Past research using a deferred imitation task has shown that 6-month-olds remember a 3-part action sequence for only 1 day. The concept of a time window suggests that there is a limited period within which additional information can be integrated with a prior memory. Its width tracks the forgetting function of the memory. This study asked if retrieving the memory of the modeled actions at the end of the time window protracts its retention, if the type of retrieval (active or passive) differentially influences retention, and if the retrieval delay influences its specificity. In Experiment 1, 6-month-olds either imitated the modeled actions (active retrieval group) or merely watched them modeled again (passive retrieval group) 1 day after the original demonstration. Both groups showed deferred imitation after 10 days. In Experiment 2, 6-month-olds who repeatedly retrieved the memory at or near the end of the time window deferred imitation for 2.5 months. In Experiment 3, 6-month-olds spontaneously generalized imitation late in the time window after 1 prior retrieval, whether it was active or passive. These studies reveal that the retention benefit of multiple retrievals late in the time window is huge. Because most retrievals are undoubtedly latent, the contribution of repeated events to the growth of the knowledge base early in infancy has been greatly underestimated.

A small amount of additional training protracts retention in both adults (for review, see R. L. Cohen, 1985) and infants (Adler, Wilk, & Rovee-Collier, 2000; Barr, 2005).
Dowden, & Hayne, 1996; Fivush & Hamond, 1989; Galluccio & Rovee-Collier, 1999; Ohr, Fagen, Rovee-Collier, Hayne, & Vander Linde, 1989). Barr et al., for example, found that 12-, 18-, and 24-month-olds who viewed a sequence of target actions demonstrated on a hand puppet three times (a total of 30 sec) could imitate those actions 24 hr later, but 6-month-olds could not. When 6-month-olds viewed the demonstration six times (a total of 60 sec), however, they too could imitate the target actions 24 hr later. In addition to the absolute duration of training, its timing is critical as well (Barr et al., 1996; Gobbo, 2000; Hayne, 1990; Hudson & Sheffield, 1998; Rea & Modigliani, 1985; Schmidt & Bjork, 1992; Vander Linde, Morrongiello, & Rovee-Collier, 1985). Barr et al. also found that if the six repetitions were distributed across two sessions separated by 24 hr, then 6-month-olds could not imitate the target actions 24 hr later. Apparently, what transpires in the first session must be remembered long enough that its memory can be retrieved at the outset of the second session. When this occurs, what transpires in the second session accumulates with it; otherwise, infants behave as if the second session is only their first (Rovee-Collier, Evancio, & Earley, 1995).

The notion that there is a limited period within which new information can be integrated with old is encapsulated in the construct of a time window (Rovee-Collier, 1995). A time window opens with the onset of an event and shuts when that event is forgotten. New information that is encountered when the time window is open is integrated with the memory of the initial event, but if the same information is encountered after the time window has shut, then it is treated as novel. Although the width of a time window tracks the forgetting function of the memory, the forgetting function is just the timing device or clock that shuts the window. Moreover, every time the memory is retrieved before the clock times out, the clock is reset to a longer interval (i.e., the time window expands). In other words, after each additional retrieval, the memory of the event can be retrieved on a subsequent occasion after an even longer delay. Finally, the effects of retrieving the memory at different points within the time window are nonuniform: Retrieving the memory near the end of a time window expands its width more than retrieving it shortly after the time window first opens (Hartshorn, Wilk, Muller, & Rovee-Collier, 1998; Schmidt & Bjork, 1992).

In an early test of the time window construct, 3-month-old infants were operantly trained to move a series of category exemplars displayed on crib mobiles by kicking (Rovee-Collier, Greco-Vigorito, & Hayne, 1993). After category training was over, independent groups were briefly exposed to a functionally similar (i.e., moving) but physically dissimilar novel object. Infants who were exposed to the novel object after delays up to 4 days (but not 5 days) treated the novel object as a member of the training category. Thus, the time window for category inclusion shut after 4 days. In a follow-up experiment, infants who were first exposed to the novel object early in the time window, immediately after category training ended, exhibited retention for 4 days, whereas infants who were first exposed to the object
4 days after category training ended, at the end of the time window, exhibited retention for 10 additional days—14 days altogether. Hence, retrieving the memory just before the time window shut protracted retention substantially longer than retrieving it shortly after the time window opened.

More recently, the nonuniform effect of memory retrieval at different points in the time window was studied with elementary school children (Pipe, Sutherland, Webster, Jones, & La Rooy, 2004). Five- to 6-year-olds who participated in a “pirate event” were interviewed immediately, 1 day, 1 week, 1 month, or 6 months afterward, and all groups were interviewed again both 1 and 2 years after the original event. Children who were first interviewed after 6 months recalled significantly more about it 1 and 2 years later than children who were first interviewed immediately, 1 day, 1 week, or 1 month afterward. This result is consistent with the predictions of the time window construct.

The time window construct presumably embodies the fundamental principles that characterize memory processing by individuals of all ages and species, and considerable evidence consistent with this construct has been reported in studies of toddlers’ event memories, infants’ language acquisition, adults’ eyewitness testimony, verbal learning and memory with adults and children, animal fear conditioning, and human anxiety and memory disorders (for review, see Rovee-Collier, 1995). To date, however, its predictions have been explicitly tested with infants only in the mobile conjugate reinforcement paradigm (Galluccio & Rovee-Collier, in press; Rovee-Collier et al., 1995; Rovee-Collier et al., 1993). In this study, we asked whether the application of this construct is unique to operant studies with infants. To answer this, we used a deferred imitation paradigm with 6-month-olds to examine one of the major predictions of the time window construct—namely, that delaying memory retrieval until the end of the time window enhances the duration of retention.

The puppet imitation task is a deferred imitation task that was developed to investigate memory development in infants between 6 and 24 months old (Barr et al., 1996; Barr & Hayne, 2000; Hayne, Boniface, & Barr, 2000; Hayne, MacDonald, & Barr, 1997). Researchers have found that the duration of infants’ memory of the modeled actions is unaffected by whether infants imitate them immediately or not (Abravanel, 1991; Barr & Hayne, 1996; Meltzoff, 1995; see also Hayne, Barr, & Herbert, 2003). This finding is consistent with the time window construct: Because immediate imitation does not necessitate retrieval from long-term memory (i.e., the representation of the demonstration is still in active short-term memory when immediate imitation occurs), it should not affect the duration of retention—and data show that it does not (Barr & Hayne, 1996; Hayne et al., 2003). In fact, the first retrieval is actually at the time of the deferred imitation test.

At 6 months of age, the time window for the puppet imitation task (i.e., the maximum duration for which infants remember the demonstration) is only 1 day (Barr et al., 1996; Barr, Vieira, & Rovee-Collier, 2001). Because the memory of the
modeled actions is so brief at this age, we thought that any retention benefit afforded by delaying its retrieval until the end of the time window would be particularly apparent. To establish the retention benefit in all experiments, we modeled the target actions on Day 1, introduced an initial retrieval trial on Day 2 (i.e., at the end of the time window), and then tested independent groups of infants after increasing delays until they no longer could imitate the actions.

In Experiment 1, we asked whether retrieving the memory of the modeled actions at the end of the time window would protract retention longer than 1 day in the future and, if so, for how long. We also examined whether the type of retrieval—active or passive—might affect the subsequent duration of retention. Active retrieval was operationally defined as the infant’s actual performance (imitation) of the target actions, whereas passive retrieval was operationally defined as the infant’s mere observation of an adult performing the target actions again. (Because the passive retrieval was latent, we inferred its occurrence from the infants’ subsequent test behavior.) In Experiment 2, we examined how long successive retrievals, each at or near the end of the expanding time window, might protract retention. In Experiment 3, we asked whether the type of retrieval might differentially influence the specificity of the original memory.

EXPERIMENT 1: THE EFFECT OF PASSIVE VERSUS ACTIVE RETRIEVAL ON DEFERRED IMITATION

In the first experiment, we examined whether passively retrieving the memory of the target actions at the end of the time window protracts deferred imitation for the same duration as active retrieval. Because 6-month-olds can defer imitation of the target actions for only 1 day whether they imitate them immediately or not (Barr & Hayne, 1996; Barr et al., 2001), 1 day defines the time window of the puppet imitation task at this age. Therefore, we gave the active retrieval group its first opportunity to imitate the target actions three times after 1 day, and we gave the passive retrieval group an opportunity to merely view these actions being modeled three additional times after 1 day. Neither procedure produces retention after a 24-hr delay at 6 months (Barr et al., 1996; Barr et al., 2001). Thereafter, we tested independent groups from each retrieval condition after progressively longer delays until they failed to imitate the target actions.

Method

Participants. The final sample consisted of 63 full-term 6-month-olds (39 boys and 24 girls) who were recruited from public birth announcements, commercial mailing lists, and word of mouth. Infants ranged in age from 180 to 212 days ($M = 193.40$, $SD = 7.21$) on the first day of the study. Participants were African
American \(n = 1\), Asian \(n = 7\), Caucasian \(n = 48\), Hispanic \(n = 4\), and mixed race \(n = 3\). Their parents’ educational attainment ranged from 12 to 18 years \(M = 15.6, SD = 1.0\), and as reported by 92% of the sample, their ranks of socioeconomic status (Nakao & Treas, 1992) ranged from 29.39 to 93.74 \(M = 71.75, SD = 18.00\). Testing was discontinued on other infants because of scheduling difficulty \(n = 3\), failure to reach for the puppet \(n = 8\), excessive fussiness \(n = 1\), or experimenter error \(n = 3\). Based on the total number of opportunities (sessions) for a given infant to be lost from the sample, the rate of attrition was 9.7%.

The base rate at which 6-month-olds spontaneously produce the target actions was provided by a pooled baseline control group containing 45 infants (26 girls and 19 boys). This group contained all 6-month-olds who had participated in a baseline control group in our previous deferred imitation studies with the same stimuli, including 7 infants in the present study (Barr, Marrott, & Rovee-Collier, 2003; Barr et al., 2001; Barr, Vieira, & Rovee-Collier, 2002; Campanella & Rovee-Collier, 2005/this issue; Heravi, Barr, & Rovee-Collier, 2001; Learmonth, Lamberth, & Rovee-Collier, 2004; Muentener, Price, Garcia, & Barr, 2004). These infants did not see a demonstration of the target actions prior to the test. Their mean age on the first (and only) day of the experiment was 195.71 days \(SD = 8.34\). They were Caucasian \(n = 30\), Asian \(n = 6\), African American \(n = 3\), Hispanic \(n = 3\), and mixed race \(n = 3\). As reported by 97.8% of the sample, parental educational attainment ranged from 12 to 16 years \(M = 15.6, SD = 1.2\), and ranks of socioeconomic status (Nakao & Treas, 1992) ranged from 26.39 to 97.16 \(M = 72.26, SD = 18.89\).

**Apparatus.** Two hand puppets, a pastel pink rabbit and a pale gray mouse (see Figure 1), were constructed for these experiments and were not commercially available. Both were 30 cm in height and made of soft acrylic fur. A removable felt mitten \((8 \text{ cm} \times 9 \text{ cm})\) in a matching color covered each puppet’s right hand. A large

![Figure 1](image-url)
jingle bell was secured to the inside of the mitten during the demonstration, but it was removed during all tests and active retrievals. Puppets were counterbalanced within groups. A Panasonic VHS-C camcorder on a tripod was placed at right angles to the infant, and all sessions were videotaped for later scoring.

**Procedure.** Infants were tested in their own homes at a time when they were most likely to be playful. This time varied across infants but remained relatively constant across sessions of the same infant. Infants were randomly assigned to one of two experimental conditions—the active retrieval group and the passive retrieval group—as they became available for study. Independent groups of infants in each condition were tested after progressively longer delays until they failed to defer imitation significantly above the control level. This strategy yielded test delays of 3, 7, 10, and 14 days after the initial demonstration for both conditions.

**Demonstration session (Session 1).** Infants in both the active and passive retrieval groups participated in this session. During the demonstration, the infant sat on the caregiver’s knees. The experimenter knelt in front of the infant, placed the puppet on her right hand, and positioned it at the infant’s eye level and out of reach, approximately 80 cm from the infant’s chest. The experimenter then removed the mitten from the puppet’s right hand, shook it three times to ring the bell inside, and replaced it on the puppet’s hand. This sequence required approximately 10 sec and was repeated five more times for a total duration of approximately 60 sec.

**Retrieval session (Session 2).** One day after the initial demonstration, the experimenter returned to the infant’s home. As before, the infant sat on the caregiver’s knee, and the experimenter knelt in front of the infant with the puppet on her right hand. For infants in the active retrieval group, the puppet was positioned within the infant’s reach, approximately 30 cm in front of the infant, and the infant was given three opportunities to reproduce the target actions. The retrieval trial lasted approximately 3 min, depending on how quickly each infant responded ($M = 205.88$ sec, $SE = 21.01$). Infants in the passive retrieval group were treated identically except that instead of imitating the target actions themselves, they merely observed the experimenter model the sequence of target actions three more times.

**Test session (Session 3).** The test session was identical for the two experimental groups and the pooled baseline control group and also was procedurally identical to the retrieval session (Session 2) of the active retrieval group. During the test session, the bell was removed from the mitten. The experimenter placed the puppet within the infant’s reach, and the infant was allowed 120 sec from the time he or she first touched the puppet in which to imitate the target actions (see Figure 2). Infants in the experimental groups were tested with the same puppet they had seen before.
Results and Discussion

An imitation score was calculated for each infant by summing the number of target behaviors (remove the mitten, shake the mitten, and attempt to put the mitten back on the puppet) that were produced during the test (range = 0–3). (Deferred imitation in this task is rarely veridical and typically consists of reproducing either one or two of the three possible actions. Six-month-olds rarely attempt to replace the mitten.) One observer scored all videotaped test sessions, and a second observer who was blind to the infants’ group assignments independently scored 94.5% of the videotapes. The interobserver reliability was 96.7% ($\kappa = .92$). When the two raters differed, the primary rater’s score was assigned. Interobserver reliability for the pooled baseline control group was calculated separately; based on dual coding of 100% of the videotapes, interobserver reliability was 99.2% ($\kappa = .92$).

An initial one-way analysis of variance (ANOVA) over the imitation test scores of all active retrieval groups 24 hr after the initial demonstration indicated that they did not differ, $F(3, 24) < 1$ (range of means = 1.16–1.83). This result eliminates group differences in original learning as a basis for group differences in retention after longer delays.

We addressed two basic questions in Experiment 1. First, did the type of retrieval after 1 day differentially benefit infants’ ability to defer imitation over the long term? Second, for how long did a single retrieval of either type protract deferred imitation? To answer the first question, we performed a two-way ANOVA over the imitation scores of infants in the two retrieval conditions (active and

FIGURE 2  A 6-month-old during a deferred imitation test. The target actions were originally modeled on either the same puppet (Experiments 1 and 2) or a different one (Experiment 3) in Session 1.
passive) who were tested after one of the four delays (3, 7, 10, and 14 days). This analysis yielded a significant main effect of delay, $F(3, 48) = 4.98$, $p < .01$, but not of retrieval condition, $F(1, 48) = 1.10$, ns; the two-way interaction was also not significant, $F(3, 48) < 1$, and there was a large effect size, $d = .42$ (Kirk, 1995; see Figure 3, left panel). A post hoc test (Duncan’s multiple range test, $p < .05$) revealed that the mean deferred imitation scores of the groups that were tested after 14 days were significantly lower than the mean imitation scores of the groups that were tested after all shorter delays, which did not differ from one another. Thus, 6-month-olds’ retention of the target actions was unaffected by whether the initial retrieval was active or passive; rather, infants in the two retrieval conditions performed equivalently after all delays. This result is surprising in view of the obvious functional differences in the two retrieval conditions. The passive retrieval condition and the initial modeling condition were veridical, but the active retrieval condition and the initial modeling condition were not. Infants in the active retrieval condition acquired additional and slightly different retrieval cues as a result of physically interacting with the puppet, yet this difference produced no apparent retention advantage.

FIGURE 3 The mean imitation scores of independent groups of infants ($n = 7$) in the active and passive retrieval groups in Experiments 1 and 3. The mean test score of the 6-month pooled baseline control group is indicated by the dashed line. Left panel: Mean imitation scores of groups in Experiment 1 who were tested with the demonstration puppet after delays of 3, 7, 10, or 14 days. The immediate active retrieval group tested on Days 1, 2, and 3 is from Barr et al. (2001). The passive control group shown the target action three times for 2 consecutive days is from Barr et al. (1996). Right panel: Mean imitation scores of groups in Experiment 3 who were tested with a novel puppet after 7 days (passive retrieval group) or 10 days (active retrieval group). An asterisk indicates that a test group exhibited significant deferred imitation (i.e., its mean imitation score was significantly greater than the mean test score of the baseline control group). Vertical bars indicate ±1 standard error.
At first blush, the finding that active and passive retrieval conditions produced equivalent retention appears to be at odds with the finding of Adler et al. (2000) that reinstatement (an interactive reminder) was more effective than a reactivation treatment (a passive reminder) in prolonging subsequent retention. Because Adler et al. used the operant mobile task, their finding undoubtedly reflects how the memory was initially acquired. In addition, their finding is consistent with the finding that reactivation is relatively ineffective when a memory is still accessible (Gordon, 1981; Hayne, 1990; Hsu, 2004).

Although the ANOVA indicated that the imitation test performance of the retrieval conditions did not differ, it did not answer the second question; namely, after what test delays did infants significantly defer imitation? To answer this question, we collapsed imitation scores across active and passive retrieval conditions and performed a one-way ANOVA over the mean imitation scores of groups that were tested after each of the four delays and the mean test score of the pooled baseline control group. This analysis indicated that the test groups differed significantly, $F(4, 96) = 19.11, p < .001$ (effect size, $d = .84$). A post hoc test (Dunnett’s $t$ test, $p < .05$) revealed that the mean imitation test scores of groups tested after all delays except 14 days (4 days: $M = 1.14, SE = 0.17$; 7 days: $M = 1.07, SE = 0.22$; 10 days: $M = 1.42, SE = 0.20$; 14 days: $M = 0.35, SE = 0.20$) were significantly higher than the mean test score of the pooled baseline control group ($M = 0.13, SE = 0.05$). (The Dunnett’s test controls for Type I errors across multiple comparisons with a control group; Dunnett, 1955.) Thus, merely retrieving the memory of the demonstration 1 day later extended its subsequent retention tenfold, from 1 day to 10 days, irrespective of whether retrieval was active or passive (see Figure 3, left panel).

Because the impact of a single retrieval on deferred imitation was so great, in Experiment 2, we asked how much longer 6-month-olds’ memory of the modeled actions would be protracted by additional retrievals.

**EXPERIMENT 2: THE EFFECT OF MULTIPLE RETRIEVALS ON DEFERRED ImitATION**

In Experiment 1, infants who first retrieved the memory of the target actions 1 day after they were modeled subsequently deferred imitation for 10 days. In Experiment 2, we asked how long infants given additional retrieval opportunities could remember the modeled actions. In a previous test of the time window construct using an operant paradigm, Galluccio and Rovee-Collier (in press) found that the duration of 3-month-olds’ retention expanded exponentially over successive retrievals. In Experiment 2, therefore, we asked whether the retention benefit of each succeeding retrieval would be absolute (e.g., 10 days) or whether it would progressively increase. To answer this question, we introduced two more retrieval opportunities after the 10-day test. If each subsequent retrieval affords an absolute future
retention benefit of 10 days, then infants who first retrieved the memory 1 day after
the original demonstration should successfully defer imitation on Days 10 (as in
Experiment 1), 20, and 30. If the retention benefit increases logarithmically over
successive retrievals, however, then infants should successfully defer imitation af-
after increasingly longer delays.

Because the retention of 3-month-olds who received a brief reinstatement (an
active retrieval) had increased exponentially with the interval between training and
retrieval (Galluccio & Rovee-Collier, in press), we tested deferred imitation after a
logarithmically (log2) increasing series of delays in which the 10-day time window
that resulted after one retrieval was successively doubled. This progression en-
sured that succeeding retrievals would continue to be nearer the end of the expand-
ing time window. Because the time window construct holds that the memory of a
prior event must be retrieved in order for subsequently encountered information to
be integrated with it, we tested infants in the active retrieval condition only. In this
way, we were able to ascertain whether infants had actually retrieved the memory
of the modeled actions from their memory performance on each succeeding test.

The two experimental groups included the 10-day active retrieval group from
Experiment 1, which we retested twice more, and an active retrieval replication
group. The two control groups included a familiarization control group that was re-
peatedly exposed to the puppet but did not initially see the modeled actions and an
8-month baseline control group, necessitated because infants in the experimental
groups were 2 to 2.5 months older at the time of their final deferred imitation test.
An experimental group was considered to exhibit significant imitation if it per-
formed significantly above the base rate at which 8-month-olds spontaneously
produced the target actions.

Method

Participants. The final sample consisted of 24 full-term infants (11 boys and
13 girls), recruited as before. Six of these infants had participated in the active re-
trieval group in Experiment 1. Six-month-olds (8 boys and 10 girls) ranged in age
from 184 to 217 days ($M = 194.33, SD = 8.53$) on the first day of the experiment;
8-month-olds (3 boys and 3 girls) ranged in age from 251 to 271 days ($M = 258.50,$
$SD = 7.67$). Participants were Asian $(n = 2)$, African American $(n = 1)$, Caucasian
$(n = 20)$, and mixed ethnicity $(n = 1)$. Their parents’ educational attainment, re-
ported by 95.8% of the sample, ranged from 14 to 16 years $(M = 15.5, SD = 0.9)$,
and their ranks of socioeconomic status (Nakao & Treas, 1992), reported by 83.3%
of the sample, ranged from 35.07 to 97.16 ($M = 71.49, SD = 18.00$). Testing was
discontinued on additional infants because of scheduling difficulty $(n = 1)$, failure
to reach for the puppet $(n = 5)$, and experimenter error $(n = 2)$. Based on the total
number of opportunities (sessions) for an infant to be lost from the sample, the rate
of attrition was 20%. 
**Apparatus and procedure.** The puppets and general procedure were identical to those that were used with the active retrieval group in Experiment 1. This group was retested for deferred imitation in Experiment 2 on Days 30 and 70. The active retrieval replication group first deferred imitation 1 day after the target actions were modeled. Subsequent active retrievals occurred on Days 10, 30, and 60.

The familiarization control group was exposed to the puppet on the same regimen and for the same amount of time as the experimental groups except that it never saw a demonstration of the target actions and did not touch the puppet until its final test session. Half of the infants in the familiarization control group were tested on Day 70, and half were tested on Day 60. To control for the age of infants in the experimental groups at the time of their final test, a baseline control group was tested at 8 months of age. Infants in this group did not see a demonstration of the target actions; rather, their first session was the test session. The baseline control group provided a measure of the base rate at which 8-month-olds spontaneously produce the target actions.

**Results and Discussion**

All infants in both experimental groups successfully deferred imitation of the target actions during each of the first three tests (retrievals) on Days 1, 10, and 30. To assess whether the two experimental groups differed after any test (retrieval) delay, we performed a 2 (70-day test group, 60-day test group) × 4 (retrieval delays: 1, 10, 30, 60 or 70 days) ANOVA, with repeated measures across retrieval delays. This analysis yielded no significant main effects and a nonsignificant interaction (all $F$s < 1). Figure 4 shows the group performance during each retrieval trial of the two repeated-retrieval groups. (Recall that exposure to the puppet and the experimenter was yoked for the familiarization group.)

A one-way ANOVA over the final imitation test scores of the four groups—70-day experimental group ($M = 1.17, SE = 0.40$), 60-day replication group ($M = 1.17, SE = 0.40$), familiarization control group ($M = 0.17, SE = 0.17$), and 8-month baseline control group ($M = 0.00, SE = 0.00$)—indicated that they differed significantly, $F(3, 20) = 7.58, p < .002$ (effect size, $d = .90$). A post hoc test (Duncan’s multiple range test, $p < .05$) revealed that the mean test scores of the experimental groups did not differ, and both were significantly higher than the mean test scores of the two control groups, which also did not differ. Thus, 6-month-olds who saw an adult model the target actions and then retrieved the memory after exponentially increasing delays still reproduced the target actions 2 to 2.5 months later. Notably, a full month (at least) intervened between the third retrieval and the final test. Thus, repeatedly retrieving the memory of the target actions after a series of expanding retention intervals enabled both experimental groups to remember the modeled actions 60 to 70 times longer than they would have remembered them otherwise. Finally, we note that 12-month-olds can remember the puppet imitation task for a maximum of 1
18-month-olds can remember it for a maximum of 3 weeks (for review, see Barr & Hayne, 2000). In Experiment 2, however, 6-month-olds given three retrievals remembered the task for 30 to 40 days—a duration of retention substantially longer than the duration for which infants two and three times older can remember. These data, then, reveal some of the conditions by which relatively short-lived memories of events can be maintained over vastly extended periods of time.

Although the final deferred imitation test occurred when infants were 8.5 months old, 2.5 months after the original demonstration, the puppet imitation task is suitable for use through 2 years of age. Hartshorn (2003) previously found that five spaced reinstatements (active retrievals) enabled operantly trained infants to remember the training event from 6 months through 2 years of age. Given that infants are highly motivated to play the puppet imitation “game” through 2 years of age as well (Barr et al., 1996), we think it highly likely that if infants in this study had been given two additional reminders, then they also would have remembered the target actions through 2 years of age (see also Hudson & Sheffield, 1998).

**FIGURE 4** The mean imitation scores of the repeated active retrieval group and the repeated active retrieval replication group after test delays of 1, 10, 30, and 60 or 70 days. The familiarization control group was tested on average after 65 days. The mean test score of the 8-month baseline control group is indicated by the dashed line. An asterisk indicates that a test group exhibited significant deferred imitation (i.e., its mean imitation score was significantly greater than the mean test score of the baseline control group). Vertical bars indicate ±1 standard error.

**EXPERIMENT 3: THE SPECIFICITY OF MEMORY**

The memories of young infants are highly specific to the stimuli presented during the original event (Hartshorn, Rovee-Collier, Gerhardstein, Bhatt, Klein, et al., 1998). At 3 and 6 months of age, for example, operantly trained infants do not
spontaneously generalize to a novel test cue 24 hr after training, but they increasingly generalize after longer test delays (Bhatt & Rovee-Collier, 1996; Borovsky & Rovee-Collier, 1990; Hartshorn, Rovee-Collier, Gerhardstein, Bhatt, Klein, et al., 1998; Rovee-Collier & Sullivan, 1980). At 6 months of age, infants also cannot imitate the target actions on a novel test puppet 24 hr after the demonstration (Hayne et al., 2000; Learmonth et al., 2004). When Hayne et al. (2000) demonstrated the target actions on a pink pastel rabbit and tested 6-month-olds 24 hr later with a pastel gray mouse (or vice versa), for example, they failed to imitate, leading to the conclusion that infants could not generalize between puppets at 6 months. Twelve-month-olds who were treated identically also did not generalize between puppets 24 hr later (Hayne et al., 1997). The brevity of 6-month-olds’ memory of the demonstration, however, has previously precluded examining generalization to a novel test puppet after longer delays.

In fact, members of most species exhibit a flattening of generalization gradients over time, irrespective of task (Riccio, Ackil, & Burch-Vernon, 1992; Riccio, Rabinowitz, & Axelrod, 1994; Thomas & Burr, 1969). Because this result is also shared by human infants in operant tasks, it seems likely that it would also be shared by 6-month-olds in a deferred imitation task if they could be tested after a longer delay. One way to study this possibility might be by increasing the number of times the target actions are modeled (Barr et al., 1996). Exactly how this parameter should be altered, however, is uncertain. The preceding experiments suggest an alternative approach: Because infants remembered the target actions for 10 days when the memory was retrieved 24 hr after the demonstration, this longer period of accessibility offers an expanded opportunity to assess whether the specificity of the active memory representation decreases over time.

In Experiment 3, therefore, the memory of the target actions was retrieved 1 day after the demonstration, and infants were tested with a novel puppet at the end of the expanded time window (7–10 days), when they were most likely to have forgotten the specific details of the demonstration puppet. Although the type of retrieval had not affected the duration for which the memory was protracted in Experiment 1, whether the type of retrieval might affect the specificity of the memory is unknown. In Experiment 3, therefore, we also asked if active and passive retrieval differentially influence specificity.

Method

Participants. The final sample consisted of 14 full-term 6-month-olds (7 boys and 7 girls), recruited as before. Infants ranged in age from 174 to 219 days ($M = 196.57, SD = 14.90$) on the first day of the study. Participants were Caucasian ($n = 12$), Asian ($n = 1$), and Hispanic ($n = 1$). Their parents’ educational attainment ranged from 12 to 16 years ($M = 15.43, SD = 1.22$), and their ranks of socioeconomic status (Nakao & Treas, 1992), reported by 93.3% of the sample, ranged from 39.59 to 92.30
Testing was discontinued on 5 additional infants who failed to reach for the puppet. Based on the total number of opportunities (sessions) for a given infant to be lost from the sample, the rate of attrition was 8.8%.

**Apparatus and procedure.** The apparatus and procedure were the same as in Experiment 1 except that the final test puppet was novel. Use of the two puppets as the demonstration puppet (puppet A) and the final test puppet (puppet B) was counterbalanced within groups. Infants were randomly assigned to active retrieval and passive retrieval groups as they became available for study.

Because both groups had remembered the target actions after 10 days in Experiment 1, we asked whether infants would generalize at the end of the time window or whether they might do so even earlier. To answer these questions, we tested the active retrieval group after 10 days and the passive retrieval group after 7 days. These test delays were selected because previous researchers had reported that active imitation facilitates the generalization of imitation across test puppets by older infants (Hayne et al., 2003; Learmonth et al., 2004). Should the passive retrieval group generalize imitation after 7 days but the active group not generalize imitation after 10 days, we planned to test an active retrieval group after 7 days as well. It seemed unlikely, however, that this would occur. On the other hand, if the passive retrieval group failed to generalize after 7 days, then we would not know if their generalization failure resulted because they had not actively imitated or because they had not been tested at the end of the time window. Should this occur, therefore, we planned to test a passive retrieval group after 10 days. As it turned out, neither additional test group was necessary.

One day after the target actions were modeled on puppet A, infants in the active retrieval group were given three opportunities to imitate the target actions on puppet A, whereas infants in the passive retrieval group observed the experimenter model the sequence of target actions on puppet A three times. For the active retrieval group, the retrieval trial lasted approximately 5 min, depending on how quickly each infant responded ($M = 298.29$ sec, $SE = 53.77$). After the specified retention interval had elapsed (7 or 10 days), each group received a deferred imitation test with puppet B.

One observer scored 100% of the videotaped test sessions, and a second observer, who was blind to the infants’ group assignments, independently scored 95% of them. The interobserver reliability was 96.7% ($\kappa = .93$). When the two raters differed, the primary rater’s score was assigned.

**Results and Discussion**

A one-way ANOVA over the mean imitation test scores of the active and passive retrieval groups and the 6-month pooled baseline control group indicated that they differed significantly, $F(2, 56) = 19.00, p < .001$ (effect size, $d = .78$). A post hoc
test (Duncan’s multiple range test, \( p < .05 \)) revealed that the imitation scores of the two retrieval groups (passive retrieval: \( M = 1.00, SE = 0.21 \); active retrieval: \( M = 0.86, SE = 0.26 \)) were higher than the test score of the pooled baseline control group (\( M = 0.18, SE = 0.05 \)) and not different from one another (see Figure 3, right panel). These results indicate that both retrieval groups exhibited significant imitation, generalizing to the novel test puppet. Because the passive retrieval group deferred imitation on a novel puppet after 7 days and because the type of retrieval had no impact on deferred imitation in either Experiment 1 or Experiment 3, the active retrieval group would undoubtedly generalize imitation after 7 days as well.

In Experiment 3, we were able to document that an accessible memory representation becomes more generalized over time by exploiting the finding that an initial retrieval extends the accessibility of the memory. Although 6-month-olds cannot defer imitation on a novel test puppet 24 hr after the original demonstration regardless of whether or not they had imitated immediately after the demonstration (Learmonth et al., 2004, Experiment 3), they can do so after a longer delay regardless of whether or not they had imitated the actions 1 day after they were modeled. These findings parallel findings from operant studies with 3- and 6-month-olds, who fail to generalize to a novel test cue 24 hr after training but do so after a longer delay (Bhatt & Rovee-Collier, 1996; Hartshorn, Rovee-Collier, Gerhardstein, Bhatt, Klein, et al., 1998; Rovee-Collier & Sullivan, 1980). They also parallel findings from studies with nonhuman species, whose generalization gradients progressively flatten over time (Feinberg & Riccio, 1990; Riccio et al., 1992; Riccio et al., 1994; Thomas & Burr, 1969).

**GENERAL DISCUSSION**

This study offers testimony to the predictive power of the time window construct. Even though retention is initially brief at 6 months, retrieving a memory on multiple occasions—each at or near the end of the time window—protracted retention exponentially. After just four increasingly spaced retrievals, 6-month-olds remembered the target actions longer than infants at least three times their age (Barr & Hayne, 2000). These data indicate that the immaturity of young infants’ brains is not the rate-limiting step in how long they remember (Liston & Kagan, 2002; Nelson, 1995, 1997, 1998). Rather, how long infants remember is determined by two experiential variables: the number of times a memory is retrieved and the time since the previous retrieval. This insight is not new: Nearly 30 years ago, Wagner (1976) argued that the retrieval process in nonverbal organisms is equivalent to rehearsal in verbally proficient individuals, and Bjork (1975; Landauer & Bjork, 1978; Schmidt & Bjork, 1992) argued that increasingly greater spacing between successive retrievals produces better retention because retrieval is cognitively more difficult or effortful.
In retrospect, these variables rather than task differences have been largely responsible for the apparent retention differences in operant and deferred imitation studies of infant memory (Rovee-Collier & Hayne, 2000). In operant studies, 6-month-olds are trained on 2 consecutive days. Thus, as in this study, they experience one retrieval after 24 hr. It is not surprising, therefore, that the maximum duration of retention in the two types of study is similar at 6 months as well—14 days in operant studies (Hartshorn, Rovee-Collier, Gerhardstein, Bhatt, Wondoloski, et al., 1998) and 10 days in this one. What is surprising is that infants in this study viewed the demonstration for a total of only 90 sec over 2 days, whereas infants in operant studies are trained for a total of 12 min over 2 days. Any residual retention advantage of operantly trained infants, therefore, can probably be attributed to their greater amount of training (Ohr et al., 1989), but this advantage is very slight considering the magnitude of the difference.

Many developmental researchers attribute the apparent retention differences of infants in operant and deferred imitation studies to differences in the underlying memory systems that the two types of task presumably tap. McDonough (cited in Bower, 1995, p. 86), for example, characterized infant memory in operant tasks as “only procedural” because the response that indicates retention is acquired via a conditioning procedure. Procedural memory is thought to be a primitive memory system that operates automatically, allows organisms to learn “how” but not “what” (N. J. Cohen & Squire, 1980), and makes few if any cognitive demands. Similarly, Bauer (1996) stated that evidence of infant retention from operant studies reveals little about infants’ ability “to construct and maintain accessible … memories … because the behaviors from which they are derived do not meet the criteria for recall” (p. 30). These and other researchers have claimed that a deferred imitation task, in contrast to an operant conditioning task, is a valid nonverbal indicator of declarative (explicit) memory, which is thought to be a higher level form of memory that emerges when infants are older (Bauer, 1996, 1997; Bauer & Hertsgaard, 1993; Bauer, Hertsgaard, & Wewerka, 1995; Mandler, 1990, 1998; McDonough, cited in Bower, 1995; McDonough, Mandler, McKee, & Squire, 1995). The findings reported here reveal that operant and deferred imitation tasks are functionally similar. Modeling a sequence of target actions on a particular object is one way of introducing infants to the target information and providing them with a nonverbal means of indicating later that they recognize it. Affixing the target information on the sides of the training mobile or in the training context and reinforcing a distinctive response to it is simply another. In short, the debate over the influence of practice and the contribution of different memory systems to retention should be reframed in terms of the specific parameters of the task in which memory is measured (for discussion, see Rovee-Collier, Hayne, & Colombo, 2001).

The applied implications of the time window construct are far reaching. Interrogators who repeatedly interview child witnesses about a prior event, for example, must ensure that the children’s memory of the event is accurate and consistent.
across multiple interviews. Peterson (1999; Peterson, Moores, & White, 2001; Peterson, Parsons, & Dean, 2004) found that four successive interviews at 6-month intervals with children aged 2 to 13 years maintained their memory of a personal injury without decreasing their accuracy, as long as they did not encounter misleading information. When false information is introduced during an interview, however, children who are younger are more likely to be misled (Sutherland & Hayne, 2001). The finding from Experiment 3 that memories become more generalized after a relatively long delay suggests why younger children are more susceptible to false information: Once the specific details of the prior event have been forgotten, the details of a similar event can be substituted for them in the memory (Boller, Rovee-Collier, Gulya, & Prete, 1996; Galluccio, 2001; Neisser, 1997; Rovee-Collier, Adler, & Borza, 1994).

In addition, practitioners can exploit knowledge of time windows to devise personalized interventions to alleviate specific learning and memory deficits of cognitively impaired individuals of all ages, whether the individuals have brain damage or not. The use of an expanding series of memory retrievals, for example, has yielded dramatic improvement in the recall of severely amnesic patients (Camp, Foss, O’Hanlon, & Stevens, 1996; Wilson, 1989). Developmental psychopathologists may be able to find a relation between the width of a time window and genetic factors, birth disorders (e.g., Down syndrome, phenylketonuria), or exposure to certain toxins or drugs (e.g., PCPs, lead, smoke, alcohol, cocaine). Finally, educators may be able to relate particular aspects of time windows (e.g., their initial width, how they expand) to factors that compromise scholastic achievement (see Cull, Shaughnessy, & Zechmeister, 1996, for a similar argument).

This study reveals that infants not only pick up information covertly, via mere observation, but they also retrieve it covertly. In turn, mere retrieval—latent or not—protracts retention. This finding raises three intriguing possibilities. First, as long as an infant periodically encounters appropriate retrieval cues, a latent memory that was formed early in development can be maintained for some time, until the infant has both an opportunity and a reason to express it. If so, then scientists greatly underestimate what preverbal infants actually know. Second, infants might even form a latent memory for behaviors that they cannot yet physically perform. If the latent memory is maintained via periodic encounters with appropriate retrieval cues, infants might perform them once their motor skills “catch up.” Campanella and Rovee-Collier (2005/this issue) recently obtained evidence that supports this possibility. Third, although nonverbal memories for single events seem not to translate into verbal memories (Simcock & Hayne, 2002), the memory of an event that has been repeatedly retrieved over a long period might be more likely to be expressed verbally after a very long delay than one that has not.

In conclusion, although the time window construct is not restricted to a particular type of content, a particular age, or even a particular stage of development, it is
especially critical for an understanding of the impact of experience early in infancy, when the knowledge base is first being formed.

ACKNOWLEDGMENTS

This research was funded by National Institute of Child Health and Human Development Grant HD043047–01 to Rachel Barr and by National Institute of Mental Health Grant MH32307 to Carolyn Rovee-Collier. We thank Yasna Heravi, Amaya Garcia, and Aurora Vieira for assistance with the data collection.

REFERENCES


Neisser, U. (1997, November). *Enabling conditions for false memories*. Colloquium presented to the Department of Psychology, Rutgers University, New Brunswick, NJ.


