Memory Retrieval by 18–30-Month-Olds: Age-Related Changes in Representational Flexibility

Jane Herbert and Harlene Hayne
University of Otago

Deferred imitation was used to trace changes in memory retrieval by 18–30-month-olds. In all experiments, an adult demonstrated 2 sets of actions using 2 different sets of stimuli. In Experiments 1A and 1B, independent groups of infants were tested immediately or after a 24-hr delay. Each infant was tested with 1 set of stimuli from the original demonstration and 1 set of stimuli that was different. Recall of the target actions when tested with different stimuli increased as a function of age, particularly after a delay. In Experiment 2, infants were provided with a unique verbal label for the stimuli during the demonstration and the test. The verbal label facilitated performance by 24-month-olds tested with different stimuli but had no effect on performance by 18-month-olds. One hallmark of memory development appears to be an age-related increase in the range of effective retrieval cues for a particular memory.

A large body of research conducted with adults has shown that memory retrieval is disrupted when the environmental context (Godden & Baddeley, 1975; Smith, Glenberg, & Bjork, 1978) or stimulus materials (for reviews see Spear, 1978; Tulving, 1983) are altered at the time of the test. The greater the discrepancy between the conditions at original encoding and the conditions at subsequent recall, the greater the disruption in memory performance. This relation between encoding and retrieval has been referred to as the “principle of encoding specificity” (Tulving, 1983, 1984; Tulving & Thompson, 1973). According to this principle, memory retrieval is dependent on a match between the attributes stored at the time of original encoding and the cues present at the time of retrieval.

Although the encoding specificity hypothesis was originally developed to account for findings obtained in studies of verbal learning with adults, some of the best empirical support for the hypothesis has come from studies conducted with preverbal infants. To date, the most comprehensive analysis of encoding and retrieval during infancy has been conducted with the mobile conjugate reinforcement paradigm. In this experimental procedure, infants between 2 and 6 months of age learn to kick their feet in an overhead mobile. Specificity of memory retrieval is subsequently assessed by manipulating the characteristics of the mobile or the environmental context present at the time of retrieval relative to those present at the time of original training.

Overall, research with the mobile conjugate reinforcement paradigm has shown that the cues present at the time of the test must be virtually identical to those present at the time of training for memory retrieval to occur. Studies conducted with 2- to 3-year-olds demonstrated that memory retrieval was precluded when infants were tested with a five-object mobile that contained more than one novel component (Hayne et al., 1986). Similarly, 3- (Butler & Rovee-Collier, 1989; Rovee-Collier & Dufault, 1991) and 6-month-olds (Amabile & Rovee-Collier, 1991; Borovsky & Rovee-Collier, 1990; Rovee-Collier, Schecter, Shyi, & Shields, 1992) exhibit no retention whatsoever when they are tested in a novel environmental context after a delay.

Taken together, studies conducted using the mobile conjugate reinforcement paradigm clearly demonstrate that changes in either the test object or in the test context preclude memory retrieval by 2–6-month-old infants. These same changes, however, have little or no effect on memory retrieval by infants tested in the deferred imitation paradigm. In tests of deferred imitation, the experimenter demonstrates a series of novel actions, and the infant’s ability to reproduce those actions is assessed either immediately or after a delay (Meltzoff, 1990, 1995). Using the deferred imitation procedure, Meltzoff and his colleagues have shown that memory retrieval by 12–18-month-old infants is much less sensitive to changes in cues at the time of the test. Infants continue to imitate at least some of the target actions when they are tested in an altered context (Barnat, Klein, & Meltzoff, 1996; Hanna & Meltzoff, 1993) or with objects that differ in size and color relative to those present during the original demonstration (Barnat et al., 1996; Klein & Meltzoff, 1999).
Two explanations have been proposed to account for the discrepancy in the pattern of results obtained with the mobile conjugate reinforcement and deferred imitation procedures. On the one hand, it has been argued that differences in the pattern of results reflect differences in the requisite memory systems involved in these two tasks (Hanna & Meltzoff, 1993). This argument has been fueled by the belief that the mobile conjugate reinforcement paradigm provides a measure of procedural memory, whereas imitation procedures provide a measure of declarative memory (Bauer, 1996; McDonough, Mandler, McKee, & Squire, 1995; Meltzoff, 1990, 1995; C. Nelson, 1995). Past research with both humans and animals has shown that retention of procedural memories is highly constrained by the original encoding context, whereas retention of declarative memories is not (for discussion, see Diamond, 1990; Schacter, 1990).

On the other hand, it may also be important to consider differences in the age of the infants tested in the two experimental procedures. The mobile conjugate reinforcement paradigm has been used exclusively with 2–6-month-olds (for reviews, see Rovee-Collier & Hayne, 1987; Rovee-Collier & Shyi, 1992), whereas deferred imitation procedures are most commonly used with infants over the age of 9 months (for reviews, see Barr & Hayne, 2000; Meltzoff, 1990). From a developmental perspective, it has been argued that the difference in findings obtained using these two experimental procedures reflects an age-related increase in infants' ability to exploit potentially effective retrieval cues, not the memory requirements of the tasks themselves (Barr & Hayne, 2000; Hayne, Boniface, & Barr, 2000; Hayne, MacDonald, & Barr, 1997). For young infants, virtually all of the cues present at the time of original encoding must be reinstated in order for memory retrieval to occur. Memory retrieval by older infants, however, is initiated on the basis of partial reinstatement of the encoding cues.

Resolution of this issue has been hindered in the past because infants of different ages have been tested with different experimental procedures. The recent finding that infants as young as 6 months of age will exhibit deferred imitation with objects (Barr, Dowden, & Hayne, 1996; Collie & Hayne, 1999) has paved the way for a developmental analysis of infant memory retrieval across an age range that encompasses the age of infants commonly tested in the mobile conjugate reinforcement (i.e., 6-month-olds) and deferred imitation (i.e., 12–18-month-olds) paradigms. In several studies of this kind, infants of different ages have been tested with the same deferred imitation procedure. In one task, for example, an experimenter used a puppet to demonstrate a series of actions. Following a delay, infants were tested with either the demonstration puppet or with a different one in either the context in which the demonstration originally occurred or in a different one.

Consistent with the findings previously obtained with the mobile conjugate reinforcement paradigm, deferred imitation by 6-month-old infants is also precluded by a change in the puppet or in the context at the time of the test. Although 6-month-olds exhibit excellent retention in the puppet task after a 24-hr delay, they exhibit no retention when tested with a puppet that differs from the puppet present during the original demonstration. Furthermore, infants of this age also exhibit no evidence of retention when the demonstration occurs in one context and the test occurs in another even when they are tested with the puppet that was present during the original demonstration session (Hayne et al., 2000).

Older infants who are tested in the puppet task successfully exploit a wider range of retrieval cues, a finding consistent with those previously obtained with other deferred imitation procedures. The range of effective cues, however, increases gradually during development. Although 12-month-olds who are tested in the puppet task imitate the target actions when tested in a novel context (Hayne et al., 2000), they exhibit no imitation when they are tested with a novel puppet (Hayne et al., 1997, 2000). Even minor changes in the color or form of the puppet have been shown to disrupt the memory performance of 12-month-olds (Hayne et al., 1997). When tested in the same procedure, however, 18-month-olds are extremely resilient to changes in the context or in the puppet. Infants of this age exhibit the same level of imitation when they are tested with the same or with a different puppet in the same or in a different context (Hayne et al., 1997, 2000). If the physical dissimilarity of the demonstration and the test puppets is increased further, however, then memory retrieval by 18-month-olds is once again precluded (Hayne et al., 1997).

A similar age-related change in memory retrieval has been obtained with an operant train task that was developed to extend the age range of infants who could be tested with the mobile conjugate reinforcement paradigm. In the train task, 6–12-month-olds learn to press a lever to make a miniature train move around a track. Consistent with studies conducted using the deferred imitation procedure, changes in the train or in the environmental context have no effect on the performance of 9- or 12-month-olds (Hartshorn et al., 1998).

Taken together, findings obtained with both operant conditioning and deferred imitation procedures converge on the same conclusion. Irrespective of the paradigm used to assess memory performance, there is a developmental progression in infants' ability to exploit effective retrieval cues. Early in development, memory retrieval is contingent upon an exact match between the cues present at the time of encoding and the cues present at the time of retrieval. As a function of age and experience, however, memory retrieval becomes more flexible, allowing past experiences to be retrieved by cues or in contexts different from those present during original encoding.

To date, our analysis of age-related changes in memory retrieval has focused primarily on studies with 6–18-month-olds. We know much less about age-related changes in the effect of novel test stimuli on memory performance during the final months of the infancy stage. Historically, the period between 18 and 24 months of age has marked an important transition in cognitive development. According to Piaget (1927/1962), this developmental epoch represents the final stage of the sensorimotor period and marks the emergence of representational thought. Although some contemporary developmental psychologists have rejected many aspects of traditional Piagetian theory, they still support the notion that the period between 18 and 24 months represents an important shift in the development of mental representation (Karmiloff-Smith, 1992; Meltzoff, 1990).

The purpose of the present study was to examine age-related changes in memory retrieval between 18 and 30 months of age. To do this, we assessed the effect of altered test stimuli on infants' performance in a deferred imitation task. The procedures used in the present experiments were modeled after those used in prior
work with 6–18-month-olds, but the stimuli and target actions were modified to meet the changing interests and motor skills of older infants.

Experiment 1A

In Experiment 1A, we assessed the effects of novel test stimuli on deferred imitation by 18–24-month-old infants who were tested either immediately or after a 24-hr delay.

Method

Participants. The final sample consisted of thirty-six 18-month-old ($M = 18.28$ months, $SD = 0.15$ months) and thirty-six 24-month-old ($M = 24.29$ months, $SD = 0.10$ months) infants who were recruited from public birth records and by word of mouth. Half of the infants at each age were girls. Infants were predominantly of European descent and came from a range of socioeconomic backgrounds. Two additional 18-month-old infants were excluded from the final sample due to a failure to touch the stimuli during the test (1 boy) and equipment failure (1 girl). One 24-month-old (girl) was excluded due to a failure to touch the stimuli during the test.

Apparatus. The stimuli were constructed specifically for the present study and were not commercially available. There were two types of stimuli (a rattle and an animal) and two versions of each type. The stimuli were constructed in such a way that exactly the same target actions could be performed with each version.

The stimuli for the green rattle (see Figure 1, top) consisted of a green stick (12.5 cm long) attached to a white plastic lid (9.5 cm in diameter), with Velcro attached to the underside of the lid; a round green block (3 cm in diameter $\times$ 2.5 cm in height); and a clear plastic square cup with Velcro around the top ($5.5$ cm in diameter $\times$ 8 cm in height). The opening of the plastic cup (3.5 cm in diameter) was covered with a 1-mm black rubber diaphragm, with 16 cuts radiating from the center. The stimuli for the red rattle (see Figure 1, bottom) consisted of a yellow D-shaped handle (gap between stick and handle $= 1.5 \times 8$ cm) attached to a yellow wooden stick (12.5 cm long) with a plug on the end, which fitted into a red plastic ball with a hole cut in the top (4 cm in diameter); and a blue plastic head (2 cm in diameter) attached to the underside of a silver metal bell (2.5 cm in width, 2 cm long).

The stimuli for the rabbit (see Figure 2, top) consisted of two plastic eyes ($3 \times 2$ cm) attached to a $9 \times 6$ cm piece of plywood with Velcro on the back, a 12-cm orange wooden carrot with green string attached to the top, and a white circle of wood (the head, 15 cm in diameter) mounted horizontally on a white rectangular wooden base ($30 \times 20$ cm). A 3-cm (in diameter) hole was drilled at the bottom of the head, and a $5 \times 15$ cm piece of white Velcro was attached to the top of the head. Two white “ears” ($20 \times 5$ cm) decorated with stripes of pink felt were hidden behind the head. A 10-cm wooden stick attached to the top of the right ear allowed the ears to be pulled up from behind the head in a circular motion to a point above the head.

The stimuli for the monkey (see Figure 2, bottom panel) consisted of two plastic eyes ($2.5$ cm in diameter) with eyelashes that were attached to a piece of brown plywood in the shape of two diamonds joined at the center ($11.5$ cm in width, 6.5 cm in height), with brown Velcro on the back; a 20.5-cm yellow plastic banana; and a brown wooden head-and-shoulders shape mounted horizontally on a brown rectangular wooden base ($22 \times 38$ cm). A 4-cm hole was drilled at the bottom of the head, and a $5 \times 18$ cm
Figure 2. The animal stimuli used in Experiments 1A, 1B, and 2. Top: rabbit. Bottom: monkey. For each set of stimuli, the experimenter modeled a three-step sequence of actions (pictured left to right; see also Table 1).

A piece of brown Velcro was attached to the top of the head. Two brown ears (3.5 × 7 cm) decorated with a piece of yellow felt were hidden behind the head. A 3-cm lever with a wooden button (3.5 cm in diameter) on the top, attached to the right ear, allowed the ears to be pulled up from behind the head in a circular motion to the side of the head.1

Procedure. Infants were tested in their own homes at a time of the day that their caregiver identified as an alert/play period. During the initial visit, the purpose of the study was explained to the caregiver, and informed consent was obtained. The female experimenter was positioned directly opposite the infant and was seated on the floor. At the beginning of the session, the experimenter interacted with the infant for 5 min or until a smile was elicited. For infants in the demonstration condition, the experimenter modeled a three-step sequence with one version of each stimulus type (see Table 1). The actions were modeled three times in succession out of reach of the infant. The experimenter did not verbally describe the actions or label the stimuli. The entire demonstration session lasted approximately 60 s. The order of presentation of the stimulus sets was counterbalanced across age and delay.

At each age, independent groups of 12 children (6 boys, 6 girls) were tested either immediately or after a delay of 24 hr (± 2 hr). Infants were tested with one set of stimuli that had been used in the original demonstration (same) and one set of stimuli that was perceptually different from the one seen during the demonstration (different) but that required the same target actions (see Table 1). The two types of stimuli (rattle or animal) and the order of presentation at test (same or different) were counterbalanced across infants. During the test, infants were given the first set of stimuli, and their behavior was videotaped for a 60-s test period. Infants were then given the second set of stimuli, and their behavior was videotaped for an additional 60-s test period. The timing of the test period began when the infant first touched the stimuli.

To assess the spontaneous production of target actions in the absence of adult demonstration, we exposed an additional control group of 12 infants at each age to the stimuli for the first time during the test. For these infants, the test was scheduled within 2 weeks of their 18- or 24-month-old birthday. The test procedure was identical to that experienced by infants in the demonstration condition.

Results and Discussion

Each videotaped test session was scored by two independent observers, one of whom was blind to the infant's group assignment. Each observer scored the presence or absence of each target action during the 60-s test period. Both the percentage of agreement (98%) and Cohen's kappa (κ = .97) were used to calculate interobserver reliability.

An imitation score was calculated for each infant by summing the number of target behaviors produced during the test with each set of stimuli (range = 0–3). Preliminary analyses indicated that there was no difference in infants' performance with the two sets of stimuli; therefore, the data were averaged across both stimulus sets for all subsequent analyses.

The mean imitation scores of infants in the demonstration and control conditions are shown in Figure 3 as a function of age.1 The rattle and the animal tasks were modeled after those previously used by Bauer and her colleagues (Bauer et al., 1995; West & Bauer, 1996).
Table 1
The Three Target Actions for Each Set of Stimuli

<table>
<thead>
<tr>
<th>Stimulus set</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green rattle</td>
<td>Push ball through diaphragm into cup</td>
<td>Put stick on jar, attaching with Velcro</td>
<td>Shake stick to make noise</td>
</tr>
<tr>
<td>Red rattle</td>
<td>Put bell through hole into red ball</td>
<td>Put plug into hole in red ball</td>
<td>Shake stick to make noise</td>
</tr>
<tr>
<td>Rabbit</td>
<td>Pull lever in circular motion to raise ears</td>
<td>Place eyes on face, attaching with Velcro</td>
<td>Put the carrot in the rabbit’s “mouth”</td>
</tr>
<tr>
<td>Monkey</td>
<td>Pull lever in circular motion to raise ears</td>
<td>Place eyes on face, attaching with Velcro</td>
<td>Put the banana in the monkey’s “mouth”</td>
</tr>
</tbody>
</table>

delay, and test stimulus (same or different). In the absence of adult demonstration, infants in the control condition rarely produced the target actions. Furthermore, as shown in Figure 3, there was no age-related change in the performance of the control groups, \( r(22) = 0.57, p > .05 \).

To assess age-related changes in infants’ performance in the demonstration condition, we subjected the data to a 2 (age) x 2 (delay) x 2 (test stimulus) analysis of variance (ANOVA) with repeated measures over test stimulus. This analysis yielded a main effect of age, \( F(1, 44) = 9.60, p < .01 \). Overall, 24-month-old infants had higher imitation scores than 18-month-old infants. There was also a main effect of delay, \( F(1, 44) = 4.27, p < .05 \), a main effect of test stimulus, \( F(1, 44) = 75.21, p < .001 \), and a Delay x Test Stimulus interaction, \( F(1, 44) = 7.24, p < .01 \). Post hoc examination of this interaction (Student Newman-Keuls tests, \( p < .05 \)) indicated that when tested with the same stimulus, infants’ imitation performance did not differ as a function of delay.

When tested with a different stimulus, however, the performance of infants tested immediately after the demonstration was significantly higher than that of infants tested after a 24-hr delay (see Figure 3).

The analyses described above indicated that there were age-related, delay-related, and stimulus-related changes in infants’ performance. The analyses did not reveal, however, whether a particular group of infants exhibited imitation. In the deferred imitation paradigm, imitation can only be inferred if a group’s performance is greater than the performance of infants in the age-matched control group (Barr & Hayne, 1999, 2000; Hayne et al., 2000; Herbert & Hayne, 2000; Meltzoff, 1985).

To assess the test conditions under which infants exhibited imitation, we combined the data from infants in the demonstration condition with the data from infants in their age-matched control condition. At each age, the data were subjected to a one-way ANOVA across all five groups (i.e., four demonstration test

![Figure 3](image-url)
groups, one control group). At each age, this analysis yielded a significant main effect of group: for 18-month-olds: $F(4, 55) = 16.52, p < .0001$; for 24-month-olds, $F(4, 55) = 32.93, p < .0001$. Post hoc Student Newman-Keuls tests ($p < .05$) indicated that at 24 months of age, the imitation scores of infants in the demonstration condition were significantly above those of their age-matched control group irrespective of their test stimulus or the retention interval (see Figure 3, right panel, bars with asterisks). At 18 months of age, the imitation scores of infants in the demonstration condition were significantly above those of their age-matched control group when they were tested with the same or a different stimulus immediately after the demonstration or when they were tested with the same stimulus after a 24-hr delay (see Figure 3, left panel, bars with asterisks). When 18-month-olds were tested with a different stimulus after a 24-hr delay, however, their imitation scores did not exceed those of their age-matched control group.

Recall that the timing of the 60-s test period did not begin until the infant first touched the test stimuli. To assess whether differences in the relative novelty of the test stimuli influenced the time it took infants to initially touch them, an observer, blind to the infants' age and group assignment, scored the latency between the presentation of the stimulus and the infant's first touch. Overall, infants touched both stimuli very quickly during the test ($M = 4.57$ s, $SE = 1.26$). A 2 (age) $\times$ 2 (delay) $\times$ 2 (test stimulus) ANOVA with repeated measures over test stimulus indicated that there was no difference in infants' latency to respond during the test as a function of age, delay, or test stimulus, and there were no interactions. Thus, differences in infants' imitation performance with the two sets of stimuli cannot be attributed to differences in their willingness to interact with them.

The results of Experiment 1A confirm our prior findings that infants' ability to exploit potentially effective retrieval cues increases as a function of age, particularly when infants are tested after a delay. Eighteen-month-olds exhibited deferred imitation in the present task when they were tested with a novel stimulus immediately after the demonstration but not after a 24-hr delay when the retrieval demands were presumably greater. Twenty-four-month-olds imitated the target actions when they were tested with a novel stimulus either immediately or after a 24-hr delay, but their performance was inferior to that when they were tested with the original demonstration stimulus. In fact, although their performance was significantly greater than that of their age-matched control group, when they were tested with a different stimulus, 24-month-olds imitated less than half of the actions they imitated when tested with the demonstration stimulus.

Experiment 1B

Research on age-related changes in memory retrieval has shown that there is an age at which infants tested with a novel stimulus recall as much as infants tested with the demonstration stimulus (Hayne et al., 1997, 2000). The purpose of Experiment 1B, therefore, was to determine the age at which infants' ability to imitate the target actions with the novel stimuli used in Experiment 1A was equivalent to their ability to imitate the same target actions with the demonstration stimuli. To do this, we tested 30-month-old infants in our imitation task immediately or after a 24-hr delay with the stimuli used in Experiment 1A.

Method

Participants. The final sample consisted of thirty-six 30-month-old ($M = 30.22$ months, $SD = 0.35$ months) infants recruited through public birth records and by word of mouth. Half the infants were girls. Four additional infants were excluded from the final sample due to failure to touch the stimuli during the test (3 boys) and maternal interference (1 girl).

Apparatus and procedure. The apparatus was identical to that described in Experiment 1A. The demonstration and test procedures were also identical to those used in Experiment 1A. As before, infants in the control condition did not observe the target actions prior to the test. During the test, infants in all conditions were given 60 s from the time they first touched the stimuli in which to respond. Infants in the control group were tested within 2 weeks of their 30-month birthday.

Results and Discussion

Each videotaped test session was scored by two independent observers, one of whom was blind to the infant’s group assignment. Each observer scored the presence or absence of each target action during the 60-s test period. Both the percentage of agreement (98%) and Cohen’s kappa ($\kappa = .98$) were used to calculate interobserver reliability.

The mean imitation scores of infants in the demonstration and control conditions are shown in Figure 4 as a function of delay and test stimulus (same or different). To evaluate infants’ performance in the demonstration condition, we subjected the data to a 2 (delay) $\times$ 2 (test stimulus) ANOVA. This analysis yielded no significant main effects or an interaction. That is, 30-month-old infants’ performance was the same regardless of their test stimulus or the retention interval.

![Figure 4](image-url)
As in Experiment 1A, significant evidence of imitation was defined relative to the performance of the control group. To assess the test conditions under which infants exhibited imitation, we combined the data from infants in the demonstration condition with the data from infants in their age-matched control condition. The data were subjected to a one-way ANOVA across all five groups (i.e., four demonstration test groups, one control group). This analysis yielded a significant main effect of group, $F(4, 55) = 9.33, p < .0001$. Post hoc Student Newman-Keuls tests ($p < .05$) indicated that the imitation scores of infants in the demonstration condition were significantly above their age-matched control group irrespective of the test stimulus or retention interval (see Figure 4, bars with asterisks).

The results of Experiment 1A and 1B indicated that 18-, 24-, and 30-month-old infants exhibited deferred imitation when they were tested with the original (same) stimuli either immediately or after a 24-hr delay. When infants were tested with different stimuli, there were both age-related and delay-related changes in their performance. Although 18-, 24-, and 30-month-old infants exhibited some evidence of imitation when tested with novel stimuli immediately after the demonstration, only the 24- and 30-month-olds exhibited imitation when they were tested with novel stimuli after a 24-hr delay. Furthermore, it was not until 30 months of age that infants' ability to imitate the target actions with the novel stimuli was equivalent to their ability to do so when tested with the original demonstration stimuli. In conclusion, the results from Experiment 1A and 1B confirm that age-related changes in memory retrieval continue to occur during the final phase of the infancy period.

Experiment 2

Past research on age-related changes in infant memory retrieval has focused on changes in the visual cues associated with the objects or with the environment present at the time of the test. The results of Experiment 1A and 1B add to a small but growing body of research indicating that, despite the high degree of specificity that is required for memory retrieval by young infants, older infants exploit visual cues that are increasingly dissimilar to those present at the time of original encoding or exploit a different set of cues than those required by younger infants (e.g., the experimenter is present during both encoding and retrieval). In Experiment 2, we continued to chart age-related changes in memory retrieval by assessing infants' ability to exploit linguistic cues at the time of the test.

The period between 18 and 24 months of age is characterized by rapid language acquisition (Fenson et al., 1994; K. Nelson, 1996). The ability to encode, store, and access information by using language may contribute in important ways to age-related changes in memory retrieval. In the course of acquiring new words, for example, children learn that the same verbal label can be used to describe objects that share few, if any, physical characteristics. Within this context, the effect of adults' language on young children's production of verbal labels has been well documented (Markman, 1987; Mervis, 1984, 1987; K. Nelson, 1974).

In Experiment 2, we assessed age-related changes in infants' ability to use an adult's language to facilitate memory retrieval. To do this, we provided 18- and 24-month-old infants with a unique verbal label for the stimuli that were used during the demonstration. We then assessed whether this same verbal label would facilitate infants' ability to recall the target actions when they were tested with the same or with novel stimuli after a 24-hr delay.

Method

Participants. The final sample consisted of twenty-four 18-month-old ($M = 18.34$ months, $SD = 0.12$ months) and twenty-four 24-month-old ($M = 24.28$ months, $SD = 0.25$ months) infants. Half the infants at each age were girls. One 18-month-old girl was excluded from the final sample due to a failure to touch the stimuli during the test.

Apparatus and procedure. The apparatus was identical to that described in Experiment 1A. The demonstration and test procedures were identical to those used in Experiment 1A except that infants were provided with novel verbal labels before the demonstration and immediately prior to the test. Novel (i.e., nonsense) verbal labels were used to ensure that infants of either age were selectively advantaged or disadvantaged by their prior exposure to the labels. Furthermore, the use of novel labels allowed us to establish a direct relation between verbal information presented during the demonstration and test sessions without potential interference from the infants' prior experience with any of the labels.

During the demonstration, the experimenter was positioned directly opposite the infant and was seated on the floor. She then placed the first set of stimuli in front of the infant but out of reach. Once the stimuli were in place, she said, "We can use these things to make a thornby" (the novel label for the animals). She then demonstrated the target actions as before. When this demonstration was complete, the experimenter removed the first set of stimuli and placed the second set of stimuli in front of the infant as before. Once these stimuli were in place, she said, "We can use these things to make a meewa" (the novel label for the rattles). She then demonstrated the target actions as before. The order of presentation was counterbalanced across participants.

All infants in the demonstration condition were tested 24 hr after the demonstration session. Each infant was tested with one set of stimuli that had been used during the demonstration and one set that had not. During the test session, the experimenter was seated on the floor directly opposite the infant. She placed the first set of stimuli in front of the infant, but out of reach. If the test stimuli were the same as those present during the original demonstration, the experimenter said, "Yesterday I showed you how to make a thornby [meewa]. These were the things we used to make a thornby [meewa]. Can you show me how we can use these things to make a thornby [meewa]?

"If the stimuli were different from those present during the original demonstration, she said, "Yesterday I showed you how to make a thornby [meewa]. These are some other things we can use to make a thornby [meewa]. Can you show me how we can use these things to make a thornby [meewa]?

The stimuli were then put within the reach of the infant and, as before, the infant was given 60 s from the time he or she first touched the stimuli in which to respond. The infant was then given the second set of stimuli, and his or her behavior was videotaped for an additional 60 s.

As before, an additional control group of 12 infants at each age was exposed to the stimuli for the first time during the test. For these infants, the test was scheduled within 2 weeks of their 18- or 24-month-old birthday.

Results

Each videotaped test session was scored by two independent observers, one of whom was blind to the infant's group assignment. Each observer scored the presence or absence of each target action during the 60-s test period. Both the percentage agreement (94%) and Cohen's kappa ($\kappa = .92$) were used to calculate interobserver reliability.
The mean imitation scores of infants in the demonstration condition are shown in Figure 5 as a function of age and test stimulus (same or different). The mean imitation scores of infants in the control condition are also shown for comparison. As in Experiment 1A, there was no age-related change in the performance of the control groups, t(22) = 1.79, p > .05.

To assess age-related changes in infants' performance in the demonstration condition, we subjected the data to a 2 (age) × 2 (test stimulus) ANOVA with repeated measures over test stimulus. This analysis yielded main effects of age, F(1, 22) = 11.95, p < .01, and test stimulus, F(1, 22) = 31.98, p < .001, and an Age × Test Stimulus interaction, F(1, 22) = 8.65, p < .01. Post hoc Student Newman-Keuls tests (p < .05) indicated that at 18 months, infants performed significantly more target actions when they were tested with the original demonstration stimuli than when they were tested with different stimuli. At 24 months, however, there was no difference in performance as a function of the test stimulus.

As in Experiment 1A and 1B, significant evidence of imitation was defined relative to the performance of the control group. To assess the test conditions under which infants exhibited imitation, we combined the data from infants in the demonstration condition with the data from infants in their age-matched control condition. The data were subjected to a one-way ANOVA across all three groups (i.e., two demonstration test groups, one control group). At each age, this analysis yielded a significant main effect of group: for 18-month-olds, F(2, 33) = 20.69, p < .0001; for 24-month-olds, F(2, 33) = 25.12, p < .0001. Post hoc Student Newman-Keuls tests (p < .05) indicated that at 24 months of age, the imitation scores of infants in the demonstration condition were significantly above their age-matched control group irrespective of the test stimulus (see Figure 5, right panel, bars with asterisks). At 18 months of age, the imitation scores of infants in the demonstration condition were significantly above those of their age-matched control group when they were tested with the same stimulus (see Figure 5, left panel, bar with asterisk). When they were tested with a different stimulus, however, their imitation scores did not exceed those of the age-matched control group.

To assess the effects of verbal labels on infants' performance directly, we compared the data from infants in the demonstration condition in Experiment 2 with data from infants in the demonstration condition in Experiment 1A who were tested after the same delay but without verbal instruction. The combined data set was subjected to a 2 (age) × 2 (label condition) × 2 (test stimulus) ANOVA with repeated measures over test stimulus. This analysis yielded significant main effects of age, F(1, 44) = 5.47, p < .05, and of test stimulus, F(1, 44) = 78.44, p < .001. These main effects were qualified by a two-way Test Stimulus × Label Condition interaction, F(1, 44) = 4.59, p < .05, and by a three-way Test Stimulus × Label Condition × Age interaction, F(1, 44) = 4.59, p < .05. Student Newman-Keuls tests (p < .05) were conducted on the highest order interaction. This post hoc analysis yielded two important findings. First, when infants were tested with the same stimulus, verbal labels had no effect on their performance. Second, when infants were tested with a different stimulus, the effectiveness of the verbal labels varied as a function of the infant's age. Although the verbal label had no effect on the performance of the 18-month-olds, the verbal label improved the performance of the 24-month-olds. In fact, when 24-month-olds were provided with the verbal label, their test performance with a different stimulus was equivalent to their test performance with the same stimulus that was present during the original demonstration.

**Discussion**

In most of our previous research on age-related changes in memory retrieval, we have assessed the relation between encoding and retrieval by altering the cues present at the time of the test (but see Hayne & Rovee-Collier, 1995). In essence, we have removed presumably effective retrieval cues (i.e., the demonstration stimuli) and have replaced them with cues (i.e., novel test stimuli) that may or may not be effective depending upon the nature of the task and the age of the infant. In the current experiment, we modified our typical approach by adding verbal cues to the demonstration and test sessions. We found that for 24-month-olds, the addition of verbal labels actually offset the effect of stimulus change such that infants provided with verbal labels responded equivalently whether the test stimulus was the same as or different from the one encountered during the demonstration.

In contrast, the test performance of 18-month-olds was not affected by the addition of a verbal retrieval cue; their memory performance was the same, regardless of whether they were exposed to the verbal labels or not. Under different test conditions, however, 18-month-olds, like 24-month-olds, might also profit from additional verbal information. In the present experiment, infants were tested after a 24-hr delay, but it is possible that 18-month-olds might exploit verbal retrieval cues when tested with...
different stimuli if the delay between the demonstration and test was shorter. This possibility remains to be tested.

The results of Experiment 2 add to a growing body of research showing that infants can use an adult’s language to guide their own behavior long before they are fluent speakers themselves. Prior work with 9–12-month-olds has shown that both visual attention (Balaban & Waxman, 1997; Waxman & Markow, 1995) and object sorting (Fuller & Haaf, 1998) are influenced by what adults say during the test. Using an elicited imitation procedure, Bauer and her colleagues (Bauer, Hertsgaard, & Wewerka, 1995) have shown that verbal cues facilitated performance by 15-month-olds tested with the demonstration stimuli after a delay. Across experiments, the nature of the task, the availability of other salient retrieval cues, and the timing of the test procedure undoubtedly influence the exact age at which infants can first exploit this kind of verbal information.

General Discussion

The results of the present study confirm that one hallmark of memory development during the infancy period is an age-related increase in the flexibility of memory retrieval. Changes in the test stimuli disrupted memory performance by 18-month-olds, particularly when they were tested after a 24-hr delay. These same changes in the test stimuli had no effect on the memory performance of 30-month-olds. The effect of altered test stimuli on the memory performance of the 24-month-olds was intermediate between these two extremes. This pattern of development is highly consistent with prior work on the effect of altered test stimuli on deferred imitation by 6–18-month-olds (Hayne et al., 1997, 2000).

The results of the present study also show that retention can be facilitated by the presentation of a verbal retrieval cue. Although changes in the test stimulus disrupted memory performance by 24-month-olds after a 24-hr delay (Experiment 1A), these same changes had no effect on infants’ performance when the experimenter provided a verbal label for the novel stimulus during the test (Experiment 2). The match between verbal cues present at the time of the demonstration and those present 24 hr later was apparently sufficient to cue retrieval of the target memory despite changes in the test stimulus per se. Children’s ability to exploit verbal retrieval cues, particularly in the absence of other salient retrieval cues, marks an important transition in memory development (Hudson, 1990; K. Nelson, 1993; see below). Our results and those of others (Bauer et al., 1995) show that children’s ability to exploit another person’s language in the service of memory begins to emerge during infancy.

Although memory retrieval at all ages is constrained by the principle of encoding specificity (Tulving, 1983, 1984; Tulving & Thompson, 1973), the nature of the match that is required apparently changes dramatically as a function of age. Collectively, our prior work and the data reported here indicate that the range of effective retrieval cues gradually broadens between 6 and 30 months of age. As such, the probability that infants will retrieve and express their memories in novel situations increases as a function of age. We hypothesize that these age-related changes in memory retrieval are made possible by age-related changes in the number or kind of attributes that are stored as part of a memory representation as well as by changes in the organization of multiple representations that share overlapping attributes. We refer to these age-related changes in the content and structure of memory as changes in representational flexibility (Hayne et al., 2000).

The term representational flexibility was originally used by Cohen and Eichenbaum (1993; Eichenbaum, 1997) to account for differences in the memory performance of adult rats with and without hippocampal lesions. In their experimental work, these investigators have shown that nonlesioned rats can learn an odor discrimination task and can apply this knowledge when confronted with new problems that involve the same rules. Rats with hippocampal lesions, however, can learn the original discrimination, but they cannot apply their knowledge to new situations. Within this context, Cohen and Eichenbaum have argued that representational flexibility is a characteristic inherent to declarative memory processing. The results of research conducted using the deferred imitation paradigm with human infants suggests that despite the precocious emergence of declarative memory, representational flexibility within this memory system continues to improve throughout the infancy period and perhaps even longer.

What factors might contribute to the age-related changes in representational flexibility that have been documented in studies conducted with human infants? We propose that both maturation and experience play an important role in this process. When tested in the same task, for example, older infants exploit a wider range of retrieval cues than younger infants. We have argued in the past that this age-related change may reflect, in part, the maturation of the higher association areas of the cortex (see Hayne et al., 2000). We hypothesized that as the cortex matures, the individual attributes that make up a memory representation become more integrated, increasing the number of retrieval paths for a particular memory.

A number of investigators have argued that the hippocampus also plays an important role in representational flexibility (Eichenbaum, 1997; O’Reilly & Rudy, 1999; Squire, 1992). According to O’Reilly and Rudy, for example, representations formed by the cortex are established slowly; flexibility in these representations emerges gradually on the basis of information that is extracted from many similar experiences. Representations formed by the hippocampus, however, are established rapidly; flexibility in these representations occurs automatically because individual attributes of the target event are bound together into a unitary representation that can be reaccessed when the participant subsequently encounters only a subset of the original experience. The relative contributions of the hippocampus and the cortex to developmental changes in the effect of novel retrieval cues remain to be determined; at present, an understanding of human brain maturation is still in its infancy.

In addition to the potential contribution made by the maturation of the brain, experience has been shown to play an important role in representational flexibility. Empirical research has shown that when infants of the same age are given additional experience with multiple stimuli or in multiple contexts during original encoding, then the range of effective retrieval cues for the memory broadens. When 3-month-old infants are trained for 3 consecutive days with the same stimulus in the mobile conjugate reinforcement paradigm, for example, memory retrieval is precluded by a change in the mobile at the time of the test. This high degree of specificity is overcome, however, if infants are exposed to a variety of mobiles during the original encoding (Fagen, Morrongiello, Rovee-Collier, & Gekoski, 1984; Greco, Hayne, & Rovee-Collier, 1990; Hayne, 1996; Hayne et al., 1986). Similarly, if original encoding occurs in
a variety of environmental contexts, subsequent retrieval is unaffected by a change in the context at the time of the test (Amabile & Rovee-Collier, 1991; Rovee-Collier & Dufault, 1991).

In studies of deferred imitation, representational flexibility is enhanced by providing infants with the opportunity to practice the target actions prior to the retention interval. In a traditional deferred imitation procedure, infants are not allowed to touch the stimuli or practice the target actions prior to the test (Barr & Hayne, 2000; Meltzoff, 1990). When 18-month-old infants are tested in this procedure, memory retrieval is precluded by large changes in the physical characteristics of the test stimuli (Experiment 1A; Hayne et al., 1997, Experiment 4). When infants of the same age are given the opportunity to practice the actions prior to the test, however, large changes in the test stimuli do not impair performance (Barr & Hayne, 2000; Bauer & Dow, 1994). Taken together, research conducted with a number of different experimental procedures has shown that at any given age, the range of effective retrieval cues for a particular memory varies as a function of the infants’ prior experience.

The studies described above indicate that additional experience with the experimental task per se can extend the range of effective retrieval cues for an infant’s memory of that task. Presumably, however, infants’ experiences outside the experimental context facilitate their ability to exploit novel retrieval cues in exactly the same way. In the course of their daily lives, for example, infants often encounter stimuli that match attributes stored as part of a memory representation. When this occurs, new information encountered at the time of retrieval may be added to the memory representation, and the range of subsequent retrieval cues is increased. We argue that as an infant’s range of experiences expands with development, so does the range of effective retrieval cues for a particular memory.

We predict that age-related changes in the effect of novel retrieval cues may be only one of the cognitive consequences of an increase in representational flexibility. Numerous studies have shown that there is a qualitative shift in a number of skills, including object permanence (Meltzoff & Gopnick, 1989), language (Gopnick & Meltzoff, 1986), and symbolic play (Leslie, 1987), that begins at approximately 18 months of age (i.e., the Stage 6 shift). Piaget argued that these changes signaled the emergence of mental representation. More recently, Gopnick and Meltzoff (1997; Meltzoff & Gopnick, 1989) have argued that these age-related changes in performance are not due to the emergence of representation per se but rather reflect a shift in the nature of the representation that infants use to guide their behavior. According to Gopnick and Meltzoff, young infants rely exclusively on representations based on their past experience, whereas older infants begin to rely on hypothetical representations made possible through deductive reasoning. We propose that representational flexibility is the mechanism that makes deductive reasoning possible. That is, as experience-based representations become more organized and interconnected, their content becomes more than merely the sum of their individual attributes, thereby allowing infants to consider possibilities for which they have no prior experience.

Finally, changes in representational flexibility may also contribute to the offset of childhood amnesia. Childhood amnesia refers to the inability of children and adults to recall events that occurred prior to their third or fourth birthday (e.g., Dudycha & Dudycha, 1941; Mullen, 1994; but see Usher & Neisser, 1993). Although there are a number of current theories regarding the source of childhood amnesia (Bachevalier, 1990; Howe & Courage, 1993; K. Nelson, 1993; Pillemert, Picariello, & Pruet, 1994), no single theory can adequately account for all the available data. On the basis of the present findings, we conclude that an age-related increase in the range of effective retrieval cues leads to an increase in the accessibility of memory over the long-term. We propose that the increased effectiveness of a wide range of retrieval cues, including but not limited to language cues (cf. Hudson, 1990; K. Nelson, 1993), contributes to the decline of childhood amnesia during the third year of life.

References


Reyna (Eds.), *Development of long-term retention* (pp. 3–55). New York: Springer-Verlag.


West, T. A., & Bauer, P. J. (1996, April). What is the effect of knowing that a rose is a rose: Effects of verbal information on young children’s recall of events. Paper presented at the biennial meeting of the International Conference on Infant Studies, Providence, RI.

Received September 1, 1999
Revision received February 1, 2000
Accepted February 2, 2000

**ORDER FORM**

Start my 2000 subscription to *Developmental Psychology*! ISSN: 0022-1649

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$87.00, APA Member/Affiliate</td>
<td>$174.00, Individual Nonmember</td>
</tr>
<tr>
<td>$370.00, Institution</td>
<td>In DC add 5.75% sales tax</td>
</tr>
</tbody>
</table>

**TOTAL AMOUNT ENCLOSED** $

Subscription orders must be prepaid. (Subscriptions are on a calendar basis only.) Allow 4–6 weeks for delivery of the first issue. Call for international subscription rates.

**SEND THIS ORDER FORM TO:**

American Psychological Association
750 First Street, NE
Washington, DC 20002-4242

Or call (800) 374-2721, fax (202) 336-5568.
TDD/TTY (202)336-6123. Email: subscriptions@apa.org

Send me a Free Sample Issue

☐ Check Enclosed (make payable to APA)

Charge my: ☐ VISA ☐ MasterCard ☐ American Express

Cardholder Name
Card No. _______________ Exp. date _______________

Signature (Required for Charge)

Credit Card
Billing Address
City State Zip
Daytime Phone

**SHIP TO:**

Name
Address
City State Zip

APA Customer # GA200

PLEASE DO NOT REMOVE – A PHOTOCOPY MAY BE USED