PAPER

Forming a stable memory representation in the first year of life: why imitation is more than child's play

Angela F. Lukowski,[1] Sandra A. Wiebe,[1] Jennifer C. Haight,[1] Tracy DeBoer,[1] Charles A. Nelson[2] and Patricia J. Bauer[1]

1. Institute of Child Development, University of Minnesota, USA
2. Institute of Child Development and Department of Pediatrics, University of Minnesota, USA

Abstract

Although 9-month-old infants are capable of retaining temporally ordered information over long delays, this ability is relatively fragile. It may be possible to facilitate long-term retention by allowing infants to imitate event sequences immediately after their presentation. The effects of imitation on immediate and delayed recognition and on long-term recall were investigated using event-related potentials (ERPs) and elicited imitation, respectively. Mnemonic facilitation resulting from the opportunity to imitate was apparent using both assessments. ERP assessments at immediate and delayed recognition tests suggested that infants who were allowed to imitate had stronger memory representations of familiar stimuli relative to infants who only viewed the presentation of the events. In addition, infants who were allowed to imitate evidenced higher levels of ordered recall after 1 month relative to infants who only watched the experimenter's demonstration. Therefore, imitation proved to have beneficial effects on explicit memory in $9^{1}/_{2}$-month-olds, providing evidence of its effectiveness as a tool to augment mnemonic capabilities in infancy.

An old proverb suggests that 'an elephant never forgets'. If true, this phenomenon stands in marked contrast to human mnemonic capabilities, which are marred with errors. Infants' memories are even more susceptible to forgetting than are those of adults. Nevertheless, 6-month-old infants evidence deferred imitation of single actions over delays of 24 hours (e.g. Collie & Hayne, 1999), and 9-month-old infants evidence retention of the individual actions of multi-step sequences over delays of 1 month (Bauer, Wiebe, Waters & Bangston, 2001; Carver & Bauer, 1999, 2001). However, even after multiple exposures to to-be-remembered material, only a subset of 9-month-olds evidence ordered recall over long delays. One possible method of promoting encoding, storage, and later retrieval of memories in infancy may be through imitation, which may be expected to aid recall by strengthening memory representations. In the present study we investigated the beneficial effects of imitation on immediate and delayed recognition and long-term recall using event-related potentials (ERPs) and elicited imitation, respectively.

Together, electrophysiological assessments of recognition and behavioral assessments of recall are powerful tools for understanding memory development. ERPs allow for recording of the spatio-temporal distribution of neural events during stimulus processing (see DeBoer, Scott & Nelson, 2004; Nelson & Monk, 2001); elicited imitation allows for an assessment of recall of temporal order information (e.g. Bauer & Mandler, 1992). The marriage of these two techniques is an important tool that allows for the collection of data that ultimately may provide insights into the neurological underpinnings of behavior.

In the elicited imitation paradigm, infants witness actions that are demonstrated by an experimenter. Either immediately after presentation (elicited imitation), after a delay (deferred imitation), or both, infants have the opportunity to reproduce the modeled actions. The technique is widely accepted as a nonverbal analogue to verbal report (e.g. Bauer, 2002; Mandler, 1990; Meltzoff, 1990; K. Nelson & Fivush, 2000; Rovee-Collier & Hayne, 2000; Squire, Knowlton & Musen, 1993; Wheeler, 2000). Because the argument for this analogy has been developed in detail elsewhere (e.g. Bauer, 1995, 1996, 1997, 2002, in press-b, in press-c;

Address for correspondence: Patricia J. Bauer, Institute of Child Development, 51 East River Road, Minneapolis, MN 55455-0345, USA; e-mail: pbauer@umn.edu

Carver & Bauer, 2001; Mandler, 1990; Meltzoff, 1990), here we provide a brief summary of four points of the argument. First, imitation after exposure to a model is one of the hallmarks of representational capacity (e.g. Piaget, 1952). Second, once children develop language skills, they talk about events experienced in the context of imitation (e.g. Bauer, Wenner & Kroupina, 2002). This is strong evidence that the representational format in which the memories are encoded is explicit or declarative, as opposed to implicit or non-declarative (as these types of memories are not amenable to verbal report). Third, adults suffering from amnesia due to temporal lobe damage are unable to perform an age-appropriate version of the task (McDonough, Mandler, McKee & Squire, 1995). This suggests that the imitation procedure taps the type of memory that gives rise to explicit recognition and recall. Finally, imitation taps recall rather than recognition. Although the available props provide perceptual support for performance, all recall is cued to some extent (Spear, 1978). Additional evidence that recall processes support performance comes from temporally ordered reproduction of modeled multi-step sequences. After modeling, no perceptual support for the order of the presented actions remains. To reproduce an ordered sequence, temporal order information must be encoded during presentation of the event and subsequently be retrieved from memory in the absence of ongoing perceptual support. In this way, the requirements of imitation-based tasks closely parallel those of verbal recall paradigms (Mandler, 1990).

Use of the elicited and deferred imitation paradigms has revealed evidence of long-term ordered recall by late in the first year. At this time, however, memory capabilities are fragile and unreliable. Whereas $9^1/_2$-month-olds are able to replicate specific novel actions immediately and after a 24-hour delay (Meltzoff, 1988), memory for multi-step sequences is apparent only after multiple exposures to to-be-remembered events. Indeed, even after multiple presentations, memory is not reliably observed: Carver and Bauer (1999) found that 45% of infants evidenced recall of the temporal order of sequences after a 5-week delay, whereas 55% of the infants did not. This approximate distribution has been replicated in two independent samples (Bauer, Wiebe, Carver, Waters & Nelson, 2003; Bauer *et al.*, 2001). These data indicate that whereas some infants encode, store, and retain temporal order information over long delays, others do not show evidence of these capabilities.

Presumably, the lack of delayed recall by half of the tested infants relates to the integrity of the underlying memory trace: those who recalled the events had intact memory traces whereas the representations of those who did not recall were degraded (Bauer, in press-a; Bauer *et al.*, 2003). Consistent with this hypothesis, relations between ERP indices of recognition and behavioral indices of recall implicate variability in the strength of the stored memory representation as the basis for individual differences. Specifically, in Carver, Bauer and Nelson (2000), infants who showed recognition of photographs of familiar events (as indicated by differential processing of familiar and novel stimuli) after a 1-week delay also demonstrated ordered recall 1 month later. In contrast, infants who did not show evidence of recognizing the familiar stimuli also did not demonstrate temporally ordered recall after 1 month. These data indicate direct relations between the underlying strength of the memory representation and the ability to access that information over the long term. In an extension of this work, Bauer *et al.* (2003) provided evidence that infants who did and did not subsequently evidence ordered recall encoded the event sequences, as evidenced by differential processing at an immediate recognition test. However, in a replication of the findings of Carver *et al.* (2000), only infants who subsequently recalled the events in the correct temporal order after 1 month evidenced recognition memory after 1 week.

Because the relations between recognition and recall memory clearly implicate trace strength as a source of variability in mnemonic capabilities late in the first year, we wanted to determine whether we could manipulate memory trace strength and thus, the probability of subsequent recognition and recall. We chose to investigate the effect of immediate imitation on indices of later recognition and recall memory for three reasons. First, current behavioral evidence is inconclusive regarding the effects of immediate imitation on memory capabilities. Re-enactment of events at the time of first experience of them has been shown to facilitate subsequent recall in 10-year-olds (e.g. Baker-Ward, Hess & Flannagan, 1990). Similarly, 15-month-olds who were allowed to imitate event sequences immediately after a single exposure to them recall more actions at a 1-month delayed recall test than did those children who only viewed the presentation of the sequences (Bauer, Hertsgaard & Wewerka, 1995). However, other data have indicated that immediate imitation does not necessarily facilitate later recall. In the same study in which they found facilitation by imitation of events experienced only once, Bauer and her colleagues (1995) found that immediate imitation did not influence later recall of events to which 15-month-olds were exposed three times before the 1-month delayed recall test. Similarly, Meltzoff (1995) did not find any difference between 14- and 16-month-old children who were and were not allowed to imitate single, object-specific actions before imposed delays of 2 or

4 months. Immediate imitation also failed to benefit 18-month-olds tested on 3-step sequences over a 1-week delay (Barr & Hayne, 1996), or 13-, 16-, or 20-month-olds' recall after delays of 1 to 12 months (Bauer, Wenner, Dropik & Wewerka, 2000). As such, the behavioral data collected to date indicate that imitation may be influential in some circumstances but not in others.

Second, all of the data on whether imitation immediately after modeling affects subsequent recall have been collected with children in the second year of life and beyond. By the middle of the second year, long-term recall abilities are both reliable and robust (Bauer, 2002, in press-c). In contrast, at 9 months of age, they are neither: only approximately 50% of 9-month-olds evidence ordered recall after a 1-month delay and recall is seemingly dependent on multiple exposures to the sequences prior to imposition of the delay (Bauer *et al.*, 2001, 2003; Carver & Bauer, 1999). This raises the possibility that younger infants may benefit from the opportunity to imitate, even though older infants may not.

Third, manipulation of the opportunity to imitate also permitted us to begin to elucidate the neural correlates of imitation, as evidenced by ERP responses. Overwhelmingly, data from the adult literature indicate overlapping patterns of neural activity for events that are imitated and events that are only watched, with some specific differences in processing between these manipulations. For example, Carr, Iacoboni, Dubeau, Mazziotta and Lenzi (2003) examined the locus and strength of neural activation using fMRI when participants either watched or watched and imitated emotional facial expressions. Across the two conditions they noted similar patterns of activation in the premotor face area, the dorsal portion of the pars opercularis, the superior temporal sulcus, the insula, and the amygdala. However, activity was greater in these areas when participants imitated the expressions compared to when they only watched them. In addition, some areas of activation were apparent only when imitating: whereas activation in the dorsal region of the pars opercularis was found in both the imitate and watch conditions, the ventral portion of this region was only activated when imitating facial expressions.

In an ERP task, Senkfor, Van Petten and Kutas (2002) also observed overlapping patterns of processing with some evidence of variation across performance and observation conditions. Adult participants were initially asked to encode three-dimensional everyday stimuli (e.g. a fork) by (a) performing typical actions with the objects ('perform-encoded', such as bringing a fork to their mouths), (b) watching an experimenter perform typical activities (such as watching the experimenter bring the fork to her/his mouth), (c) imagining themselves performing typical activities, or (d) estimating the costs of

the items. At test, ERPs were recorded as the adults reported how they initially encoded the item (i.e. perform, watch, imagine, or cost estimation). There were similar ERP responses in the perform and watch conditions at a variety of electrode sites from 800 to 1400 ms after stimulus presentation. Between 600 and 800 ms after stimulus onset, however, perform-encoded objects elicited more positive electrophysiological responses at posterior and medial electrode sites than did any of the other encoding categories. Differences in the encoding of performed versus watched items were also seen in the behavioral data: participants were fastest and most accurate at naming the encoding mechanism used for performed items. These data, together with those from Carr *et al.* (2003), suggest there are both similarities and differences in neural processes when participants watch and when they perform actions. The differences in processing may be associated with variability in behavioral performance.

In the present study we investigated the effects of immediate imitation on the formation of stable, long-term event memories in $9\frac{1}{2}$-month-old infants. Infants were presented with three novel two-step sequences at each of three exposure sessions. At each session, one group of infants was allowed to imitate the experimenter's actions immediately after the presentation of the events; another group only viewed the presentation of the sequences. In addition, each infant participated in two ERP recognition assessments: one immediately after the second exposure session and one directly preceding the third. Finally, recall memory was tested behaviorally approximately 1 month after the third session.

The use of both electrophysiological and behavioral measures allowed us to address multiple questions. The primary question was whether the opportunity to imitate sequences after their presentation facilitated the formation of stronger memory representations as evidenced by recognition, recall, or both. We also asked whether the effects were apparent over the short term (i.e. at encoding), over the long term (i.e. at the delayed assessments), or both. If the beneficial effects of imitation are sustained over time, differential responses should be apparent at both the recognition and behavioral assessments between the groups of infants who were and were not given the opportunity to imitate the sequences. If the facilitating effects of imitation are derived only from practice in the elicited imitation paradigm, then no effect on recognition memory would be expected during the ERP assessments, since the behavioral and electrophysiological protocols are distinct (i.e. the skills garnered from one task would not be expected to generalize to the other). That is, infants who are allowed to imitate the event sequences should evidence higher levels of recall on familiar versus novel event sequences only at the

delayed recall assessment as compared to those infants who only view the presentation of the sequences. If different responses are evident between the two groups of infants in both the electrophysiological and behavioral analyses, conceptual correlations can be made regarding brain activity and behavioral capabilities. That is, evidence suggesting the benefits of imitation in two unrelated tasks indicates the veracity of the findings and lessens the possibility that the results could be attributed to some unaccounted, extraneous variable.

Method

Participants

Seventy-nine $9^1/_2$-month-old infants (mean age = 9 months, 15 days; range 9 months, 2 days to 9 months, 24 days) participated in the study. An additional 13 infants were enrolled but did not complete all required testing sessions due to scheduling conflicts. All participants were recruited from the infant participant pool of a large midwestern university. The pool includes infants whose parents were contacted by mail shortly after their infants' births and subsequently returned informational cards stating their willingness to be contacted about participation in research. All of the participants in the current study were full-term (40 ± 2 weeks gestational age) and had no known mental or physical conditions or disorders. Seventy-two of the infants were of Caucasian descent, two were of mixed Caucasian and Asian descent, three were of mixed Caucasian and African-American descent, one infant was of Caucasian, Asian, and African-American descent, and one infant was of Hispanic, Pacific Islander, Asian, and Caucasian descent. Most of the infants were living in middle- to upper-middle-class families.

Materials

Behavioral stimuli

Infants were presented with three novel two-step sequences drawn from a pool of 12 sequences.[1] The props used in the sequences were designed to be unlike materials that 9-month-olds normally encounter. Each sequence consisted of two actions that culminated in an interesting end-state, and the sequence orders were constrained by enabling relations (as in Bauer *et al.*, 2003; Carver & Bauer, 1999, 2001; Carver *et al.*, 2000), such that the two actions had to be completed sequentially for the end-state to be realized. The presentation of the sequences was block-randomized such that each participant saw one of four sets of three sequences; one other block of three sequences was used at the recall test as a within-subjects control.

The choice to use sequences constrained by enabling relations, rather than arbitrarily ordered sequences, was made so as to allow the infants the best opportunity to imitate. Bauer and her colleagues have demonstrated that recall is enhanced when sequences are constrained by enabling relations, relative to when they are absent (e.g. Bauer, 1992; Wenner & Bauer, 1999). In fact, prior to 20 months of age, infants and toddlers perform at chance when tested on sequences that lack enabling relations (Wenner & Bauer, 1999). Therefore, the presence of enabling relations should allow $9^1/_2$-month-olds the best chance to immediately reproduce what they see before them and recall the presented information at a later date.

Recognition stimuli

Infants also participated in two recognition memory assessments in which they saw still photographs of one familiar (previously presented) and one novel (not previously presented) sequence. Photographs depicted a woman's hand completing Step 1 and Step 2 of the sequences, and all of the props used to complete the event sequence.

Questionnaires

Parents completed MacArthur-Bates Communicative Development Inventories (Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994) and Infant Behavior Questionnaires (IBQ; Rothbart, 1981) before their final visits to the laboratory. The MacArthur-Bates is a validated and respected measure of language development that indicates (a) infant comprehension and (b) infant comprehension and production of words and phrases. The IBQ contains six scales that include common behaviors and reactions expressed by infants: smiling and laughter, duration of orientation, distress to limitations, distress to novelty, soothability, and activity level. These measures are typically used to evaluate possible relations between language capabilities, temperament, and recognition and recall (e.g. Carver & Bauer, 1999). In the present report, this measure was examined as a means of testing for uncontrolled differences between groups of infants who (a) were and were not permitted to imitate, and (b) did and did not provide useable ERP data.

---

[1] Such a large pool of sequences was necessary because half of the participants (namely, those in the watch group) were participating in the first wave of a longitudinal study that necessitated different sequences at each session (a partial report of the longitudinal component of the study is provided in Bauer *et al.*, in press).

Table 1 *Outline of presentation of event sequences across sessions*

Session	Mean delay and range in days	Testing phase					
		Baseline	Modeling	Immediate ERP recognition test	Delayed ERP recognition test	Modeling	Delayed behavioral recall
Session 1	–	A, B, C	A, B, C	–	–	–	–
Session 2	2 (range 1–4)	–	A, B, C	A, D	–	–	–
Session 3	7 (range 7–14)	–	–	–	B, E	A, B, C	–
Session 4	29 (range 21–37)	–	–	–	–	–	A, B, C, D, E, F

Note: At Sessions 1, 2, and 3, infants were presented with event sequences A, B, and C. At Session 2, infants experienced a behavioral re-exposure to event sequences A, B, and C before the recognition assessment; at Session 3, infants experienced the recognition assessment before event re-exposure. At recognition, event sequences A (at Session 2) and B (at Session 3) represent any one of the three sequences that were experienced behaviorally at Session 1. Event sequences were counterbalanced across infants so that each one was used as familiar and novel equally as often. At recall, event sequences D, E, and F represent another block of three events chosen to serve as a within-subjects control.

Procedure

Infants participated in three exposure sessions, two ERP sessions, and one recall session (see the schematic in Table 1). Each infant was tested by one of seven female experimenters; infants were tested by the same experimenter at each session. The experimenters were trained on procedures similar to those used in Carver and Bauer (1999) and adherence to the protocol was checked periodically. All experimenters remained faithful to the protocol.

Exposure sessions

At the first exposure session, infants experienced a warm-up period during which they were familiarized with the experimenter by playing with toys unrelated to the study. At the same time, the parents were reminded of the procedure, were encouraged to ask questions of the experimenter, and signed informed consent statements. Once the session began, infants were presented with the materials for each of three sequences in turn. During this baseline assessment, the infants were allowed to explore each set of materials for approximately $1^1/_2$–2 minutes. The experimenter then modeled the sequence, using verbal descriptions of each step. For example, for the event sequence 'Find Bubbles', the experimenter initially said, 'I can use this stuff to find Bubbles. Watch how I find Bubbles with this stuff.' While pushing a button on a track so as to allow a block to move down a pathway unobstructed, she said, 'Push the button.' While pushing the block down the track to trigger a switch that made a puppet appear, the experimenter said, 'Slide it.' After each sequence was presented, infants in the 'imitate' group ($n = 46$) were given the opportunity to reproduce it immediately. The imitation period ended after infants had interacted with the props for $1^1/_2$–2 minutes. Infants in the 'watch' group ($n = 30$) were not permitted to interact with the materials after sequence demonstration.

The second session (mean delay = 2 days; range 1–4 days) began with a short warm-up period, after which infants were exposed to the three event sequences that they had seen previously. As before, infants in the imitate group were allowed the opportunity to immediately reproduce the sequences modeled by the experimenter; infants in the watch group were allowed to play with distracter toys. The distracter toys were included in the protocol for two reasons. First, allowing infants to play with toys unrelated to the test sequences reduced any possible frustration that could have resulted from being required to view many event presentations without any opportunity for interaction with the materials. Second, the use of distracter toys was necessary to equate the amount of time infants in the two groups spent in the laboratory (i.e. if infants in the watch group had not received distracter toys, their sessions would have been substantially shorter than those of the infants in the imitate group). After all three sequences had been presented, the immediate recognition test was administered (see *Recognition assessments*).

At the beginning of the third session (mean delay = 7 days; range 7–14 days), infants completed the delayed recognition assessment (see *Recognition assessments*) followed by a third exposure to the event sequences. The sequences were presented exactly as they were at the first two sessions. Once again, infants in the imitate group were allowed to imitate the experimenters' actions whereas the infants in the watch group received distracter toys.[2] The modeled events were counterbalanced across sessions so that each sequence was presented in a different serial position at each session.

Recognition assessments

ERPs were recorded at the end of the second session and at the beginning of the third session to assess recognition memory for the familiar sequences. Specifically, ERPs were used at the end of the second exposure session to
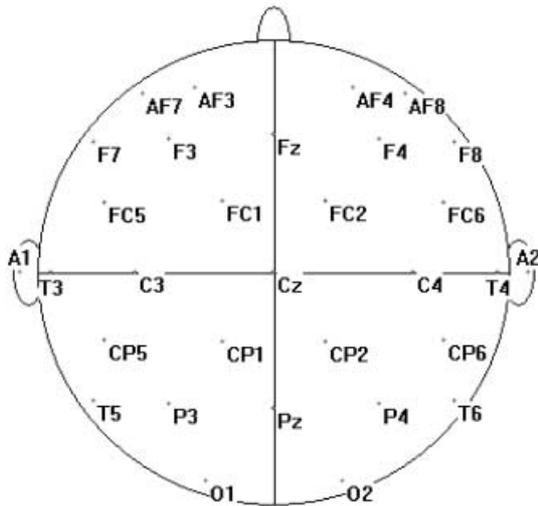
Figure 1 *Headplot indicating the placement and names of the measured electrode sites.*

determine whether infants had encoded the events to which they had been exposed. We reasoned that if the infants exhibited differential processing of the familiar and novel stimuli, they must have encoded the events. The immediate ERP test was administered as soon as possible after the exposure to the sequences at the second session. In practice, testing began an average of 31 minutes after modeling was complete (range 18–43 minutes required to fit the electrode cap). The delayed ERP test was administered at the third session and allowed for a measure of delayed recognition memory (as such, infants experienced the ERP test before event re-exposure).

Immediately before the ERP test, infants were fitted with a nylon Electrocap© that was fastened under their chins with Velcro straps. Data were recorded from 25 scalp locations (Fz, Cz, Pz, AF3, AF4, F3, F4, F7, F8, FC5, FC6, C3, C4, CP1, CP2, CP5, CP6, P3, P4, T3, T4, T5, T6, O1, and O2) placed according to the international 10–20 system (Jasper, 1958; Figure 1). In addition, electro-ocular electrodes were placed above and below the right eye; one electrode was also placed on

each mastoid bone. Electrodes were filled with a conductive gel and a mildly abrasive cleanser, and impedances were usually below 10 kΩ and were often less than 5 kΩ. Impedance readings were balanced across electrode pairs and hemispheres.

Infants were tested individually while seated on their parents' laps approximately 75 cm from a computer screen. The screen was set within a black barrier so that infants could only view a portion of the room during testing. The first trial began when the infant fixated the screen. Infants were presented with randomized digitized pictures of the first step, second step, and all of the props for each of two sequences, one to which they had been familiarized and one that was novel to them. The presentation of the first picture was predetermined so that familiar and novel slides were not disproportionately represented. The remaining slides were randomly presented, and the sequences were counterbalanced across participants so that each one served as familiar and novel equally often.

During the presentation of the stimuli, the infant's gaze was monitored by an observer from a small hole within the black barrier. If the infant looked away from the screen or became inattentive during the recording period, the observer redirected the infant's attention by tapping on the wall or by squeaking an infant toy, and the trial was repeated. Data were not recorded when the infant was not attending to the screen, when the parent or experimenter was speaking, or when the infant's attention was being redirected to the stimulus. A maximum of 100 trials was presented to the infants. The mean number of completed trials at the immediate ERP recognition test was 72 (range 41–100) for infants in the imitate group and 69 (range 43–100) for infants in the watch group. The mean number of completed trials at the delayed ERP recognition test was 67 (range 31–100) for infants in the imitate group and 61 (range 31–100) for infants in the watch group.

EEG was recorded using a Grass Neurodata Acquisition System with Model 12A5 amplifiers. EEG gain was set at 20,000 Hz and EOG gain was set at 5000 Hz. Bandpass filters were set between .1 and 30 Hz. A notch filter was also in place at 60 Hz. Each trial consisted of a 100-ms baseline followed by stimulus presentation for 500 ms. Data were recorded for 1200 ms after the end of the stimulus presentation. Throughout the recording epoch, EEG was sampled every 10 ms (100 Hz) and was referenced to Cz online. The inter-trial interval varied randomly between 400 and 900 ms, including the 100-ms baseline of the following stimulus. After collection, the data were digitized online and stored on a hard disk. An average reference was applied before data analysis (see *Data reduction*).

---

[2] This procedure was followed in order to reduce the number of sessions parents and infants were required to complete. Previous research made clear that infants who were required to defer imitation for 1 month needed three exposures to sequences in order to recall them after the delay (Bauer *et al.*, 2001). Had we included three exposure sessions, followed by a 1-week delayed recognition session, then followed by a 1-month delayed recall assessment, five visits to the laboratory would have been necessary; for infants in the longitudinal sample, the required number of visits would have totaled 10. To reduce the total number of sessions, yet still provide three exposure sessions and both an immediate and delayed recognition test, we adopted the procedure used here.

Recall assessment

At the fourth session (mean delay = 29 days; range 21–37 days), infants were tested for their recall memory of the three previously seen sequences and they were presented with three novel control events. After a short warm-up period, infants were presented with the props for each sequence in turn, along with a verbal reminder of the desired end-state ('reminders' were given for both familiar and novel control events for infants in the imitate and watch groups; for control events, the verbalizations served as suggestions of activities that could be performed with the props). For example, for the sequence 'Find Bubbles', the experimenter said, 'You can use this stuff to find Bubbles. How do you find Bubbles with this stuff?' The infants were then given approximately $1\frac{1}{2}$–2 minutes to perform the target actions. Performance during this phase served as a measure of long-term recall for previously presented sequences and as a baseline measure for the novel control sequences.

Data reduction

Behavioral data

Overall, 79 infants contributed behavioral data. However, the data from three infants were not included in the behavioral analysis due to procedural error or recording equipment failure. Thus, the recall performance of 76 infants was included in the behavioral analyses.

All of the sessions were videotaped for the infants in the imitate group; only the first (baseline) and the fourth (recall) sessions were videotaped for infants in the watch group. An experienced behavioral coder who was unaware of the hypotheses of the study recorded each infant's behavior (including both the performance of target actions and the order in which they were completed). For each sequence, an infant could produce two target actions in one correct order (Figure 2). For example, in the sequence 'Turn on the light', one must put a car into the vertical portion of an L-shaped clear apparatus (Step 1) before pushing a plunger-like stick (Step 2) to propel the car down the track, ultimately allowing it to trip a switch that causes a small light to illuminate (end-state). However, infants received credit for performing the target actions in the incorrect order (i.e. infants could push the plunger and then put the car into the apparatus, thereby receiving credit for 2 target actions and 0 ordered pairs). Only the first occurrence of each target action was coded so as to reduce credit that might be received due to chance or trial and error, thereby providing for a more conservative measure of memory performance. Across the groups, a second coder independently recoded the

Figure 2 *Photographs of infant completing the event sequence 'Turn on the light'. To complete the first step, the infant places a small car into the track of an L-shaped apparatus. To ultimately realize the end-state of the event, she pushes a plunger, thereby causing the car to travel down the track, tripping a small switch that causes a light to illuminate. Used with permission: Bauer, P.J., Wiebe, S.A., Carver, L.J., Waters, J.M., & Nelson, C.A. (2003). Developments in long-term explicit memory late in the first year of life: behavioral and electrophysiological indices.* Psychological Science, **14**, *629–635.*

tapes for 20 (25%) of the 76 infants. Mean percent agreement between the coders on both the performance of target actions and their order was 86% (range from 66% to 100%). Although this range of reliability values is lower than that typically seen in studies conducted with older children (e.g. Bauer & Mandler, 1989; Bauer *et al.*, 2000), these values parallel those found in other investigations with 9-month-old participants (Bauer *et al.*, 2003; Bauer, Wiebe, Carver, Lukowski, Haight, Waters & Nelson, in press; Carver & Bauer, 1999). When disagreements occurred between the coders, the observations of the primary coder were used.

Recognition data

The ERP data were re-referenced offline using an average reference technique; the reference was applied following the procedures outlined in Bauer *et al.* (in press). Our choice of the average reference was motivated by several factors. As discussed in Bauer *et al.* (in press), a reference should ideally be a neutral site devoid of electrical activity. In practice, such a site does not exist on the human body (Dien, 1998; Geselowitz, 1998; Junghöfer, Elbert, Tucker & Braun, 1999). Many studies of infant ERPs, including some of our own previous work (Bauer *et al.*, 2003; Carver *et al.*, 2000), have used linked mastoids as a reference. However, the signal recorded from the mastoids is potentially influenced by activity at nearby brain regions, such as the temporal lobes. Additionally, references based on linked mastoids are known to attenuate the amplitude of components at nearby sites (Dien,

1998). Perhaps because of our past reference choice, we have observed effects at the midline leads with little activity related to memory apparent at the temporal leads (Bauer *et al.*, 2003; Carver *et al.*, 2000). In contrast, in Bauer *et al.* (in press), we observed effects both at the midline and at lateral-posterior leads.

In the present study, to expand the range of sites over which activity might be observed, we selected an average reference constructed from 24 scalp leads (Fz, Pz, AF3, AF4, F3, F4, F7, F8, FC5, FC6, C3, C4, CP1, CP2, CP5, CP6, P3, P4, T3, T4, T5, T6, O1, and O2). In at least one study with typical adults, an average reference constructed from 19 electrodes yielded results similar to those found using other reference techniques (e.g. Onofrj, Fulgente, Thomas, Locatelli & Comi, 1995). Because we used more channels with infants (who have smaller heads and thus, smaller inter-electrode distances), we were afforded even better coverage. Consequently, our calculation of the average reference could reasonably be expected to result in a reference devoid of electrical activity due to sufficient measurement of positive and negative currents across various scalp locations (insufficient coverage of the scalp may lead to artifact, such that the reference point may not be devoid of electrical activity and will, as a result, provide contaminated output; Dien, 1998). Even with appropriate application of the average reference, however, we cannot draw conclusions about the location of the neural generators of the ERP components under investigation due to insufficient spatial coverage (e.g. Junghöfer *et al.*, 1999; Nunez, 1990). Also, because different references will result in different absolute amplitudes of ERP components, there are limits on the extent to which we are able to compare the findings of the present research with those of previous studies in which other reference techniques were used. Nevertheless, we are able to make qualitative comparisons, since ERP waveforms and topographic maps constructed from different references display similar features (Geselowitz, 1998).

In constructing the grand means, data were excluded if the EEG signal exceeded analog to digital values in any 100-ms window or if the EOG signal exceeded 250 μV in any 250-ms window. Specifically, if the EEG signal for any channel exceeded −225 to 225 μV for longer than 100 ms, that channel was marked as bad; if three or more bad channels were marked for one trial, that trial was excluded from further analysis. Individual averages were obtained for each infant, with the constraint that an equal number of trials was included for each condition. If 10 or more trials were included in the averages for an infant, the average waveform was visually inspected to exclude data contaminated by motion artifact. The mean number of cross-averaged trials at the first ERP

assessment was 22 (range 9–38) for infants in the imitate group and 21 (range 10–35) for infants in the watch group; at the second ERP assessment, the mean number of cross-averaged trials was 16 (range 5–28) for infants in the imitate group and 21 (range 8–35) for infants in the watch group. Grand means were then created from the uncontaminated data for each group of infants (i.e. imitate and watch) for each event type (i.e. familiar and novel) at each of the two phases of the recognition assessment (i.e. immediate and delayed).

For the immediate ERP recognition test, 17 infants in the imitate group (mean age = 9 months, 14 days; range 9 months, 8 days to 9 months, 11 days; 11 girls) and 11 infants in the watch group (mean age = 9 months, 14 days; range 9 months, 8 days to 9 months, 18 days; 6 girls) provided useable data. The data from 50 infants were excluded because (a) the EEG or EOG reading exceeded the set limits (13 infants), (b) of failure to complete a sufficient number of trials (13 infants), (c) of procedural error (1 infant), or (d) of equipment failure (21 infants). The data lost to equipment failure was due to the recording of ERPs using a faulty amplifier. Once the fault in the amplifier was realized, we discontinued use of it and discarded all data recorded with it. Only data recorded with properly functioning equipment were analyzed.

For the delayed recognition test, 18 infants in the imitate group and 10 infants in the watch group provided analyzable data. The data from 50 infants were excluded because (a) the EEG or EOG reading exceeded the set limits (17 infants), (b) of failure to complete a sufficient number of trials (11 infants), (c) of procedural error (2 infants), or (d) of equipment failure (19 infants). Again, the large amount of data lost to equipment failure was due to use of a faulty amplifier, and these data were excluded from analysis.

Overall, 19 infants (13 from the imitate group and six from the watch group) provided ERP data at both recognition memory assessments; nine infants (four from the imitate group and five from the watch group) provided data at the immediate recognition memory assessment only; and nine infants (five from the imitate group and four from the watch group) provided data at the delayed recognition memory assessment only. Except for one infant, all participants who provided useable ERP data also provided behavioral data at delayed recall; 40 infants who provided behavioral data did not provide useable ERP data.

Two major components of the ERP waveform were analyzed statistically. These components are present in infants less than 1 year of age and are believed to reflect different cognitive processes (Courchesne, Ganz & Norcia, 1981). The first component was a middle-latency deflection that occurs approximately 400–800 ms after

stimulus onset (Courchesne *et al.*, 1981; Nelson, 1994). At frontal and central midline leads, this deflection is negative and is referred to as the Nc (Nelson, 1994). This component is thought to reflect attentional processes (Courchesne *et al.*, 1981; de Haan & Nelson, 1997; Nelson, 1994; Nelson & Dukette, 1998; Nelson, Henschel & Collins, 1993; Richards, 2003), with a larger deflection indicating a greater allocation of attention to the presented stimuli (Nelson *et al.*, 1993). The component is also influenced by the familiarity or probability of the stimulus presentation, suggesting that it also indexes certain aspects of long-term memory for stimuli that are well established in memory (e.g. Bauer *et al.*, 2003, in press; Carver *et al.*, 2000; de Haan & Nelson, 1997, 1999; Nelson *et al.*, 1993). At lateral-posterior sites, the middle-latency deflection is positive (Bauer *et al.*, in press) and may or may not be homologous to the anterior-based negative-going Nc. After comparing the data for the imitate and watch groups, the time window containing a prominent middle-latency component was identified as 260–700 ms after stimulus onset. As in de Haan and Nelson (1997), the grand means of each group and condition were visually inspected to make certain that the specified window included the component under investigation.

The other component of interest, the slow wave, was analyzed in a long-latency window from 900 to 1500 ms. Two potential variations in slow wave activity include a deflection from baseline (typically positive at anterior sites and negative at posterior sites) and a return to baseline. Whereas the deflections from baseline are thought to indicate memory updating of a partially encoded stimulus (Nelson, 1994), a return to baseline indicates that the stimulus has been well encoded and no further processing is necessary (see DeBoer *et al.*, 2004, for more detailed explanations of potential components of interest in developmental research).

Two dependent measures were analyzed for the middle-latency component; one dependent measure was analyzed for slow wave activity. For the middle-latency component, the minimum and maximum peak amplitudes (for the midline-anterior and lateral-posterior electrode sites, respectively) and the latency to peak amplitude were computed. Because slow wave activity is devoid of readily apparent peaks and exists over a relatively large time window, the dependent measure for this component is the area under the curve relative to the baseline measure.

Results

Three sets of analyses were conducted to determine the effects of imitation on recognition and recall memory. The results of the immediate recognition memory test at

the second session indicated the effects of imitation on encoding processes, whereas the results of the delayed recognition assessment at the third session illustrated the effects of imitation over the long term. These data are presented in the first section. The behavioral data presented in the second section indicate the effects of imitation on recall memory over a 1-month delay. The analyses found in the third section indicate that the differences in recognition and recall between the imitate and watch groups are attributable to the imposed between-group manipulations and are not the result of uncontrolled, confounding factors (i.e. differential length of delays, language capabilities, temperamental characteristics, selective attrition).

Recognition data

Immediate recognition test

In previous related research (Bauer *et al.*, 2003; Carver *et al.*, 2000), analyses were concentrated on the midline-anterior leads Fz, Cz, and Pz due to the relatively small number of electrodes used and the need to employ a linked-mastoids reference. The rationale was that, when using a linked-mastoids reference, little activity was apparent at leads T3, T4, T5, T6 (Bauer *et al.*, 2003; Carver *et al.*, 2000), and C3 and C4 (Bauer *et al.*, 2003). In more recent research using a greater number of electrodes and average reference techniques, however, there is evidence of differential brain activity at lateral-posterior leads in response to photographs of familiar and novel objects (event sequences: Bauer *et al.*, in press; toys: Snyder, 2002; Snyder & Nelson, 2001). Accordingly, in the present research we analyzed both midline-anterior and lateral-posterior electrode sites. Based on the results of Bauer *et al.* (in press), de Haan and Nelson (1997), and Dawson, Carver, Meltzoff, Panagiotides, McPartland, and Webb (2002), we expected a reversal of polarity between the anterior and posterior sites. Specifically, at anterior sites, we expected a negative middle-latency component and a positive slow wave; at posterior sites, we expected a positive middle-latency component and a negative slow wave. The anticipated polarity reversal made it advisable to examine the midline-anterior and lateral-posterior leads in separate analyses: in a combined analysis, main effects would have been obscured. Separate analyses of the midline-anterior and lateral-posterior sites also permitted qualitative comparisons between data from the present research and data collected previously (e.g. Bauer *et al.*, 2003; Carver *et al.*, 2000).

To determine the effects of imitation on immediate recognition at the midline-anterior leads, we conducted 2 (group: imitate, watch) × 2 (condition: familiar, novel)

× 2 (lead: Fz, Cz) mixed analyses of variance (ANOVAs) with repeated measures on event type and lead, for each of three dependent variables: peak amplitude of the middle-latency component, latency to the peak amplitude of the middle-latency component, and area under the curve for the slow wave activity. We did not include electrode site Pz because it was the point at which the polarity of effects shifted from negative (at anterior sites) to positive (at posterior sites). Parallel analyses were computed for the six lateral-posterior leads (T5, T6, P3, P4, O1, and O2).

At the midline-anterior leads, there was a main effect of condition on the peak amplitude [$F(1, 26) = 4.54, p < .05$] of the middle latency component (Figure 3). This effect indicated that the infants differentially processed the photographs of the familiar and novel sequences. Specifically, the peak amplitude of the middle-latency component was more negative to novel ($M = -10.41$ μV, SD = 8.05 μV) than to familiar ($M = -7.88$ μV, SD = 6.81 μV) sequences. This effect is consistent with the pattern observed in Bauer *et al.* (2003, in press). That the effect was not qualified by an interaction with group indicates that both groups of infants encoded the sequences. No significant effects were found for the other dependent measure in the middle-latency window, or for the midline-anterior leads in the long-latency window.

Differential processing of the familiar and novel sequences across the groups was also apparent at the lateral-posterior leads. Specifically, a Group × Condition interaction was found for slow wave activity: $F(1, 26) = 6.08, p < .02$ (Figure 4). Analyses of the interaction revealed a main effect of group on the familiar sequences: $F(1, 26) = 4.18, p = .05$. Infants in the imitate group had larger area scores in response to the familiar stimuli ($M = -5110.98$ ms × μV, SD = 7963.86 ms × μV) relative to infants in the watch group ($M = -2612.00$ ms × μV, SD = 6300.88 ms × μV). Therefore, whereas all of the infants encoded the familiar events, differential processing apparent in the slow wave at the lateral-posterior leads indicates a stronger memory representation for infants in the imitate group relative to the infants in the watch group. For the novel stimuli, the effect of group was not significant. Thus, the difference in processing was confined to the events to which the infants had been familiarized.

Delayed recognition test

The analyses for the delayed recognition test paralleled those for the immediate recognition test. At the midline-anterior leads, there were no significant effects for any of the dependent measures. The data from lateral-posterior electrode sites, however, indicated the beneficial effects of imitation in the formation of a stable memory trace over the long term. Specifically, in the middle-latency

window there was a Group × Condition interaction on the variable of maximum amplitude: $F(1, 26) = 5.76, p < .03$. When the data were analyzed by group (Figures 5 and 6), trends towards main effects of condition were apparent for both groups: for the imitate group, $F(1, 17) = 4.17, p = .06$; for the watch group, $F(1, 9) = 4.37, p = .07$. Although they are below the conventional level of statistical significance, these effects suggest differential processing of the familiar and novel stimuli by both groups. When the data were analyzed by condition (Figures 7 and 8), a significant main effect of group emerged only for the familiar event sequences: $F(1, 26) = 5.83, p < .02$. Infants in the imitate group showed a larger middle-latency component to the familiar stimuli ($M = 17.00$ μV, SD = 12.29 μV) than did infants in the watch group ($M = 11.83$ μV, SD = 10.20 μV). Consequently, although the recognition effect waned over the delay, differential processing of the familiar stimuli was still apparent for infants in the imitate group when compared to those in the watch group. As observed at the immediate recognition test, effects of group were confined to the familiar stimuli: processing of the novel stimuli did not differ as a function of group (Figure 8).

Behavioral data

One-month delayed recall

Descriptive statistics for the behavioral data are found in Table 2. The data in the table are for average performance across the three sequences to which infants had been exposed (familiar sequences) and three sequences that served as novel controls (novel sequences). Inspection of the table suggests higher performance at delayed recall relative to baseline, thereby indicating memory for the sequences over the delay. However, such a comparison is not the strongest test of memory. During the

Table 2 *Mean levels of production of single actions and pairs of actions for familiar (panel A) and novel sequences (panel B)*

		Measure			
		Single actions		Pairs of actions	
Group/Condition		Mean	(SD)	Mean	(SD)
Panel A: Familiar sequences – baseline and delayed recall					
Baseline	Imitate	.68	(.45)	.05	(.14)
	Watch	.75	(.38)	.09	(.18)
Delayed recall	Imitate	1.11	(.38)	.23	(.21)
	Watch	.98	(.41)	.09	(.15)
Panel B: Novel sequences – delayed recall					
Novel	Imitate	.93	(.34)	.13	(.18)
	Watch	.99	(.39)	.12	(.16)
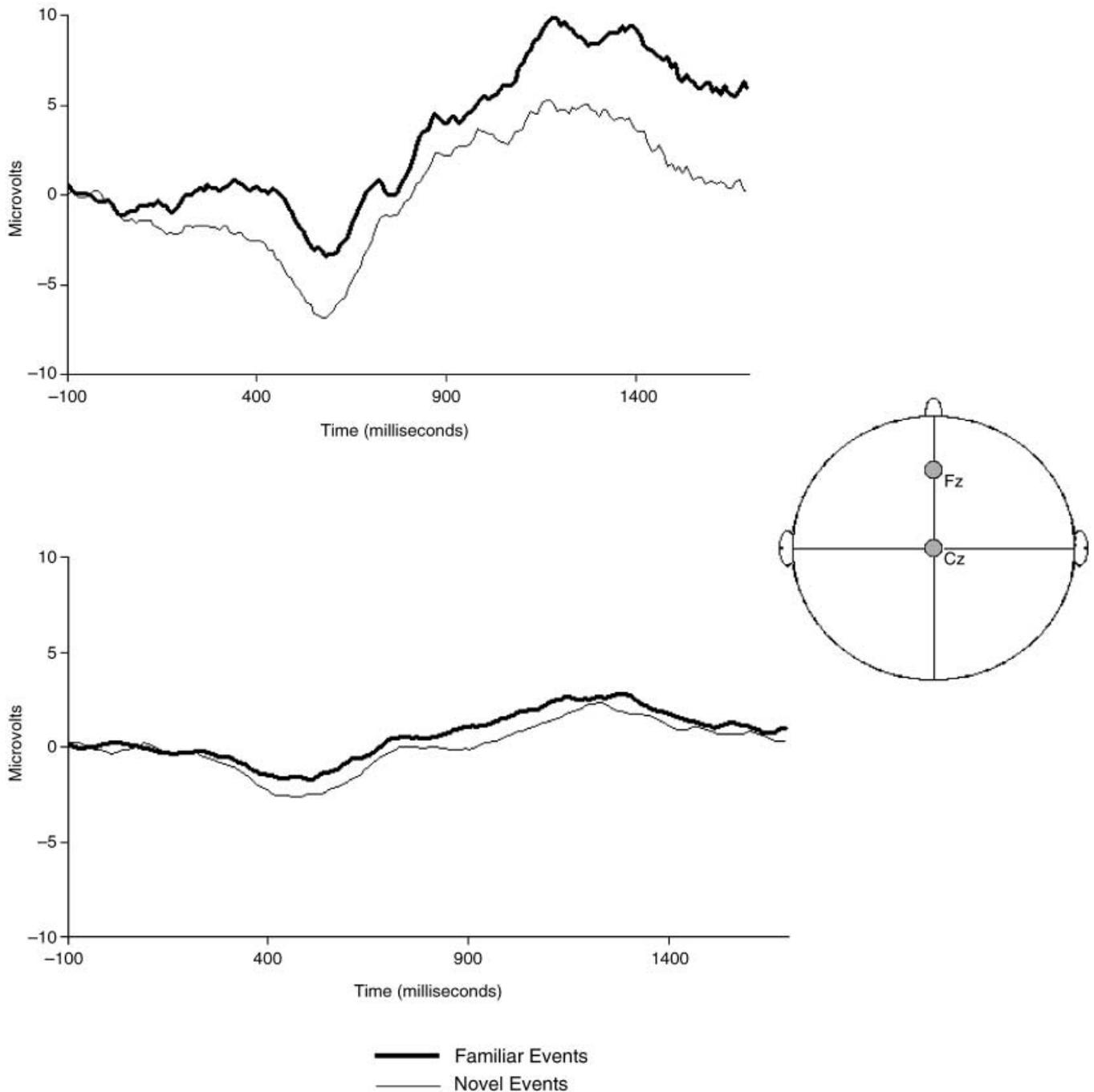
Figure 3 *Waveforms for the midline-anterior leads (Fz, top and Cz, bottom) at the immediate recognition memory test.*

delay, the infants could be expected to have undergone non-mnemonic developmental change, resulting in a larger repertoire of problem-solving skills than was available to them at baseline. This would allow them to deduce or infer a greater number of target actions and pairs of actions at delayed recall relative to baseline. As a result, significant differences between phases may be apparent not because infants remembered the familiar sequences at the delayed recall assessment but, perhaps, because they were better able to spontaneously generate actions as a result of increased problem-solving skills. The true test of recall memory, then, is found in comparing performance on familiar and novel sequences at the same point in time (e.g. Bauer *et al.*, 2000; Carver &
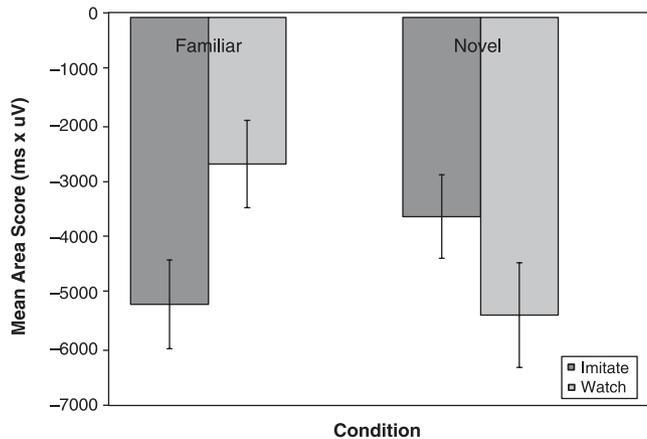
Figure 4 *Histogram indicating the Group × Condition interaction found for slow wave activity at the lateral-posterior leads (T5, T6, P3, P4, O1, and O2) at the immediate recognition test.*

Table 3 *Number of infants in the imitate group who produced target actions (panel A) and pairs of actions (panel B) at baseline (n = 45), the first, second, and third exposure sessions (n = 46)*

	Number of infants			
	Baseline	Session 1	Session 2	Session 3
Panel A: Target actions				
0	4	2	1	0
1	14	7	4	3
2	14	12	6	5
3	4	14	16	14
4	7	5	7	13
5	2	6	10	9
6	0	0	2	2
Panel B: Pairs of actions				
0	39	26	20	17
1	5	16	19	19
2	1	4	7	10
3	0	0	0	0

Bauer, 1999, 2001). To accomplish this comparison for performance at the delayed recall assessment, we conducted 2 (group: imitate, watch) × 2 (condition: familiar, novel) mixed univariate ANOVAs for each dependent variable (the number of target actions and the number of pairs of actions). For the variable of number of target actions produced, there were no significant effects. Thus, there was no evidence of differential production of single actions either as a function of group or of familiarity.

For the variable of pairs of actions produced in the target order, the Group × Condition interaction was reliable: $F(1, 74) = 3.90$, $p = .05$. For infants in the imitate group, the main effect of condition was significant: $F(1, 45) = 5.08$, $p < .03$. Infants who had been permitted to imitate prior to imposition of the delay performed more pairs of actions on familiar relative to novel sequences. No difference in performance was evident across conditions for infants in the watch group.

The beneficial effects of imitation when forming a stable memory representation also were apparent in data concerning the frequency with which infants produced ordered pairs on the familiar sequences. Specifically, 60.87% of the infants in the imitate group recalled at least one sequence in the correct temporal order at the 1-month delay, compared to only 26.67% of the infants in the watch group: $\chi^2 = 8.52$, $p < .004$.

Imitation before the delay

Attribution of the observed group effects to the opportunity to imitate requires that (a) no differences between the group existed prior to the 'opportunity to imitate' manipulation, and (b) that the infants in the imitate group in fact imitated the target actions and sequences prior to imposition of the delay. The analysis of baseline performance across the two groups revealed no significant effects for target actions ($p = .49$) or temporally ordered pairs of actions ($p = .27$; see Table 2, Panel A, for means). Consequently, variation in long-term recall between the groups cannot be attributed to differences in spontaneous problem-solving skills.

To determine whether the infants in the imitate group actually imitated any of the target actions or sequences, for the baseline period and for each exposure session, we tallied the number of infants who performed 0, 1, 2, 3, 4, 5, or 6 target actions (2 target actions possible for each of the 3 sequences) and 0, 1, or 2 ordered pairs of actions (1 pair of actions possible for each of the 3 sequences). As is apparent in Table 3, the infants interacted with the props and performed the modeled actions. At the first exposure session, only 2 infants out of 46 failed to perform any of the target actions; at the second exposure session, only one infant failed to produce any of the target actions. At the third exposure session, all infants performed at least one target action; 83% of the infants performed half or more of the possible actions. The infants also combined the actions to complete the modeled sequences. The percentages of infants who produced one or more ordered pairs of actions were 43%, 57%, and 63%, at exposure sessions 1, 2, and 3, respectively. Importantly, the infants did not generate the target behaviors on their own during the baseline period: although most infants performed at least one target action before experimenter demonstration, only 13% of the infants performed one or more pairs of actions prior to demonstration. Thus, the infants who were given the opportunity to imitate after modeling
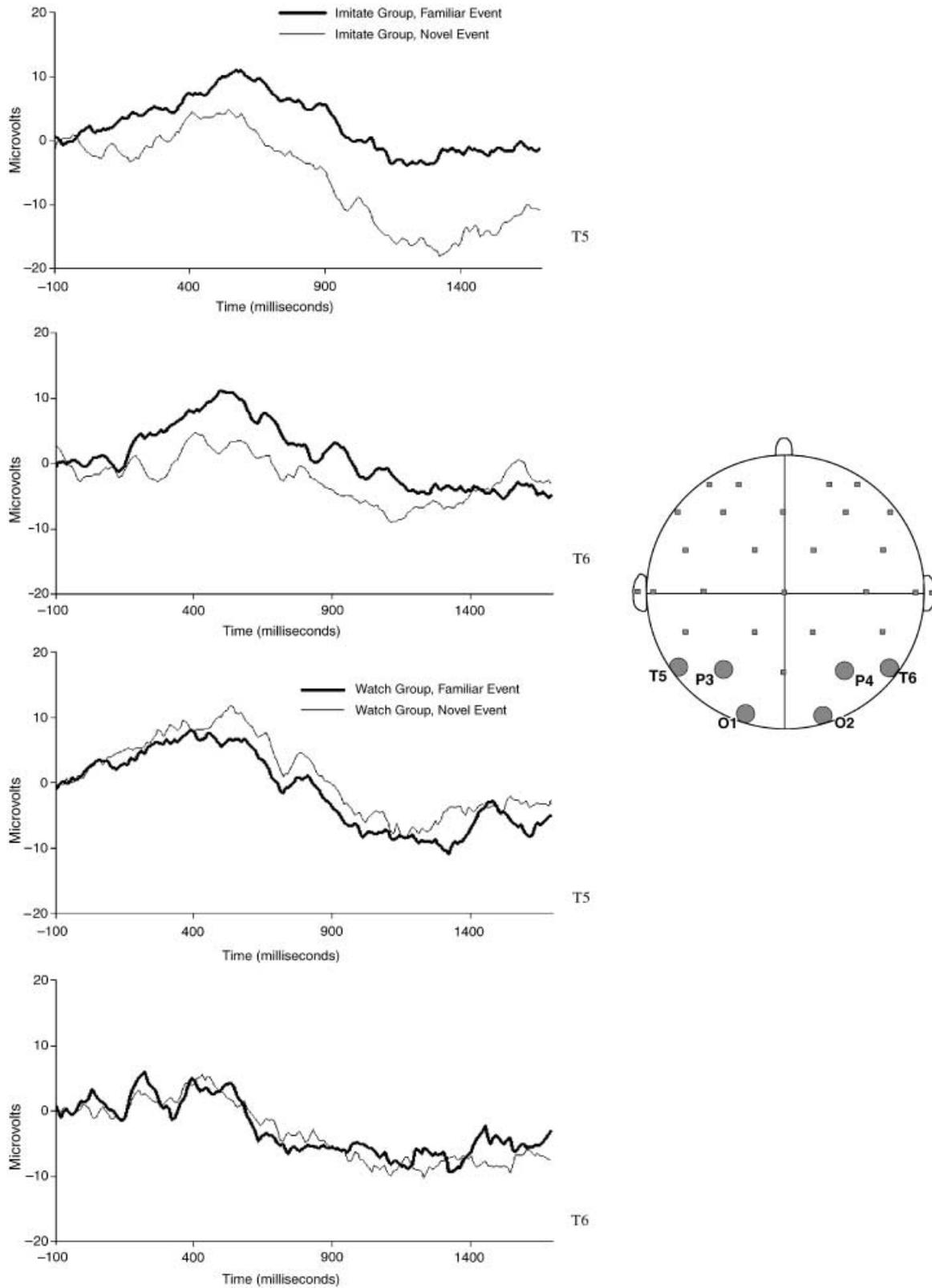
Figure 5   *Group × Condition interaction (shown by group) found for the maximum middle-latency component at the lateral-posterior leads (illustrated by the representative waveforms for T5 and T6) at the delayed recognition memory test.*
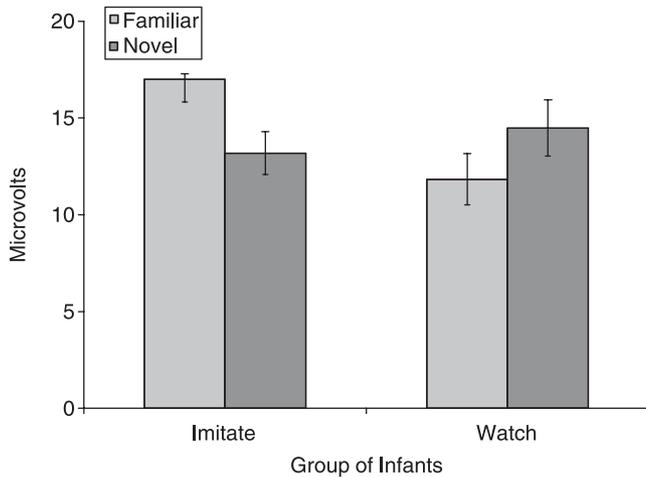
Figure 6 *Histogram indicating the Group × Condition interaction (shown by group) found for the maximum middle-latency component at lateral-posterior leads (T5, T6, P3, P4, O1, and O2) at the delayed recognition memory test.*

took it: they interacted with the props and produced the actions and sequences modeled for them, and their performance increased across sessions as a result.

Tests of uncontrolled differences between the groups

Additional analyses were conducted to determine whether the two groups of infants differed on any variables that could influence or allow for the pattern of results obtained herein.

Length of delays between sessions

We examined the possibility that the lengths of the delays between the sessions varied as a function of group by computing three independent samples *t*-tests. The lengths of the delays between Sessions 1 and 2, Sessions 2 and 3, and Sessions 3 and 4 did not vary systematically by group ($ts > .18$). Pearson product-moment correlations revealed no significant relation between the length of the delay between Sessions 3 to 4 and infants' subsequent recall performance ($rs < .25$). Therefore, although the lengths of the delays between sessions varied among infants, delay length did not differ by group, and the performance of those infants who endured longer delays between Sessions 3 and 4 did not suffer as a result.

Language

Analyses were completed to ensure that the infants who were randomly assigned to the imitate and watch groups did not differ in terms of language capabilities at recall.

Using the data from the MacArthur-Bates Communicative Development Inventories, we computed two between-groups ANOVAs to infants' (a) comprehension and (b) comprehension and production of specific words. No significant differences were apparent for the number of words comprehended and produced as a function of group. Consequently, variability in (a) language comprehension and (b) language comprehension and production between the groups cannot explain the obtained pattern of results.

Temperament

To determine whether the infants assigned to the imitate and watch groups differed in their temperaments, we conducted one-way between-subjects ANOVAs on each of the subscales of the IBQ. Infants in the imitate and watch groups were similar on all measured scales, including smiling and laughter, duration of orientation, distress to limitations, distress to novelty, soothability, and activity level. Consequently, the significant group differences at long-term delayed recall cannot be attributed to temperamental differences between the groups. We also conducted analyses to determine whether infants who provided useable ERP data were temperamentally different from infants who did not. None of the analyses yielded significant effects. Thus, as measured by the IBQ, the infants who provided ERP data were not qualitatively different from those who did not.

Selective attrition

Chi square analyses were computed to test whether attrition rates for the recognition memory assessments varied as a function of group. At the immediate ERP test, 37% of the infants in the imitate group provided useable data compared with 34% of the infants in the watch group; at the delayed ERP assessment, 39% of the infants in the imitate group provided useable data compared with 31% of the infants in the watch group. For both the immediate and delayed ERP assessments of recognition memory, the chi square analyses were non-significant, indicating that differential rates of attrition were not apparent as a function of group.

Discussion

The primary impetus for the present research was to determine whether the opportunity to imitate event sequences prior to the imposition of a delay affected (a) immediate recognition memory, delayed recognition memory, or both, as evidenced by ERPs; and (b) long-term recall, as evidenced by imitation. Analyses of the
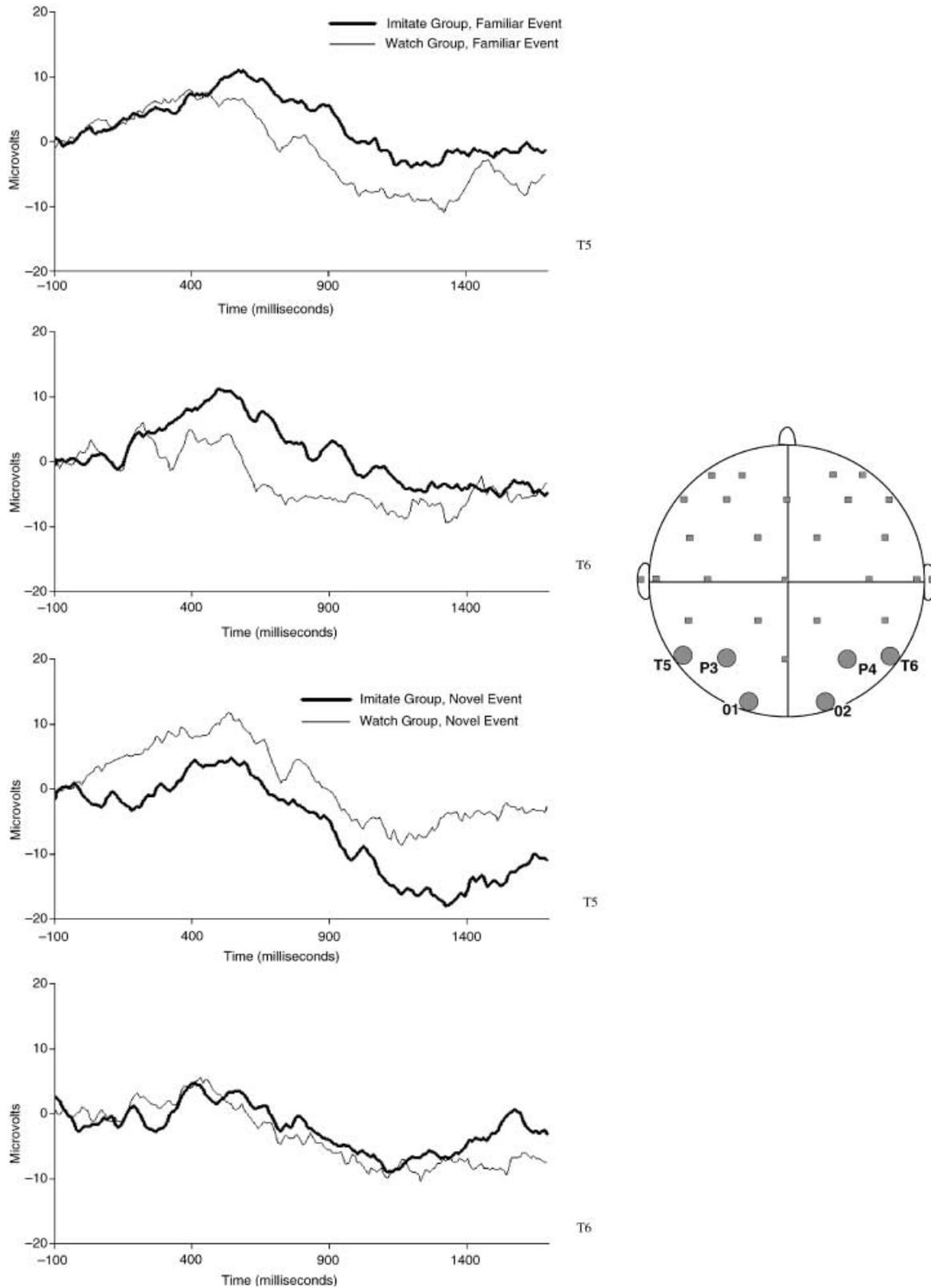
Figure 7   *Group × Condition interaction (shown by condition) found for the maximum middle-latency component at the lateral-posterior leads (illustrated by the representative waveforms for T5 and T6) at the delayed recognition memory test.*
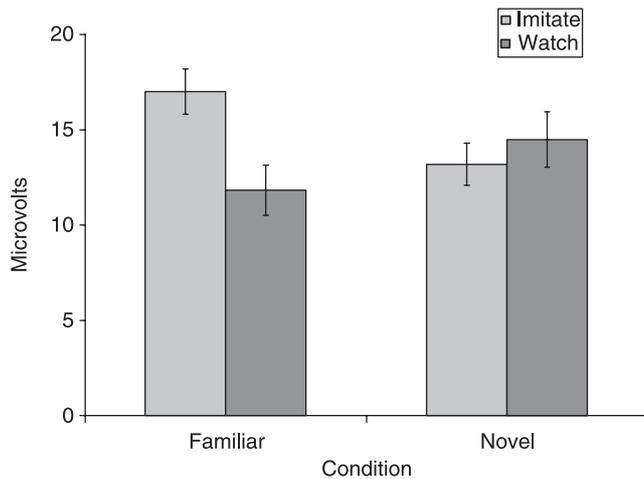
Figure 8   *Histogram indicating the Group × Condition interaction (shown by condition) found for the maximum middle-latency component at lateral-posterior leads (T5, T6, P3, P4, O1, and O2) at the delayed recognition memory test.*

electrophysiological data at the immediate recognition test indicated that infants in both the imitate and watch groups encoded the sequences. Examination of the slow wave activity at the lateral-posterior leads indicated that infants in the imitate group processed the familiar stimuli differently than infants in the watch group. Although over the 1-week delay the recognition effect waned – becoming a trend for both groups evident at the middle-latency component – differential processing of the familiar events was again apparent for those infants in the imitate group compared with those in the watch group. The data indicate that imitation has a beneficial effect on recognition memory, as infants in the imitate group evidenced larger responses to familiar events compared with infants in the watch group both immediately after event presentation and after a 1-week delay.

The ERP results converge with the behavioral data, which also indicated mnemonic facilitation related to manual interaction with the props. Specifically, infants who had been permitted the opportunity to imitate performed more temporally ordered pairs of actions for familiar events relative to novel ones at 1-month delayed recall testing. There was no suggestion of differential production of ordered pairs of actions across event types by infants in the watch group. Additionally, 60.87% of infants in the imitate group performed at least one familiar sequence in the correct order at the delayed recall test, a percentage that stands in stark contrast to the 26.67% of infants in the watch group who achieved the same level of performance.

The present study was conducted to investigate the effects of imitation on the formation of memory traces

late in the first year of life. Past research indicates that long-term recall capabilities are reliable and robust after an infant's first (e.g. Bauer & Hertsgaard, 1993; Meltzoff, 1995) birthday. After this time, the opportunity to imitate is not necessary to ensure memory for an event (Abravanel, 1991; Barr & Hayne, 1996; Bauer *et al.*, 1995, 2000; Meltzoff, 1995). However, in the present study infants in the imitate group evidenced differential processing of familiar stimuli compared with infants in the watch group and they also completed more ordered pairs of actions at the delayed recall test. As such, the opportunity to imitate the events facilitated the mnemonic performance of these infants.

There are at least two possible explanations for the beneficial effects of imitation in this sample. First, the infants may have gained practice with and formed expectations about the elicited imitation paradigm as a result of interacting with the props at each session. This practice effect may have culminated in a higher level of performance at the delayed recall test, relative to infants who only viewed the presentation of the event sequences. Although this is a reasonable interpretation, it does not account for the data. The beneficial effects of imitation were evident not only on the recall memory test, but on the recognition memory test as well, even though these assessments were quite different. There is no reason to expect that practice on a behavioral test would generalize to an ERP test. Rather than to a task-specific practice effect, we attribute the beneficial effects of imitation to formation of stronger memory representations when event sequences are imitated relative to when they are only watched. Although we have no direct evidence to support the claim, we suggest that stronger memory representations are a byproduct of the activation of multiple brain regions resulting from manual interaction with the props.

The data presented herein add to the current developmental literature by describing the effects of immediate imitation on the memory performance of infants late in the first year of life. Although the behavioral effects of immediate imitation have been investigated in samples of older children and the electrophysiological correlates have been studied in adult participants, the possible facilitating effects of this manipulation have not been examined in infancy. Our data indicate that imitation benefits recall late in the first year of life. This finding is consistent with some, but not all, previously published research. For example, Bauer *et al.* (1995) found that immediate imitation facilitated the recall capabilities of 15-month-olds who were allowed only one exposure to event sequences prior to a 1-month delay. However, interaction with the props did not benefit these children when they were given three exposures to the

to-be-remembered information before the same delay. We attribute the presence or absence of the facilitative effects of imitation to the strength of the neural circuit underlying explicit memory abilities in combination with the difficulty of the task the infants are asked to perform. That is, infants and children seemingly benefit from imitation when the task is difficult (e.g. when they must remember information over a long delay with only one exposure) in comparison to when the task is easier (e.g. when infants are allowed three exposures to an event before the delayed recall test; Bauer *et al.*, 1995).

The opportunity to imitate may allow for effects similar to those seen when infants are allowed multiple exposures to the sequences before an imposed delay. In the present study, infants were allowed two exposures to the sequences before delayed recognition. Although differential processing was apparent across both groups of infants at immediate and delayed ERP recognition memory assessments at lateral-posterior leads, significant effects were not found at midline-anterior leads. Conversely, when infants are allowed three exposures to the event sequences before delayed recognition is assessed, significant effects are found at midline-anterior electrode sites (e.g. Bauer *et al.*, 2003; Carver *et al.*, 2000). Behaviorally speaking, the beneficial effects of multiple exposures to to-be-remembered information are well documented: infants produce more ordered pairs of actions after three exposures than they do after one or two (Bauer *et al.*, 2001). Additionally, behavioral recall parallels delayed recognition memory: those infants who performed at least one pair of actions in the correct temporal order at delayed recall are those same infants who exhibited differential processing of familiar and novel sequences 1 month earlier (Bauer *et al.*, 2003; Carver *et al.*, 2000). Similarly, in the present research, infants who were allowed to imitate the event sequences immediately after their presentation evidenced higher levels of behavioral recall; 1 month earlier, these infants also had larger responses to photographs of the familiar sequence relative to infants in the watch group. Consequently, both immediate imitation and multiple exposures to the sequences facilitate long-term recognition and recall memory, although the underlying neural mechanism allowing for these effects may be different.

How might the opportunity to imitate affect recall? In accord with results from the adult literature, our findings indicate overlapping patterns of activity across imitate and watch manipulations with some specific differences. For example, amidst other similarities and differences in activity between the imitate and watch groups, variations in processing were not apparent at the midline-anterior electrodes for the slow wave at the immediate ERP but differences were found at the midline-anterior sites for the middle-latency deflection at the same assessment.

Similarly, Senkfor *et al.* (2002) reported similar processing of perform- and watch-encoded objects in an adult sample at a variety of electrode sites from 800 to 1400 ms after stimulus onset. However, the perform-encoded objects elicited differential processing from all other stimulus categories at medial and posterior electrode sites from 600 to 800 ms after stimulus presentation. These variations in processing paralleled the behavioral findings: participants reacted fastest and were most accurate in stating the encoding condition for the perform-encoded objects. Across this study and the present work, then, the processing of events that were performed overlaps with the processing of events that were only watched, but differences in activity are also apparent. Indeed, these differences may be the very ones that lead to the variation in behavioral performance evident between the two groups.

It is interesting to speculate on the nature of the differences in neural activity that might account for the observed differences between the groups. One intriguing possibility is that mirror neurons might be involved. Mirror neurons are a subset of cells found in area F5 of monkey cortex that become activated when a primate views a goal-directed action being completed and when the animal completes the same action (e.g. Gallese, Fadiga, Fogassi & Rizzolatti, 1996). These same neurons are also activated when monkeys view the end-state of a previously seen action (Umiltà, Kohler, Gallese, Fogassi, Fadiga, Keysers & Rizzolatti, 2001). Evidence for an action observation-execution matching system (Wohlschläger & Bekkering, 2002) has been identified in adult humans (e.g. Fadiga, Fogassi, Pavesi & Rizzolatti, 1995), but there is not a relevant developmental literature to date. Nonetheless, one could speculate as to how imitation might impact later recall through this system, ultimately facilitating retention over the long term. If a mirror neuron system exists in the human infant, participants in the present study may have experienced increased neuronal activation as the sequences were modeled; to the extent that viewing still photographs of the sequences cues the memory representation for the action sequence, mirror neuron activation may be apparent. Those infants who were allowed to imitate the sequences may have had an additional opportunity for increased neuronal activation as a result of engaging in manual interaction with the props. To the extent that this differential neural activation between the groups affects long-term retention, the infants in the imitate group may have been better prepared to evidence recall over the 1-month delay than were the infants in the watch group.

In conclusion, the opportunity to imitate multi-step sequences immediately after their presentation facilitates recognition and recall memory in $9^1/_2$-month-old infants.

The integration of electrophysiological and behavioral techniques in this study allowed for results that were informative as to the relation between brain activity and behavioral capabilities. More specifically, these data suggest that certain expressions of neural activity occur in tandem with measurable behavioral advances. As such, these methodologies are useful tools for investigating the development and stability of explicit memory capabilities during the first year of life and beyond.

Acknowledgements

References

Abravanel, E. (1991). Does immediate imitation influence long-term memory for observed actions? *Journal of Experimental Child Psychology*, **51** (2), 235–244.

Baker-Ward, L., Hess, T.M., & Flannagan, D.A. (1990). The effects of involvement on children's memory for events. *Cognitive Development*, **5** (1), 55–69.

Barr, R., & Hayne, H. (1996). The effect of event structure on imitation in infancy: practice makes perfect? *Infant Behavior and Development*, **19** (2), 253–257.

Bauer, P.J. (1992). Holding it all together: how enabling relations facilitate young children's event recall. *Cognitive Development*, **7** (1), 1–28.

Bauer, P.J. (1995). Recalling past events: from infancy to early childhood. *Annals of Child Development*, **11**, 25–71.

Bauer, P.J. (1996). What do infants recall of their lives? Memory for specific events by 1- to 2-year-olds. *American Psychologist*, **51**, 29–41.

Bauer, P.J. (1997). Development of memory in early childhood. In N. Cowan (Ed.), *The development of memory in childhood: Studies in developmental psychology* (pp. 83–111). Hove, Sussex: Psychology Press.

Bauer, P.J. (2002). Building toward a past: construction of a reliable long-term memory system. In N.L. Stein, P.J. Bauer & M. Rabinowitz (Eds.), *Representation, memory, and development: Essays in honor of Jean Mandler* (pp. 17–42). Mahwah, NJ: Erlbaum.

Bauer, P.J. (in press-a). Developments in declarative memory: decreasing susceptibility to storage failure over the second year of life. *Psychological Science*.

Bauer, P.J. (in press-b). New developments in the study of infant memory. In D.M. Teti (Ed.), *Handbook of research methods in developmental psychology*. Oxford: Blackwell Publishers.

Bauer, P.J. (in press-c). *Remembering the times of our lives: Memory in infancy and beyond*. Mahwah, NJ: Erlbaum.

Bauer, P.J., & Hertsgaard, L.A. (1993). Increasing steps in recall of events: factors facilitating immediate and long-term memory in 13.5- and 16.5-month-old children. *Child Development*, **64** (4), 1204–1223.

Bauer, P.J., Hertsgaard, L.A., & Wewerka, S.S. (1995). Effects of experience and reminding on long-term recall in infancy: remembering not to forget. *Journal of Experimental Child Psychology*, **59** (2), 260–298.

Bauer, P.J., & Mandler, J.M. (1989). One thing follows another: effects of temporal structure on 1- to 2-year-olds' recall of events. *Developmental Psychology*, **25** (2), 197–206.

Bauer, P.J., & Mandler, J.M. (1992). Putting the horse before the cart: the use of temporal order in recall of events by one-year-old children. *Developmental Psychology*, **28** (3), 441–452.

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

Bauer, P.J., Wenner, J.A., & Kroupina, M.G. (2002). Making the past present: later verbal accessibility of early memories. *Journal of Cognition and Development*, **3**, 21–47.

Bauer, P.J., Wiebe, S.A., Carver, L.J., Lukowski, A.F., Haight, J.C., Waters, J.M., & Nelson, C.A. (in press). Electrophysiological indices of encoding and behavioral indices of recall: examining relations and developmental change late in the first year of life. *Developmental Neuropsychology*.

Bauer, P.J., Wiebe, S.A., Carver, L.J., Waters, J.M., & Nelson, C.A. (2003). Developments in long-term explicit memory late in the first year of life: behavioral and electrophysiological indices. *Psychological Science*, **14**, 629–635.

Bauer, P.J., Wiebe, S.A., Waters, J.M., & Bangston, S.K. (2001). Reexposure breeds recall: effects of experience on 9-month-olds' ordered recall. *Journal of Experimental Child Psychology*, **80** (2), 174–200.

Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J.C., & Lenzi, G.L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences*, **100** (9), 5497–5502.

Carver, L.J., & Bauer, P.J. (1999). When the event is more than the sum of its parts: 9-month-olds' long-term ordered recall. *Memory*, **7** (2), 147–174.

Carver, L.J., & Bauer, P.J. (2001). The dawning of a past: the emergence of long-term explicit memory in infancy. *Journal of Experimental Psychology: General*, **130** (4), 726–745.

Carver, L.J., Bauer, P.J., & Nelson, C.A. (2000). Associations between infant brain activity and recall memory. *Developmental Science*, **3** (2), 234–246.

Collie, R., & Hayne, H. (1999). Deferred imitation by 6- and 9-month-old infants: more evidence for declarative memory. *Developmental Psychobiology*, **35** (2), 83–90.

Courchesne, E., Ganz, L., & Norcia, A. (1981). Event-related brain potentials to human faces in infants. *Child Development*, **52**, 804–811.

Dawson, G., Carver, L., Meltzoff, A.N., Panagiotides, H., McPartland, J., & Webb, S.J. (2002). Neural correlates of face and object recognition in young children with autism spectrum disorder, developmental delay and typical development. *Child Development*, **73** (3), 700–717.

DeBoer, T., Scott, L., & Nelson, C.A. (2004). Event-related potentials in developmental populations. In T. Handy (Ed.), *Event-related potentials: A methods handbook.* Cambridge, MA: MIT Press.

de Haan, M., & Nelson, C.A. (1997). Recognition of the mother's face by six-month-old infants: a neurobehavioral study. *Child Development*, **68** (2), 187–210.

de Haan, M., & Nelson, C.A. (1999). Brain activity differentiates face and object processing in 6-month-old infants. *Developmental Psychology*, **35** (4), 1113–1121.

Dien, J. (1998). Issues in the application of the average reference: review, critiques, and recommendations. *Behavior Research Methods, Instruments, & Computers*, **30** (1), 34–43.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation through action observation: a magnetic stimulation study. *Journal of Neurophysiology*, **73** (6), 2608–2611.

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, **59** (5), v–173.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, **119** (2), 593–609.

Geselowitz, D.B. (1998). The zero of potential. *IEEE Engineering in Medicine and Biology*, **17**, 128–132.

Jasper, H.H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, **10**, 371–375.

Junghöfer, M., Elbert, T., Tucker, D.M., & Braun, C. (1999). The polar average reference effect: a bias in estimating the head surface integral in EEG recording. *Clinical Neurophysiology*, **110**, 1149–1155.

Mandler, J.M. (1990). Recall of event sequences by preverbal children. In A. Diamond (Ed.), *The development and neural bases of higher cognitive functions* (pp. 485–516). New York: New York Academy of Science.

McDonough, L., Mandler, J.M., McKee, R.D., & Squire, L.R. (1995). The deferred imitation task as a nonverbal measure of declarative memory. *Proceedings of the National Academy of Sciences*, **92** (16), 7580–7584.

Meltzoff, A.N. (1985). Immediate and deferred imitation in fourteen- and twenty-four-month-old infants. *Child Development*, **56** (1), 62–72.

Meltzoff, A.N. (1988). Infant imitation after a 1-week delay: long-term memory for novel acts and multiple stimuli. *Developmental Psychology*, **24** (4), 470–476.

Meltzoff, A.N. (1990). The implications of cross-modal matching and imitation for the development of representation and memory in infants. In A. Diamond (Ed.), *The development and neural bases of higher cognitive functions* (pp. 1–37). New York: New York Academy of Science.

Meltzoff, A.N. (1995). What infant memory tells us about infantile amnesia: long-term recall and deferred imitation. *Journal of Experimental Child Psychology*, **59** (3), 497–515.

Nelson, C.A. (1994). Neural correlates of recognition memory in the first postnatal year. In G. Dawson & K.W. Fischer (Eds.), *Human behavior and the developing brain* (pp. 269–313). New York: Guilford Press.

Nelson, C.A., & Dukette, D. (1998). A cognitive neuroscience perspective on the relation between attention and memory development. In J.E. Richards (Ed.), *Cognitive neuroscience of attention: A developmental perspective* (pp. 327–362). Mahwah, NJ: Erlbaum.

Nelson, C.A., Henschel, M., & Collins, P.F. (1993). Neural and behavioral correlates of cross-modal recognition memory by 8-month-old infants. *Developmental Psychology*, **29**, 411–420.

Nelson, C.A., & Monk, C. (2001). The use of event-related potentials in the study of cognitive development. In C.A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 125–136). Cambridge, MA: MIT Press.

Nelson, K., & Fivush, R. (2000). Socialization of memory. In E. Tulving & F.I.M. Craik (Eds.), *The Oxford handbook of memory* (pp. 283–295). London: Oxford University Press.

Nunez, P.L. (1990). Physical principles and neurophysiological mechanisms underlying event-related potentials. In J.W. Rohrbaugh, R. Parasuraman & R. Johnson Jr. (Eds.), *Event-related brain potentials: Basic issues and applications* (pp. 19–36). New York: Oxford University Press.

Onofrj, M., Fulgente, T., Thomas, A., Locatelli, T., & Comi, G. (1995). P300 asymmetries in focal brain lesions are reference dependent. *Electroencephalography and Clinical Neurophysiology*, **94**, 432–439.

Piaget, J. (1952). *The origins of intelligence in children.* New York: International Universities Press.

Richards, J.E. (2003). Attention affects the recognition of briefly presented visual stimuli in infants: an ERP study. *Developmental Science*, **6** (3), 312–328.

Rothbart, M.K. (1981). Measurement of temperament in infancy. *Child Development*, **52**, 569–578.

Rovee-Collier, C., & Hayne, H. (2000). Memory in infancy and early childhood. In E. Tulving & F.I.M. Craik (Eds.), *The Oxford handbook of memory* (pp. 267–282). London: Oxford University Press.

Senkfor, A.J., Van Petten, C., & Kutas, M. (2002). Episodic action memory for real objects: an ERP investigation with perform, watch, and imagine action encoding tasks versus a

non-action encoding task. *Journal of Cognitive Neuroscience*, **14** (3), 402–419.

Snyder, K.A. (2002, April). Neurophysiological correlates of habituation and novelty preferences in human infants. Paper presented at the International Conference on Infant Studies, Toronto, Canada.

Snyder, K.A., & Nelson, C.A. (2001). How neural processes underlying age-related changes in face recognition can inform about cortical specialization. Paper presented at the biennial meeting of the Society for Research in Child Development, Minneapolis, Minnesota.

Spear, N.E. (1978). *The processing of memories: Forgetting and retention.* Hillsdale, NJ: Erlbaum.

Squire, L.R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology*, **44**, 453–495.

Umiltà, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: a neurophysiological study. *Neuron*, **31**, 155–165.

Wenner, J.A., & Bauer, P.J. (1999). Bringing order to the arbitrary: one- to two-year-olds' recall of event sequences. *Infant Behavior and Development*, **22** (4), 585–590.

Wheeler, M.A. (2000). Episodic memory and autonoetic awareness. In E. Tulving & F.I.M. Craik (Eds.), *The Oxford handbook of memory* (pp. 597–608). London: Oxford University Press.

Wohlschläger, A., & Bekkering, H. (2002). Is human imitation based on a mirror-neurone system? Some behavioural evidence. *Experimental Brain Research*, **143** (3), 335–341.