The Performance of Human Infants on a Measure of Frontal Cortex Function, the Delayed Response Task

ADELE DIAMOND
BERTHA DOAR
Department of Psychology
Washington University and
the McDonnell Center
for Studies of Higher Brain Function
Washington University School of Medicine

The Delayed Response task is the best-established marker of frontal lobe function in nonhuman primates. This article reports the developmental progression of human infants on that task. It is proposed that maturation of prefrontal cortex may make possible these age-related improvements in Delayed Response performance. This would suggest the importance of prefrontal cortex functioning very early in life.

Twelve infants (6 male, 6 female) were tested longitudinally every two weeks from 6–12 months of age. Another 36 infants (18 male, 18 female) were tested only once: 12 each at 8, 10, and 12 months. We predicted that infants would improve on Delayed Response over these ages because infants' performance on A$\bar{B}$ improves during this time, and Delayed Response is very similar to A$\bar{B}$. The A$\bar{B}$ task, devised by Piaget, is used to study cognitive development in infants. The ages over which A$\bar{B}$ performance improves are well established. In both A$\bar{B}$ and Delayed Response, the subject watches as the experimenter hides a desired object in one of two identical wells. After a brief delay, the subject is allowed to reach. In A$\bar{B}$, the object is hidden in the same well on subsequent trials until the subject reaches to the correct well; then side of hiding is reversed and the procedure repeated. In Delayed Response, side of hiding is varied randomly over trials. In the present study of Delayed Response each testing session consisted of 16 trials (eight to the right, eight to the left). We found: (1) the developmental progression for Delayed Response is almost identical to that for A$\bar{B}$. (2) Infants of 7½–9 months fail Delayed Response under the same conditions and in the same ways as do monkeys with lesions of dorsolateral prefrontal cortex. It is therefore suggested that A$\bar{B}$ and Delayed Response require the same cognitive abilities, and that improved performance on these tasks provides an index of maturation of frontal cortex function.

Reprint requests should be sent to Adele Diamond, University of Pennsylvania, Department of Psychology, 3815 Walnut Street, Philadelphia, PA 19104-6196, U.S.A.

Received for publication 30 March 1988
Revised for publication 30 May 1988
Accepted at Wiley 23 November 1988

Developmental Psychobiology, 22(3):271–294 (1989)

CCC 0012-1630/89/030271-24$04.00

Delayed Response is *the* classic test for frontal lobe function in nonhuman primates, specifically the dorsolateral subdivision of prefrontal cortex (for reviews see Fuster, 1980; Goldman-Rakic, 1987; Rosenkilde, 1979). This has been confirmed by virtually every anatomical, physiological, and pharmacological technique in existence (e.g., *lesions*: Jacobsen, 1936; Goldman & Rosvold, 1970; *localized electrical stimulation*: Stamm, 1969; *localized cooling*: Fuster & Alexander, 1970; *single unit recording*: Fuster, 1973; Niki, 1974; *scalp recording*: Stamm & Rosen, 1972; *localized dopamine depletion*: Brozoski, Brown, Rosvold, & Goldman, 1979; *2-deoxyglucose metabolic labelling*: Bugbee & Goldman-Rakic, 1981). Indeed, even in human adults, success on Delayed Response seems to depend upon the integrity of frontal cortex (Freedman & Oscar-Berman, 1986). Damage to frontal cortex does not impair performance on other cognitive tasks (e.g., visual discrimination: Goldman, 1971; Harlow & Dagnon, 1943; Jacobsen, 1935; 1936; Passingham, 1972) and damage to other regions of the brain does not impair performance on Delayed Response (e.g., parietal cortex: Harlow, Davis, Settlage, & Meyer, 1952; Jacobsen, 1936; Meyer, Harlow, & Settlage, 1951). Damage to areas that form part of a neural circuit with dorsolateral prefrontal cortex impairs performance on Delayed Response, albeit less severely than do lesions of dorsolateral prefrontal cortex itself (e.g., the caudate: Battig, Rosvold, & Mishkin, 1960; mediodorsal nucleus of the thalamus: Isseroff, Galkin, Rosvold, & Goldman-Rakic, 1982).

This article reports the developmental progression of human infants on the Delayed Response task. The significance of such testing is that since Delayed Response has been unequivocally linked to the frontal lobe, the ages at which performance on Delayed Response improves may provide important information relating frontal cortex maturation to human cognitive development. Frontal cortex is by far the largest and most prominent functional subdivision of human cerebral cortex. We know that it matures postnatally (Goldman, Rosvold, & Mishkin, 1970; Miller, Goldman, & Rosvold, 1973; Goldman, 1974; Brown & Goldman, 1977; Huttenlocher, 1979), but the precise timetable for this maturation is not yet known. Frontal association cortex is probably not fully mature until puberty, and until recently it was not thought to play any role at all in cognitive changes during the early years of life.

We predicted that human infants would show a developmental progression on Delayed Response between $7\frac{1}{2}$–12 months of age because they show such a progression on the $A\bar{B}$ task, and Delayed Response is very similar to $A\bar{B}$. $A\bar{B}$ (pronounced "A, not B") is a classic test originally devised by Piaget (1954 [1937]) used for studying developmental changes in cognition in human infants (for reviews see Gratch, 1975; Schuberth, 1982; Harris, 1983; Wellman, Cross, & Bartsch, 1987). In both $A\bar{B}$ and Delayed Response, the subject watches as a reward is hidden in one of two identical wells. The wells are then covered by identical covers, a brief delay is imposed, and then the subject is allowed to reach. The procedures within a trial are identical in the two tasks. The primary difference between Delayed Response and $A\bar{B}$ is that in Delayed Response the side of hiding varies randomly over trials, whereas in $A\bar{B}$ the reward is consistently hidden on a given side until the subject is correct, then side of hiding is reversed and the procedure repeated (see Table 1). Both $A\bar{B}$ and Delayed Response were used for almost 50 years before anyone realized that these tasks were virtually identical. Because of the striking similarity between the $A\bar{B}$ and Delayed Response tasks,

TABLE 1.

$A\bar{B}$			DR	
Trial No.	Side of Hiding	Response	Trial No.	Side of Hiding
1	L	√	1	L
2	L	√	2	L
3	R	X	3	R
4	R	X	4	R
5	R	√	5	L
6	R	√	6	R
7	L	X	7	L
8	L	√	8	L
9	L	√	9	R
10	R	X	10	L
11	R	X	11	L
12	R	X	12	R
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

Typical Testing Session for $A\bar{B}$ and Delayed Response. Note that side of hiding varies in a regular fashion in the $A\bar{B}$ task, whereas side of hiding varies randomly during Delayed Response. Note also that in $A\bar{B}$ side of hiding is dependent upon the subject's response (side of hiding changes when the subject is correct twice in a row), whereas in Delayed Response side of hiding is independent of subject's response.

we predicted that the performance of human infants on Delayed Response would be comparable to their performance on $A\bar{B}$. This would help create a bridge between the large research literature on human infants' performance on $A\bar{B}$ and the large research literature on the role of frontal cortex in Delayed Response performance.

This study completes a line of research begun in 1982 (see Table 2). Before 1982, Delayed Response had been used primarily to study brain function in monkeys, and $A\bar{B}$ had been used primarily to study cognitive development in human infants. It was known that monkeys with lesions of dorsolateral prefrontal cortex are impaired on Delayed Response, and that human infants show a developmental progression on $A\bar{B}$ between 7½–12 months of age. (Infants will not reach for a hidden object before 7½ months and so cannot be tested on $A\bar{B}$ at younger ages.)

The first direct link between Delayed Response and $A\bar{B}$ was made by Diamond and Goldman-Rakic (1989) who demonstrated (1) that the same lesions that disrupt performance on Delayed Response (lesions of dorsolateral prefrontal cortex) also disrupt performance on $A\bar{B}$ in adult monkeys, and (2) that the $A\bar{B}$ performance of monkeys with dorsolateral prefrontal cortex lesions is fully comparable to the $A\bar{B}$ performance of 7½–9 month old human infants. Both prefrontal monkeys and 7½–9 month old human infants fail $A\bar{B}$ with delays as brief as 2–5 sec, and both fail in the same manner, i.e., their errors are confined to the same types of trials, rather than being randomly distributed. At delays of 10 sec the performance of both prefrontal monkeys and 7½–9 month old human infants is poor on all trials and they show signs of severe distress such as long perseverative error strings, no reach to either hiding well, agitation and fussing.

TABLE 2.

	Human Infants: Developmental Progression Between 6–12 Months	Adult Rhesus Monkeys: Lesions of Prefrontal Cortex Impair Performance	Infant Rhesus Monkeys: Developmental Progression Linked to Prefrontal Maturation
A$\overline{\text{B}}$	Diamond, 1985; for reviews see: Gratch, 1975; Schuberth, 1982; Harris, 1983; Wellman, et al., 1987	Diamond & Goldman-Rakic, 1989	Diamond & Goldman-Rakic, 1986
Object Retrieval	Diamond, submitted	Diamond & Goldman-Rakic, 1985	Diamond & Goldman-Rakic, 1986
Delayed Response	this paper	Diamond & Goldman-Rakic, 1989; for reviews see: Fuster, 1980; Goldman-Rakic, 1987; Rosenkilde, 1979	Diamond & Goldman-Rakic, 1986

Diamond and Goldman-Rakic (1989) demonstrated that lesions of parietal cortex do not impair performance on A$\overline{\text{B}}$ at delays of at least 10 sec. Similarly, Diamond, Zola-Morgan, & Squire (in press) demonstrated that lesions of the hippocampus do not impair performance on A$\overline{\text{B}}$ at delays of at least 10 sec. At delays of 30 sec, the performance of hippocampal monkeys on A$\overline{\text{B}}$ was poor, but neither at that delay nor at briefer delays did hippocampal monkeys ever show the pattern of differential performance across types of trials characteristic of human infants and prefrontal monkeys. Thus, only lesions of prefrontal cortex produced a pattern of performance on A$\overline{\text{B}}$ similar to that seen in human infants.

Infant monkeys improve on A$\overline{\text{B}}$ over the same ages and at the same rate as they improve on Delayed Response (Diamond & Goldman-Rakic, 1986). By 4 months of age they succeed on both tasks with delays at least as long as 10 sec.

Lesions in the infant do not always produce the same effect as lesions in the adult. If a neural region is late maturing, lesions of that region may produce deficits in the adult, but not in the infant (e.g., Goldman, 1971; 1974). It has been suggested that lower areas of the brain might mediate infants' performance on a task, even though performance of that task by adults is mediated by a higher area of the brain. Thus, although prefrontal cortex would seem to mediate A$\overline{\text{B}}$ performance in the adult, it might not be implicated in improved A$\overline{\text{B}}$ performance in the infant. For that reason, Diamond and Goldman-Rakic (1986) went on to demonstrate that dorsolateral prefrontal cortex lesions in 4-month-old infant monkeys produce the same effect on A$\overline{\text{B}}$ performance as they do in the adult. That is, 5-month-old infant monkeys with prefrontal cortex lesions failed A$\overline{\text{B}}$ with delays as brief as 2–5 sec and showed the same pattern of performance by type of trial.

Finally, Diamond and Goldman-Rakic (1985) also demonstrated that performance on another task on which human infants show a developmental progression between 7½–12 months is also sensitive specifically to lesions of dorsolateral prefrontal cortex. The task was Object Retrieval, a detour task with transparent barrier (Diamond, submitted). Monkeys with prefrontal cortex lesions showed the same errors on Object Retrieval as do human infants of 7½–9 months of age. Monkeys with lesions of parietal cortex (Diamond & Goldman-Rakic, 1985) or the hippocampus (Diamond et al., in press) did not. Infant monkeys improve on Object Retrieval over the same age period in which they improve on A$\overline{B}$ and Delayed Response (Diamond & Goldman-Rakic, 1986).

Thus, by the time of the present study, monkeys with lesions of dorsolateral prefrontal cortex had been tested on Delayed Response, A$\overline{B}$, and Object Retrieval, and human infants had been tested on A$\overline{B}$ and Object Retrieval. The one piece of evidence needed to complete this line of inquiry was the performance of human infants on Delayed Response. That is provided here.

Methods

Subjects

Twelve infants (6 male, 6 female) were tested longitudinally every two weeks from 6 months of age (when none could uncover a hidden object) until 12 months of age. Delayed Response testing began for these infants as soon as they could uncover a hidden object. One girl dropped out of the study after the age of 10 months because her family moved away.

Another group of 36 infants (18 male, 18 female) were tested only once. Twelve of these infants (6 male, 6 female) were tested at 8 months of age (mean = 36.58 weeks, range = 34.7—37.0 weeks); 12 (6 M, 6 F) were tested at 10 months (mean = 44.93 weeks, range = 44.4–45.4 weeks) and 12 (6 M, 6 F) were tested at 12 months (mean = 53.63 weeks, range = 52.8–54.1 weeks).

All infants were healthy and full-term. Most were from middle or upper-middle class intact families. They were located through the St. Louis City and County birth records, and the parent called to request participation. All were given a present for taking part in the study.

Materials

The Delayed Response testing table stood 70 cm high, 90 cm long, and 30 cm wide. Two wells, each lined with green felt, 9.4 cm in diameter, and 5 cm deep were embedded in the table. The wells were 30 cm apart, center to center. Two light blue cotton cloths (25 × 21.3 cm) served as covers. A collection of toys, including keys, bell rattle, squeak toys, and small stuffed animals, were on hand to hide in the wells.

Procedure

All testing was done at the Infant Study Laboratory, Washington University. The infant was seated on the parent's lap facing the testing table. The experimenter sat on the other side of the table, opposite parent and child. No infant was tested if sleepy, irritable, or heavily medicated.

Pretesting

Infants were first tested for their ability to uncover a hidden object. The experimenter placed a toy in which the infant showed a clear, strong interest on the table between the two hiding wells, and partially covered it with one of the blue cloths. If the child began reaching before the cover was in place, the parent held the child back. As soon as the cover was in place, the child was allowed to reach. If the child did not try to retrieve the toy, the experimenter made noise with the toy to encourage the infant to reach. If the child showed any hesitation in reaching for the partially covered object the first time, the trial was repeated. All infants succeeded in retrieving the partially covered toy with no help from the experimenter on the first or second presentation.

The procedure was then repeated with the toy totally covered. If the infant failed to reach, the experimenter made noise with the toy under the cover, and lifted the corner of the cover to reveal the toy and then replaced the cover. The criterion for success here was retrieval of the totally covered toy on two consecutive trials without the aid of the experimenter. Infants were given up to four trials to meet this criterion. If there were any doubt about whether an infant was reaching for the toy or for the cloth cover, the trial was repeated. Only infants who met the criterion were included in the cross-sectional sample (one 8-month-old did not meet the criterion and so was excluded). Infants in the longitudinal sample were not tested on Delayed Response until the session in which they passed the criterion for retrieving a totally covered object, one hiding place.

Delayed Response Testing

For Delayed Response, the infant remained on the parent's lap, and parent and infant were positioned so they were equidistant from the two hiding wells. The experimenter held up a desired toy to attract the infant's attention and slowly hid it in one of the two wells as the infant watched. If there were any doubt about whether the infant had seen where the toy was placed, the hiding was repeated. The experimenter then covered both wells simultaneously with identical cloth covers. A delay of 0–12 sec followed. During delays greater than 0 sec the experimenter called to the infant and counted aloud to break the infant's visual fixation on the wells. If the infant failed to look up, the experimenter snapped her fingers or clapped her hands to attract the infant's attention. From the time the experimenter held up the toy until the end of the delay, the parent restrained the infant's arms, holding them firmly against the infant's body, preventing the infant from reaching, straining, or leaning toward the correct well. Following the delay, the infant was allowed to reach. Reward for correct retrieval was a play period with the toy. If an infant reached incorrectly, the experimenter showed the infant where the toy had been hidden, but did not permit the infant to play with it.

Where the toy was hidden was determined by a random sequence of 16 trials (L, L, R, R, L, R, L, L, R, L, L, R, R, R, L, R; or the mirror-image beginning with hiding on the right). Thus, all infants received eight trials with hiding at the left, and eight trials with hiding at the right. The toy was not hidden in the same well on more than three consecutive trials.[1] For infants tested longitudinally, initial side of hiding was counterbalanced across children for the first testing session and counterbalanced across testing sessions for each child. If testing began with the toy hidden on the right in one session, the toy was hidden in the well on the left for the first trial of the next session. Hiding proceeded according to this predetermined schedule regardless of the infant's performance on any trial.

All trials within a single testing session were at a constant delay. During the first Delayed Response testing session, no delay and no distraction were used. The procedure for incrementing delay over sessions was as follows: Delay was incremented if (1) the infant reached correctly on at least 14 of the 16 trials during the preceding session (88% correct), or if (2) the same delay had been used on at least the last two preceding visits and performance during the last session had been at the 70% level or better (at least 11 trials correct). Otherwise, the same delay as on the preceding visit was used. Delay was incremented to the following levels: 3, 5, 8, 10, and 12 sec.

We tried to test all infants every 2 weeks ± 3 days. Except for the girl who dropped out of the study after 10 months of age, only one subject missed any of our semimonthly testing intervals (one boy was not tested at 10½ months).

For infants in the cross-sectional sample: At 8 months, half of the infants were tested at 0 sec (no delay and no distraction), and half were tested with a delay of 3 sec. At 10 months, half were tested with a delay of 3 sec and half with a delay of 8 sec. At 12 months, half were tested with a delay of 8 sec, and half with a delay of 12 sec. Equal numbers of males and females were assigned to each condition, and initial side of hiding was counterbalanced within each age *X* delay condition.

Infants tested at the 3 sec delay received two practice trials at 0 sec, one with the hiding at the left and one with hiding at the right. Infants tested at delays of 8 and 12 sec received these practice trials and two more with a delay of 3 sec, again with the hiding first at the left and then at the right.

All sessions were recorded on videotape for detailed analysis.

Predictions

Because of the marked similarity between the Delayed Response and $A\overline{B}$ tasks, we predicted that infants tested longitudinally on Delayed Response would show the same developmental progression as Diamond (1985) had previously found with longitudinal testing on $A\overline{B}$. That is, we did not simply predict that children would perform better on Delayed Response with age. Rather we predicted that infants would improve on Delayed Response over the very same ages they improve on $A\overline{B}$ and at the same rate.

In addition, we predicted that infants would show the same *pattern* of performance on Delayed Response as infants have previously shown on $A\overline{B}$ (Diamond, 1985). That is, we predicted they would reach correctly on trials where the toy was hidden in the same well as on the preceding trial and the infant had reached correctly on the preceding trial (REPEAT FOLLOWING CORRECT trials), but that infants would tend to err if either of those variables were changed. Thus, infants would make significantly more errors when side of hiding was reversed and the infant had reached correctly on the preceding trial (REVERSAL FOLLOWING CORRECT trials) or when the toy was hidden in the same well as on the preceding trial and the infant had reached *incorrectly* on the preceding trial (REPEAT FOLLOWING ERROR trials).

A$\overline{\text{B}}$ derives its name from this pattern, first observed by Piaget (1954 [1937]). Infants correctly find a toy at the first place it is hidden (A), and can repeat that correct response on subsequent trials (REPEAT FOLLOWING CORRECT trials), but when side of hiding is reversed to B (REVERSAL FOLLOWING CORRECT trials) infants err, reaching back to A, even though they have observed the hiding at B. Thus, infants can find the toy at A, but not at B. Infants typically repeat this error if the toy continues to be hidden at B (REPEAT FOLLOWING ERROR trials).

Our predictions for the cross-sectional sample were as follows: Infants of 8 months would succeed (i.e., reach correctly on at least 14 out of 16 trials) with a 0 sec delay (i.e., no delay). However, they would fail at the 3 sec delay. Infants of 10 months, on the other hand, would succeed with a 3 sec delay, but fail with an 8 sec delay. Infants of 12 months would succeed at the 8 sec delay, but fail at the 12 sec delay. These predictions, too, are based on the performance of infants on A$\overline{\text{B}}$ (e.g., Diamond, 1985; Fox, Kagan, & Weiskopf, 1979; Gratch, Appel, Evans,

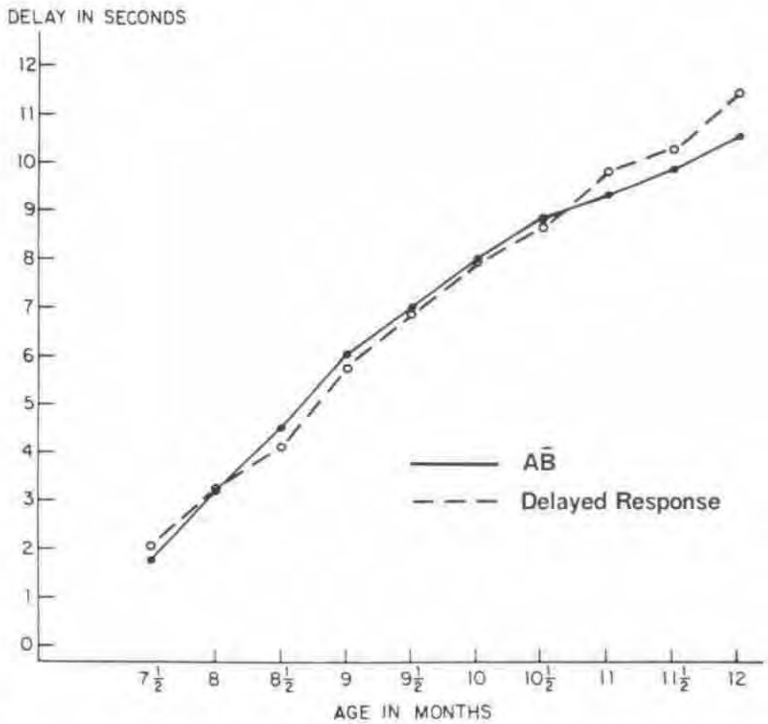

Fig. 1. Developmental progression in the delay infants could tolerate on Delayed Response and A$\overline{\text{B}}$. A$\overline{\text{B}}$ results are usually reported in terms of the age at which the A$\overline{\text{B}}$ *error* occurs. In an attempt to use a comparable measure for Delayed Response, results are reported here in terms of the delay at which errors occurred (i.e., the delay at which performance was below the criterion of 88% correct). Delayed Response results for the infants studied longitudinally (shown by the dashed line) are presented superimposed over the A$\overline{\text{B}}$ results for infants studied longitudinally by Diamond (1985) (shown by the solid line).

LeCompte, & Wright, 1974; Gratch & Landers, 1971). The cross-sectional sample was included for two reasons: (1) to provide evidence that improvement on Delayed Response with increasing age is not simply a result of experience or practice, and (2) to yield results that could be compared with the majority of A$\bar{B}$ studies, which are cross-sectional.

Results

Longitudinal Sample

Prediction: Infants tested longitudinally on Delayed Response will display the same developmental progression as infants display during longitudinal testing on A$\bar{B}$. This prediction was strongly confirmed, as can be seen by Figure 1, where the developmental progression for Delayed Response performance is superimposed over the developmental progression for A$\bar{B}$ performance. The curves are virtually identical (slope for Delayed Response = 2.10; slope for A$\bar{B}$ = 1.91), even though (1) A$\bar{B}$ testing was conducted at Harvard University with babies from the Boston area and Delayed Response testing was conducted at Washington University with babies from the St. Louis area, (2) different testers administered A$\bar{B}$ and Delayed Response, and (3) the Delayed Response testers were blind to the earlier A$\bar{B}$ results.

In all respects, the results for Delayed Response are comparable to those for A$\bar{B}$: As with A$\bar{B}$, infants could tolerate increasingly long delays on Delayed Response as they got older ($t[11] = 10.64$, $p = .0001$, linear regression coefficient of age). There was no evidence of a sudden discontinuity in the delays infants could tolerate. The delay infants could tolerate increased at a constant rate of approximately 2 sec per month (same rate as for A$\bar{B}$). Age accounted for roughly half of the variance in the delay infants could tolerate ($R^2 = .49$ for Delayed Response, similar to $R^2 = .46$ for A$\bar{B}$ [Diamond, 1985]). Individual differences between children of the same age were very large, as they had been for A$\bar{B}$. The standard deviation in the delay infants could tolerate ranged from 1.98 at $7\frac{1}{2}$ months to 4.65 at 12 months. At each age, the difference in the delay for the A$\bar{B}$ error and the delay at which performance on Delayed Response was below criterion was never greater than 1 sec (mean = .34 sec, see Table 3).

As had been true for A$\bar{B}$, girls could consistently tolerate longer delays than boys (sex difference significant at .01, $t[10] = 2.5$). Knowing both age and sex significantly improved the ability to predict the delay an infant could tolerate on Delayed Response: Regression models that included both sex and age as independent variables accounted for the data significantly better than did the regression model that included age as the only independent variable ($p = .01$; just as had been true for A$\bar{B}$).

Girls did not appear to be improving at a faster rate than boys, rather they got off to an earlier start. The slope for delay × age in Delayed Response performance for girls was not significantly different from the slope of the comparable function for boys ($t[10] = 0.56$, NS). On the other hand, girls could uncover a hidden object at a younger age than could boys ($t[10] = 2.5$, $p = .01$). The mean age at which girls first reached for a hidden object was $7\frac{1}{2}$ months (range = $6\frac{1}{2}$–$7\frac{1}{2}$ months), while the mean age for boys was 8 months (range = $6\frac{1}{2}$–9 months). While

TABLE 3. Mean Delay at which Performance was Less than 88% Correct on Delayed Response Compared with Mean Delay for the $A\bar{B}$ Error

	All Subjects		Difference	Boys		Girls		Male-Female Difference	
Age in Months	Delayed Response (N = 12)	$A\bar{B}$ (N = 25)	between DR & $A\bar{B}$ Means	Delayed Response (N = 6)	$A\bar{B}$ (N = 11)	Delayed Response (N = 6)	$A\bar{B}$ (N = 14)	Delayed Response	$A\bar{B}$
7½	2.1	1.7	.4	1.2 (n = 3)	1.5 (n = 6)	3.0	1.9 (n = 12)	−1.8	−.4
8	3.2	3.2	0	1.8 (n = 4)	2.0 (n = 7)	4.6	3.8	−2.8	−1.8
8½	4.0	4.5	−.5	2.9 (n = 5)	2.6	5.1	6.0	−2.2	−3.4
9	5.9	6.1	−.2	5.6	4.5	6.2	7.4	−.6	−2.9
9½	6.9	7.0	−.1	6.7	5.8	7.2	8.0	−.5	−2.2
10	8.0	8.0	0	7.7	6.6	8.4	9.2	−.7	−1.2
10½	8.8	8.9	−.1	8.7 (n = 5)	7.9	8.9 (n = 5)	9.8 (n = 13)	−.2	−1.9
11	9.9	9.3	−.6	9.5	8.0	10.2 (n = 5)	10.2	−.7	−1.2
11½	10.3	9.8	.5	9.5	8.6	11.0 (n = 5)	11.0 (n = 12)	−1.5	−2.4
12	11.6	10.6	1.0	10.1	9.8 (n = 10)	13.1 (n = 5)	11.4 (n = 10)	−3.0	−1.6

Results for $A\bar{B}$ are from Diamond, 1985.

five out of the six girls tested longitudinally could uncover a hidden object by 7 months of age, only two of the six boys could do so. Thus, girls began testing on Delayed Response (which requires the ability to uncover a hidden object) at a younger age than boys and so started off at an advantage which they maintained throughout the months of testing. The same had been found for $A\bar{B}$ (Diamond, 1985).

Prediction: Infants will show the same pattern of performance on Delayed Response as infants have previously shown on $A\bar{B}$. This prediction, too, was confirmed. Infants performed very well on REPEAT FOLLOWING CORRECT trials (92% correct), but poorly on REVERSAL FOLLOWING CORRECT trials (54% correct) and poorly on REPEAT FOLLOWING ERROR trials (39% correct). That is, infants were generally correct if the toy was hidden in the same well where they had just reached correctly, but they tended to err if side of hiding was reversed after they had just reached correctly and they tended to repeat that error if the toy continued to be hidden in the new place. This is the same pattern as found for $A\bar{B}$ (Fig. 2).

An analysis of variance comparing percent correct by whether side of hiding remained at the same well as on the previous trial or was reversed and by whether the infant had been correct or wrong on the previous trial revealed a significant difference by type of trial ($F [3,44] = 50.05$, $p = .0001$). When the variance was partitioned into orthogonal linear contrasts it was revealed that the difference in performance between REPEAT FOLLOWING CORRECT trials and REVERSAL FOLLOWING CORRECT trials was significant ($F [1,44] = 57.51$, $p = .0001$) and the difference between REPEAT FOLLOWING CORRECT trials and REPEAT FOLLOWING ERROR trials was significant ($F = [1,44] = 115.13$, $p = .0001$) (Table 4). As expected, however, the difference between REPEAT trials (whether following correct reaches or errors) and REVERSAL trials (whether following correct reaches or errors) was not significant ($F = 0.26$, NS), nor was the difference between trials FOLLOWING CORRECT REACHES (whether hiding remained in the same place or was reversed) and trials FOLLOWING ERRORS (whether hiding remained in the same place or was reversed) ($F = 0.12$, NS). That is, as expected there were no significant main effects for hiding location (same as previous trial or reversed) or response on previous trial (correct or wrong), but there was a significant interaction.

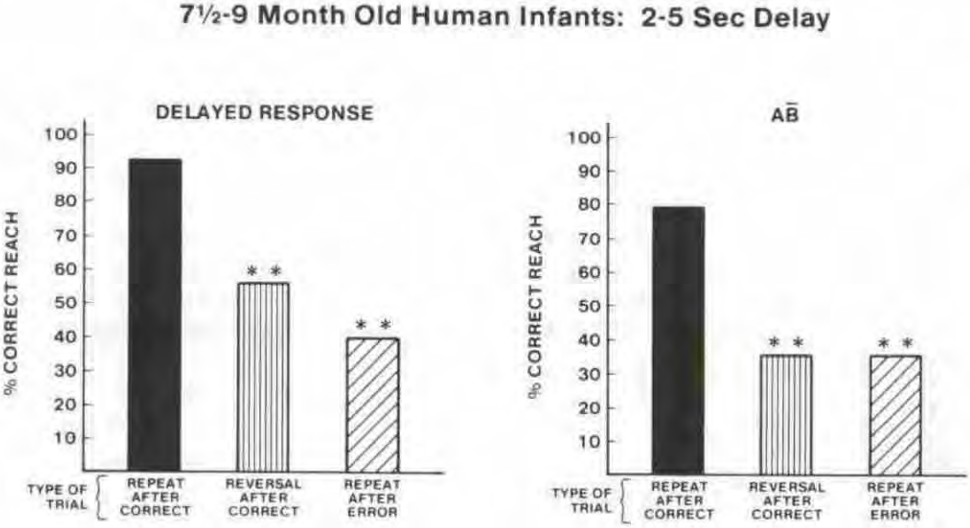

Fig. 2. Performance by Type of Trial on Delayed Response and AB̄. ** = performance on that type of trial differs from performance on REPEAT FOLLOWING CORRECT trials at $p < .0001$. Note that the same pattern of differential performance by type of trial was found on both tasks even though the hiding procedures and length of delay were identical on all trials. The AB̄ results are from Diamond (1985). Results for only three types of trials are pictured here because in the AB̄ task no reversal is administered unless the child has reached correctly on the preceding trial. Thus, reversal trials in this figure refer only to REVERSAL FOLLOWING CORRECT trials. Results for the fourth type of trial on Delayed Response, REVERSAL FOLLOWING ERROR trials, are provided in Table 4.

Both boys and girls showed this same pattern of results (see Table 4 again). There were no age differences in performance on any type of trial, nor in the difference in performance between any two types of trials. Although Janowsky, Schwartz, & Stiles-Davis (1987) found less perseveration in older infants, we found that older infants made as many errors on REPEAT FOLLOWING ERROR trials as younger infants, which would seem to indicate no change in

TABLE 4. Percent Correct on Delayed Response by Type of Trial.

	Repeat Following Correct Trials	Reversal Following Correct Trials	Repeat Following Error Trials	Reversal Following Error Trials	Repeat vs. Reversal Following Correct	Repeat Following Correct vs. Error
Longitudinal Sample						
All	92	54	39	83	$t = 9.95, p = .0001$	$t = 6.47, p = .0001$
Boys	93	53	44	85	$t = 8.09, p = .005$	$t = 6.66, p = .001$
Girls	90	55	33	81	$t = 5.83, p = .002$	$t = 9.19, p = .0001$
Cross-Sectional Sample						
All	92	52	37	82	$t = 7.21, p = .0001$	$t = 7.65, p = .0001$
Boys	89	52	43	89	$t = 5.06, p = .0001$	$t = 3.88, p = .001$
Girls	95	51	31	74	$t = 5.06, p = .0001$	$t = 6.39, p = .0001$
8 Months	89	54	33	81	$t = 3.12, p = .01$	$t = 3.60, p = .004$
10 Months	92	54	25	71	$t = 4.14, p = .002$	$t = 5.88, p = .0001$
12 Months	95	47	53	93	$t = 5.30, p = .0003$	$t = 3.12, p = .01$
Shorter Delays	98	72	57	70	$t = 4.91, p = .0001$	$t = 3.42, p = .003$
Longer Delays	86	32	19	94	$t = 6.11, p = .0001$	$t = 7.56, p = .0001$

Shorter delays = 0 sec for 8-month olds, 3 sec for 10-month olds, and 8 sec for 12-month olds.
Longer delays = 3 sec for 8-month olds, 8 sec for 10-month olds, and 12 sec for 12-month olds.

perseverative tendencies over age (F [1,11] = 1.10, NS, for linear regression of percent correct on REPEAT FOLLOWING ERROR trials over age).

Another way to compare the pattern of performance on Delayed Response with that on A$\overline{B}$ is to look at performance on the first reversal trial. Most studies of A$\overline{B}$ have used brief testing sessions and only one reversal trial. The studies have consistently found that about half the infants are correct on this trial and about half are not (see discussion of this in Butterworth [1975]). Similarly, in the study of A$\overline{B}$ by Diamond (1985), which included multiple reversals, percent correct on the first reversal trial was 42%, which did not differ significantly from chance. Reversal trials are administered in A$\overline{B}$ only after the subject has reached correctly on the preceding trial. To find a comparable measure in Delayed Response, we looked at percent correct on the first reversal trial for those sessions where the infant had reached correctly on the preceding trial. Again the results for Delayed Response are in agreement with those for A$\overline{B}$. For infants followed longitudinally, 83% of the sessions met the criterion for inclusion (101 sessions), and percent correct on the first reversal trial in those sessions was 52% (versus chance: t = 0.50, NS). Results were the same whether the first reversal was to the right (51% correct) or to the left (54% correct).

Infants tested only once also displayed the pattern of differential performance by type of trial (Table 4). This pattern was found for both girls and boys and at all three ages (Table 4). There was no significant age difference in performance on any type of trial nor in the difference between performance on any two types of trials.

Looking only at the first reversal trial, infants tested cross-sectionally performed comparably to the infants tested longitudinally on Delayed Response and to the infants tested cross-sectionally on A$\overline{B}$. As we did for the longitudinal subjects, only sessions where the child was correct on the trial preceding the reversal are considered here. That left 30 sessions (83%). Percent correct on the first reversal trial in these sessions was 48% (versus chance: $t = 0.18$, NS). Results were the same whether the first reversal was to the right (44% correct, $t = 0.49$, NS) or to the left (54% correct, $t = 0.27$, NS).

One can compare this to performance on the first trial of the session, where percent correct for infants tested longitudinally was 77% (versus chance: $z = 6.99$, $p < .0001$; versus performance on first reversal [59%]: $x^2 = 5.97$, $p = .02$, McNemar) and percent correct for infants in the cross-sectional sample was 89% (versus chance: $z = 4.83$, $p < .0001$; versus performance on the first reversal [53%]: $x^2 = 4.35$, $p = .05$, McNemar). (For these comparisons, *all* reversals, not simply those following a correct reach, are included.)

Cross-Sectional Sample

Prediction: 8-month-olds will fail Delayed Response at 3 sec, but succeed at 0 sec; 10-month-olds will fail Delayed Response at 8 sec, but succeed at 3 sec; 12-month-olds will fail Delayed Response at 12 sec, but succeed at 8 sec. All predictions specifying that infants would fail to pass criterion were confirmed. No 8-month-old infant succeeded on more than 69% of the trials at the 3 sec delay. Mean percent correct here was significantly lower than the 88% criterion ($t[5] = 7.04$, $p = .001$). Only one 10-month-old infant performed at criterion with a delay of 8 sec. Again, mean percent correct was significantly below criterion ($t[5]$ =

TABLE 5. Performance on Delayed Response of Infants Tested Only Once as a Function of Age and Length of Delay.

	0-Sec Delay			3-Sec Delay		
8-Month-Olds	All ($N = 6$)	Boys ($N = 3$)	Girls ($N = 3$)	All ($N = 6$)	Boys ($N = 3$)	Girls ($N = 3$)
Percent passing criterion*	67	33	100	0	0	0
Percent Correct on First Reversal Trial**	67 (6)	67 (3)	67 (3)	33 (3)	50 (2)	0 (1)
Mean percent correct	85	79	90	61	65	56
Difference between mean percent correct at the two delays: t (10) = 3.25, $p < .01$						
	3-Sec Delay			8-Sec Delay		
10-Month-Olds	All ($N = 6$)	Boys ($N = 3$)	Girls ($N = 3$)	All ($N = 6$)	Boys ($N = 3$)	Girls ($N = 3$)
Percent passing criterion	83	67	100	17	0	33
Percent Correct on First Reversal Trial	83 (6)	67 (3)	100 (3)	25 (4)	33 (3)	0 (1)
Mean percent correct	86	80	92	63	60	65
Difference between mean percent correct at the two delays: t (10) = 3.09, $p = .01$						
	8-Sec Delay			12-Sec Delay		
12-Month-Olds	All ($N = 6$)	Boys ($N = 3$)	Girls ($N = 3$)	All ($N = 6$)	Boys ($N = 3$)	Girls ($N = 3$)
Percent passing criterion	50	33	67	0	0	0
Percent Correct on First Reversal Trial	40 (5)	0 (3)	100 (2)	20 (5)	33 (3)	0 (2)
Mean percent correct	79	77	81	60	66	54
Difference between mean percent correct at the two delays: t (10) = 3.33, $p < .01$						

* Criterion = 88% correct (correct on 14 out of 16 trials)

** This measure is included to compare performance here with the standard dependent measure reported in studies of AB. Because reversals are only administered in AB after a subject has reached correctly, only infants who reached correctly on the trial preceding the first reversal are included here. The number in parenthesis gives the number of infants on which percentage is based.

4.47, $p = .01$). No 12-month-old infant succeeded on more than 75% of the trials at the 12 sec delay. Mean percent correct was significantly below criterion ($t[5] = 7.47$, $p = .001$). (Table 5.)

However, not all infants passed criterion in the conditions where we had predicted success (8-months-old, 0 sec delay: 67% passed; 10-months-old, 3 sec delay: 83% passed; 12-months-old, 8 sec delay: 50% passed). At each age, however, performance was significantly better at these briefer delays than at the longer delay (statistics provided in Table 5). In addition, older infants performed better at the same delay than did younger infants. Thus, 10-month-old infants performed significantly better at the 3 sec delay than did 8-month-olds ($t[10] = 4.01$, $p = .003$), and 12-month-old infants performed significantly better at the 8 sec delay than did 10-month-olds ($t[10] = 2.33$, $p = .04$).

Infants failed at the 0 sec delay when they interrupted their visual fixation of the correct well. There was no distraction at the 0 sec delay and no attempt to try to make the infant look up. However, most infants became distracted, and looked away, on their own on one or more trials at this delay. For example, the six children in the cross-sectional sample tested with a 0 sec delay were *never* wrong when they maintained their visual fixation on the correct well unbroken (15 incorrect reaches total, all when they broke visual fixation), and on 92% of the trials where they reached correctly they looked at the correct well throughout the trial (81 correct reaches in all). The correlation between whether visual fixation of the correct well was maintained and whether the infant reached correctly was $\phi = .79$, $p = .05$ (phi coefficient [Lahey, Downey, & Saal, 1983]; for infants in the cross-sectional sample tested at 0 sec delay.) At delays over 0 sec this relation no longer held because visual fixation of the correct well was broken on almost all trials, as distraction was intentionally introduced at these delays.

Our predictions were actually more precise than simply that infants would fail at the longer delays. We predicted they would show the pattern of differential performance by type of trial at the longer delays, while performing well on all types of trials at the shorter delays.

The pattern of differential performance by trial type was found at both short and long delays (Table 4). However, this pattern was significantly more pronounced at longer than at shorter delays. The difference between performance on REPEAT FOLLOWING CORRECT trials and on REVERSAL FOLLOWING CORRECT trials was significantly greater at the longer delays than at the shorter delays (54% vs. 26%, $t[34] = 2.64$, $p = .01$). The difference between REPEAT FOLLOWING CORRECT trials and REPEAT FOLLOWING ERROR trials was also greater at the longer delays than at the shorter delays, although it did not quite reach statistical significance (68% vs. 41%, $t[34] = 1.79$, $p = .08$).

Thus, our predicitons here were confirmed in general form, but not in as precise detail as we had hoped. Performance was better at each age at the briefer delay and better in older children at a given delay. However, performance was not as good at the briefer delays as we had predicted. Although significantly more children passed criterion at the shorter delays than at the longer delays (67% vs. 5%), not all children passed criterion at the shorter delays as we had predicted. This may reflect wide individual differences in the delay infants could tolerate at a given age, which had been observed in longitudinal testing of both Delayed Response and A$\overline{\text{B}}$. Although differential performance by type of trial was found at the longer delays as we had predicted, it was still found in a more attenuated form at the shorter delays, where we had expected excellent performance on all trials.

There is also some evidence that infants may have found the longer delays more difficult than we had predicted. Percent correct on REPEAT FOLLOWING ERROR trials was extremely low at the longer delay at 8 and 10 months (only 8% and 6% correct respectively). Poor performance on this type of trial may be indicative of perseveration. Here the child is given feedback that the last response was wrong, and yet the child repeats that response anyway. Perseveration would indicate that the delay was too difficult for the infant (e.g., see Diamond, 1985).

Another indication that the long delays may have been too taxing is the marked tendency of infants in this condition to reach predominantly to the right or left side. Only two infants tested at the shorter delays (11%) showed a side preference (defined as reaching $\geq$ 69% of the time to one side, $p \leq .10$, binomial), but 16 of the infants tested at longer delays (89%) did so ($p = .0007$, binomial). Percent of reaches to the preferred side was significantly higher at longer delays than at shorter delays (84% vs. 61%, $t[34] = 3.70$, $p = .002$). A position habit, or a tendency to reach to the same side, is a common response to a difficult or unsolvable problem.

In sum, while our predictions were generally supported, there is evidence that perhaps both the shorter and longer delays were more difficult than we had predicted.

As noted above, the prediction concerning differential performance across types of trials was confirmed as robustly for the cross-sectional sample as it was for the longitudinal sample.

The results from the longitudinal study suggested that girls could tolerate longer delays than boys of their age. Girls did not differ from boys, however, in overall percent correct or in percent correct on any type of trial. Similarly, girls in the cross-sectional study did not differ from boys in overall percent correct or in percent correct by trial type. Could girls in the cross-sectional sample tolerate longer delays than their male age-mates? The longer delays were difficult for everyone; only one child passed criterion. There was no sex difference in performance at the longer delays. At the shorter delays there was a tendency for the girls to perform better than the boys. At each of three ages, three infants of each sex were tested at the shorter delay. Eight of these nine girls (89%, $p = .02$, binomial) passed criterion, while only four of the nine boys passed (44%, NS). At each age, girls performed better than boys at the shorter delays (Table 5), but this difference was only significant at the .13 level.

Did infants show a tendency to reach significantly more often to the right or left? Can any of the results reported above be accounted for by such a tendency? Since infants err on about half the reversal trials, it would be possible for infants to achieve such a score by reaching to the same well on most trials. In that case we would simply be observing a side preference on this task.

An infant followed longitudinally might show a preference for one side in one testing session and the other side in another testing session. Therefore, calculating percent of reaches to the right or left over sessions would not have been particularly meaningful. Instead we calculated side preference individually for each testing session. The criterion we used for determining that an infant had shown a tendency to reach significantly more often to the right or to the left in a given testing session was that at least 69% of the child's reaches were to that side ($p = .10$, binomial distribution). Using that criterion, infants studied longitudinally showed a side preference on 62 of their 122 sessions (51%; 21% to the right; 30% to the left). Only two of the twelve children followed longitudinally showed a side preference on more sessions than one might expect by chance. One girl (F3) showed a preference to reach to the left on 9 of her 12 sessions ($p = .07$, binomial) and one boy (M3) showed a preference on 9 of his 11 sessions ($p = .03$) (three sessions with right preference; six sessions with left preference) (Table 6).

There is no evidence that side preference can account for the pattern of differential performance across trials. If F3 and M3 are omitted from the analysis,

TABLE 6. Analysis of Infants' Tendency to Reach Preferentially to the Right or Left.

Subjects	Number of DR Sessions*	Number of Sessions where at least One Error Occurred	Number of Sessions where more Reaches Were to *Right* than to Left**	Number of Sessions where more Reaches Were to *Left* than to Right**	Percent of Sessions where a Preference was Shown to Either Side†	Percent of Sessions where Error(s) occurred on . . .‡		
						Reversal(s) to Left, but Not to Right	Reversal(s) to Right, but Not to Left	Reversal(s) to Both Sides
F1	12	10	0	2	17	20	20	60
F2	7	6	2	0	17	50	17	33
F3	12	11	0	9	75	0	64	37
F4	12	12	0	8	67	0	67	33
F5	12	11	4	2	50	36	9	55
F6	11	11	3	3	55	36	9	55
M1	9	7	5	0	56	72	14	14
M2	10	9	4	3	70	11	44	44
M3	11	11	3	6	82	18	36	46
M4	10	10	1	3	40	10	60	30
M5	8	8	3	0	38	25	0	75
M6	8	6	1	0	13	50	0	50

* Infants were tested on Delayed Response from the earliest age at which they could find a hidden object until the age of 12 months. Since the ability to find a hidden object comes in at different times in different infants, some children received more Delayed Response testing sessions than others.

** The criterion for determining that an infant had shown a tendency to reach to the right or left in a given testing session was that at least 69% of the child's reaches were to that side ($p = .10$, binomial distribution).

† Percent of sessions where a preference was shown to either side = (columns 3 + 4) divided by column 1.

‡ Percents are based on those sessions where infant made at least one error.

the same pattern of performance is still observed: percent correct equals 92% on REPEAT FOLLOWING CORRECT trials, 40% on REVERSAL FOLLOWING CORRECT trials, 57% on REPEAT FOLLOWING ERROR trials, and 83% on REVERSAL FOLLOWING ERROR trials. Indeed, if the seven children who showed a side preference on at least 50% of their sessions are omitted from the analysis, the same pattern still remains: 92%, 51%, 63%, and 80% for each type of trial respectively (REPEAT FOLLOWING CORRECT trials versus REVERSAL FOLLOWING CORRECT trials: $t[4] = 4.72$, $p = .009$; REPEAT FOLLOWING CORRECT trials versus REPEAT FOLLOWING ERROR trials: $t[4] = 4.91$, $p = .008$). Likewise, if all the children are included in the analysis, but all sessions on which a side preference was shown (sessions where $\geq$ 69% of reaches to one side) are omitted, the pattern still remains: 85%, 74%, 57%, and 48% for each type of trial respectively (REPEAT FOLLOWING CORRECT trials versus REVERSAL FOLLOWING CORRECT trials: $t[11] = 2.41$, $p = .02$; REPEAT FOLLOWING CORRECT trials versus REPEAT FOLLOWING ERROR trials: $t[11] = 3.92$, $p = .0003$; note, however, the low score on REVERSAL FOLLOWING ERROR trials).

The results for the infants studied cross-sectionally are similar. Infants reached significantly more often to the right or left on 50% of the sessions (28% to the right, 22% to the left). If these 18 sessions are omitted from the analysis, however, the pattern of differential performance across trials remains. Percent correct by type of trial is 97%, 71%, 53%, and 70% respectively (REPEAT FOLLOWING CORRECT trials versus REVERSAL FOLLOWING CORRECT trials: $t[17] = 4.89$, $p = .0001$; REPEAT FOLLOWING CORRECT trials versus REPEAT FOLLOWING ERROR trials: $t[17] = 3.59$, $p = .002$). Looking at reversal trials only, 28% of the infants erred on reversals to both the right and the left; 28% erred only on reversals to the right; 39% erred only on reversals to the left; 5% never erred on a reversal trial.

Piaget: Location A = "The Place Where I Find Things"

Piaget felt there was something special about the first hiding place for the infant. Infants, Piaget theorized, would continue reaching back to where they first found the toy. Piaget looked at performance when side of hiding is only changed once; it is not clear what he would have predicted over multiple reversals. We investigated, however, whether infants would reach back to that first side over the entire session significantly more often than chance.

In Delayed Response, unlike A$\overline{\text{B}}$, a reversal can occur even if the child has never found the toy at the first hiding place. Therefore, to investigate the Piagetian hypothesis we looked at only that subset of sessions on which infants were correct on the first trial (77% of the longitudinal sessions, 89% of the cross-sectional sessions). Percent of reaches during a testing session to the first side of hiding in that session, if the infant was correct on the first trial, was significantly greater than chance for both the longitudinal and cross-sectional samples (59%, $t[93] = 3.83$, $p = .002$; 61%, $t[31] = 2.27$, $p = .03$, respectively). This is true despite the fact that initial side of hiding was alternated from right to left across longitudinal testing sessions for the same child, and was counterbalanced across children in the cross-sectional sample.

Percent of reaches over a testing session to the first side of hiding (whether infant was correct on first trial or not) was 54% and that was not significantly different from chance ($t[157] = 1.38$, NS).

Discussion

There have been two other attempts to study performance on Delayed Response in the first year of life. The results reported here are in full agreement with those two studies. Harris (1973, Experiment 2) tested 9½–10-month-old infants on four trials at a 0 sec delay and four trials at a delay of 5 sec, with the toy hidden randomly at the left or right over trials. He found that infants performed significantly better at the 0 sec delay than at the 5 sec delay. This compares well with the finding of the present study that infants perform significantly better on Delayed Response at briefer delays than at longer delays.

Brody (1981) tested infants on an Indirect Delayed Response problem. A light appeared at one of two locations, a delay was imposed, and then infants were allowed to reach. Infants were rewarded for reaching to where the light had been. She found that 8-month-olds succeeded with a 0 sec delay (250 msec), but failed with delays of 3 sec or more. This compares well with the finding that the mean delay at which the 8-month-old infants make the A$\overline{\text{B}}$ error is 3 sec (Gratch & Landers, 1971; Fox et al., 1979; Diamond, 1985), and the present finding that the mean delay for below criterion performance on Delayed Response at 8 months is 3 sec. Brody also found that 12 month olds succeeded with delays of 3, 6, and 9 sec. This compares well with the finding that 12 month olds succeed on A$\overline{\text{B}}$ with delays as long as 10 sec (Diamond, 1985), and the finding of the present study that 12 months olds succeed on Delayed Response with delays of at least 10 sec.

We found the developmental progression in performance on Delayed Response to be almost identical to that previously shown for A$\overline{\text{B}}$ (Diamond, 1985), even though Delayed Response testing was conducted in a different laboratory by different testers with babies from a different part of the country than that for A$\overline{\text{B}}$ testing. Not only were the delays infants could tolerate at each age comparable on A$\overline{\text{B}}$ and Delayed Response, but the results were in agreement in all other respects as well. For example, as on A$\overline{\text{B}}$, large individual differences were found at each age and the same differential pattern of performance by type of trial was found. This pattern of performance indicates that errors did not occur randomly over a session, rather they occurred much more often on certain kinds of trials than on other kinds of trials even though the hiding procedure and delay were identical throughout. Infants tended to reach to where they had reached on the previous trial, despite observing the hiding on each trial. They tended to be correct on the first trial, and to reach back to that well on subsequent trials. Once they were correct at the other well, they tended to continue to reach there.

This evidence of a developmental progression on Delayed Response between 7½–12 months links changes during this age period with the best established marker of frontal lobe function in nonhuman primates. The performance of the younger infants (7½–9 months of age) on Delayed Response is comparable in all respects to the performance of prefrontally operated monkeys on Delayed Response. Monkeys with dorsolateral prefrontal cortex lesions succeed on Delayed Response if there is no delay (Harlow et al., 1952; Battig, Rosvold, &

Mishkin, 1960; Goldman et al., 1970; Fuster & Alexander, 1971) or if they are allowed to keep looking at, or orienting their body toward, the correct well throughout the delay (Fulton & Jacobsen, 1935; Battig et al., 1960; Miles & Blomquist, 1960; Pinsker & French, 1967). They fail, however, if a delay is introduced, even if it is as brief as 2–5 sec (Harlow et al., 1952; Battig et al., 1960; Goldman et al., 1970; Fuster & Alexander, 1971). If a distractor is introduced, or if the subject spontaneously shifts orientation, the performance of monkeys with dorsolateral prefrontal cortex lesions falls to chance levels, even if the disturbance is only momentary (Fulton & Jacobsen, 1935). By 12 months, human infants succeed on Delayed Response at delays of at least 10 sec, just as do unoperated monkeys and monkeys with lesions elsewhere in the brain (e.g., Goldman et al., 1970; Jacobsen, 1935; 1936).

Researchers have never investigated whether monkeys show the differential pattern of performance on Delayed Response described here for human infants, but when tested on $A\overline{B}$, monkeys with dorsolateral prefrontal cortex lesions show exactly the same pattern of performance by type of trial reported here in human infants (Diamond & Goldman-Rakic, 1989).

Fragile memory, as indicated by difficulty in spanning the delay between hiding and retrieval, can account for some of the results on Delayed Response and $A\overline{B}$. Human infants and prefrontal monkeys perform much better at briefer delays than at longer delays, and they perform perfectly with no delay if they do not take their eyes off of the correct well. As they get older infants can withstand longer and longer delays.

Forgetting, however, cannot account for the pattern of differential performance across trials, nor can side preferences. Forgetting cannot account for the pattern because delay, and hence presumably memory demands, are equal across trials, yet error rate is not equal across trials. The pattern of performance indicates a pull to reach back to where the subject had previously been correct. Success at the first well (A) strengthens the response of reaching to A. The tendency to repeat the rewarded response must be inhibited, or overridden, if the subject is to reach correctly. The influence of the success experience at well A seems to be more powerful than more recent failure experiences at the second well (B), as infants and prefrontal monkeys tended to repeat errors over subsequent trials. We think the pattern of differential performance across trials indicates that infants and prefrontal monkeys have difficulty inhibiting the tendency to repeat the rewarded response when they are given more recent information that a different response is now appropriate (i.e., when they see the toy hidden at the other well). The very first trial does not require inhibition because no response tendency has been created by previous success experiences. Note that infants performed very well on trial 1. Note, also, that performance on trial 1 was much better among infants tested only once than for infants tested every two weeks. This suggests perhaps that experience, even over a gap of two weeks, was still exerting an influence over infants' behavior.

There is some evidence that infants may sometimes know which well is correct, even when they reach to the incorrect well. For example, sometimes they reach to A, uncover but do not look into the well, reach immediately to B, uncover it and look in for the toy. It is as if they know the toy is not at A, even though they reach there first. Occasionally, an infant will look at B as he or she reaches to A, again it is as if the infant knows the toy is at B even as the reach goes to A (see Diamond, 1988; in press). If looking rather than reaching were the dependent measure, the infant would be judged correct. This is reminiscent of the behavior of adult patients with damage to frontal cortex. For example, on the Wisconsin Card Sort test, frontal patients will sometimes tell the experimenter the new, correct sorting criterion even as they continue to sort by the old sorting criterion, and will reprimand themselves (saying, "Wrong") each time they sort a card by the old criterion (Luria & Homskaya, 1964; Milner, 1964). Evidence such as this is consistent with a problem in inhibiting an engrained response. It is also consistent with the wealth of data which Baillargeon is accumulating which shows that when looking rather than reaching is used as the dependent measure, infants appear to remember quite well where the toy has been hidden (Baillargeon, 1987; Baillargeon & Graber, in press).

Thus, it is argued that success on $A\overline{B}$ or Delayed Response requires both memory and inhibitory control. If either ability alone is taxed, some errors will occur (e.g., a few errors occur at well A when a delay is used, taxing only memory; a few errors occur at well B when transparent covers are used, taxing only inhibitory control), however, the vast majority of errors occur when both abilities are taxed (when the toy is hidden at well B with opaque covers and a delay is imposed).

A word is in order about the kind of memory required by $A\overline{B}$ and Delayed Response. It might seem contradictory to argue that infants have difficulty remembering where the toy has been hidden over a delay of 3–12 secs, and yet can remember where they last found the toy on previous trials (which happened perhaps minutes ago) and perhaps can remember where they found the toy on the previous testing session (2 weeks ago). We do not see this as contradictory because we think two different kinds of memory are involved, which rely on different neural systems.

The kind of memory that shows up as a response bias is the kind of memory that has traditionally been assessed using conditioning paradigms. Studies that have used conditioning to assess memory in infants (the dependent measure being how long a response is retained) have typically found quite long memory in very young infants. For example, infants of only 2 months can remember a conditioned response for at least 3–5 days (Rovee-Collier, 1984). We would argue that this is the kind of memory called "implicit" or "procedural" (Cohen, 1984; Mishkin, Malamut, & Bachevalier, 1984; Squire & Cohen, 1984; Schacter, 1987). It is the kind of memory that can be demonstrated in behavior without any conscious awareness of the "memory" on the part of the person. Adults with amnesia demonstrate similar robust memory on conditioning paradigms, even though they have no conscious recollection of having seen or performed the task and even though their conscious recall and recognition are very poor after a few minutes (Weiskrantz & Warrington, 1979). The areas of the brain required for implicit or procedural memory, i.e., required to show the effects of conditioning, are clearly subcortical and mature very early. For example, Thompson and colleagues (McCormick & Thompson, 1984; Thompson, Clark, Donegan, Lavond, Lincoln, Madden, Mamounas, Mauk, McCormick, & Thompson, 1984) have demonstrated the crucial involvement of the cerebellum in retention of the classically conditioned eyeblink response. Frontal cortex is not required to remember or demonstrate a conditioned association (e.g., Allen, 1943).

Very different results are obtained if a conditioning paradigm is used to determine how long a delay between response and reward the subject can withstand within a trial (as opposed to how long a response once learned is retained). For example, Millar and Watson (1979) demonstrated that infants of 6-8 months could acquire a conditioned response if the delay between response and reinforcement were 0 sec, but not if it were 3 sec. These results are quite comparable to those found for A$\bar{B}$ and Delayed Response. Infants of 8 months succeed on Delayed Response or A$\bar{B}$ when the delay between hiding and response is 0 sec, but not when the delay is 3 sec. The Millar and Watson task, like A$\bar{B}$ and Delayed Response, requires that memory be maintained on-line either to relate the response to the reward (Millar & Watson) or to relate the cue (site of hiding) to the response (site of retrieval) (A$\bar{B}$ and Delayed Response). These tasks all look at the ability to bridge a delay within a trial, the ability to integrate information over a temporal separation. Similar results are found for monkeys with frontal cortex lesions, not only on A$\bar{B}$ and Delayed Response, but with conditioning paradigms as well. For example, using a conditioning paradigm, Passingham (1985) demonstrated that frontally lesioned monkeys can make use of a cue if it is still present at the time they are allowed to respond, but are not aided by the cue if there is a delay of 1 or 2 sec between cue and response.

Thus, infants of 8 months do not appear to be able to span a delay within a trial of 3 sec or more, whether the delay is between hiding and retrieval in A$\bar{B}$ or between response and reward in Millar and Watson's operant conditioning task. By 12 months, infants appear to be able to integrate information over delays as long as 10 sec. On the other hand, once they have learned a response (e.g., kicking to make a mobile move) they can retain that response over weeks so that if they see the mobile again after a two week absence they will kick to it at a level significantly above baseline (Rovee-Collier, 1984).

In conclusion, we have found that human infants show a developmental progression on Delayed Response similar to the one they show on A$\bar{B}$. They improve on Delayed Response over the same ages ($7\frac{1}{2}$–12 months) and at the same rate at which they improve on A$\bar{B}$. They show the same pattern of performance by type of trial. Moreover, the performance of human infants of $7\frac{1}{2}$–9 months on Delayed Response is fully comparable to the Delayed Response and A$\bar{B}$ performance of monkeys with lesions of dorsolateral prefrontal cortex. We suggest that a maturational change in frontal cortex may underlie the improvement in A$\bar{B}$ and Delayed Response performance between $7\frac{1}{2}$–12 months in human infants, and that the cognitive abilities subserved by frontal cortex and required for success on these tasks are (1) memory (specifically, the ability to hold something in mind for use in guiding action) and (2) inhibitory control (the ability to resist a bias to make the prepotent response). The ability to hold a goal in mind in the absence of external cues and to use that remembered goal to guide behavior despite the pull of previous reinforcement to act otherwise confers flexibility and the freedom to choose and control one's actions.

Notes

This research was supported by BRSG grant #S07RR07054-22 and NIMH grant #RO1 MH/HD41842 to the first author. We would like to thank Johanna Grant Nicholas for her excellent work in testing subjects, Jerome Kagan, Michael Posner, and especially Richard Abrams for their helpful comments on an earlier draft, and all the parents and infants who participated in the study.

[1] The rule for determining where the toy was hidden, plus a slightly simplified rule for incrementing delay over sessions during longitudinal testing, were the *only* differences between Delayed Response testing and $A\bar{B}$ testing as conducted by Diamond (1985). Procedures within a trial were absolutely identical for Delayed Response reported here and for $A\bar{B}$ (Diamond, 1985).

The procedures used within a trial by Diamond (1985) for $A\bar{B}$ testing were standard, but testing was conducted over more trials than are commonly used. Most $A\bar{B}$ testing includes only one reversal. Diamond includes more reversals to accumulate more data per subject on this critical type of trial. Performance on the first reversal trial is reported below, however, to maximize comparability between the present results and the literature on $A\bar{B}$.

References

Allen, W. F. (1943). Results of prefrontal lobectomy on acquired and on acquiring correct conditioned differential responses with auditory, general cutaneous and optic stimuli. *The Am. J. Physiol., 139:* 525–531.

Baillargeon, R. (1987). Object permanence in 3½-4½-month old infants. *Dev. Psychol., 23:* 655–664.

Baillargeon, R., & Graber, M. (1988). Evidence of location memory in 8-month-old infants in a non-search task. *Dev. Psychol., 24:* 502–511.

Battig, K., Rosvold, H. E., & Mishkin, M. (1960). Comparison of the effects of frontal and caudate lesions on delayed response and alternation in monkeys. *J. Comp. and Physiol. Psychol., 4:* 400–404.

Brody, L. R. (1981). Visual short-term cued recall memory in infancy. *Child Devel., 52:* 242–249.

Brown, R. M., & Goldman, P. S. (1977). Catecholamines in neocortex of rhesus monkeys: Regional distribution and ontogenetic development. *Brain Res., 124:* 576–580.

Brozoski, T., Brown, R. M., Rosvold, H. E., & Goldman, P. S. (1979). Cognitive deficit caused by depletion of dopamine in prefrontal cortex of rhesus monkey. *Science, 205:* 929–931.

Bugbee, N. M., & Goldman-Rakic, P. S. (1981). Functional 2-deoxyglucose mapping in association cortex: Prefrontal activation in monkeys performing a cognitive task. *Soc. Neurosci. Abst. 7:* 416.

Butterworth, G. (1975). Object identity in infancy: The interaction of spatial location codes in determining search errors. *Child Devel., 46:* 866–876.

Cohen, N. J. (1984). Preserved learning capacity in amnesia: Evidence for multiple memory systems. In: L. R. Squire & N. Butters (Eds.), *Neuropsychology of Memory* NY: Guilford.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on $A\bar{B}$. *Child Devel., 56:* 868–883.

Diamond, A. (1988). Differences between adult and infant cognition: is the crucial variable presence or absence of language? In L. Weiskrantz (ed.), *Thought Without Language*. Oxford: Oxford University Press.

Diamond, A. (in press). Frontal lobe involvement in cognitive changes during the first year of life. In K. Gibson, M. Konner, and A. Patterson (eds.), *Brain and Behavioral Development*. New York: Aldine Press.

Diamond, A. (submitted). Retrieval of an object from an open box: The development of visual-tactile control of reaching in the first year of life. *Mon. Soc. Res. Child Devel.*

Diamond, A., & Goldman-Rakic, P. S. (1985). Evidence for Involvement of Prefrontal Cortex in Cognitive Changes during the First Year of Life: Comparison of Performance of Human Infants and Rhesus Monkeys on a Detour Task with Transparent Barrier. *Soc. Neurosci. Abst. (Part II), 11:* 832.

Diamond, A., & Goldman-Rakic, P. S. (1986). Comparative development in human infants and infant rhesus monkeys of cognitive functions that depend on prefrontal cortex. *Soc. Neurosci. Abst., 12:* 742.

Diamond, A., & Goldman-Rakic, P. S. (in press). Comparison of human infants and rhesus monkeys on Piaget's $A\bar{B}$ Task: Evidence for dependence on dorsolateral prefrontal cortex. *Exper. Brain Res.*

Diamond, A., Zola-Morgan, S., & Squire, L. (in press). Successful performance by monkeys with lesions of the hippocampal formation on A$\overline{\text{B}}$ and object retrieval, two tasks that mark developmental changes in human infants. *Behav. Neurosci.*

Fox, N., Kagan, J., & Weiskopf, S. (1979). The growth of memory during infancy. *Genet. Psychol. Monog., 99:* 91–130.

Freedman, M. & Oscar-Berman, M. (1986). Bilateral frontal lobe disease and selective delayed response deficits in humans. *Behav. Neurosci., 100:* 337–342.

Fulton, J. E., & Jacobsen, C. F. (1935). The functions of the frontal lobes, a comparative study in monkeys. *Adv. Mod. Biol., 4:* 113–123.

Fuster, J. M. (1973). Unit activity in prefrontal cortex during delay-response performance: Neuronal correlates of transient memory. *J. Neurophysiol., 36:* 61–78.

Fuster, J. M. (1980). *The Prefrontal Cortex*. New York: Raven Press.

Fuster, J. M., & Alexander, G. E. (1970). Delayed response deficit by cryogenic depression of frontal cortex. *Brain Res., 20:* 85–90.

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science, 173:* 652–654.

Goldman, P. S. (1971). Functional development of the prefrontal cortex in early life and the problem of neuronal plasticity. *Exper. Neurol., 32:* 366–387.

Goldman, P. S. (1974). Recovery of function after CNS lesions in infant monkeys. *Neurosci. Res. Prog. Bull., 12:* 217–222.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational knowledge. In: Plum, F. & Mountcastle, V. (Eds.), *Hand. Physiol., 5:* 373–417.

Goldman, P. S. & Rosvold, H. E. (1970). Localization of function within the dorsolateral prefrontal cortex of the rhesus monkey. *Exper. Neurol., 27:* 291–304.

Goldman, P. S., Rosvold, H. E., & Mishkin, M. (1970). Selective sparing of function following prefrontal lobectomy in infant monkeys. *Exper. Neurol., 29:* 221–226.

Gratch, G. (1975). Recent studies based on Piaget's view of object concept development. In: L. B. Cohen & P. Salapatek (Eds.), *Infant Perception: From Sensation to Cognition (Vol. 2)*. New York: Academic.

Gratch, G., Appel, K. J., Evans, W. F., LeCompte, G. K., & Wright, N. A. (1974). Piaget's stage IV object concept error: Evidence for forgetting or object conception? *Child Devel., 45:* 71–77.

Gratch, G. & Landers, W. F. (1971). Stage IV of Piaget's theory of infant's object concepts: A longitudinal study. *Child Devel., 42:* 359–372.

Harlow, H. F. & Dagnon, J. (1943). Problem solution by monkeys following bilateral removal of prefrontal areas: I. Discrimination and discrimination reversal problems. *J. Exper. Psychol., 32:* 351–356.

Harlow, H. F., Davis, R. T., Settlage, P. H., & Meyer, D. R. (1952). Analysis of frontal and posterior association syndromes in brain-damaged monkeys. *J. Comp. Physiol. Psychol., 54:* 419–429.

Harris, P. L. (1973). Perseverative errors in search by young infants. *Child Devel., 44:* 29–33.

Harris, P. L. (1983). Infant cognition. In: M. M. Haith & J. J. Campos (Eds.), *Handbook of Child Psychology: Vol. 2. Infancy and Developmental Psychobiology*. New York: Wiley.

Huttenlocher, P. R. (1979) Synaptic density in human frontal cortex—developmental changes and effects of aging. *Brain Res., 163:* 195–205.

Isseroff, A., Galkin, T., Rosvold, H. E., & Goldman-Rakic, P. S. (1982). Spatial memory impairments following damage to the mediodorsal nucleus of the thalamus in rhesus monkeys. *Brain Res., 232:* 97–113.

Jacobsen, C. F. (1935). Functions of frontal association areas in primates. *Arch. of Neurol. Psychiat., 33:* 558–560.

Jacobsen, C. F. (1936). Studies of cerebral functions in primates. I. The function of the frontal association areas in monkeys. *Comp. Psychol. Monog., 13:* 1–60.

Janowsky, J. S., Schwartz, L., & Stiles-Davis, J. (1987). The A$\overline{\text{B}}$ error: A reanalysis of its reinstatement at 10 months. *Abst. Soc. Res. Child Devel.*, 28.

Lahey, M. A., Downey, R. G., & Saal, F. E. (1983). Intraclass correlations: There's more there than meets the eye. *Psychol. Bull., 93:* 586–595.

Luria, A. R. & Homskaya, E. D. (1964). Disturbance in the regulative role of speech with frontal lobe lesions. In: J. M. Warren & K. Akert (Eds.), *The Frontal Granular Cortex and Behavior*. New York: McGraw-Hill.

McCormick, D. A. & Thompson, R. F. (1984). Cerebellum: Essential involvement in the classically conditioned eyelid response. *Science, 223:* 296–299.

Meyer, D. R., Harlow, H. F., & Settlage, P. H. (1951). A survey of delayed response performance by normal and brain-damaged monkeys. *J. Comp. and Physiol. Psychol., 44:* 17–25.

Miles, R. C. & Blomquist, A. (1960). Frontal lesions and behavioral deficits in monkey. *J. Neurophysiol., 23:* 471–484.

Miller, E., Goldman, P. & Rosvold, H. (1973). Delayed recovery of function following orbital prefrontal lesions in infant monkeys. *Science, 182:* 304–307.

Millar, W. S., & Watson, J. S. (1979). The effect of delayed feedback on infant learning reexamined. *Child Devel., 50:* 747–751.

Milner, B. (1964). Some effects of frontal lobectomy in man. In: J. M. Warren & K. Akert (Eds.), *The Frontal Granular Cortex and Behavior*. New York: McGraw-Hill.

Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In: G. Lynch, J. L. McGaugh, & N. M. Weinberger (Eds.), *Neurobiology of Learning and Memory*. New York: Guilford.

Niki, H. (1974). Differential activity of prefrontal units during right and left delayed response trials. *Brain Res., 70:* 346–349.

Passingham, R. E. (1972). Non-reversal shifts after selective prefrontal ablations in monkeys. *Neuropsychologia, 10:* 41–46.

Passingham, R. E. (1985). The memory of monkeys (*macaca mulatta*) with lesions in prefrontal cortex. *Behav. Neurosc., 99:* 3–21.

Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Basic Books. Original French edition, 1937.

Pinsker, H. M., & French, G. M. (1967). Indirect delayed reactions under various testing conditions in normal and midlateral frontal monkeys. *Neuropsychologia, 5:* 13–24.

Rosenkilde, C. E. (1979). Functional heterogeneity of the prefrontal cortex in the monkey: A review. *Behav. Neural Biol., 25:* 301–345.

Rovee-Collier, C. (1984). The ontogeny of learning and memory in human infancy. In: R. Kail & N. E. Spear (Eds.), *Comparative Perspectives on the Development of Memory* (pp 103–134). Hillsdale, NJ: Lawrence Erlbaum.

Schacter, D. L. (1987). Implicit memory: History and current status. *J. Exp. Psychol.: Learn., Mem. & Cogn., 13:* 501–518.

Schuberth, R. E. (1982). The infant's search for objects: Alternatives to Piaget's theory of object concept development. In: L. P. Lipsitt & C. K. Rovee-Collier (Eds.), *Advances in Infancy Research* (Vol. 2). Norwood, NJ: Ablex.

Squire, L. R. & Cohen, N. J. (1984). Human memory and amnesia. In: G. Lynch, J. L. McGaugh, & N. M. Weinberger (Eds.), *Neurobiology of Learning and Memory* (p. 3–64), New York: Guilford.

Stamm, J. S. (1969). Electrical stimulation of monkeys' prefrontal cortex during delayed response performance. *J. Comp. Physiol. Psychol., 67:* 535–546.

Stamm, J. S., & Rosen, S. C. (1972). Cortical steady potential shifts and anodal polarization during delayed response performance. *Acta Neurobiol. Exp., 32:* 193–209.

Thompson, R. F., Clark, G. A., Donegan, N. H., Lavond, D. G., Lincoln, J. S., Madden IV, J., Mamounas, L. A., Mauk, M. D., McCormick, D. A., & Thompson, J. K. (1984). Neuronal substrates of learning and memory: A "multiple-trace" view. In: G. Lynch, J. L. McGaugh, & N. M. Weinberger (Eds.), *Neurobiology of Learning and Memory*. New York: Guilford Press.

Weiskrantz, L. & Warrington, E. K. Conditioning in amnesic patients. *Neuropsychologia, 17:* 187–194.

Wellman, H. M., Cross, D., & Bartsch, K. (1987). A meta-analysis of research on Stage 4 object permanence: The *A*-not-*B* error. *Mon. Soc. Res. Child Devel., 5(3)*.