varied in features. Rats encountered the objects arranged in one position (AB) in context 1, and those two objects set in opposite positions (BA) in context 2 (see Figure 1). At test, rats entered one of the enclosures, which now held two duplicate objects (AA or BB). Note that the rats had seen both objects in both locations. Thus, if they possessed only a general (semantic) representation, they would not be expected to notice the novel placement of one of the objects. However, if rats had specific representations of each object's location given each context, they should notice the object that was novel in the specific context. Rats with lesions to the fornix (the output pathway from the hippocampus) did not distinguish between the two objects, spending equal amounts of time exploring each. However, intact rats spent a significantly greater proportion of time exploring the novel object/context than the familiar object/context, suggesting that they had formed specific associations between the objects, their spatial locations, and the context in which they appeared.

Adapting this work for use with human children, we explored contextual episodic memory in toddlers in a series of three experiments. In all experiments, children learned different configurations of the same set of objects, arranged in unique arrays and paired with different hidden toys, different people and different rooms. In Experiment 1, children were provided with an explicit memory cue as well as contextual cues, but in Experiments 2 and 3, they were only cued by context. Unlike the rats, children did not receive extensive familiarization, so we did not predict a preference for novelty. Instead, we probed for contextually dependent memory for the location of a toy in one container in each room. To preview, data from these experiments suggest that the ability to make arbitrary object-to-object associations first emerges at about 20 months, based on findings with explicit recall cues. However, when there are less direct linkages, evidence of episodic associations is actually seen earlier, although it is fragile at first, and strengthens gradually across the preschool years.

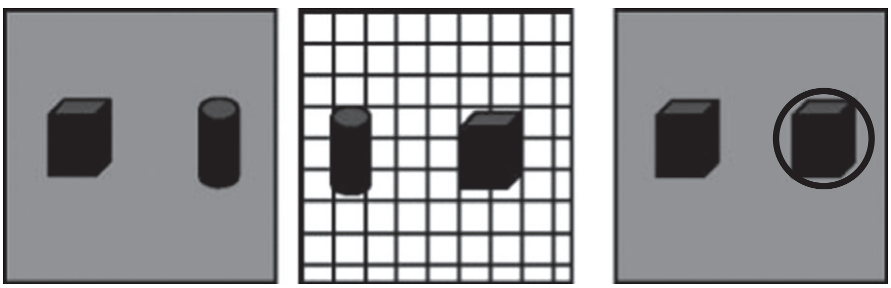

Figure 1 *Rats were familiarized with both objects in both rooms, and at test were presented with duplicate objects in one of the rooms. Note that both objects had been seen in both positions, and thus only the cube on the right was novel, given its context. Adapted from 'Integrated memory for objects, place, and context in rats: a possible model of episodic-like memory?', by M.J. Eacott & G. Norman (2004) Journal of Neuroscience, 24, p. 1949. Copyright 2004 by The Society for Neuroscience.*

Experiment I

Methods

Participants

The participants were 15- to 20-month-old children (n = 32, M = 18.85, SD = 1.40) and 21–26-month-old children (n = 32, M = 23.26, SD = 1.72).

Materials

Each experiment used two rooms, called the rainbow room (RR) and the cloud castle (CC), two toys, two experimenters, and two sets of four containers (cylinder, box, basket and bag), with one set of containers per room. Both rooms held the same four containers, which were arranged differently in each room. There were four different toys used for the experiment (bubbles, shape sorter, blocks, and crayons). All of these toys had multiple component parts (e.g. the bubble bottle and the bubble wand). Only two toys were used with each participant, chosen based on the parent's recommendation. One toy was hidden in a different container in each room (e.g. the crayons in the basket in the rainbow room, and the blocks in the cylinder in the cloud castle). For each participant, the toys, target containers, and experimenters were randomly assigned to RR and CC. Each participant therefore experienced two contexts that shared the four containers, but differed in room context, container arrangement, toys, and experimenter. In addition, each room differed in regard to the container in which the toy was hidden (see Figure 2).

Procedures

There were six trials: two familiarization trials (one per room) and four test trials (two per room). Every trial began with an experimenter walking into a room with the child, and every trial ended with the other experimenter knocking on the door. The child then exited the room with the new experimenter and walked across a short hallway to the other room.

Familiarization. The two familiarization trials lasted 6 minutes each, one in CC and one in RR. Familiarization began by identifying the room and experimenter, and asking the child to search for a hidden toy (e.g. 'My name is Frances and we are going to the Cloud Castle! Can you find something to play with in here?'). For 4 minutes, children were allowed to explore the room freely, opening the containers in any order and as many times as they wished. After 4 minutes, if children had not opened all four containers, the experimenter cued the children to approach each container and open it to look inside. After finding the toys, the experimenter and children played in the immediate vicinity of the target container. Note that by the end of the familiarization trials, children had witnessed, in each room, the retrieval of the toy from the target container, as well as the fact that the other three containers were empty. Familiarization ended when there was a knock on the door. Children were asked to help clean up, including replacing the toy within the container in which they found it, and then the door was opened, revealing the second experimenter (e.g. 'Look! It's Katrina!'). The second experimenter then asked the children to join her in the other room (e.g. 'Let's go to the Rainbow Room!'), and together they walked to the second room to begin the second familiarization trial.

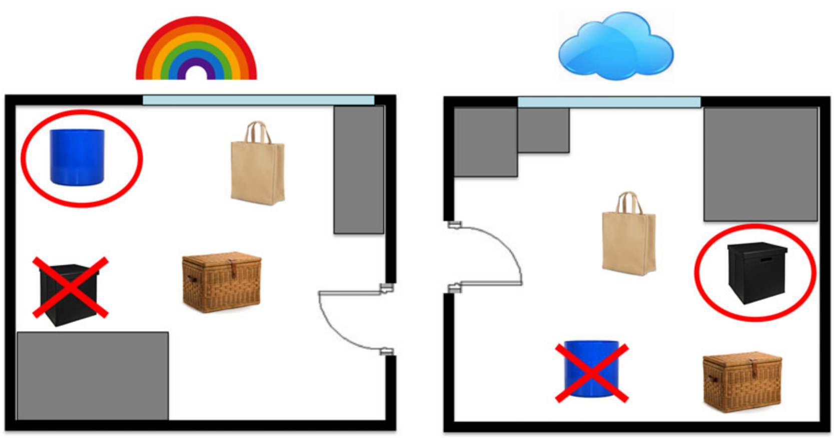

Figure 2 *Schematic diagram of test rooms, with features and objects. The circled object denotes the target container in each room, while the 'X' denotes the container that was the target in the other room.*

Test trials. There were four test trials lasting 2.5 minutes each. Children were assigned to either the Toy Cue condition, or the No Toy Cue condition. In the Toy Cue condition, the experimenter and the child entered the room and the experimenter handed the child a part of the toy (e.g. the bubble wand, if the toy was bubbles) and prompted the child to search for the toys by saying, 'Do you remember where the [bubbles] are?' In the No Toy Cue condition, the experimenter did not hand the child a part of the toy, but asked, 'Do you remember where the toys are in here?'

Regardless of condition, children searched for 2.5 minutes, opening any containers as many times as they liked. If, after 2.5 minutes, they had not found the toy, the experimenter provided a hint by opening the correct container, then re-closing it, and encouraging the child to come and discover the toy. Thus, the length of each trial occasionally extended beyond 2.5 minutes when children did not find the toys on their own, but each trial ended by the child approaching the target container and retrieving the toy. The test trial was over when there was a knock on the door, and children left the room with the second experimenter. Together they walked to the second room to begin the next test trial.

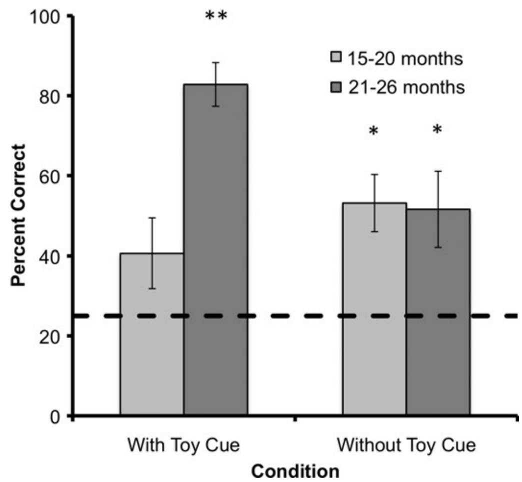

Figure 3 *Percent of correct searches with and without a toy cue made by the younger (15–20 months) and older (21–26 months) children in Experiment 1. Chance at 25% is indicated by the dotted line. Error bars represent standard error of the mean. Asterisks indicate significance relative to chance (*) or significance relative to chance and the other age group (**).*

Data coding

All six trials were videotaped and coded offline. Children's first searches were coded 1 if correct, and 0 if incorrect. To determine overall accuracy, scores were summed for the four searches conducted across four test trials and the total percent correct was calculated (total correct). In addition, scores for the first two test trials (first test in RR and first test in CC) were summed as the first test trial score.

Errors were identified as being either semantic errors or random errors. A semantic error occurred when the child selected the container that marked the correct location of the toy in the *other* room, and a random error occurred when the child selected a container that was not correct in either room.

Results

We began by testing for gender differences. There was no significant difference in performance between boys and girls in either of the age groups tested. Hence, the analyses are collapsed across gender, both in this experiment and also in the two subsequent studies, because there were no gender differences in any of the data (all ps > .05).

An overall ANOVA (age group by cue) showed a main effect of age ($F(1, 60) = 6.61$, $p = .01$) and an age by cue interaction ($F(1, 60) = 7.67$, $p < .01$), as displayed in Figure 3. Younger children's performance did not significantly differ between the No Toy Cue and Toy Cue conditions ($t(30) = 1.1$, $p = .28$). For older children, performance was significantly higher with the toy cue ($t(30) = 2.84$, $p = .008$).

We explored the data further by conducting analyses within each of the two cue conditions. There were three phases to this exploration: (1) examining age differences in correct choice and comparing performance at each age to chance levels of 25%; (2) testing for semantic memory in two ways: determining whether choices of the two containers that sometimes contained toys exceeded choices of the two containers that never contained toys, and determining whether semantic errors exceeded random errors (with the latter divided by 2 so that the maximum possible was always 25%); (3) testing for episodic memory by comparing correct choices to semantic errors.

Toy Cue condition

Age differences and comparisons to chance. There was a significant positive correlation between age and total correct ($r = .70$, $p < .001$). Similarly, an independent samples t-test showed that older children (21–26 months) were significantly more accurate ($M = 83\%$, $SD = 22\%$) across the four test trials (total correct) than younger children ($M = 41\%$, $SD = 35\%$, $t(30) = 4.07$, $p < .001$, two-tailed, $d = 1.48$; see Figure 3). To ensure that this effect was not merely due to a carry-over of learning on the first test trials, performance on the first two test trials alone (the first test in CC and in RR) was also compared between age groups. Again, older children were significantly more accurate (M = 81%, SD = 31%) than younger children (M = 34%, SD = 35%, $t(30)$ = 3.99, $p < .001$, d = 1.47; see Figure 3). Older children performed significantly better than chance levels of 25% (based on the presence of four containers), ($t(15)$ =10.59, $p < .001$, d = 2.65), but younger children did not quite exceed chance ($t(15)$ = 1.78, p = .10, d = .44).

Semantic memory. Although the younger children were not quite above chance overall, there is evidence that they had learned something about which two containers of the four were the ones that held the toys. Responses were coded as 'semantic correct' if the child picked either the correct location in the test room or the correct location in the other room. This total was compared to choices of the two containers that never contained toys. Paired t-tests revealed that both younger children (semantic correct: M = 70%, SD = 28%; incorrect: M = 28%, SD = 24%; $t(15)$ = 3.51, p = .003, d = 1.63) and older children (semantic correct: M = 97%, SD = 09%; incorrect: M = 2%, SD = 6%; $t(15)$ = 28.04, $p < .001$, d = 12.73) made significantly more semantic correct choices than completely incorrect choices. In addition, the younger children made significantly more semantic errors (M = 29%, SD = 23%) than random errors (M = 14%, SD = 12%, $t(15)$ = 2.48, p = .03, d = .82). That is, they often chose the container in which the object was hidden in the other room. This pattern was also evident for the older children (semantic: M = 14%, SD = 18%, random: M = 1%, SD = 3%, $t(15)$ = 2.96, p = .01, d = 1.01). Thus, it seems that even very young children can form explicit memories of a semantic nature, but they seem to struggle to form specific item-context pairs, i.e. contextualized memories.

Additional analyses examined whether errors (semantic or random) differed between age groups. Overall, younger children made more errors, and this occurred for both random errors (M = 14%, SD = 12% for younger and M = 1%, SD = 3% for older, $t(30)$ = 4.30, $p < .001$,, d = 1.51), and semantic errors (M = 30%, SD = 23% for younger and M = 14%, SD = 18% for older, $t(30)$ = 2.15, p = .04,, d = .79).

Episodic memory. When we evaluated episodic memory by comparing correct responses to semantic errors, paired samples t-tests revealed that older children made significantly more correct choices than semantic errors (correct: M = 83%, SD = 22%; semantic errors: M = 14%, SD = 18%; $t(15)$ = 7.00, $p < .002$, d = 3.42), but younger children did not (correct: M = 41%, SD = 35%; incorrect: M = 30%, SD = 23%; $t(15)$ = .84, p = .42, d = .38).

No Toy cue

Age differences and comparisons to chance. There was no correlation between age and total number of choices of the correct container (r = .23, p = .12), and older children did not differ significantly in search accuracy (M = 51%, SD = 38%) from younger children (M = 53%, SD = 29%, $t(30)$ = .13, p = .90,, d = .03; see Figure 3). When performance on the first two test trials was isolated, search accuracy did not differ significantly between older (M = 47%, SD = 43%) and younger children (M = 50%, SD = 37%, $t(30)$ = .22, p = .83,, d = .08; see Figure 3). Comparison to chance at 25% showed that both younger ($t(15)$ = 3.92, p = .001, d = .98) and older children ($t(15)$ = 2.79, p = .01, d = .70) performed significantly above chance.

Semantic memory. As in the Toy Cue condition, there was evidence of strong semantic memory. First, semantic errors predominated over random errors at both ages: younger children made significantly more semantic errors (M = 28%, SD = 22%) than random errors (M = 9%, SD = 15%, $t(15)$ = 2.40, p = .03,, d = 1.01), as did older children (semantic: M = 31%, SD = 30%, random: M = 8%, SD = 9%, $t(15)$ = 3.38, p = .004, d = 1.04). Second, paired t-tests comparing the percent of semantic correct responses to incorrect responses revealed that both younger children (semantic correct: M = 81%, SD = 30%; incorrect: M = 19%, SD = 30%; $t(15)$ = 4.23, p = .001, d = 2.10) and older children (semantic correct: M = 83%, SD = 18%; incorrect: M = 16%, SD = 18%; $t(15)$ = 7.67, $p < .001$, d = 3.78) made significantly more semantic correct choices than completely incorrect choices. Thus, even when no toy cue was provided, there was evidence of fundamental understanding of the task, and an ability to encode and retain information about the containers that held toys.

Younger children were not significantly more likely to make random errors (M = 9%, SD = 15%) than older children (M = 8%, SD = 9%, $t(30)$ = .361, p = .72,, d = .08), nor were they significantly more likely to make semantic errors (M = 28%, SD = 22%) than older children (M = 31%, SD = 30%, $t(30)$ = .34, p = .74,, d = .16).

Episodic memory. Oddly, evaluating episodic memory with paired samples t-tests comparing correct responses to semantic errors revealed that older children did not make significantly more correct choices than semantic errors (correct: M = 52%, SD = 38%; semantic errors: M = 31%, SD = 30%; $t(15)$ = 1.23, p = .24, d = .59), but younger children did (correct: M = 53%, SD = 29%; semantic errors: M = 28%, SD = 22%; $t(15)$ = 2.39, p = .03, d = .98). However, the means are very similar across the two ages, and the lower effect size for the older age group seems mainly due to the somewhat greater variability.

Discussion

The results from the Toy Cue condition suggest that children show significant improvement in the second year of life in their abilities to make arbitrary associations between objects (i.e. the toys and their containers), and to recall specifically which container held which toy in which context, which requires individuating two highly similar memories. However, analysis of error types demonstrates that the children across this age range had abstracted a more general kind of information about which containers were likely to contain toys. This analysis also shows that the children understood the task and were motivated to find the toys. These patterns of performance suggest that early explicit memory is primarily semantic, but with a noticeable improvement in the binding of elements basic to episodic memory in the second half of the second year, at around the age of 20 or 21 months. Interestingly, this change occurs during an age range when the hippocampus is known to undergo significant developmental changes (Lavenex & Banta Lavenex, 2013).

When there is a lower level of explicit cueing, provided by room and person context alone rather than from the sight and name of the toy associated with the container, a different picture emerged. The contextual information appeared to support fragile episodic recall in the younger but not the older children. However, although the younger group showed barely significant differences and the older group just missed showing such differences, there was no significant age-related difference, and in fact mean performances for the two age groups were almost identical. It is somewhat puzzling why the younger children seemed to do better without the toy cue; the safest conclusion at present is perhaps that even young children, in the age range of 15 to 20 months, have a weak ability to form episodic memories. Perhaps there are different kinds of binding, with somewhat different developmental trajectories.

One possible alternative explanation for the choice of the two containers that sometimes contained toys is that choices are based on implicit memory. Children might develop a general feeling that they ‘like’ some containers more than others. However, while we cannot definitely exclude this possibility, it seems unparsimonious, because it is known from studies of deferred imitation that children of these ages form explicit memories of some kind, and because an explicit episodic response can be elicited, at least in the older group of children, by the toy cue.

Experiment 2

Experiment 1 yielded only fragile evidence that 15–26-month-old children could display episodic-like memories in the absence of tightly associated cues (i.e. prompting regarding the hidden toy). Thus, we were interested in finding out exactly *when* this ability emerges. In Experiment 2, we tested two groups of older children to determine when they would be capable of displaying more robust episodic memory in our paradigm.

Method

Participants

The participants were 34- to 40-month-olds (n = 19, M = 37.27, SD = 1.33) and 64–72-month-olds (n = 8, M = 67.01, SD = 2.43). None of the children had participated in Experiment I.

Materials

The materials were the same as in Experiment I, except that more age-appropriate toys were used to maintain the interest and engagement of the older children (paint markers and marbles).

Procedures

The procedures were the same as the No Toy Cue condition in Experiment I, with one exception: since the children were older, we reduced the duration of the familiarization and testing phases to avoid boredom (3 minutes for familiarization, 1.5 minutes for each test trial).

Data coding

Data coding was the same as in Experiment I.

Results

Age differences and comparisons to chance

There was a significant correlation between age and total correct (r = .57, p = .002). Similarly, older children (64–72-month-olds) were significantly more accurate (M = 97%, SD = 9%) across the four test trials than younger children (34–40-month-olds; $M = 53\%$, $SD = 32\%$, $t(23.12) = 5.52$, $p < .001$, $d = 1.67$). To ensure that this effect was not merely due to learning, performance on the first two test trials (the first test in CC and in RR) was also analyzed. Again, older children were significantly more accurate ($M = 94\%$, $SD = 18\%$) than younger children ($M = 45\%$, $SD = 37\%$, $t(25) = 3.56$, $p = .002$, $d = 1.55$). With chance set at 25%, both older ($t(7) = 23.00$, $p < .001$, $d = 8.17$) and younger children ($t(18) = 3.75$, $p = .001$, $d = .86$) performed significantly above chance.

Semantic memory

When combining correct and semantic errors for older children accuracy was at 100%, so a paired *t*-test comparing the percent of semantic correct responses to incorrect responses was only conducted for the younger children. The analysis revealed that younger children made significantly more semantic correct choices than incorrect choices (semantic correct: $M = 79\%$, $SD = 30\%$; incorrect: $M = 20\%$, $SD = 30\%$; $t(18) = 4.25$, $p < .001$, $d = 1.94$). Younger children made significantly more semantic errors ($M = 26\%$, $SD = 24\%$) than random errors ($M = 10\%$, $SD = 15\%$, $t(18) = 2.17$, $p = .04$, $d = .80$). Older children performed nearly at ceiling ($M = 97\%$, $SD = 9\%$), and the difference between error types was hence non-significant (semantic $M = 3\%$, $SD = 9\%$; random $M = 0\%$, $SD = 0\%$, $t(7) = 1.00$, $p = .35$, $d = .47$).

Younger children were significantly more likely to make semantic errors ($M = 26\%$, $SD = 24\%$) than older children ($M = 3\%$, $SD = 9\%$, $t(25) = 2.61$, $p = .015$, $d = 1.14$), but were not more likely to make random errors ($M = 10\%$, $SD = 15\%$) than the older children ($M = 0\%$, $SD = 0\%$, $t(25) = 1.80$, $p = .08$, $d = .82$).

Episodic memory

Younger as well as older children showed evidence of episodic memory, albeit at lower levels. Paired samples *t*-tests comparing correct responses to semantic errors revealed that both age groups made significantly more correct choices than semantic errors: older children (correct: $M = 97\%$, $SD = 9\%$; semantic errors: $M = 3\%$, $SD = 9\%$; $t(7) = 15.00$, $p < .001$, $d = 10.66$) and younger children (correct: $M = 53\%$, $SD = 32\%$; semantic errors: $M = 26\%$, $SD = 24\%$; $t(15) = 2.39$, $p = .03$, $d = .92$).

Discussion

These results show that children at 34–40 months are forming episodic memories and accessing them based on contextual cues in a way that was not reliably evident in the 21–26-month-olds in the no toy condition of Experiment 1 (although note that the effect was reliable for the 15- to 20-month-olds in Experiment 1). However, the 34–40-month-olds (as well as the younger children in Experiment 1) performed at substantially lower levels than the older children; by 64–72 months, children performed nearly perfectly on this task. Overall, the message of the two experiments seems to be that contextual associations not tightly bound to the container are sufficient to support only fragile episodic memory across the age range from 15 to 40 months.

Experiment 3

In Experiment 2, the 34–40-month-olds performed at low but above-chance levels and the 64–72-month-olds were at ceiling. In Experiment 3, we sought to chart the development in episodic memory performance across the age gap between 3.5 and 5.5 years. We additionally explored two variations in our procedure that might have an effect on children's ability to bind together aspects of their experiences in an episodic fashion: labels and familiarization time. Past research has suggested the importance of labels and verbal reminders in facilitating young children's memories (Bauer, Wenner, Dropik & Wewerka, 2000; Imuta, Scarf & Hayne, 2013; Simcock & Hayne, 2002), so it is possible that the labels given to the rooms (e.g. 'Rainbow Room') and experimenters (e.g. 'Katrina') might serve as verbal cues that could improve performance. Thus, we were interested to see whether the use of labels would influence the age at which the developmental shift in performance on this episodic task occurs. In addition, although we had shortened the times in Experiment 2 to avoid boredom, we wondered whether providing children with a longer time to examine and search the room would enable them to better encode episodic memories for the task, and thus boost performance. While the 64–72-month-olds quickly performed the task in the shortened time frame (3 minutes for familiarization, 1.5 minutes for test), often with time to spare before moving on to the next room, younger children (34–40 months) might benefit from a longer familiarization period during which to interact with the items in the space.

Method

Participants

The participants were 34–40-month-olds ($n = 32$, $M = 36.92$, $SD = 1.48$), 42–48-month-olds ($n = 32$,

$M = 44.20$, $SD = 1.57$), and 50–56-month-olds ($n = 16$, $M = 52.85$, $SD = 1.92$), who had not participated in Experiments 1 or 2.

Materials

The materials were the same as in Experiment 2.

Procedures

The procedure was the same as Experiment 2, except that children participated in one of two conditions. In the Label condition (34–40-month-olds: $n = 16$; 42–48-month-olds: $n = 16$; 50–56-month-olds: $n = 8$), the procedure was identical to Experiment 2, where the rooms and experimenters were named (e.g. ‘Look, it’s Katrina!’/’Let’s go back to the Rainbow Room!’). In the No Label condition (34–40-month-olds, $n = 16$; 42–48-month-olds, $n = 16$; 50–56-month-olds, $n = 8$), the procedure was the same, but the rooms and experimenters were not named (e.g. ‘Look who’s at the door!’/’Let’s go back to the other room!’).

In addition, we used two familiarization times: a short time (3 minutes for familiarization, 1.5 minutes for test trial) and a long time (6 minutes for familiarization, 2.5 minutes for test trial). We tested 34–40- ($n = 32$) and 42–48-month-olds ($n = 32$) using the long familiarization time. Note that in Experiment 2, 34–40–month-olds were also tested using the short familiarization time, thus enabling us to compare performance within this age group across experiments based on length of familiarization. For 50–56-month-olds ($n = 16$), we used the short time to avoid boredom. (See Table 1 for an overview of which children were tested in which ways across the three experiments.)

Table 1 *Sample size (N) by age and condition for all experiments*

Age (months)	Long Familiarization Label	Long Familiarization No Label	Short Familiarization Label	Short Familiarization No Label
Experiment 1				
15–20	32			
21–26	32			
Experiment 2				
34–40			19	
64–72			8	
Experiment 3				
34–40	16	16		
42–48	16	16		
50–56			8	8
Total *N*	96	32	35	8

Note: 34– 40-month-olds in Experiment 2 were included in analysis in Experiment 3.

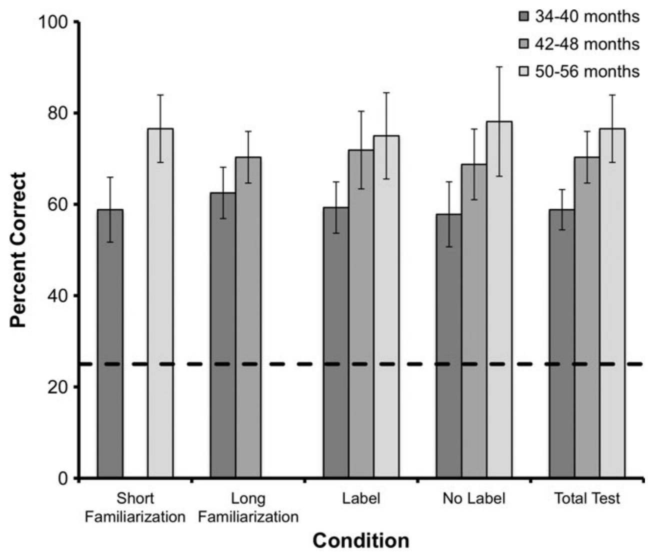

Figure 4 *Percent of correct searches by 34–40, 42–48, and 50–56-month-olds in Experiment 3 by condition (Short Familiarization, Long Familiarization, Label, and No Label), and overall (Total Test, which is collapsed across conditions). Chance at 25% is indicated by the dotted line. Note: 42–48-month-olds were only tested in the long familiarization condition, and 50–56-month-olds were only tested in the short familiarization condition. Error bars represent standard error of the mean.*

Data coding

Data coding was the same as in Experiments 1 and 2.

Results and discussion

Figure 4 and Table 1 provide an overview of the data from this experiment as well as the prior two experiments.

Effect of labels

We first looked at whether labels improved performance on the task. The was no significant difference in search accuracy between the Label and No Label condition for 34–40-month-olds (Label: $M = 67\%$, $SD = 34\%$; No Label: $M = 58\%$, $SD = 28\%$, $t(30) = .85$, $p = .40$,, $d = .30$), 42–48-month-olds (Label: $M = 72\%$, $SD = 34\%$; No Label: $M = 69\%$, $SD = 31\%$; $t(30) = .27$, $p = .79$,, $d = .10$), or 50–56-month-olds (Label: $M = 75\%$, $SD = 27\%$; No Label: $M = 78\%$, $SD = 34\%$; $t(14) = .21$, $p = .84$, $d = .10$), nor was there an effect of label collapsing across age groups (p = .32). These results suggest that although other research has emphasized the importance of language and labels in binding (Simcock & Hayne, 2002), labels did not aid children in the present study. Because there were no differences between the Label and No Label conditions, we collapsed them for further analyses.

Effect of familiarization time

We were also interested in whether familiarization time would affect performance. For this analysis, we compared the 34–40-month-olds in the short familiarization period from Experiment 2 to the 34–40-month-old children in Experiment 3 (see Table 1). The children with the short familiarization time (n = 19, M = 53%, SD = 32%) were not significantly more accurate in their searches than those with the long familiarization time (n = 32, M = 63%, SD = 31%, $t(49)$ = 1.08, p = .29, d = .33). Because there were no differences between these two groups, we also collapsed across familiarization time for further analyses (n = 51).

Age effects and comparisons to chance

There was a marginally significant correlation between age and total correct (r = .20, p = .05), and a one-way ANOVA of the effect of age showed only a weak trend ($F(2, 96)$ = 2.52, p = .09). As in the previous two experiments, we also compared performance on the first test trials only, finding no significant difference in search accuracy on the first test trials across age groups ($F(2, 96)$ = 1.02, p = .37). With chance set at 25% (given the four containers), all age groups performed significantly better than chance: 34–40-month-olds ($t(50)$ = 7.66, $p <$ .001, d = 1.07), 42–48-month-olds ($t(31)$ = 8.00, $p <$ .001, d = 1.42), and 50–56-month-olds ($t(15)$ = 6.98, $p <$ .001, d = 1.75).

Semantic memory

There was strong evidence of semantic memory. First, paired t-tests comparing the percent of semantic correct responses to incorrect responses revealed that 34–40-month-olds (semantic correct: M = 85%, SD = 25%; incorrect: M = 15%, SD = 25%; $t(50)$ = 10.22, $p <$ .001, d = 8.55), 42–48-month-olds (semantic correct: M = 88%, SD = 19%; incorrect: M = 12%, SD = 19%; $t(31)$ = 11.38, $p <$ .001, d = 4.03), and 50–56-month-olds (semantic correct: M = 97%, SD = 13%; incorrect: M = 3%, SD = 13%; $t(15)$ = 15.00, $p <$ .001, d = 7.50) made significantly more semantic correct choices than incorrect choices. Second, within each age group, we once again found a difference in the type of error committed. The 34–40-month-olds made significantly more semantic (M = 26%, SD = 25%) than random errors (M = 7%, SD = 12%, $t(50)$ = 4.34, $p <$ .001, d = .94), as did 42–48-month-olds (semantic: M = 18%, SD = 26%; random: M = 6%, SD = 10%, $t(31)$ = 2.52, p = .02, d = .64) and 50–56-month-olds (semantic: M = 20%, SD = 23%; random: M = 2%, SD = 6%, $t(15)$ = 3.50, $p <$ .01, d = 1.16).

There were no age differences for semantic errors ($F(2, 96)$ = 1.07, p = .35) or random errors ($F(2, 96)$ = 1.79, p = .17).

Episodic memory

Paired samples t-tests comparing correct responses to semantic errors revealed that all three age groups made significantly more correct choices than semantic errors: 34–40-month-olds (correct: M = 59%, SD = 32%; semantic errors: M = 26%, SD = 25%; $t(50)$ = 4.53 $p <$ .001, d = 1.14), 42–48-month-olds (correct: M = 70%, SD = 32%; semantic errors: M = 18%, SD = 26%; $t(31)$ = 5.41, $p <$ .001, d = 1.80), and 50–56-month-olds (correct: M = 77%, SD = 30%; semantic errors: M = 20%, SD = 23%; $t(15)$ = 4.39, p = .001, d = 2.14).

Overall analyses

It is possible to combine the data from these three studies to examine the overall developmental picture. The following analyses include all participants across all three experiments, except for those from Experiment 1 in the unique Toy Cue condition. A bivariate correlation revealed an overall significant linear trend across age groups, with children improving as they got older (r = .33, $p <$ .001). Similarly, a Brown-Forsythe F^* test revealed a significant difference in accuracy across age groups ($F^*(5, 85.50)$ = 4.24, p = .002). Post-hoc comparisons using the Games-Howell test to account for the heterogeneity of variances across age groups showed that the 64–72-month-olds (n = 8) were significantly different from the 15–20-month-olds (n = 16, $p <$ .001), the 21–26-month-olds (n =16, p = .003), the 34–40-month-olds (n = 51, $p <$ .001), and the 42–48-month-olds (n = 32, p = .003), but not from the 50–56-month-olds (n = 16). No other comparisons were significant (all ps > .15). These data suggest a gradual developmental transition in task performance across the preschool period; above-chance performance on this task was seen as early as 15 months, but, given the variance in performance, only 5-year-olds clearly differed from the younger groups, with 4-year-olds

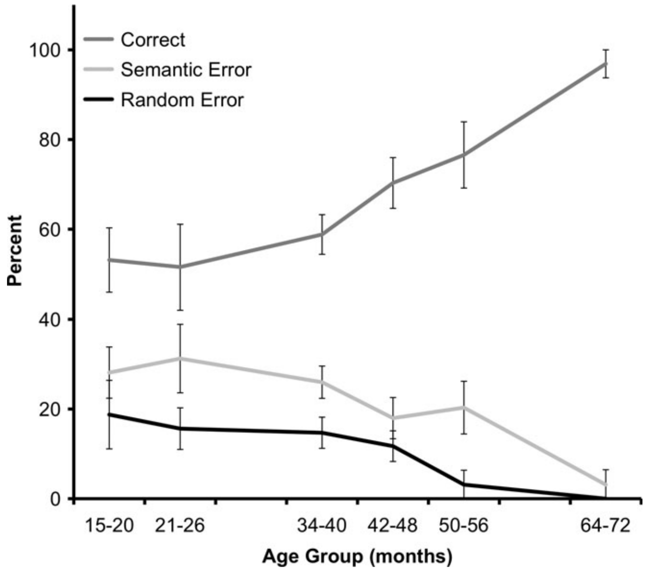

Figure 5 *Percent of correct responses, semantic errors, and random errors across all age groups tested in all three experiments. Error bars represent standard error of the mean.*

intermediate (i.e. not differing significantly either from 3- or from 5-year-olds).

Next, multiple regression analysis was used to examine whether age, familiarization time, and room label significantly predicted search accuracy across age groups. The overall model including all three predictor variables was significant ($F(3, 135) = 5.98$, $p = .001$, $R^2 = .12$). However, only age was a significant predictor of search accuracy ($t(138) = 3.92$, $p < .001$, $b = .01$), whereas room label ($t(138) = .83$, $p = .41$, $b = .05$) and familiarization time ($t(138) = 1.17$, $p = .25$, $b = .08$) were not. As was previously noted, increasing familiarization time or using room labels does not appear to affect performance.

A final analysis was performed to examine overall trends in errors across age groups. Polynomial trend contrasts adjusted for unequal variances and unequal spacing of groups showed a significant decreasing linear trend in both semantic ($t(46.13) = -4.08$, $p < .001$) and random errors ($t(23.03) = -3.46$, $p = .002$) as age increased. Only random errors, however, are completely absent in the oldest children tested (see Figure 5).

General discussion

The results of these three experiments amplify our understanding of the early development of episodic memory, giving an impression of both qualitative and quantitative change. On the qualitative side, Experiment 1 suggests that children's episodic memory may fundamentally change towards the end of the second year of life, to encompass a robust ability to bind together two closely associated objects. This age is a time when both behavioral (Newcombe *et al.*, 1998; Sluzenski *et al.*, 2004) and imaging (Utsunomiya *et al.*, 1999) research suggests maturation of the hippocampus, a crucial structure in episodic memory (Cohen & Eichenbaum, 1993; Squire, Stark & Clark, 2004). However, this qualitative change is initially only evident for the closest of associations, the link between a toy and its container. On the quantitative side, binding more incidental features of context (such as features of the rooms themselves) to a memory of what container hides a toy has a protracted developmental course, which was found to be present as early as 15 months and to increase quite gradually, reaching maturity in this paradigm at about 5 years. It is clear that more research needs to be done to elucidate the differences among different kinds of associations, and the reason why object-to-object associations seemed especially difficult in the youngest toddlers. As pointed out before, it may be that there are different kinds of binding, with somewhat different developmental trajectories.

A number of recent behavioral studies on the development of episodic memory have also suggested a shift in performance accuracy at the end of the second year. For example, there seems to be memory for details of an event and long-term recollection of that event among 20-month-olds, but not 16-month-olds (Bauer & Lukowski, 2010) and a robust ability to encode the temporal order of actions from one-time exposures by 20 months but not before (Bauer & Leventon, 2013). Richmond and Nelson (2009) argue for relational memory as early as 9 months using eye movement monitoring as babies look at slides of faces superimposed on scenes. However, some doubt is cast on what such data mean, given the fact that analogous results are not found in 4-year-olds (Koski, Olson & Newcombe, 2013). It is, of course, possible that later-emerging cognitive processes disrupt older children's ability to bind faces with background scenes, but replication and further exploration of the infancy data, along with specification of what causes the effect to disappear in older children, is needed to support such an argument.

In considering any possible change at the end of the second year, it is important to note that the predominance of semantic over random errors at all ages in our experiments underlines the well-established fact that even the youngest children are capable of forming explicit semantic memories (e.g. Richmond & Nelson, 2007). Thus, the transition does not involve explicit memory in general. Rather, it is specific to the binding of elements in situations that form the essential core of episodic memory: associations that are contingent and specific, rather than enduring and general.

In terms of quantitative change, the present experiments join a considerable body of evidence that the ability of children to bind items in memory shows improvement over the preschool years and into elementary school. For example, in tasks where children are asked to recall individual items and item–background combinations, 4-year-olds are successful at recalling individual items but do not perform as well as 6-year-olds on memory for combinations of items on particular backgrounds (Sluzenski *et al.*, 2006). In addition, although children as young as 2.5 years sometimes show surprising monitoring abilities in source memory tasks (Hala, Brown, McKay & San Juan, 2013), many other studies show substantial change in these abilities across an age range extending into elementary school (e.g. Sluzenski *et al.*, 2004). Although the task demands and designs of each of these studies vary, the overall picture is one of early fragility and strengthening across a wide age range. It is possible that children are not capable of successfully forming episodic memories until an array of basic cognitive functions, including linguistic skills, inhibitory control, autonoetic consciousness and theory of mind (Perner & Ruffman, 1995) are firmly in place.

The paradigm we have developed allows identification of the first signs of an emergent episodic memory system, which continues to develop across childhood in parallel with other supportive cognitive skills. Future research using this paradigm could examine the limits and potentials of the system. For example, we do not know whether children can recall the associations made in a task like this after a period of hours, or days. It may be that the stability of these early episodic memories is weak, as also suggested by recent data from a task requiring young children to remember to take a key to a situation in which they would encounter a locked treasure chest (Scarf, Gross, Colombo & Hayne, 2013). On the other hand, perhaps children would do better in this paradigm if they themselves made active decisions about where to hide the toys, analogous to the behavior of birds that choose caching sites for their food. Using this paradigm may be helpful in future research conducted in comparative perspective, given that that there are parallels to paradigms used with non-human animals (Eacott & Norman, 2004).

Research on the development of episodic memory is not only pertinent to furthering our understanding of autobiographical memory, but also bears relevance to higher-level cognitive abilities such as planning for future events. Episodic memory has been linked to one's predictions for the future (Schacter, Addis & Buckner, 2007; Wheeler *et al.*, 1997), a link which is made clearly evident in tasks such as remembering an encounter with a locked chest in order to plan to take a key along in the future (Scarf *et al.*, 2013). Development of episodic memory may also signify an important change in young children's thinking, moving from a stage at which they are still forming general categories and expectancies about the world, to one in which they can remember both overall distributional characteristics and distinguish individual events (Schacter & Moscovitch, 1984).

Acknowledgements

We thank Amanda Funk, Sarah Hittinger, Jule Grant, Deepali Patel, Kerri McShane, Jonathan Welch, Shannon Ens, and Sarah Dunn for help in conducting this study, and the parents and children for participating. This research was supported by the National Science Foundation #SBE-0541957 and #SBE-1041707.

References

Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6- to 24-month-old infants. *Infant Behavior and Development*, **19**, 159–170.

Bauer, P.J., Doydum, A.O., Pathman, T., Larkina, M., Güler, O.E., & Burch, M. (2012). It's all about location, location, location: children's memory for the 'where' of personally experienced events. *Journal of Experimental Child Psychology*, **113**, 510–522.

Bauer, P.J., & Leventon, J.S. (2013). Memory for one-time experiences in the second year of life: implications for the status of episodic memory. *Infancy*, **18**, 755–781.

Bauer, P.J., & Lukowski, A.F. (2010). The memory is in the details: relations between memory for the specific features of events and long-term recall during infancy. *Journal of Experimental Child Psychology*, **107**(1–14), 2010. doi: 10.1016/j.jecp.04.004

Bauer, P.J., Wenner, J.A., Dropik, P.L., & Wewerka, S.S. (2000). Parameters of remembering and forgetting in the transition from infancy to early childhood. *Monographs of the Society for Research in Child Development*, **65** (4, Serial No. 263).

Bindschaedler, C., Peter-Favre, C., Maeder, P., Hirsbrunner, T., & Clarke, S. (2011). Growing up with bilateral hippocampal atrophy: from childhood to teenage. *Cortex*, **47**(8), 931–944. doi: 10.1016/j.cortex.2010.09.005

Clayton, N.S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, **395**(6699), 272–273.

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

Drummey, A.B., & Newcombe, N.S. (2002). Developmental changes in source memory. *Developmental Science*, **5**(4), 502–513. doi: 10.1111/1467-7687.00243

Eacott, M.J., & Crawley, R.A. (1998). The offset of childhood amnesia: memory for events that occurred before age 3. *Journal of Experimental Psychology: General*, **127**, 22–33.

Eacott, M.J., & Norman, G. (2004). Integrated memory for object, place, and context in rats: a possible model of episodic-like memory? *Journal of Neuroscience*, **24**(8), 1948–1953. doi: 10.1523/JNEUROSCI.2975-03.2004

Feldman, A., & Acredolo, L. (1979). The effect of active versus passive exploration on memory for spatial location in children. *Child Development*, **50**(3), 698–704.

Ghetti, S., & Angelini, L. (2008). The development of recollection and familiarity in childhood and adolescence: evidence from the dual-process signal detection model. *Child Development*, **79**(2), 339–358.

Ghetti, S., & Bunge, S.A. (2012). Neural changes underlying the development of episodic memory during middle childhood. *Developmental Cognitive Neuroscience*, **2**(4), 381–395. doi: 10.1016/j.dcn.2012.05.002

Ghetti, S., DeMaster, D.M., Yonelinas, A.P., & Bunge, S.A. (2010). Developmental differences in medial temporal lobe function during memory encoding. *Journal of Neuroscience*, **30**(28), 9548–9556.

Gogtay, N., Nugent, T.F. 3rd, Herman, D.H., Ordonez, A., Greenstein, D., Hayashi, K.M., Clasen, L., Toga, A.W., Giedd, J.N., Rapoport, J.L., & Thompson, P.M. (2006). Dynamic mapping of normal human hippocampal development. *Hippocampus*, **16**, 664–672. doi: 10.1002/hipo.20193

Güler, O.E., & Thomas, K.M. (2013). Developmental differences in the neural correlates of relational encoding and recall in children: an event-related fMRI study. *Developmental Cognitive Neuroscience*, **3**(1), 106–116. doi: 10.1016/j.dcn.2012.07.001

Hala, S., Brown, A.M.B., McKay, L.-A., & San Juan, V. (2013). Two-and-a-half-year-olds' memory for sources of actions: contextual support facilitates recall. *Journal of Cognition and Development*, **14**, 343–358.

Hayne, H., Gross, J., McNamee, S., Fitzgibbon, O., & Tustin, K. (2011). Episodic memory and episodic foresight in 3- and 5-year-old children. *Cognitive Development*, **26**(4), 343–355.

Hayne, H., & Imuta, K. (2011). Episodic memory in 3- and 4-year-old children. *Developmental Psychobiology*, **53**(3), 317–322.

Hupbach, D.A., Gomez, R., & Nadel, L. (2009). Episodic memory reconsolidation: updating or source confusion? *Memory*, **17**(5), 502–510.

Imuta, K., Scarf, D., & Hayne, H. (2013). The effect of verbal reminders on memory reactivation in 2-, 3-, and 4-year-old children. *Developmental Psychology*, **49**, 1058–1065.

Koski, J., Olson, I.R., & Newcombe, N.S. (2013). Tracking the eyes to see what children remember. *Memory*, **21**, 396–407.

Lavenex, P., & Banta Lavenex, P. (2013). Building hippocampal circuits to learn and remember: insights into the development of human memory. *Behavioural Brain Research*, **254**, 8–21. doi: 10.1016/j.bbr.2013.02.007

Lloyd, M.E., Doydum, A.O., & Newcombe, N.S. (2009). Memory binding in early childhood: evidence for a retrieval deficit. *Child Development*, **80**(5), 1321–1328. doi: 10.1111/j.1467-8624.2009.01353.x

Mandler, J.M. (2002). On the foundations of the semantic system. In E.M. Forde & G.W. Humphreys (Eds.), *Category-specificity in brain and mind* (pp. 315–374). New York: Psychology Press.

Meltzoff, A.N. (1988). Infant imitation and memory: nine-month-olds in immediate and deferred tests. *Child Development*, **59**, 217–225.

Newcombe, N., Huttenlocher, J., Drummey, A.B., & Wiley, J.G. (1998). The development of spatial location coding: place learning and dead reckoning in the second and third years. *Cognitive Development*, **13**, 185–200.

Newcombe, N.S., Lloyd, M.E., & Ratliff, K.R. (2007). Development of episodic and autobiographical memory: a cognitive neuroscience perspective. In R.V. Kail (Ed.), *Advances in child development and behavior*, Vol. **35** (pp. 37–85). San Diego, CA: Elsevier.

Ofen, N. (2012). The development of neural correlates for memory formation. *Neuroscience and Biobehavioral Reviews*, **36**(7), 1708–1717. doi: 10.1016/j.neubiorev.2012.02.016

Perner, J., & Ruffman, T. (1995). Episodic memory and autonoetic consciousness: developmental evidence and a theory of childhood amnesia. *Journal of Experimental Child Psychology*, **59**, 516–548. doi: 10.1006/jecp.1995.1024

Picard, L., Cousin, S., Guillery-Girard, B., Eustache, F., & Piolino, P. (2012). How do the different components of episodic memory develop? Role of executive functions and short-term feature-binding abilities. *Child Development*, **83**(3), 1037–1050. doi: 10.1111/j.1467-8624.2012.01736.x

Povinelli, D.J., Landau, K.R., & Perilloux, H.K. (1996). Self-recognition in young children using delayed versus live feedback: evidence of a developmental asynchrony. *Child Development*, **67**, 1540–1554. doi: 10.1111/j.1467 8624.1996.tb01813.x

Povinelli, D.J., Landry, A.M., Theall, L.A., Clark, B.R., & Castille, C.M. (1999). Development of young children's understanding that the recent past is causally bound to the present. *Developmental Psychology*, **35**, 1426–1439. doi: 10.1037/0012-1649.35.6.1426

Rhodes, S.M., Murphy, D., & Hancock, P.J.B. (2011). Developmental changes in the engagement of episodic retrieval processes and their relationship with working memory during the period of middle childhood. *British Journal of Developmental Psychology*, **29**(4), 865–882. doi: 10.1111/j.2044-835X.2010.02014.x

Ribordy, F., Jabès, A., Banta Lavenex, P., & Lavenex, P. (2013). Development of allocentric spatial memory abilities in children from 18 months to 5 years of age. *Cognitive Psychology*, **66**(1), 1–29. doi: 10.1016/j.cogpsych.2012.08.001

Richmond, J., & Nelson, C.A. (2007). Accounting for change in declarative memory: a cognitive neuroscience perspective. *Developmental Review*, **27**(3), 349–373. doi: 10.1016/j.dr.2007.04.002

Richmond, J., & Nelson, C.A. (2009). Relational memory during infancy: evidence from eye tracking. *Developmental Science*, **12**(4), 549–556. doi: 10.1111/j.1467- 7687.2009. 00795.x

Riggins, T., Miller, N.C., Bauer, P.J., Georgieff, M.K., & Nelson, C.A. (2009). Electrophysiological indices of memory for temporal order in early childhood: implications for the development of recollection. *Developmental Science*, **12**(2), 209–219.

Russell, J., Cheke, L.G., Clayton, N.S., & Meltzoff, A.N. (2011). What can What–When–Where (WWW) binding tasks tell us about young children's episodic foresight? Theory and two experiments. *Cognitive Development*, **26**, 356–370.

Russell, J., & Hanna, R. (2012). A minimalist approach to the development of episodic memory. *Mind & Language*, **27**(1), 29–54.

Scarf, D., Gross, J., Colombo, M., & Hayne, H. (2013). To have and to hold: episodic memory in 3- and 4-year-old children. *Developmental Psychobiology*, **55**, 125–132.

Schacter, D.L., Addis, D.R., & Buckner, R.L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience*, **8**(9), 657–661. doi: 10.1038/ nrn2213

Schacter, D.L., & Moscovitch, M. (1984). Infants, amnesics, and dissociable memory systems. In M. Moscovitch (Ed.), *Infant memory: Its relation to normal and pathological memory in humans and other animals* (pp. 173–216). New York: Plenum Press.

Shing, Y.L., & Lindenberger, U. (2011). The development of episodic memory: lifespan lessons. *Child Development Perspectives*, **5**(2), 148–155. doi: 10.1111/j.1750- 8606.2011. 00170.x

Simcock, G., & Hayne, H. (2002). Breaking the barrier? Children fail to translate their preverbal memories into language. *Psychological Science*, **13**, 225–231. doi: 10.1111/ 1467-9280.00442

Sluzenski, J., Newcombe, N.S., & Kovacs, S.L. (2006). Binding, relational memory, and recall of naturalistic events: a developmental perspective. *Journal of Experimental Psychology*, **32**(1), 89–100.

Sluzenski, J., Newcombe, N.S., & Satlow, E. (2004). Knowing where things are in the second year of life: implications for hippocampal development. *Journal of Cognitive Neuroscience*, **16**(8), 1443–1451.

Squire, L.R., Stark, C.E., & Clark, R.E. (2004). The medial temporal lobe. *Annual Review of Neuroscience*, **27**, 279–306. doi: 10.1146/annurev.neuro.27.070203. 144130

Townsend, E.L., Richmond, J.L., Vogel-Farley, V.K., & Thomas, K. (2010). Medial temporal lobe memory in childhood: developmental transitions. *Developmental Science*, **13**(5), 738–751. doi: 10.1111/j.1467-7687.2009.00935.x

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 382–403). New York: Academic Press.

Tulving, E. (1983). *Elements of episodic memory.* Oxford: Clarendon Press.

Utsunomiya, H., Takana, K., Okazaki, M., & Mitsudome, A. (1999). Development of the temporal lobe in infants and children: analysis by MR-based volumetry. *American Journal of NeuroRadiology*, **20**, 717–723.

Vargha-Khadem, F., Gadian, D.G., Watkins, K.E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, **277**(5324), 376–380.

Wheeler, M.A., Stuss, D.T., & Tulving, E. (1997). Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychological Bulletin*, **121**(3), 331–354. doi: 10.1037/0033-2909.121.3.331

Received: 2 January 2013
Accepted: 22 October 2013