Developmental Psychology
2003, Vol. 39, No. 5, 825–847

0012-1649/03/$12.00 DOI: 10.1037/0012-1649.39.5.825

Early Success in Using the Relation Between Stimuli and Rewards to Deduce an Abstract Rule: Perceived Physical Connection Is Key

Adele Diamond
University of Massachusetts Medical School

EunYoung Lee
University of Pennsylvania

Michael Hayden
University of Massachusetts Medical School

Are spatial proximity (0.10–12.5 cm), temporal proximity (0-, 2-, and 5-s gaps), and/or perceived connectedness of stimulus and reward key to infants' ability to deduce an abstract nonmatching rule from reward feedback? In this investigation, 3 conditions of the delayed nonmatching to sample task were administered to infants 9, 12, and 15 months old, and 5 more conditions were administered just to 12-month-olds. Results showed that connectedness is key. In its presence, neither close spatial or temporal proximity was needed. In the absence of the perception that stimulus and reward were components of a single thing, even close spatial and temporal proximity were insufficient for infants in the 1st year to grasp the rule-based association between stimuli and rewards.

In the delayed nonmatching to sample (DNMS) task, reaching to the novel stimulus is always rewarded. Participants are not told that rule; they must deduce it on the basis of feedback. In the standard testing procedure, during the first part of each trial (the "sample presentation" or "familiarization" phase), a new, nonsense object is presented at the midline over a shallow, baited well. The participant displaces the stimulus and retrieves the reward. After a 5-s delay, the second part of each trial (the "test phase") occurs: The sample object and a new, nonsense object are presented, one to the right and one to the left, each over a shallow well. The well under the familiar sample is empty; the well under the new stimulus is baited. Hence the name of the task, delayed nonmatching to sample—after a delay, the participant is rewarded for displacing the stimulus that does not match the sample.

Infants generally cannot deduce the rule governing correct performance on the DNMS task until they are almost 2 years old (roughly 21 months; Diamond, Towle, & Boyer, 1994; Overman, Bachevalier, Turner, & Peuster, 1992). This is true whether a child is tested only once (Diamond, 1990) or tested repeatedly, 5 days a week, every week, beginning at 12 months of age (Overman, 1990).

The DNMS task was originally devised as a test of recognition memory. Adult monkeys and adult human patients who have damage to the medial temporal lobe system critical for recognition memory but in whom other cognitive abilities are mature and intact fail the DNMS task because of their impaired memory. They perform well at the brief 5-s delay but are increasingly impaired as the length of the delay increases (e.g., Meunier, Bachevalier, Mishkin, & Murray, 1993; Mishkin, 1978; Murray & Mishkin, 1998; Squire, Zola-Morgan, & Chen, 1988; Zola et al., 2000; Zola-Morgan, & Squire, 1986; Zola-Morgan, Squire, Amaral, & Suzuki, 1989; Zola-Morgan, Squire, & Mishkin, 1982).

Infants younger than 21 months fail the standard DNMS task even at the brief 5-s delay, though there is abundant evidence that their recognition memory span is far longer than 5 s (e.g., Pascalis, de Haan, Nelson, & de Schonen, 1998). Moreover, in the same session at which an infant succeeds at the 5-s delay, that same infant usually succeeds at longer delays as well (Diamond, Churchland, Cruess, & Kirkham, 1999; Diamond et al., 1994). If performance on the task were indexing memory development, one might expect success first at shorter delays and then, as infants get older, emerging success at longer delays. There is now widespread agreement that inadequate recognition memory does not appear to be the reason infants fail the standard DNMS task. It appears that they fail because they are unable to deduce the nonmatching rule. They are unable to use the feedback provided on each trial about which response is correct because they are evidently unable to understand that the feedback they are receiving (receipt or nonreceipt of the reward) is related to which stimulus they have chosen.

Adele Diamond and Michael Hayden, Center for Developmental Cognitive Neuroscience, Eunice Kennedy Shriver Center Campus, University of Massachusetts Medical School; EunYoung Lee, Department of Psychology, University of Pennsylvania.

This research was supported by National Institute of Mental Health Grant R01 MH41842 and National Institute for Child Health and Human Development Grant R01 HD35453.

We gratefully acknowledge the outstanding help of Kathy Mason, Kristin Shutts, Lisa Briand, and Kim Dilda with the testing and of Stephen Baker, Matt Davidson, and Lorrie Gehlbach with the data analyses. We also thank Kristin Shutts and Jerome Kagan for astute comments on a draft of the manuscript and all the parents and infants, without whose participation this research would not have been possible.

Correspondence concerning this article should be addressed to Adele Diamond, Center for Developmental Cognitive Neuroscience, Eunice Kennedy Shriver Center Campus, University of Massachusetts Medical School, 200 Trapelo Road, Waltham, Massachusetts 02452. E-mail: adele.diamond@umassmed.edu

If the reward is attached with Velcro ("velcroed") to the base of the stimulus, however, infants can succeed on the DNMS task at only 9–12 months of age with delays of 5 s (Diamond et al., 1999). Here the reward is connected to (though detachable from) the stimulus. When the stimulus is atop its well, the reward is concealed within its well (as in standard DNMS testing). When an infant picks up the stimulus, however, the reward moves with it, rather than remaining in the well.

Evidently, infants are able to grasp the relation between the stimulus and the reward in the Velcro condition but not in the standard condition. Why? Is it the closer spatial proximity of the stimulus and reward (reward attached to stimulus [in the Velcro condition] vs. reward a fraction of an inch below the stimulus [in the standard DNMS condition])? Is it the closer temporal proximity of acting on the stimulus and seeing the reward? (In the Velcro condition, as soon as the stimulus is picked up and turned over, the reward is visible, whereas in the standard DNMS condition, the infant usually turns over and explores the stimulus for a few moments before looking in the well and seeing the reward.) Or is it the physical connection between the stimulus and the reward in the Velcro condition that is key?

We hypothesized that temporal proximity would be more important than spatial proximity. For example, Millar and Schaffer (1972) showed with an operant conditioning task that if the reward and the response lever were spatially displaced but in the same visual field (so that infants could see the visual–auditory reward as soon as they depressed the lever), even infants as young as 6 months could learn the rule that pushing the lever caused the reward to appear. However, when the site of responding and the site of the reward were separated by the same distance in space but positioned so that both could not be seen simultaneously, infants of 6 months could no longer learn the conditioned response. Strong evidence also comes from the substantial body of work by Rovee-Collier (e.g., Fagen & Rovee-Collier, 1982; Rovee-Collier, 1984, 1990, 1995, 1999), in which infants had to learn that kicking their feet made a mobile move that was at least 60 cm overhead. The spatial gap between the response site and the reward is substantial in the mobile paradigm, but the reward (movement of the mobile) happens almost instantaneously with the infant's response. Under those circumstances, even infants of only 2 months can acquire the operant response. Similarly, Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, and Hopkins (1989) reported that rhesus monkeys easily mastered precise control of a joystick to respond to computer-generated targets even though the joystick was 9–18 cm from the stimulus on the computer screen. The reason: Because the monkeys were not looking at their hands moving the joystick (contrary to the way monkeys normally attend to the movement of their hands in the Wisconsin General Test Apparatus), they were looking at the stimulus while responding and while receiving the reward; there was no temporal gap.

We hypothesized that physical connection, especially if it led to the perception that the stimulus and reward were parts of a single object, might be as helpful as close temporal proximity. Consider the results of Aguiar and Baillargeon (2000), who placed two cloths, one twice as long as the other, directly in front of 9-month-old infants. On top of the longer cloth, near its far end, sat a desired toy. Pulling the cloth would bring the toy within reach. Equally far from the infant sat an identical toy behind the shorter cloth; pulling that cloth, however, would not bring the toy within reach. Aguiar and Baillargeon found that after an infant's initial success, if the locations of the shorter and longer cloths were reversed, 9-month-olds continued to succeed if the toy was attached to the longer cloth but not if the toy and cloth were not physically attached, even though the toy was still on top of the cloth, spatially contiguous, and there was no temporal gap between pulling on the cloth and movement of the toy toward the infant. In the former case (with physical connection), infants correctly switched from pulling the cloth on Side A to pulling the cloth on Side B. In the latter case (with physical contiguity but not connection), 9-month-olds made the A-not-B error by continuing to pull the cloth on Side A. Evidently, the existence of a physical attachment made a big difference to the infants.

Other evidence consistent with the centrality of physical connectedness comes from the work of Jarvik. It can take monkeys 100–200 trials to learn a simple color discrimination that is tested by placing a green plaque over one well and a red plaque over the other. The left–right placement of the plaques is varied randomly over trials, but the same color is always rewarded. Jarvik (1956) varied one simple thing—whether the reward was placed *in the well* under a plaque or *taped to the underside* of the plaque. When the reward was attached to the plaque, Jarvik found one-trial learning (much as we found that velcroing the reward to the base of the stimulus dramatically improved infants' DNMS performance; Diamond et al., 1999). Was it because the reward and plaque were contiguous or because they were physically connected? Jarvik addressed that question in a 1953 study with monkeys, using bread as the reward (one slice treated with a disagreeable flavor). He found that monkeys learned the color discrimination in one trial if there was a physical connection between the stimulus and reward (the stimuli being red and green translucent celluloid *pasted* on top of the bread) but performed as poorly as in the standard procedure when the same stimuli were *placed* on top of, but not attached, to the bread. In the latter condition, monkeys still performed at chance after 75 trials. Here, as in the Aguiar and Baillargeon (2000) work with infants, physical connection was critical.

It is possible that when the objects were attached, their synchrony of movement was exact, whereas when one was on top of the other unattached, the correlation between their movements was not exact. Synchrony of movement has long been known to be a powerful cue for infants in determining whether two things are part of one whole or are separate objects (Spelke, 1985; Vishton & Badger, 2003).

Other work with children 1–2 years older may also be relevant. Rudel (1955) found that when the reward was placed *inside* the stimulus boxes, children of 1½–3½ years learned to choose on the basis of relative size in far fewer trials than did even older children who were tested with the reward *underneath* the stimulus (Alberts & Ehrenfreund, 1951; Kuene, 1946). DeLoache and Brown (1983) did something reminiscent of Rudel's work. They varied whether toys were hidden in, or near, a piece of furniture. They found that infants of 18–22 months performed significantly better when the reward was hidden *in* the piece of furniture rather than *near* it; by 24–30 months, infants performed equally well in both conditions. DeLoache (1986) varied whether a reward was hidden in one of four distinctive containers or whether the distinctive containers were mounted on top of plain boxes into which the rewards were placed. When the boxes were scrambled, 21-month-olds were 80%

correct when the rewards were *in* the distinctive containers but only 35% correct when the distinctive containers marked where the rewards were hidden (the reward being in the box *underneath*).

In several studies, DeLoache and colleagues (DeLoache, 1989, 1995, 2000) have presented preschoolers with the task of finding a toy and a small replica of the toy in a full-size room and in a small scale-model of the room, respectively. In these studies, the child sees only the toy or only the tiny replica being hidden. To find both, the child must appreciate the relation between the two spaces. In the standard condition, children are told that the small model room is a scale-model of the larger room, and they go through elaborate training on that. In one of the variations on the standard condition, children are told that there is only one space but that there is a magic machine that can shrink or enlarge things. The children are to find the toy in the full-size room where they saw it hidden and then to find the shrunken version of the toy in the shrunken room. Adults understand the standard condition and this variant to be equivalent: The task in both is to find the full-size toy in the full-size space and the tiny toy in the tiny space after having seen only one of the toys hidden. For children of 2–2½ years, however, the conditions are not equivalent at all. They fail in the first condition and succeed in the second. Even in the first condition, children of 2–2½ years can find the full-size (or tiny) toy in the space where they saw it hidden (so they remember where it was placed), but they fail to find its counterpart in the other space. However, if exactly the same objective problem is framed differently, framed as involving only one space that changes size, then it is transparent to children. We reasoned that if the stimulus and reward appeared to be part of one entity, infants might be able to appreciate the connection between the two but that if the stimulus and reward were two unrelated objects, even if closely juxtaposed in space and time, infants might not be able to understand that the two were meant to be related to one another.

Overview of Experimental Conditions in Studies 1 and 2

Normally, a temporal separation accompanies a spatial separation, and typically connectedness and close spatial proximity co-occur. Hence, most situations do not permit these variables to be disambiguated. (In the standard DNMS task, the reward is separated spatially and temporally from acting on the stimulus, and the stimulus and reward are not physically connected. In the Velcro variant, the reward is spatially proximal and physically contiguous with the stimulus, and sight of the reward occurs temporally proximal with acting on the stimulus.) The use of a jack-in-the-box reward enabled us to circumvent the typical problems encountered with independently manipulating these variables and enabled us to use a 0-s delay between acting on the stimulus and receipt of the reward, something not possible in the standard DNMS procedure.

For five of our experimental conditions, a large, white, rectangular apparatus was used that housed the jack-in-the-boxes and to which stimuli were affixed. The infant could not remove or pick up the reward or any stimulus once it was affixed to the apparatus. The infant could, however, try to retrieve a stimulus, or pull it toward him or her. In so doing, as soon as an infant moved the sample during familiarization or the novel object during test, even ever so slightly, the jack-in-the-box behind that stimulus popped up. It gave the appearance that pulling the stimulus was like pulling a lever that caused the jack-in-the-box to pop up. (During test, the familiar sample object could not be moved, and trying to move it caused no effect.) Here, the stimulus and reward were spatially farther apart than in the standard procedure (12.5 cm apart vs. < 1 cm apart), but displacing the stimulus could immediately produce the reward. We could also delay the appearance of the reward and thus were able to use conditions of 0, 2, and 5 s between an infant's action on the stimulus and appearance of the reward, the stimulus and reward being spatially displaced but apparently physically connected to the same large white box in all those conditions.

Three other experimental conditions used the more traditional DNMS testing procedure. In one, the standard DNMS condition was run with small objects (such as tiny rattles, blinking rings, tiny animals, and coins) serving as rewards. They were placed in shallow wells directly beneath the stimuli. Two jack-in-the-box conditions were run that closely approximated that condition. In those two conditions, each jack-in-the-box was housed within its own little box, and stimuli were positioned directly in front of their associated boxes. Here, stimuli and rewards were spatially proximal but clearly not parts of a single apparatus. In one condition, the jack-in-the-box popped up immediately upon the infant's choosing a stimulus (close temporal proximity and close spatial proximity). In the other condition, the jack-in-the-box popped up 2 s later, which corresponds to the typical lag in standard DNMS testing between choosing a stimulus and seeing its associated reward.

Thus, to summarize, to examine whether spatial proximity, temporal proximity, and/or perceived connectedness are key to infants' ability to understand the relation between a stimulus and its associated reward in the DNMS task, we ran eight conditions that enabled us to assess the independent contribution of each of those three factors (see Table 1). In five of the eight conditions, the stimuli and rewards were parts of a single large white box, a jack-in-the-box served as the reward, and the reward was farther from the stimulus than in the standard DNMS condition. In one of those five conditions, Condition A, the reward popped up the moment an infant moved the stimulus (close temporal proximity), and the sample object was presented for 20–25 s. Condition B was the same as Condition A, but the sample object was presented for only 5 s (as in standard DNMS testing). Condition D was the same as Condition B, but the reward was delayed 2 s (response at stimulus not close in time to receipt of reward). Condition E was the same as Condition B, but the reward was delayed 5 s (response at stimulus and receipt of reward further apart in time), and Condition F was the same as Condition B but there was no cheering from the experimenter (unlike standard DNMS testing, which includes cheering from the experimenter).

In three of the eight conditions, the standard DNMS testing procedure was followed. In Condition C, the standard procedure was followed exactly (stimuli and rewards closer spatially, not close in time, and not part of the same entity). Condition G differed from Condition C only in substituting a jack-in-the-box reward for the small-object rewards (same temporal and connectedness values as for Condition C). Condition H was the same as Condition G except that the jack-in-the-box popped up as soon as the infant chose a stimulus (close temporal and spatial proximity, but no physical connection). For Conditions A, B, and C, infants of 9, 12, and 15 months were tested (Study 1); for all other conditions, only 12-month-olds were tested (Study 2).

Table 1
Characteristics of the Eight Experimental Conditions in Terms of Spatial and Temporal Proximity and Perceived Physical Connection

	Relation of stimulus and reward in terms of		
Condition	Spatial proximity	Temporal proximity	Both being parts of a single thing
Study 1			
A: Jack-in-the-box reward—stimuli and rewards connected to the same white box and apparently connected to one another	Relatively far (12.5 cm apart)	Close (0-s delay)	Yes
B: Jack-short—control for possible confound in A	Relatively far (12.5 cm apart)	Close (0-s gap)	Yes
C: Standard DNMS procedure—small objects and/or bite-size food as rewards	Close (< 1 cm apart)	Relatively far (2-s gap)	No
Study 2			
D: Same as B except 2-s delay between acting on stimulus and jack-in-the-box popping up	Relatively far (12.5 cm apart)	Relatively far (2-s gap)	Yes
E: Same as B except 5-s delay between acting on stimulus and jack-in-the-box popping up	Relatively far (12.5 cm apart)	Farther (5-s gap)	Yes
F: Same as B but no verbal feedback from experimenter	Relatively far (12.5 cm apart)	Close (0-s gap)	Yes
G: Same as C but with jack-in-the-box rewards	Relatively close (5 cm apart)	Relatively far (2-s gap)	No
H: Same as G except no delay between acting on stimulus and jack-in-the-box popping up	Relatively close (5 cm apart)	Close (0-s gap)	No

Note. DNMS = delayed nonmatching to sample.

Study 1

Two jack-in-the-box conditions as well as the standard DNMS condition were tested in Study 1 (three conditions in all). For both jack-in-the-box conditions, a large white box containing all rewards and to which all stimuli were affixed was used. In these conditions, the stimulus and reward were spatially farther apart than in the standard procedure (12.5 cm apart vs. < 1 cm apart), but displacing the stimulus immediately produced the reward.

Although the jack-in-the-box reward was farther from the stimulus than was the small-toy reward in the well beneath the stimulus in the standard DNMS procedure, we predicted that infants would perform better in the jack-in-the-box conditions than in the standard DNMS condition because of the closer temporal proximity between acting on the stimulus and seeing the reward and because of the perceived connectedness of the stimulus and reward. We predicted that infants would be able to grasp the relation between a stimulus and its associated reward in the DNMS situation (even if the stimulus and reward were spatially separated) as long as they were closely temporally linked and physically connected.

In the first jack-in-the-box condition, the infant watched as the sample object was attached to the center drawer at the outset of each trial. Sample presentation time was 20–25 s (measured from the time the infant could first see the sample stimulus until that stimulus was removed from view before the delay). Sample presentation time in the standard DNMS condition was the usual 5 s. Hence, the jack-in-the-box and standard conditions differed in the length of time the sample was presented during the initial part of each trial. If 5 s is too short a time for infants to satisfactorily encode the sample stimuli and grow bored with them, infants might prefer to reach back to the familiar sample during the test phase rather than to reach for the new, nonmatching object. This would predispose infants in the standard DNMS condition to perform worse than infants tested with the longer sample presentation time in the jack-in-the-box condition. Studies of visual paired comparison have often found that if the sample is presented only briefly, infants do not look preferentially to the novel stimulus (e.g., Caron, Caron, Minichiello, Weiss, & Friedman, 1977; Hunter & Ames, 1975, 1988; Lasky, 1980; Rose, Gottfried, Melloy-Carminar, & Bridger, 1982). This is analogous to infants not reaching preferentially to the novel stimulus on the DNMS task. Now infants have not been found to reach preferentially to the novel stimulus in previous DNMS studies with 5-s presentation times (Diamond et al., 1994, 1999); instead, when infants have erred, their reaches tended to be random. However, if 5 s is too brief for only some stimuli, then that might yield an apparently random selection of the familiar sample on those trials and of the novel object on other trials.

To rule out an explanation of superior performance in the jack-in-the-box condition as due to the longer sample presentation time, we administered a second jack-in-the-box condition in which the presentation time was comparable to that in the standard DNMS condition. Henceforth we refer to these two jack-in-the-box conditions as jack-in-the-box long presentation (or jack-long; Condition A) and jack-in-the-box short presentation (or jack-short; Condition B).

Method

Jack-in-the-Box Conditions (Conditions A and B)

Materials. The jack-in-the-box apparatus was a large white box 70 cm long × 40 cm wide × 22.5 cm high. The top surface consisted of three 15 × 15 cm drawers. The left and right drawers were angled in toward the center so that an equal distance was maintained from each drawer to the infant. Each of the three drawers contained a jack-in-the-box. These jack-in-the-boxes were identical pink and blue stuffed birdies that sprang up when the drawer was moved ever so slightly by the infant's reaching for the stimulus object on top of the drawer. See Figure 1. Each drawer had a hole in the center; this hole was filled by either a white plug or the base of a stimulus object. A long brass pin protruded at the center of the back of each drawer. The back of that pin was used to secure the stimulus object to the apparatus so that when an infant pulled the stimulus object, the drawer moved forward with the object. All drawers always had this pin, so there was no visible cue distinguishing one drawer from another. In the back left corner of each drawer was another hole that measured 0.63 cm in diameter. When the stimulus object on the drawer was the wrong choice, a long peg was in this hole that fastened the drawer in place so that when an infant pulled the stimulus object, the drawer did not move. When the correct stimulus object was on the drawer, a short peg was in the hole. The short peg did not attach the drawer to the box; when the infant pulled this object, the drawer slid forward, releasing the jack-in-the-box. The short and long pegs looked identical on the surface. Again, no visible cue was provided that distinguished the drawers. The apparatus was constructed so that the jack-in-the-box would pop up immediately at the slightest pull on the correct stimulus object. The reward popped up 12.5 cm from the stimulus.

The apparatus sat on a 70 × 70 × 50 cm brown table with two wheels on the front legs and no wheels on the back legs. The table and apparatus were pushed forward, within the infant's reach, and back out of reach, for each trial.

The stimulus objects consisted of 62 "junk" objects made out of wood, plastic, and/or rubber. They were brightly colored and constructed to be

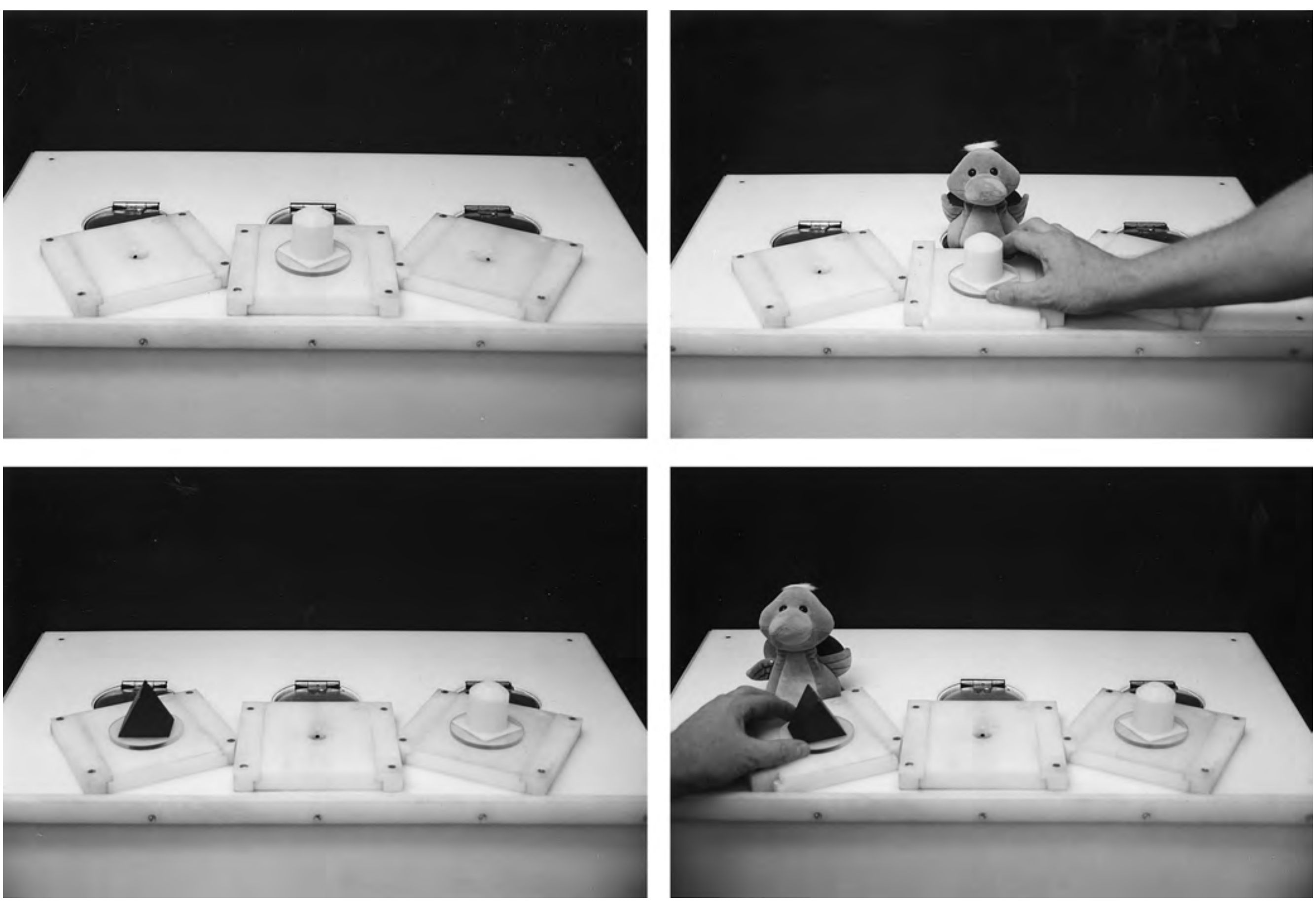

Figure 1. Photos showing the jack-in-the-box testing procedure used in Study 1 with the large white apparatus. Top left frame shows the presentation of the sample stimulus. Pulling on the stimulus even slightly enabled the jack-in-the-box penguin to pop up behind it (top right frame). Bottom left frame shows the presentation of the sample stimulus paired with a novel object. Pulling even slightly on the stimulus not previously seen enabled the jack-in-the-box penguin directly behind it to pop up (bottom right frame). (Pulling on the familiar sample stimulus in the paired presentation would have produced no effect.)

unusual so that the infants would not have seen them before. Heights of the objects ranged from 3 to 15 cm, and their widths ranged from 3 to 7 cm. They were arranged in 31 pairs in a wooden case facing the experimenter. The objects paired together on a given trial had similar dimensions and were roughly equal in their attractiveness to infants, although they always differed in color and shape. For half of the infants in each condition, Stimulus A of each pair served as the sample object, and for half of the infants Stimulus B was used as the sample. Each stimulus object had a peg beneath its clear, round Plexiglas base so that it could be attached to the testing apparatus.

Procedure. Each infant sat on his or her parent's lap on one side of the table, and the experimenter sat opposite them. An assistant sat next to the experimenter.

Training: The experimenter attached a sample training stimulus to the top of the center drawer and demonstrated that pulling the stimulus made the jack-in-the-box pop up. The apparatus was then moved within reach and the infant was encouraged to pull the stimulus. Next, the stimulus was moved to one side and then the other, and the training was repeated at these two locations to let the infant know that there was a jack-in-the-box under each of the other drawers as well and to give the infant experience in acting at all three locations. Each infant received two consecutive practices at each drawer.

Testing: The only difference in procedure between the two jack-in-the-box conditions occurred during the period after a trial through the beginning of the next trial. In the jack-in-the-box long presentation (jack-long) condition, the order of events was as follows: After each trial, the apparatus was rolled back, away from the infant, to prepare for the next trial. At the beginning of the next trial, the experimenter made sure the infant watched as the sample object was attached to the center drawer. In the jack-in-the-box short presentation (jack-short) condition, the order of events was as follows: After each trial, the apparatus was rolled back, away from the infant, an opaque screen was put in place that blocked the infant's view of the testing apparatus, and the sample stimulus was fastened to the center drawer out of view of the infant.

Thus, a testing session started with the table containing the testing apparatus pulled back, out of the infant's reach. In the jack-long condition, the experimenter made sure the infant watched as the sample object was attached to the center drawer (20- to 25-s sample presentation time). In the short presentation condition, the sample object was already attached to the center drawer and the trial began as soon as the experimenter had gotten the infant's attention and directed the infant's gaze toward the sample object (5-s sample presentation time).

The apparatus was then pushed forward, within reach and centered for the infant, and the experimenter asked the infant if she or he could find the birdie. The infant was then allowed to pull and make the jack-in-the-box pop up. Pretesting had indicated that infants got bored with the jack-in-the-box after a few trials. To maintain infants' interest in the jack-in-the-box over trials, the experimenter attached elastic bands, buttons, and various head and face coverings to the birdies so they would look a bit different from trial to trial and so infants would have things to pull on and explore when the jack-in-the-box appeared. After the jack-in-the-box popped up and the infant had explored it briefly, an opaque screen was put in place and the apparatus was moved back.

The opaque screen (a piece of foamboard) was kept in place for the delay period. This opaque screen was used within each trial, between the sample presentation and the test phase (when the sample was paired with a new stimulus). During the delay between sample presentation and the test phase, the parent was encouraged to entertain the infant by drawing the infant's attention to the pictures on the screen or to other things in the room. During this time, the experimenter and an assistant detached the sample stimulus from the center drawer, attached it to the left or right drawer, attached a new stimulus to the other lateral drawer, and moved the pins so that only the drawer under the new stimulus could move.

Just before the end of the delay period, the experimenter asked the parent to close his or her eyes so that the parent could not influence the infant's choice of a stimulus. This instruction was repeated on every trial, and its importance was emphasized to the parent before testing and throughout the session. Then the screen was removed, the experimenter drew the infant's attention to the stimulus to the right and to the stimulus to the left (half the time that stimulus was the familiar sample, and half the time it was the nonmatching object), and then drew the infant's attention back to the center. The apparatus was moved forward and centered for the infant so that it was just barely within the infant's reach; thus the infant needed to strain to reach the apparatus and was therefore unlikely to reach for both stimuli at the same time.

If the infant reached for the new stimulus, a jack-in-the-box displaced behind that stimulus popped up the moment the infant pulled the stimulus. The experimenter cheered enthusiastically as well, as we have done in all our DNMS conditions (Diamond et al., 1994, 1999) except the no-reward condition (Diamond et al., 1994) and as we also did in the standard condition in the present experiment. The infant was allowed to look at and play with the jack-in-the-box birdie as much as desired (which was never more than 2–3 s). The infant was not allowed to play with the stimulus. When an infant chose incorrectly, the experimenter sounded very disappointed to try to make it clear to the infant that the response had been wrong. The experimenter showed the infant that pulling the other stimulus made the jack-in-the-box pop up, but the infant was not allowed to play with the jack-in-the-box or the stimulus. Immediately, the jack-in-the-box was pushed back down and the drawer closed. After each trial, either the apparatus was rolled back, away from the infant (jack-long condition), or the apparatus was rolled back, away from the infant, and the opaque screen was lowered to block the infant's view of it (jack-short condition). For both conditions, the time from when the sample stimulus was presented until it was removed from view was 5 s.

Different novel objects were used on every trial. All infants received the same stimuli in the same order and in the same locations. However, for half the participants within each Age × Gender cell, one member of each pair served as the sample, and the reverse was true for the other infants. For example, on Trial 1, a multicolored eraser case was always on the left and an orange Lego with a black and white target on top was always on the right. For half of the participants, the eraser served as the sample (hence the correct choice was on the right); for the other participants, the Lego served as the sample (the correct choice was on the left). A Gellermann (1933) series was used to pseudo-randomize the side of the correct choice (either right [R], left [L], R, R, L, L or L, R, L, L, R, R).

Each testing session began with delays of 5 s. Once an infant was correct on 5 consecutive trials at the 5-s delay, the delay was incremented to 30 s. Infants were given up to 25 trials to reach criterion at the 5-s delay. Only infants who passed that criterion, or who participated for a minimum of 15 trials without passing that criterion, were included in our analyses.

The number of trials that could be administered at 30 s was determined by how many trials it took an infant to pass criterion at 5 s (we had a total of 30 pairs of stimuli for testing) and the willingness of the infant to continue participating. Testing at the 30-s delay ideally continued until an infant passed criterion (5 correct trials in a row) but often ended sooner because all 30 pairs of stimuli had been administered or the infant refused to try any longer. Performance at the 30-s delay was tabulated only for infants who received at least 8 trials at the 30-s delay. No infant received more than 15 trials at the 30-s delay.

Standard Delayed Nonmatching to Sample Condition (Condition C)

Materials. The same stimulus pairs were used, but they were affixed to wooden bases and had no peg protruding from beneath the base. The testing table was 70 cm long × 65 cm wide × 70 cm high.

Each reward was hidden in a "well" (4 cm in radius and 1.6 cm deep) embedded in the center of the top surface of a wooden block

(7.3 × 7.3 × 3.5 cm). During the experiment, a reward was placed in the appropriate well, and the stimulus objects were placed on top of the wells, completely covering the wells. The reward in the well was < 1 cm from the stimulus. The rewards were marbles (which could be collected and rattled in a cup or rolled down a ramp), blinking rings, small balls, tiny rattles, pennies (which could be placed in a wind-up bank), or tiny plastic or wooden vehicles or animals (which could be collected or given a ride in a truck). For infants tested at the 30-s delay, either a maze was used to let the infant enjoy the marble reward or a large ball or puppet was brought out to entertain the infant during the delay. Participants could play interactively with these toys and delighted in doing so.

Procedure. The procedure was formally quite similar to that for the jack-in-the-box conditions (sample presentation at the midline was 5 s in duration [as in the jack-short condition], delay was initially 5 s and then was 30 s if the infant was correct on 5 trials in a row at 5 s, and in the test phase, the sample was presented to one side, and the new stimulus was presented to the other side), as has been described elsewhere (Diamond et al., 1994, 1999).

The infant sat on the parent's lap, opposite the experimenter. An assistant sat next to the experimenter. The stimuli were presented to the infant on the testing table. Before each trial, out of view of the infant, the experimenter hid a reward in each of the two wells. The object serving as the sample on that trial, with its baited well underneath it, was then positioned at the midline at the rear of the testing table. At the beginning of each trial, the opaque screen was removed and the experimenter pushed the sample object atop its well forward toward the infant, encouraging the infant to reach and retrieve the reward. The infant was allowed to examine the stimulus for only a brief period. The infant's total exposure time to a sample stimulus was 5 s; the infant held and examined the stimulus for 2–3 s of that 5-s period.

After the infant displaced the sample, he or she had to then reach into the well and retrieve the reward. The sample object and the well were then removed, the opaque screen was lowered, and a delay was imposed. During the delay, the assistant held a white foamboard barrier between the infant and experimenter so that the infant's view of the testing surface was obscured. The parent was encouraged to entertain the infant by drawing the infant's attention to the pictures on the foamboard or to other things in the room.

At the end of the delay period, any objects on the table or in the infant's hands were removed. The experimenter instructed the parent to close his or her eyes. Two stimulus objects (the sample and a new object) atop their respective wells were first positioned at the midline, at the rear of the table, and then the opaque screen was removed. As in the jack-in-the-box condition, the experimenter drew the infant's attention first to the stimulus to the right and then to the stimulus to the left, not pushing the stimuli forward until the infant had clearly seen both. (Because the left–right location of the novel stimulus was randomly varied over trials, half the time the infant's attention was drawn to it first, and half the time the infant's attention was drawn to the familiar stimulus first.) The stimuli, atop their own wooden bases, were then pushed diagonally forward at a constant rate, one to the left and one to the right (7.5 cm from the midline), so that they were equidistant from the infant and just barely within reach. The stimuli were kept at that distance to discourage the infant from reaching simultaneously for both objects.

As in the jack-in-the-box condition, a correct response was defined as choosing (displacing) the new stimulus object, the one that did not match the sample presented during familiarization. The reward for doing so was the opportunity to discover and play with the hidden reward and praise and applause from the experimenter. The other stimulus was immediately removed. Displacing the nonmatching stimulus object did not automatically produce the hidden reward or make it visible. The infant had to look in the well and reach in to retrieve the reward. Trials on which infants made incorrect responses were not rewarded. The experimenter explained in a sad and disappointed voice that the infant was incorrect, and showed the infant that the other stimulus had been the correct choice but did not allow the infant to play with the reward object that was thus revealed.

The same procedures concerning length of delays and number of trials were used for both the jack-in-the-box and standard conditions. Infants and their parents were never told the principle determining which response was correct until after the experiment was over. The idea was to see if infants could use the information on reward contingency provided on each trial to deduce the rule and, once they had done so, to see if they could remember which stimulus had been the sample after a 30-s delay. Different novel objects were used on every trial. All infants received the same stimuli in the same order in the same locations. Which stimulus of each pair was the sample stimulus was counterbalanced within each Age × Gender cell, as were the left–right locations of the sample and novel stimuli. All sessions were videotaped, which allowed detailed analysis, including verification of the length of each sample presentation and delay period.

The dependent measures were whether an infant passed criterion at the 5-s delay (criterion was 5 consecutively correct trials), percentage of correct responses at the 5-s delay, percentage of correct responses on the last 5 trials at the 5-s delay (infants had to deduce the correct rule on the basis of feedback on each trial within their single testing session; hence one would expect their performance at the outset to be at the level of chance and their performance later in the session to be more informative of whether they had learned the rule to always select the nonmatching stimulus), number of trials needed to pass criterion, percentage of correct responses at the 30-s delay, and percentage of correct responses on the last 5 trials at the 30-s delay. The between-subjects independent variables were condition (3 levels: jack-long, jack-short, and standard), age (3 levels: 9, 12, and 15 months), and gender (2 levels: male and female). The within-subject variable was length of delay (2 levels: 5 s and 30 s).

Because of the frequent occurrence of nonnormal distributions and unequal variances across groups, nonparametric statistics were employed using StatXact (Mehta & Patel, 2000). For all dependent measures (except for whether or not the infant passed criterion), when comparisons involved independent variables with three levels (condition and age), the Kruskal–Wallis test (Siegel & Castellan, 1988) was performed. If the Kruskal–Wallis test was significant, Wilcoxon Mann–Whitney tests (Wilcoxon, 1945), with Bonferroni corrections, were conducted as planned comparisons. For the independent variable of gender, which has only two levels, the Wilcoxon Mann–Whitney test was used.

For analyses of the dichotomous outcome of whether or not an infant passed criterion at the 5-s delay, logistic regression was used (Hosmer & Lemeshow, 2000). Nested models were fit in which the main effects (condition, age, and gender) and all their interaction terms were entered into the analysis at the start and then removed individually so that likelihood ratio chi-square tests could be used to evaluate the significance of each term while controlling for the other terms. To examine differences by condition in the percentage of infants passing criterion at the 5-s delay when the value for one condition was 100% or 0%, we used Fisher's exact test with Bonferroni correction.

To investigate within-subject differences in performance at the 5- and 30-s delays, on the advice of Stephen G. Baker of the Biostatistics Department of the University of Massachusetts Medical School (previously of the Biometry Branch, National Cancer Institute), we performed mixed within-subject analyses of variance (ANOVAs) with condition, age, and gender as between-subjects variables. To investigate whether a significant difference existed between performance at the 5- and 30-s delays within a specific condition and/or age, we used Wilcoxon signed ranks tests.

Participants

A total of 118 infants (59 boys and 59 girls) provided usable sessions in Study 1. These included 42 infants (21 boys and 21 girls) in each of the two jack-in-the-box conditions (14 infants at each age [9, 12, and 15 months] in each of the jack-in-the-box conditions). In the standard condition, 34 infants (17 boys and 17 girls) provided usable sessions (6 [3 boys and 3 girls] at 9 months and 14 [7 boys and 7 girls] each at 12 and 15 months). Thus, there were 14 infants in each condition at each age except in the standard condition at 9 months, in which there were only 6 infants. There were equal numbers of boys and girls in every condition and at every age except for the jack-long condition (in which 8 boys and 6 girls were tested at 9 months and 6 boys and 8 girls were tested at 15 months). Information on the mean age in weeks, range of ages, and demographic characteristics of the infants in each condition at each age is provided in Table 2.

All participants were healthy and full-term. Most were from middle-class backgrounds and of European Caucasian descent. Across conditions, the mean age of fathers at the infant's birth ranged from 30 to 33 years, the mean age of mothers at the infant's birth ranged from 28 to 32 years, the mean years of the fathers' education ranged from 15 to 17 years, the mean years of the mothers' education ranged from 15 to 16 years, the percentage of mothers working ranged from 58% to 75%, and the percentage of infants with at least one sibling ranged from 50% to 67%. All infants received a toy present valued at approximately $5, and all parents received reimbursement for travel and parking expenses. Informed, written consent was obtained from a parent before each infant was tested.

In addition to these 118 infants, another 68 infants came into the laboratory but could not be used in the study (see Table 3). The main reason for unusable sessions in the standard condition was that the condition was too difficult. Indeed, many more 9-month-old infants (65%) than 12-month-old infants (36%) refused to continue participating in the standard condition. In our previous work, we obtained similar results and found that the incidence of unusable sessions for the DNMS task decreased with age until 18–21 months, an age at which the task was clearly within the infants' ability (Diamond et al., 1994, 1999).

Many 9-month-olds who we tried to test in the standard condition grew too frustrated and refused to continue. For that reason, we settled for only 6 usable sessions for 9-month-old infants. It was clear that 9-month-old infants could not succeed in that condition, and it seemed pointless to frustrate still more of them. For other infants in the standard condition, the most common reason they refused to continue to participate was frustration at our removing the rewards before they were ready to relinquish them.

The main reason for unusable sessions in the jack-in-the-box condition was boredom with the reward. Unlike the standard condition, in which we could vary the reward to maintain interest and choose rewards preferred by the particular infant being tested, in the jack-in-the-box condition we were constrained to use the same reward on each trial. For all 15–30 trials, the same pink and blue stuffed penguin popped up; indeed, it popped up twice on each trial (after the sample presentation and after the paired presentation). Even though the experimenter placed one adornment or another (such as elastic necklaces, buttons, hats, and funny noses) on the stuffed animal on various trials, for many infants that was not sufficient.

Previous work had shown that when the stimulus itself is the reward, even infants as young as 6 months can succeed at long delays (Diamond, 1995). Therefore, any sessions in any condition in which infants showed marked interest in the stimuli and little or no interest in the rewards were considered unusable in the present study. We tried to prevent the stimulus objects themselves from serving as the reward in the present experiment by removing them quickly after one was displaced and by directing the infant's attention to the reward.

Table 2
Characteristics of the Infants Tested in Study 1 by Age and Condition

	9-month-olds			12-month-olds			15-month-olds		
Characteristic	Jack-in-the-box long presentation	Standard condition	Jack-in-the-box short presentation	Jack-in-the-box long presentation	Standard condition	Jack-in-the-box short presentation	Jack-in-the-box long presentation	Standard condition	Jack-in-the-box short presentation
n	14	6	14	14	14	14	14	14	14
Mean age (in weeks)	40.4	39.3	40.3	53.7	53.3	53.1	66.3	65.7	66.2
Age range (in weeks)	39.3–42.1	38.3–40.3	38.9–41.4	52.4–55.7	51.7–54.6	51.1–54.4	64.4–67.6	62.9–68.8	64.7–67.3
Fathers' mean age (in years) at infant's birth	33.25			32.75		31	31.92		32.24
Mothers' mean age (in years) at infant's birth	31.92			30.75		28.77	29.08		29.65
Fathers' mean years of education	17.00			16.04		15.25	15.67		16.1
Mothers' mean years of education	16.08			15.46		14.86	15.25		15.5
Mean number of siblings	1.25			0.75		0.63	0.67		0.42
Percentage of infants who had no siblings	33			50		50	58		50
Percentage of mothers									
Not working	42			25		33	33		42
Working part time	42			67		17	25		50
Working full time	16			8		50	42		8

Table 3
Reasons Sessions in Study 1 Were Not Usable by Age, Sex, and Condition

Age and sex	Experimenter error	Equipment failure	Afraid of puppet	Too tired, cranky, or sick	Lost interest in the task	Always reached to same side	Other	Total
Jack-in-the-box long presentation								
9-month-olds								
Boys	1	0	1	1	0	0	0	3
Girls	0	0	0	0	2	2	0	4
Subtotal	1	0	1	1	2	2	0	7
12-month-olds								
Boys	1	0	0	0	2	1	0	4
Girls	0	0	0	1	3	0	0	4
Subtotal	1	0	0	1	5	1	0	8
15-month-olds								
Boys	0	1	0	2	1	0	0	4
Girls	0	0	0	0	3	0	0	3
Subtotal	0	1	0	2	4	0	0	7
Total								
Boys	2	1	1	3	3	1	0	11
Girls	0	0	0	1	8	2	0	11
Total	2	1	1	4	11	3	0	22
Jack-in-the-box short presentation								
9-month-olds								
Boys	1	0	0	1	3	0	0	5
Girls	0	0	0	1	2	0	0	3
Subtotal	1	0	0	2	5	0	0	8
12-month-olds								
Boys	0	0	0	0	2	0	1	3
Girls	0	0	1	1	2	0	0	4
Subtotal	0	0	1	1	4	0	1	7
15-month-olds								
Boys	0	0	0	0	3	1	1	5
Girls	0	0	0	1	1	0	0	2
Subtotal	0	0	0	1	4	1	1	7
Total								
Boys	1	0	0	1	8	1	2	13
Girls	0	0	1	3	5	0	0	9
Total	1	0	1	4	13	1	2	22
Standard condition								
9-month-olds								
Boys	0			1	5	0	0	6
Girls	0			1	4	0	0	5
Subtotal	0			2	9	0	0	11
12-month-olds								
Boys	0			1	3	1	0	5
Girls	1			0	2	1	0	3
Subtotal	1			1	5	2	0	8
15-month-olds								
Boys	0			0	2	0	0	2
Girls	0			0	2	1	0	3
Subtotal	0			0	4	1	0	5
Total								
Boys	0			2	10	1	0	13
Girls	1			1	8	2	0	11
Total	1			3	18	3	0	24

Another common reason for a session's being unusable was that the infant had missed a nap or was getting a cold and was therefore too cranky or irritable. Other reasons for difficulty in getting infants to participate were that they were more interested in interacting with people than in our stimuli or rewards, were afraid of the jack-in-the box, or did not like sitting on a parent's lap; all of these were rare occurrences. All of the above reasons resulted in sessions with too few trials (because the infant refused to participate further) for them to be considered usable. Other reasons for unusable sessions, even though the sessions were completed, were experimenter error, equipment failure, and an extremely strong side bias or hand preference that we were unable to correct for by centering the presentation for the preferred hand; all of these were also rare occurrences.

Results

Degrees of freedom are not reported below because degrees of freedom are parameters for the probability distributions that *non-exact* methods rely on; exact tests (such as those used here) do not use degrees of freedom. There were no significant sex differences or interactions of sex with age. Girls tended to reach criterion in the jack-long condition in fewer trials than did boys, but that difference was not significant.

Effect of Condition

Infants performed significantly better in each of the jack-in-the-box conditions than in the standard condition. This superiority of performance in the jack-in-the-box conditions was clear on all dependent measures at all ages (see Tables 4 and 5) with the single exception that the superiority in performance in percentage of correct responses at the 30-s delay in the jack-in-the-box conditions was not statistically significant among 15-month-olds. For example, all 12- and 15-month-old infants passed criterion for successful performance at the 5-s delay in the jack-long condition, as did all 15-month-old infants in the jack-short condition; but only 36% of 12-month-olds and 43% of 15-month-olds were able to succeed at the 5-s delay in the standard condition (see Figure 2). The standard DNMS condition was so difficult for 9-month-olds that none passed criterion at that age, whereas by 15 months of age, the jack-in-the-box conditions were so easy for the infants that all passed criterion.

There were no significant differences between performance in the jack-long and jack-short conditions on any dependent measure at any age or with all ages combined. Although on every dependent measure, 9- and 12-month-old infants performed marginally better in the jack-long condition than in the jack-short condition (see Table 4), even when these two age groups were combined in the analyses, no significant difference emerged between performance in the two jack-in-the-box conditions on any dependent measure. By 15 months of age, even this marginal effect of the duration of sample presentation (long presentation vs. short) was gone.

Boys showed no difference in performance in the two jack-in-the-box conditions. However, girls performed significantly better in the jack-long condition than in the jack-short condition on two of the dependent measures (percentage correct at 5 s, $p < .02$; number of trials needed to pass criterion at the 5-s delay, $p < .003$).

Age Effects

Over the age span investigated (9–15 months), there was no statistically significant evidence of superior performance by older versus younger infants in the standard condition on any dependent measure except one: Among infants who passed the standard condition at 5 s and went on to testing at 30 s, 15-month-olds were correct on more trials at the 30-s delay than were 12-month-olds ($p < .04$). (No 9-month-old infants passed criterion in the standard condition at the 5-s delay; hence comparisons of performance at the 30-s delay could not be made between 9- and 12-month-old infants or between 9- and 15-month-old infants.) Differences in performance in the standard condition at the 5-s delay between 9-month-old infants and older infants may have failed to reach significance because so few 9-month-olds could be tested in the standard condition. Infants throughout the 9–15-month age range found the standard DNMS task very difficult; few succeeded (0% and 36% passing criterion at 9 and 12 months, respectively), and performance was generally still quite poor even by 15 months of age (43% passing criterion).

In the jack-long condition, 12-month-old infants performed better than 9-month-old infants when the delay was 5 s (trials to criterion, $p < .006$; percentage of correct responses at the 5-s delay, $p < .02$). There were no significant differences between the performance of 12- and 15-month-old infants in the jack-long condition; by 12 months, infants were already approaching ceiling on the task.

As was true for the jack-long condition, 12-month-old infants performed significantly better than 9-month-old infants in the jack-short condition (trials to criterion, $p < .004$; percentage of correct responses at the 5-s delay, $p < .0003$). There were no significant differences between the performances of 12- and 15-month-old infants in the jack-short condition; the performance of both age groups in this condition was excellent.

Effect of Delay

Readers are reminded that the percentages given in Table 4 at the 5-s delay are for all infants tested, but when within-subject comparisons were made between performances at delays of 5 and 30 s, only those infants who were tested at 30 s were included in the analyses. Infants tested in either the jack-long or jack-short condition showed no significant difference in overall percentages of correct responses at the 5- and 30-s delays, but they did perform better on the last 5 trials at the 5-s delay than on the last 5 trials at the 30-s delay ($p < .002$ for each condition).

There was more of an effect of delay in the standard condition. Far fewer infants were able to succeed in the standard condition even at the brief 5-s delay. Those few who were able to pass criterion at 5 s could not sustain that level of performance when the delay increased to 30 s (percentage correct at 5 s vs. at 30 s, $p < .004$; percentage correct on last 5 trials at 5 s vs. 30 s, $p < .004$). For example, the 5 infants of 15 months who received at least 8 trials at 30 s in the standard condition were correct on 100% of the last 5 trials at 5 s but on only 72% of the last 5 trials at 30 s.

The effect of condition on the difference between performance at the 5- and 30-s delays was significant, as assessed by percentage of correct responses at each delay ($p < .001$) and percentage of correct responses on the last 5 trials at each delay ($p < .001$). That overall difference across conditions was significant because the effect of delay was significantly greater in the standard condition than in the jack-in-the-box conditions. The difference between performance at the 5- and 30-s delays in the standard DNMS condition was significantly greater than that difference in the jack-long condition ($p < .002$) or in the jack-short condition ($p < .003$). The difference between performance at the 5- and 30-s delays in the jack-short condition was not significantly greater on any dependent measure than that difference in the jack-long condition.

Table 4
Study 1: Mean Performance on Each Dependent Measure in Each Condition

	5-s delay					30-s delay[c]	
	Mean number of trials to criterion[a]						
Age and sex	All participants[b]	Only those who passed criterion	% passing criterion	% correct	% correct on the last 5 trials	% correct	% correct on the last 5 trials
Jack-in-the-box long presentation (Condition A)							
9-month-olds							
Boys	11.9 (*n* = 8)	10.0 (*n* = 7)	87.5	67.4	95.0	90.0 (*n* = 3)	93.3 (*n* = 3)
Girls	8.0 (*n* = 5)	4.6 (*n* = 5)	83.3	81.5	86.7	80.8 (*n* = 4)	80.0 (*n* = 4)
All	10.2 (*n* = 14)	7.8 (*n* = 12)	85.7	73.4	91.4	84.6 (*n* = 7)	85.7 (*n* = 7)
12-month-olds							
Boys	2.7 (*n* = 7)	2.7 (*n* = 7)	100	88.3	100	91.9 (*n* = 7)	97.1 (*n* = 7)
Girls	2.7 (*n* = 7)	2.7 (*n* = 7)	100	85.4	100	80.0 (*n* = 6)	90.0 (*n* = 6)
All	2.7 (*n* = 14)	2.7 (*n* = 14)	100	86.9	100	86.4 (*n* = 13)	93.8 (*n* = 13)
15-month-olds							
Boys	4.7 (*n* = 6)	4.7 (*n* = 6)	100	80.0	100	88.6 (*n* = 5)	92.0 (*n* = 5)
Girls	0.9 (*n* = 8)	0.9 (*n* = 8)	100	92.0	100	89.0 (*n* = 7)	94.3 (*n* = 7)
All	2.5 (*n* = 14)	2.5 (*n* = 14)	100	86.8	100	88.8 (*n* = 12)	93.3 (*n* = 12)
Jack-in-the-box short presentation (Condition B)							
9-month-olds							
Boys	16.0 (*n* = 7)	9.3 (*n* = 4)	57.1	69.1	82.9		
Girls	9.3 (*n* = 7)	6.7 (*n* = 6)	85.7	66.6	94.3		
All	12.6 (*n* = 14)	7.7 (*n* = 10)	71.4	67.9	88.6		
12-month-olds							
Boys	5.6 (*n* = 7)	2.3 (*n* = 6)	85.7	82.3	91.4	86.7 (*n* = 6)	90.0 (*n* = 6)
Girls	4.3 (*n* = 7)	4.3 (*n* = 7)	100	79.6	100	79.6 (*n* = 7)	82.9 (*n* = 7)
All	4.9 (*n* = 14)	3.4 (*n* = 13)	92.9	80.1	95.7	82.8 (*n* = 13)	86.2 (*n* = 13)
15-month-olds							
Boys	2.9 (*n* = 7)	2.9 (*n* = 7)	100	88.0	100	90.0 (*n* = 7)	94.3 (*n* = 7)
Girls	3.3 (*n* = 7)	3.3 (*n* = 7)	100	85.6	100	89.4 (*n* = 7)	96.4 (*n* = 7)
All	3.1 (*n* = 14)	3.1 (*n* = 14)	100	86.8	100	89.7 (*n* = 14)	95.4 (*n* = 14)
Standard condition (Condition C)							
9-month-olds							
Boys	25.0 (*n* = 3)		0.0	48.0	46.7		
Girls	25.0 (*n* = 3)		0.0	51.3	53.3		
All	25.0 (*n* = 0)		0.0	50.0	50.0		
12-month-olds							
Boys	19.4 (*n* = 7)	5.5 (*n* = 2)	28.6	58.1	65.7	50.0 (*n* = 2)	50.0 (*n* = 2)
Girls	16.7 (*n* = 7)	5.7 (*n* = 3)	42.9	62.1	74.3	55.0 (*n* = 2)	50.0 (*n* = 2)
All	18.1 (*n* = 14)	5.6 (*n* = 5)	35.7	60.1	70.0	52.5 (*n* = 4)	50.0 (*n* = 4)
15-month-olds							
Boys	15.3 (*n* = 7)	2.3 (*n* = 3)	42.9	64.6	71.4	72.3 (*n* = 3)	66.7 (*n* = 3)
Girls	17.7 (*n* = 7)	8.0 (*n* = 3)	42.9	60.3	71.4	79.0 (*n* = 2)	80.0 (*n* = 2)
All	16.5 (*n* = 14)	5.2 (*n* = 6)	42.9	62.4	71.4	75.0 (*n* = 5)	72.0 (*n* = 5)

[a] The number of trials to criterion was calculated as is done in neuroscience studies; that is, it equals the number of trials administered at the 5-s delay minus the string of 5 consecutively correct trials used to satisfy the criterion. [b] Those who failed criterion were assigned a score of 25 here (the maximum number of trials administered at the 5-s delay). [c] Only infants who passed the 5-s delay were tested at the 30-s delay. Infants who refused to perform the task any longer after the testing at 5 s, who took more than 18 trials to pass criterion at the 5-s delay, or who received fewer than 8 trials at the 30-s delay were also excluded from analyses at the 30-s delay.

Discussion

Infants 9–12 months of age learned the rule to reach to the nonmatching stimulus and succeeded on the DNMS task with a jack-in-the-box reward even when the sample was presented as briefly as is done in the standard DNMS task. Importantly, the spatial separation of 12.5 cm between the stimulus and the reward in the jack-in-the-box conditions did not preclude infants from learning the nonmatching rule. Hence, the reason that infants of 9, 12, and 15 months fail the standard DNMS task is not because of the brief sample presentation time (consistent with the results of Diamond et al., 1994), nor is it because the stimulus and reward are not spatially contiguous.

Moreover, infants 9–12 months of age performed at least as well in the jack-in-the-box conditions here (12.5-cm spatial separation

Table 5
Results of Statistical Analyses Comparing the Three Conditions Tested in Study 1

Performance	Jack-short vs. Jack-long vs. Standard[a]	Jack-short vs. Standard[b]	Jack-long vs. Standard[b]
All ages combined			
At delays of 5 s			
No. of trials to pass criterion	$p < .0001$	$p < .0001$	$p < .0001$
% passing criterion	$p < .001$	$p < .001$	$p < .001$
% correct	$p < .0001$	$p < .0001$	$p < .0001$
% correct on last 5 trials	$p < .0001$	$p < .0001$	$p < .0001$
At delays of 30 s			
% correct	$p < .002$	$p < .002$	$p < .001$
% correct on last 5 trials	$p < .0004$	$p < .0004$	$p < .0002$
9-month-old infants			
At delays of 5 s			
No. of trials to pass criterion	$p < .004$	$p < .01$	$p < .001$
% passing criterion	—[c]	$p < .03$	$p < .004$
% correct	$p < .002$	$p < .0003$	$p < .001$
% correct on last 5 trials	$p < .0002$	$p < .001$	$p < .001$
At delays of 30 s			
% correct			
% correct on last 5 trials			
12-month-old infants			
At delays of 5 s			
No. of trials to pass criterion	$p < .0001$	$p < .0003$	$p < .0001$
% passing criterion	$p < .03$	$p < .01$	*ns*
% correct	$p < .0001$	$p < .0001$	$p < .0001$
% correct on last 5 trials	$p < .0002$	$p < .003$	$p < .0004$
At delays of 30 s			
% correct	$p < .03$	$p < .02$	$p < .003$
% correct on last 5 trials	$p < .002$	$p < .02$	$p < .003$
15-month-old infants			
At delays of 5 s			
No. of trials to pass criterion	$p < .001$	$p < .003$	$p < .0004$
% passing criterion	—[c]	$p < .01$	$p < .01$
% correct	$p < .001$	$p < .001$	$p < .001$
% correct on last 5 trials	$p < .0002$	$p < .002$	$p < .002$
At delays of 30 s			
% correct	*ns*		
% correct on last 5 trials	$p < .03$	$p < .02$	$p < .03$

[a] For no. of trials to pass criterion, % correct, and % correct on last 5 trials, the *p* values are from Kruskal–Wallis tests. For % passing criterion, Wald test *p* values from logistic regression are presented. [b] For no. of trials to pass criterion, % correct, and % correct on last 5 trials, the *p* values are from Wilcoxon Mann–Whitney tests. For % passing criterion, Wald test *p* values from logistic regression are presented. [c] Logistic regression was unable to handle analyses in which one value was 100% or 0%. Pairwise comparisons in the next two columns were done using Fisher's exact test with Bonferroni correction.

between the stimulus and reward) as in the Velcro condition of Diamond et al. (1999; 0-cm separation, stimulus and reward contiguous and attached). Thus, infants' success in the Velcro condition was probably not due to the spatial contiguity of the stimulus and reward. The present results show that even infants in the 1st year of life can succeed if the stimulus and reward are spatially separated.

Presentation time may not be completely irrelevant for 9- and 12-month-old infants because they performed marginally better in the jack-long condition (with sample presentation times of 20–25 s) than they did in the jack-short condition (with sample presentation times of only 5 s). These results are consistent with those of Diamond et al. (1994, Study 4), who compared long and short stimulus presentation times (20 s vs. 2–5 s) in the standard DNMS task. As we did here, they found that performance in the long presentation condition was not significantly better than that in the short presentation condition, although performance on every dependent measure was marginally better when infants were given more time to encode the sample stimulus.

Study 1 demonstrates that the reward does not need to be contiguous with the stimulus, as it was in the Velcro condition (Diamond et al., 1999), for infants 9–15 months of age to succeed.

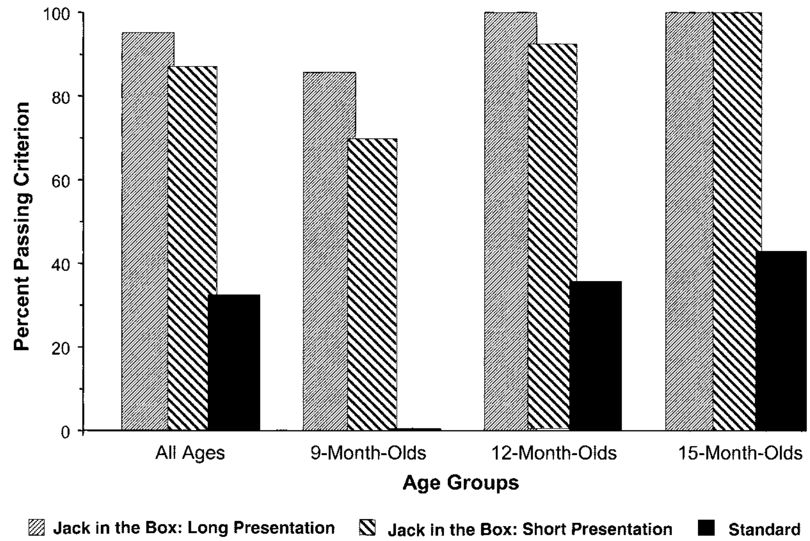

Figure 2. Percentage of infants passing criterion with delays of 5 s between sample presentation and test in each of the three conditions in Study 1.

However, Study 1 does not allow us to determine whether temporal proximity or physical connectedness, or their conjunction, is the critical factor, because those two factors covaried in Study 1. In the jack-in-the-box conditions, the stimulus and reward were temporally proximal and physically connected; in the standard condition, they were neither.

There were also two other differences between the jack-in-the-box conditions and the standard condition. One difference was in the rewards used. In the jack-in-the-box conditions, the reward was a cute animated event (pop-up birdie), and the same reward was used on all trials. In the standard condition, the rewards were small objects with which the infant could make noise but that were taken from the infant after only a minimal play period. Perhaps infants performed better in the jack-in-the-box conditions because the reward used there was a better motivator. We think not, because after infants had seen the pop-up birdie a few times, it became “old” and boring to them, especially because they could not remove it or hold it. Infants appeared to much prefer the varied rewards in the standard condition to the unchanging reward in the jack-in-the-box conditions. Of course, it is possible that the frustration of having the reward removed so quickly each time in the standard condition impeded infants’ performance there or that infants performed better when they were less interested in the reward (hence performing better in the jack-in-the-box conditions).

Another difference between the jack-in-the-box and standard conditions was that the act of displacing the stimulus caused the reward to appear in the jack-in-the-box conditions (one simple action required), whereas in the standard condition a two-action means–end sequence was required (displace the stimulus, reach for the reward). Perhaps that additional complexity in the standard condition was the critical difference between the conditions in which infants succeeded and those in which they failed. We think not, because infants also had to execute two successive reaches (one to the stimulus, one to the reward) in the Velcro condition (Diamond et al., 1999), and infants succeeded there. Hence the reason infants fail the standard DNMS condition and succeed in the jack-in-the-box conditions is unlikely to be because of the two reaches required in the standard condition, because the Velcro condition required two reaches as well and infants succeeded there.

Finally, infants who touched the correct stimulus in either jack-in-the-box condition saw the birdie pop up and also received enthusiastic cheers from the experimenter. Perhaps the verbal reward in conjunction with the jack-in-the-box reward was critical. We thought this was unlikely because the same enthusiastic verbal reward was given in all conditions, standard as well as jack-in-the-box. Seemingly, therefore, the presence of verbal reward could not account for the differences in performance observed in the different conditions. Indeed, if the jack-in-the-box conditions had not included enthusiastic cheers from the experimenter, that would have been a difference between them and the standard condition. However, our previous work has also shown that when given a verbal reward alone, 9- and 12-month-old infants perform well on the DNMS task (Diamond et al., 1999). Perhaps infants focused on the small-object rewards in the standard condition, ignoring experimenter feedback, but were able to make use of experimenter feedback in the jack-in-the-box conditions.

Study 2

In Study 2, five conditions were tested, which enabled us to (a) disambiguate temporal proximity and physical connectedness and (b) explore each of the alternative explanations and possible confounds mentioned above. All jack-in-the-box conditions in Study 2 used the same stimulus presentation time as in the standard DNMS procedure and as in the jack-short condition of Study 1. Because in Study 1 we found minimal age differences and because age was not central to the questions under exploration, only infants of the middle age (12 months) were tested in Study 2.

To start to disambiguate temporal proximity and physical connectedness, in Study 2 we used jack-in-the-box conditions in which the reward did not appear immediately. These delayed-reward conditions used the same reward (a jack-in-the-box), the same apparatus, and the same testing conditions as were used in the immediate-reward jack-in-the-box conditions in Study 1. Because we had predicted that close temporal proximity was critical, we predicted that infants would not be able to succeed in a jack-in-the-box condition if the close temporal proximity was broken. We ran a pilot jack-in-the-box condition with a 2-s gap between pulling on the stimulus and the jack-in-the-box popping up. To our surprise, infants succeeded. We noticed, however, that during the 2 s before the jack-in-the-box popped up, infants had a great time moving the stimulus and its base back and forth in its track. Thus, although the jack-in-the-box reward was delayed, infants were getting a reward of sorts by being able to play with the movement of the stimulus. Therefore, we revised the testing procedure for the jack-in-the-box 2-s gap condition, adding a locking mechanism inside the apparatus (not visible from the outside) that could be set so that when the base of the stimulus was moved forward, it locked in place (Condition D). Again, to our surprise, the infants succeeded. Another jack-in-the-box delayed-reward condition was run with a 5-s gap between acting on the stimulus and receipt of the reward, with the locking mechanism preventing the stimulus or its base from moving back and forth (Condition E).

To explore the effect of experimenter cheering on performance in the jack-in-the-box conditions, we reran the jack-short procedure of Study 1 without any experimenter feedback (no applause or cheering at correct responses and no sounding disappointed at incorrect choices; Condition F). We predicted that performance here would be comparable to what we had found in the jack-short condition of Study 1 and significantly better than performance in the standard DNMS condition, because we did not think experimenter feedback accounted for any of the effects observed in Study 1. After all, the very same experimenter feedback had been used in all three conditions in Study 1, yet performance was significantly worse in the standard condition.

To control for many of the differences between the jack-in-the-box conditions of Study 1 and the standard DNMS procedure, and to implement a condition of close temporal proximity but with a lack of physical connectedness, we used jack-in-the-box rewards in the standard DNMS procedure (Condition G). We reasoned that infants should succeed here if their strikingly better performance in the jack-in-the-box conditions of Study 1 compared with their performance on the standard DNMS task was due to (a) temporal proximity being key, (b) infants preferring the jack-in-the-box rewards to the little objects used as rewards in standard DNMS testing, or (c) the lack of a means–end requirement in the jack-in-the-box conditions. However, if the key variable was physical connectedness, then infants should fail here, as they did with the standard DNMS procedure, even though jack-in-the-box rewards were used.

To match the jack-in-the-box testing to the standard DNMS condition even more closely, we also ran a condition with a jack-in-the-box reward using the standard DNMS procedure with a 2-s gap (vs. no gap at all) between acting on the stimulus and receipt of the reward (Condition H). In the standard DNMS procedure there is about a 2-s gap between the infant's picking up a stimulus and noticing the reward in the well.

Thus, five jack-in-the-box conditions were run in Study 2 (Conditions D through H; see Table 1). Three used procedures very similar to the procedure of the jack-short condition in Study 1 (but with temporal gaps inserted between acting on the stimulus and receiving a reward or with no experimenter feedback). Two conditions used procedures closely resembling those of the standard DNMS procedure but with jack-in-the-box rewards.

Method

Materials

All five conditions used the same pop-up jack-in-the-box rewards. By the time of Study 2, the blue penguins used in the jack-in-the-box conditions of Study 1 were faded and haggard-looking. In all conditions in Study 2, identical brown bears served as the pop-up rewards. These stuffed, soft brown bears were 11.63 cm in height, 7.13 cm wide, and 7.25 cm from front to back. For the no-cheer condition, the materials were the same as those for the jack-in-the-box conditions in Study 1 in all other respects.

For the two delayed-reward conditions, the materials were also the same except for the following: (a) One of three identical, clear Plexiglas bars 22 cm long, 4.5 cm wide, and 0.5 cm thick sat behind each well, extending from the back of the apparatus to the rear of each jack-in-the-box lid. This bar over the rear of the lid kept the jack-in-the-box from popping up when the infant displaced the stimulus. The experimenter could simply pull the bar back to allow the jack-in-the-box to appear. (b) A mechanism was inserted inside the apparatus (invisible from outside) that could be set so that when the base of a stimulus was moved forward, it locked in place. This mechanism consisted of a spring-loaded pin that extended from underneath the base of the top surface of the apparatus to the movable base (or drawer) onto which the stimulus was affixed (one pin per base, three total for the apparatus). When a drawer was moved forward, the spring automatically forced the pin up into the locking position. Thus, a drawer could be pulled forward, but then it was locked in that position, which prevented it or the stimulus from moving at all.

For the two conditions with the standard DNMS procedure that used jack-in-the-box rewards, the large white apparatus housing all jack-in-the-boxes and to which all stimuli were attached was gone. Instead, the teddy bear jack-in-the-box rewards were contained in independent, identical, traditional jack-in-the-box boxes (13.44 cm high $\times$ 13.44 cm wide $\times$ 13.75 cm deep). The exterior of each box was covered with plain sky blue paper.

Procedure

For the three conditions in which the large white apparatus was used, the procedures were the same as that for the jack-short condition of Study 1 except for the following:

In the condition without experimenter feedback, the experimenter and the assistant were cheerful and pleasant throughout but did not change intonation, facial expression, or volume in response to the infant's performance.

In the 2-s delayed-reward condition, the clear Plexiglas bars stayed in place for a period of 2 s after the infant pulled on a stimulus. The locking device prevented the drawer or the stimulus from moving after an infant had pulled it forward. Thus, the infant received no reward during the 2-s period before the jack-in-the-box appeared. Then the experimenter pulled the clear Plexiglas bar back from behind the stimulus the infant had chosen, allowing the jack-in-the-box behind that stimulus to pop up.

The 5-s delayed-reward condition was the same except that the temporal gap between pulling on the stimulus and appearance of the jack-in-the-box was 5 s.

For the two conditions with separate and identical boxes for each of the jack-in-the-boxes, the stimuli were not attached to anything. For sample presentation, when the screen was lifted, an infant saw an independent stimulus adjacent to, and directly in front of, a closed jack-in-the-box box at the rear of the table. The experimenter moved the stimulus and box forward as a unit. In the immediate-reward condition, as soon as the infant touched the stimulus, the teddy bear popped up. In the delayed-reward condition, there was a 2-s delay before the lid was released. In both conditions, the experimenter and assistant cheered enthusiastically, as they did in all conditions except the no-cheer condition noted above. When the screen was lifted after the delay period between sample presentation and test, the infant saw two stimuli at the midline, each adjacent to, and directly in front of, a closed jack-in-the-box box at the rear of the table. As in the standard DNMS procedure, the experimenter made sure the infant saw both stimuli and then moved both diagonally forward so that they were to the infant's left and right, just barely within reach. Each jack-in-the-box box was moved with each stimulus as a unit. In the immediate-reward condition, as soon as an infant chose a stimulus, if that choice was correct, the teddy bear popped up. In the delayed-reward condition, there was a 2-s delay before the lid was released (see Figure 3). When the infant's choice was incorrect, the lid did not open. The experimenter called the infant's attention to the other stimulus and showed the infant that that lid would have opened and the teddy bear would have popped up had the infant chosen that other stimulus.

Analyses for Study 2 were conducted with the same statistical procedures described for Study 1. In Study 2, sample sizes were generally too small at the 30-s delay to allow statistical comparisons between performances at the 5- and 30-s delays. For the only condition in which sample sizes were large enough for that comparison (Condition F), the Wilcoxon signed ranks test (Sprent, 1993) was used. For analyses comparing performance at the 30-s delay in Condition F with the conditions in Study 1, Kruskal–Wallis tests were used. If the Kruskal–Wallis test was significant, Wilcoxon Mann–Whitney tests (Wilcoxon, 1945) with Bonferroni correction were conducted as planned comparisons.

Participants

A total of 52 infants, all 12 months old (average age = 51.02 months), provided usable data for Study 2: 12 infants (5 girls and 7 boys) in the 2-s gap condition, 7 infants (3 girls and 4 boys) in the 5-s gap condition, 13 infants (6 girls and 7 boys) in the no-cheer condition, 6 infants (3 girls and 3 boys) in the standard DNMS with separate jack-in-the-boxes 2-s gap condition, and 14 infants (7 girls and 7 boys) in the standard DNMS with separate jack-in-the-boxes no-gap condition (see Table 6). Fewer infants were tested in the standard DNMS procedure with jack-in-the-box rewards and a 2-s gap between acting on the stimulus and receipt of the reward because so few infants had succeeded in the easier, comparable condition (the same condition but with no gap between action on the stimulus and receipt of the reward), and so few infants were succeeding in this condition that the results seemed clear enough without subjecting more infants to a failure experience. Fewer infants were tested in the 5-s gap condition

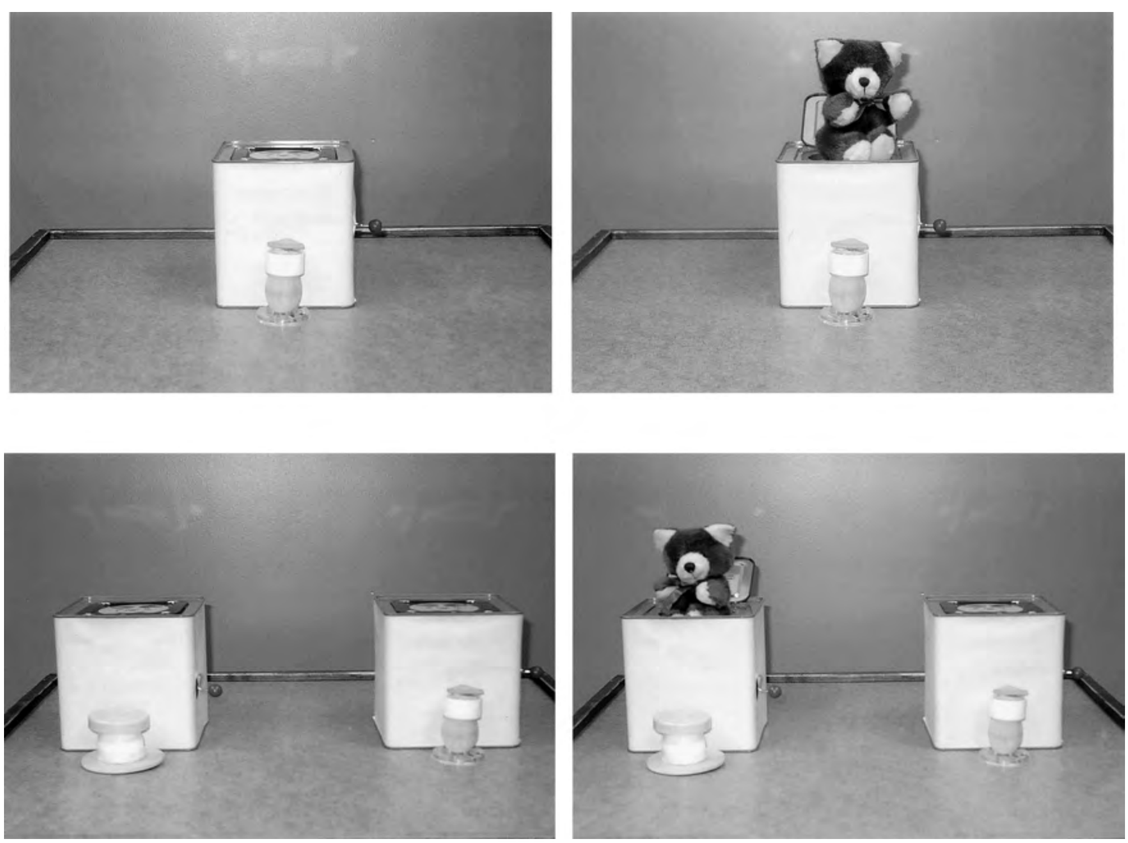

Figure 3. Photos showing the standard delayed nonmatching to sample procedure with jack-in-the-box rewards used in Study 2. Top left frame shows the presentation of the sample stimulus with the box containing the jack-in-the-box immediately behind it. Touching the stimulus resulted in the jack-in-the-box bear popping up behind it (top right frame). Bottom left frame shows the presentation of the sample stimulus paired with a novel object, each with a jack-in-the-box box behind them. Choice of the stimulus not previously seen resulted in the bear popping up directly behind it (bottom right frame). (Pulling on the sample stimulus would have produced no effect.) These same bears were used in all jack-in-the-box conditions in Study 2, including Conditions D, E, and F, in which the single large white apparatus was used.

Table 6
Age Means and Range for Infants Tested in the Five Conditions of Study 2

Condition	*n*	Age (in weeks) *M*	Range
Conditions using large white Plexiglas apparatus that housed all rewards and to which all stimuli were affixed			
2-s gap between acting on stimulus and bear popping up			
All	12	50.8	49.6–53.0
Boys	7	50.8	49.6–53.0
Girls	5	50.8	50.0–52.6
5-s gap between acting on stimulus and bear popping up			
All	7	52.5	50.9–54.9
Boys	4	52.0	50.9–53.3
Girls	3	53.0	50.9–54.9
No-cheer condition (no feedback from experimenter)			
All	13	51.6	47.9–55.3
Boys	7	51.2	49.1–53.7
Girls	6	52.0	47.9–55.3
Conditions run with standard DNMS procedure but with jack-in-the-box rewards			
2-s gap between acting on stimulus and bear popping up			
All	6	51.8	48.9–54.1
Boys	3	51.3	48.9–54.1
Girls	3	52.4	49.3–54.0
0-s gap between acting on stimulus and bear popping up			
All	14	49.7	48.0–51.7
Boys	7	50.2	48.0–51.7
Girls	7	49.1	48.0–51.0

Note. DNMS = delayed nonmatching to sample.

because the people who had been doing the testing were leaving; the results seemed clear enough without adding the additional complication of a difference in testers across conditions. The proportion of unusable subjects was roughly comparable across conditions.

Another 20 infants came into the lab but could not be used in the study: 6 infants (4 girls and 2 boys) in the 2-s gap condition, 3 infants (1 girl and 2 boys) in the 5-s gap condition, 4 infants (3 girls and 1 boy) in the no-cheer condition, 3 infants (1 girl and 2 boys) in the DNMS with separate jack-in-the-boxes 2-s gap condition, and 4 infants (2 girls and 2 boys) in the DNMS with separate jack-in-the-boxes no-gap condition. The main reason for unusable sessions was that individual infants were too fussy, squirmy, or restless. This was true for 15 out of 20 (75%) of the unusable sessions. Sometimes this restlessness developed over the course of a session as an infant lost interest; other times the fussiness was present from the start. For a few sessions, it was clear that some infants were fussy because they were not feeling well. We could not get these 15 infants to participate for 15 trials, the minimum needed for a session to be considered usable if an infant had not passed criterion. For three other sessions, the data were too difficult to interpret either because the infant showed a persistent side bias or because the infant always reached for both stimuli. Two sessions were unusable because of equipment problems.

All participants were healthy and full-term. Most were from middle- to upper-middle-class backgrounds and of European Caucasian descent. Across conditions, most parents (both fathers and mothers) were in their early 30s, and all had graduated from college. Many also had postgraduate training. All infants received a toy present or souvenir tee-shirt valued at approximately $5. For all infants, informed consent was obtained from a parent before testing.

Results

There were no significant sex differences. Hence, in all analyses reported below the results were collapsed across sex.

Effect of Physical Connectedness

Infants performed significantly better in the conditions in which the stimulus and reward were both attached to the same apparatus than in the conditions in which the stimulus and reward were clearly unconnected (see Figure 4). Most infants (84%) passed criterion at the 5-s delay in the three conditions in which the stimulus and reward were affixed to the large white box. Very few infants (25%) passed criterion in the two conditions in which separate, independent boxes were used for each jack-in-the-box and the stimulus and reward were clearly not attached to the same apparatus. Hence, infants in Conditions D, E, and F (the conditions in which the stimulus and reward were affixed to the same apparatus) performed significantly better than infants in Conditions G and H (in which the stimulus and reward were clearly not attached) on every dependent measure (see Table 7; total percentage correct at the 5-s delay, $p < .0004$; percentage correct on the last 5 trials at the 5-s delay, $p < .0001$; trials to criterion, $p < .0001$; whether passed the criterion or not, $p < .0001$).

Similarly, in the two conditions of Study 1 in which the stimulus and reward were attached to the same white apparatus (Conditions A and B), infants performed significantly better on *every* dependent measure than did infants of the same age tested in the two conditions of Study 2 (Conditions G and H) in which jack-in-the-box rewards were still used but the stimulus and reward were clearly not attached ($p < .001$ for whether passed criterion or not; $p < .0001$ for each of the other dependent measures). Likewise, in the three conditions of Study 2 in which the stimulus and reward were attached to the same white apparatus (Conditions D, E, and F), infants performed significantly better on every dependent measure than did infants of the same age tested in the condition of Study 1 (Condition C) in which the stimulus and reward were clearly not attached ($p < .002$ for number of trials to criterion; $p < .003$ for each of the other dependent measures).

There were no significant differences, on any dependent measure, among the five conditions in which the stimulus and reward were both attached to the same apparatus (Conditions A and B [Study 1] and Conditions D, E, and F [Study 2]). Similarly, there were no significant differences, on any dependent measure, among the three conditions in which the stimulus and reward were clearly not attached (Condition C [Study 1] and Conditions G and H [Study 2]).

Effect of Temporal Proximity

There was no effect of the size of the temporal gap between acting on the stimulus and receipt of the reward within the range

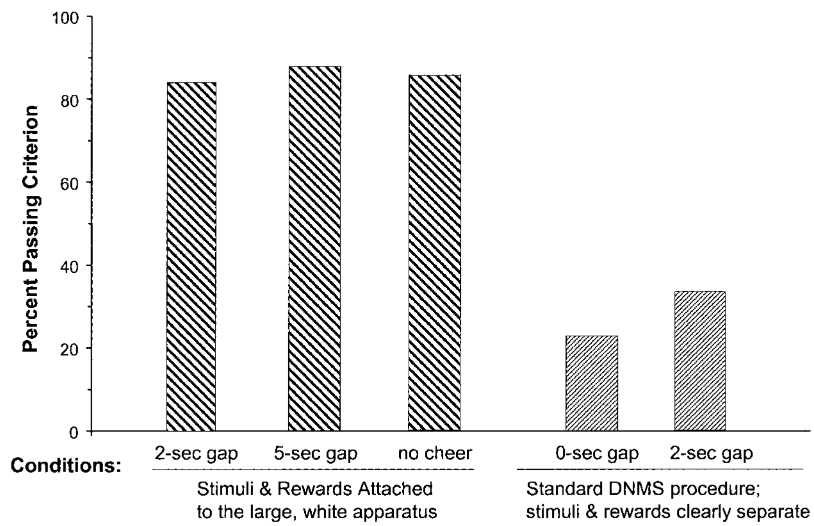

Figure 4. Percentage of infants passing criterion at the 5-s delay in each of the five conditions tested in Study 2. DNMS = delayed nonmatching to sample.

of temporal gaps used in this study. An analysis combining all conditions with a 0-s temporal gap between the response at stimulus and receipt of the reward in Studies 1 and 2 (Conditions A, B, and H), all conditions with a 2-s temporal gap (Conditions C, D, F, and G), and the condition with a 5-s gap (Condition E), for a total of three levels of temporal gap, yielded no significant effect for size of the temporal span between response at stimulus and receipt of reward for any dependent measure.

When the large white apparatus to which stimuli and rewards were all attached was used, infants performed no worse with a 5-s gap between acting on the stimulus and seeing the reward than they did with a 2-s gap (Condition E vs. Condition D). On no dependent measure was there any significant difference between performance in these two conditions. Further, infants performed no worse in either of those conditions than in the condition with no temporal gap between stimulus and reward (jack-short; Condition B in Study 1). On no dependent measure was there any significant difference between the performance of infants in Conditions D or E (or the two combined) and that of infants of the same age tested in Condition B of Study 1.

Infants performed no better in the standard DNMS procedure with jack-in-the-box rewards when the reward appeared immediately (no temporal gap between stimulus and reward, Condition H) than they did when there was a 2-s gap between acting on the stimulus and seeing the reward (Condition G). On no dependent measure was there any significant difference between performances in these two conditions.

Effect of the length of delay (5 vs. 30 s) between sample and test could be examined only for the no-cheer condition (Condition F), because too few infants were tested at the 30-s delay in the other conditions. Performance in Condition F did not differ in overall percentage of correct responses at the two delays, but infants did perform better on the last 5 trials at the 5-s delay than they did on the last 5 trials at the 30-s delay ($p < .02$).

Effect of Reward Type

Infants performed no better with the standard DNMS procedure when jack-in-the-box rewards were used (Conditions G and H) than other infants of the same age had performed with the same procedure when small objects were the rewards in Study 1 (Condition C). On no dependent measure was there any significant difference between performance in these conditions (Conditions G and H vs. C, G vs. C, or H vs. C). Because even when jack-in-the-box rewards were used, performance was significantly worse in the standard DNMS procedure than when the stimulus and reward appeared to be physically connected, that difference cannot be attributed to the type of reward used.

Though the same jack-in-the-box rewards were used in all conditions of Study 2, infants performed significantly better in three of those conditions (D, E, and F) than they did in the other two conditions (G and H). This was true for every dependent measure (total percentage correct at the 5-s delay, $p < .0004$; percentage correct on the last 5 trials at the 5-s delay, $p < .0001$; trials to criterion, $p < .0001$; whether passed criterion or not, $p < .0001$). Because reward was held constant across those conditions, the observed differences in performance across conditions cannot be due to the type of reward used.

Infants performed no differently in the no-cheer condition (Condition F) than they did when exactly the same procedure was used with experimenter cheers and applause (jack-short condition, or Condition B of Study 1) or than they did in any condition that used the same apparatus (Conditions A, B, D, and E). On no dependent measure was there any significant difference between performance in the no-cheer condition and performance in any, or all, of those other conditions. Moreover, performance in the no-cheer condition was significantly better than performance in each of the conditions using the standard DNMS procedure (Conditions C, G, and H) on all dependent measures (p values for all comparisons $< .02$). It appears clear that the presence or absence of experimenter cheering in this experiment did not affect infants' performance given

Table 7
Study 2: Mean Performance on Each Dependent Measure in Each Condition

Condition and sex	5-s delay: Mean number of trials to criterion[a]: All participants[b]	5-s delay: Mean number of trials to criterion[a]: Only those who passed criterion	5-s delay: % passing criterion	5-s delay: % correct	5-s delay: % correct on last 5 trials	30-s delay[c]: % correct	30-s delay[c]: % correct on last 5 trials
Conditions using large white Plexiglas apparatus that housed all rewards and to which all stimuli were affixed							
Condition D: 2-s gap between acting on stimulus and appearance of jack-in-the-box							
Boys	8.3 (*n* = 7)	1.6 (*n* = 5)	71.4	72.6	88.6	83.7 (*n* = 3)	86.7 (*n* = 3)
Girls	5.6 (*n* = 5)	5.6 (*n* = 5)	100	73.8	100	50.0 (*n* = 1)	80.0 (*n* = 1)
All	7.2 (*n* = 12)	3.6 (*n* = 10)	83.3	73.1	93.3	75.3 (*n* = 4)	85.0 (*n* = 4)
Condition E: 5-s gap between acting on stimulus and appearance of jack-in-the-box							
Boys	4.3 (*n* = 4)	4.3 (*n* = 4)	100	86.3	100	75.0 (*n* = 1)	80.0 (*n* = 1)
Girls	12 (*n* = 3)	5.5 (*n* = 2)	66.7	54.3	73.3	75.0 (*n* = 1)	80.0 (*n* = 1)
All	7.6 (*n* = 7)	4.7 (*n* = 6)	85.7	72.6	88.6	75.0 (*n* = 2)	80.0 (*n* = 2)
Condition F: No-cheer (No feedback from experimenter, jack-in-the-box rewards, 0-s gap)							
Boys	5.1 (*n* = 7)	1.8 (*n* = 6)	85.7	82.1	91.4	58.7 (*n* = 3)	66.7 (*n* = 3)
Girls	4.8 (*n* = 6)	0.8 (*n* = 5)	83.3	81.7	86.7	85.2 (*n* = 5)	80.0 (*n* = 5)
All	5 (*n* = 13)	1.4 (*n* = 11)	84.6	81.9	89.2	75.3 (*n* = 8)	75.0 (*n* = 8)
Conditions run with standard DNMS procedure but with jack-in-the-box rewards							
Condition G: 2-s gap between acting on stimulus and appearance of jack-in-the-box							
Boys	25 (*n* = 3)		0	53	40		
Girls	11.7 (*n* = 3)	5.0 (*n* = 2)	66.7	59.3	86.7	75.0 (*n* = 1)	60.0 (*n* = 1)
All	18.3 (*n* = 6)	5.0 (*n* = 2)	33.3	56.2	63.3	75.0 (*n* = 1)	60.0 (*n* = 1)
Condition H: 0-s gap between acting on stimulus and appearance of jack-in-the-box							
Boys	21.4 (*n* = 7)	0 (*n* = 1)	14.3	59.9	65.7		
Girls	19.9 (*n* = 7)	7.0 (*n* = 2)	28.6	57.8	62.9		
All	20.6 (*n* = 14)	4.67 (*n* = 3)	21.4	58.8	64.3		

Note. All infants in all conditions were 12 months old.
[a] The number of trials to criterion was calculated as is done in neuroscience studies; that is, it equals the number of trials administered at the 5-s delay minus the string of 5 consecutively correct trials used to satisfy the criterion. [b] Those who failed criterion were assigned a score of 25 here (the maximum number of trials administered at the 5-s delay). [c] Only infants who passed the 5-s delay were tested at the 30-s delay. Infants who refused to perform the task any longer after the testing at 5 s, who took more than 18 trials to pass criterion at the 5-s delay, or who received fewer than 8 trials at the 30-s delay were also excluded from analyses at the 30-s delay.

that in all conditions infants did receive feedback in the form of receipt or nonreceipt of a jack-in-the-box reward or a small object.

General Discussion

Study 1 eliminated close spatial proximity between stimulus and reward as a key factor in enabling infants to grasp the relation between stimuli and rewards or as the reason why infants can learn the nonmatching rule in the Velcro condition (Diamond et al., 1999) but not in the standard condition. In both jack-in-the-box conditions of Study 1, stimuli and rewards were separated in space by 12.5 cm, yet infants succeeded. However, in Study 1, all conditions in which infants were able to deduce the nonmatching rule were characterized by both (a) close temporal proximity between acting on the stimulus and receipt of the reward and (b) an apparent physical connection between the stimulus and reward.

One of our goals in Study 2 was to determine whether temporal proximity or physical connection was the critical factor in enabling infants to grasp the relation between the stimulus and reward. The results clearly show that physical connectedness appears to be the critical factor. In all conditions of physical connectedness, infants were able to understand the connection between the stimulus and reward sufficiently well to deduce the DNMS rule on the basis of the presence or absence of reward feedback. In no condition in which physical connectedness was absent were infants able to do that (see Table 8). Close temporal proximity was present in some conditions in which infants succeeded and absent in others. Similarly, close temporal proximity was present in some conditions in which infants failed and absent in others. The same was true for spatial proximity—the reward was spatially separated from the stimulus and site of response in the jack-in-the-box conditions in which infants succeeded and in the jack-in-the-box conditions in which infants did not. Importantly, even in conditions in which neither close temporal proximity nor close spatial proximity was present (the 2-s and 5-s gap conditions with the large white apparatus) but both stimuli and rewards were part of a single apparatus, infants succeeded.

Physical connectedness appears to be a sufficient and necessary condition. Neither spatial proximity, nor temporal proximity, nor their conjunction appears to be necessary or sufficient.

Our predictions had been that spatial proximity would not matter but that temporal proximity and physical connectedness would. We were right about spatial proximity and physical connectedness but dead wrong about temporal proximity. Indeed, physical connectedness appears to be a far more powerful factor than we had anticipated. We had predicted its influence would be about equal to that of temporal proximity. Its influence was far greater. In its absence, even the conjunction of close spatial proximity and close temporal proximity was insufficient for infants to succeed at the DNMS task.

Table 8
When Do 12-Month-Old Infants Pass Criterion With a 5-s Delay Between Stimulus Presentation and Test?

Critical factor and condition	Do 12-month-olds pass?
Temporal proximity	
Close	
Condition A: Jack-long	Yes
Condition B: Jack-short	Yes
Condition H: Same as G but no delay	No
Relatively far	
Condition C: Standard DNMS	No
Condition D: 2-s delay	Yes
Condition F: No-cheer	Yes
Condition G: Standard DNMS with jack-in-the-box rewards	No
Farther	
Condition E: 5-s delay	Yes
Physical connection	
Present	
Condition A: Jack-long	Yes
Condition B: Jack-short	Yes
Condition D: 2-s delay	Yes
Condition E: 5-s delay	Yes
Condition F: No-cheer	Yes
Absent	
Condition C: Standard DNMS	No
Condition G: Standard DNMS with jack-in-the-box rewards	No
Condition H: Same as G but no delay	No

Note. DNMS = delayed nonmatching to sample.

The work reported here makes it possible to rule out several other possible factors as central to infants' ability to grasp the relation between stimuli and rewards in the DNMS task. *It cannot be length of presentation time* because there were conditions in which infants succeeded despite a short (5-s) presentation time (Conditions B, D, E, and F: jack-short, 2-s gap, 5-s gap, and no cheer; see also Diamond et al., 1999) and a condition in which infants failed even with a long (20-s) presentation time (see Study 4 in Diamond et al., 1994). Not surprisingly, the reverse was also true: There were conditions with only 5-s stimulus presentation times in which infants failed (Conditions C, G, and H: standard DNMS procedure with or without jack-in-the-box rewards) and a condition with long presentation times in which infants succeeded (20–25 s; Condition A: jack-long).

It cannot be the requirement of making a two-action means–end sequence that causes infants to fail the DNMS task because there were conditions without such sequences in which infants failed anyway (Conditions G and H: standard DNMS procedure with jack-in-the-box rewards) and a condition with such means–end sequences in which infants succeeded (the Velcro condition; see Diamond et al., 1999). Not surprisingly, there were also DNMS conditions requiring such means–end sequences in which infants failed (Condition C: standard DNMS procedure; see also Diamond et al., 1994, 1999) and DNMS conditions not requiring such sequences in which infants succeeded (Conditions A, B, D, E, and F: all jack-in-the-box conditions in which stimuli and rewards were affixed to the white apparatus).

It cannot be the type of reward, or that infants prefer pop-up jack-in-the-box rewards more (or less) than small objects, that accounts for when infants can or cannot demonstrate through their actions that they have deduced the DNMS rule. There were jack-in-the-box conditions in which infants succeeded (Conditions A, B, D, E, and F: all jack-in-the-box conditions in which stimuli and rewards were affixed to the white apparatus) and in which they failed (Conditions G and H: standard DNMS procedure with jack-in-the-box rewards). Similarly, there were DNMS conditions with small objects in which infants succeeded (Velcro condition; Diamond et al., 1999) and in which they failed (Condition C: standard DNMS procedure; see also Diamond et al., 1994, 1999).

It cannot be the presence of verbal feedback (cheering to correct responses, disappointed face and voice to incorrect responses) and/or applause from the experimenter that accounts for when infants can or cannot succeed at the DNMS task. In all conditions save one, the experimenter gave verbal feedback and applause in addition to the jack-in-the-box or small-object reward. In some of those conditions infants succeeded; in some they failed. In one jack-in-the-box condition, verbal feedback and applause from the experimenter were omitted, yet infants succeeded anyway (Condition F: no-cheer).

It cannot be the memory requirements of the DNMS task that keep infants roughly 12 months old from succeeding, for they succeeded here with temporal gaps of 2 s and 5 s between response and reward (Conditions D and E), and here, as in other studies, they succeeded with temporal delays of 5 s between stimulus presentation and test (Conditions A, B, D, E, and F; stimulus = reward condition, Diamond, 1995; verbal reward and Velcro conditions, Diamond et al., 1999). Indeed, infants who succeeded at the 5-s delay generally went on to perform fairly comparably at the longer 30-s delay (in all the studies cited above and in Diamond et al. [1994] as well). Their performance did not show a relation to length of delay, as it should have if insufficient memory was the limiting factor.

Thus, the present body of data on DNMS variations allows us to rule out the following as reasons for why infants fail the standard DNMS task until almost 2 years of age: (a) the presence of too brief a stimulus presentation time for infants to have sufficiently encoded the sample stimulus, (b) the need to execute a means–end action sequence, (c) verbal feedback as to whether the infant's response is right or wrong, (d) a spatial separation between the stimulus and reward, (e) a temporal gap between the infant's response to the stimulus and receipt of the reward, and (f) a delay between sample presentation and test.

How Do the Present Findings Concerning Spatial Proximity Compare With Those in the Literature?

The results of the present study indicate that close spatial proximity of the stimulus and the reward is not important to infants' ability to grasp the relation between stimulus and reward in the DNMS task. In all jack-in-the-box conditions, stimuli and rewards were spatially displaced. Independent of that, infants succeeded when stimuli and rewards were attached to the large white apparatus and failed when they were not attached. Similarly, infants succeed in the Velcro condition (Diamond et al., 1999), in which stimuli and rewards are contiguous, and fail in the standard DNMS condition, in which stimuli and rewards are also spatially close.

Most research on the importance of close spatial proximity to learning has focused on a spatial separation between the stimulus and response. For example, Murphy and Miller (1959) found that 50% of the fourth-grade children they tested achieved criterion on a discrimination learning task when the cue, reward, and response were contiguous, but when the cue was spatially displaced only 33% of the fourth graders were able to reach criterion. Jeffrey and Cohen (1964) found that nursery school children could learn a visual discrimination problem if the stimulus and response were contiguous but not if they were spatially separated. Ramey and Goulet (1971) found that children in Grades 4 and 6 performed significantly worse on a discrimination task when the stimulus was separated from the site of the response and reward than when they were all spatially contiguous.

The results are similar with rhesus monkeys. For example, as the distance between the relevant visual cues and the site of the response and the reward increases, the percentage of correct responses by monkeys decreases (McClearn & Harlow, 1954; see also Wunderlich & Dorff, 1965). When stimulus objects were displaced 15 cm vertically from the site of the response and the reward, monkeys failed to solve an object-quality discrimination problem, and the performance of those monkeys who had previously learned the discrimination fell to near-chance levels (Murphy & Miller, 1955). Indeed, monkeys performed much better on a discrimination problem when the response was made to the stimulus panel than when the response was made to a manipulandum only 5 cm below the panel (Meyer, Polidora, & McConnell, 1961) or when the cue was where the monkeys responded rather than only 2.5 cm away (Schuck, Polidora, McConnell, & Meyer, 1961). More recently, Iwai, Yaginuma, and Mishkin (1986) found that even an increment in spatial separation of the cue and response as small as 0.5 cm markedly reduced the rate at which rhesus monkeys acquired a pattern discrimination. In a study with chimpanzees, Jarvik (1956) found that they could acquire a visual discrimination in 1 trial if the reward was attached to the stimulus but that they needed over 100 trials if the reward was 0.1 cm below in a shallow food well.

However, although studies show that it is clear that a spatial separation between *stimulus and response* impedes learning, studies also show that a spatial separation between *stimulus and reward* or between *response and reward* does not. (In the DNMS task, participants respond at the stimulus, but the reward is separated by being in the well below [standard condition] or by popping up several inches behind the stimulus [jack-in-the-box condition].) Thus, for example, Millar and Schaffer (1972) found that infants could tolerate a spatial separation between the response site and the reward as long as both were in the same visual field. Rovee-Collier (1984, 1990, 1995, 1999) showed that infants can learn the operant response of kicking for the reward of seeing a mobile move overhead even though the response and reward sites are several feet apart. Jeffrey and Cohen (1964) found that nursery school children could not learn a visual discrimination problem if the stimulus and response were spatially separated, but they could learn when the reward was separated from the stimulus and response site. Indeed, children's performance was as good when only the cue and response were contiguous and the reward was spatially separated as it was when the cue, response, and reward were all spatially contiguous. Similarly, the spatial separation that Ramey and Goulet (1971) found to be problematic for children was separation between the stimulus, on the one hand, and the site of the response and reward, on the other hand. In all of the studies cited in the paragraph above, showing that rhesus monkeys perform worse under conditions of spatial separation, the spatial separation investigated was that between the stimulus and response (McClearn & Harlow, 1954; Meyer et al., 1961; Murphy & Miller, 1955; Iwai et al., 1986; Schuck et al., 1961; Wunderlich & Dorff, 1965). Three studies with rhesus monkeys (Miller & Murphy, 1958, 1964; Polidora, Thompson, & Wayne, 1965) have systematically varied spatial displacements among the stimulus, response site, and reward. The findings of these studies are consistent: Spatial displacement of the stimulus and response presents major difficulties; spatial displacement of the reward from the stimulus and/or response site does not. Because in the jack-in-the-box condition the response is on the stimulus, and it is the reward that is spatially displaced, our finding that infants succeeded in the jack-in-the-box condition despite spatial displacement of the reward is fully consistent with the literature.

How Do the Present Findings Concerning Temporal Proximity Compare With Those in the Literature?

We found here that delays of 2 s, or even 5 s, between acting on a stimulus and receiving reward feedback did not impede infants' ability to succeed at the DNMS task as long as the stimulus and reward were affixed to the same white apparatus. Conversely, even the absence of any temporal gap between acting on a stimulus and seeing the jack-in-the-box reward was insufficient to enable infants to succeed at the DNMS task if the stimulus and reward were not affixed to the same thing. Hence, temporal proximity of the stimulus and reward did not appear to be important to infants' success in the DNMS paradigm.

Evidence from some paradigms indicates that the presence or absence of a temporal gap between a stimulus and the infant's response can mean the difference between failure and success to infants in the 1st year of life (e.g., the A-not-B paradigm: Gratch, Appel, Evans, LeCompte, & Wright, 1974; Harris, 1973). Diamond (1985) demonstrated, for example, that infants' performance in the A-not-B paradigm improves when the delay between the stimulus and response is decreased and worsens when the delay is increased. Using an operant conditioning paradigm, Millar (1972; Millar & Watson, 1979) found that when there was no delay between an infant's response and the onset of the reward (a colored light display), even infants of 6 months could master the task. When there was a delay of 3 s between the response and reward, however, Millar (1972; Millar & Watson, 1979) found that even 8-month-old infants failed to learn the relationship between the response and reward.

Evidence from visual paired comparison paradigms (e.g., Fagan, 1973), however—even when assessed by infants' reaching for stimuli rather than just looking at them (Diamond, 1995)—indicates that infants in the 1st year of life can, under these circumstances, tolerate quite long delays (delays of minutes or more) between stimulus presentation and response/reward (the stimulus looked at, or reached for, being the reward, hence response and reward are intertwined).

Why We Conclude That If Infants Perceive That the Stimulus and Reward Are Components of a Single Thing, They Are Able to Understand the Relation Between the Stimulus and Reward and to Use That Relation to Deduce the Abstract Nonmatching Rule in the DNMS Task

Infants succeeded on the DNMS task in all conditions in which stimuli and rewards were attached to the same apparatus. In those conditions, infants were not able to remove either the stimulus or reward from that apparatus. Without a doubt, any adult watching the procedure in the jack-long, jack-short, or no-cheer conditions came away with the strong impression that it looked as if pulling on the stimulus *caused* the jack-in-the-box to pop up, as if the stimulus and reward were attached by a lever. We had predicted that if we broke the close temporal connection between pulling on the stimulus and the appearance of the reward, we would disrupt the impression that pulling on the stimulus directly caused the jack-in-the-box to pop up, and infants would no longer succeed. We were wrong; infants still succeeded. With a gap of 2 s between pulling the stimulus and the appearance of the jack-in-the-box, it still appeared to some adults as if the apparatus were "cranky" or in need of a good oiling or that the jack-in-the-box had to overcome resistance to make its way up. Pulling the stimulus still seemed perhaps to be directly, causally linked to the appearance of the jack-in-the-box. With a gap of 5 s, however, adults no longer perceived the stimulus and reward to be directly, causally linked, yet infants still succeeded. Apparently, having the stimulus and reward as components of a single piece of apparatus made the connection between stimuli and rewards transparent to infants in a way that having the stimuli on the table directly in front of the boxes housing the jack-in-the-boxes did not or that having the stimuli directly on top of the shallow wells containing the rewards did not.

In the conditions in which infants succeeded (when stimuli and rewards were parts of the same apparatus), it was also true that the stimulus could not be picked up, and neither could the reward. Could picking up the stimulus, perhaps, pull infants' attention to the stimulus as interesting in its own right, distracting infants from the intended reward feedback? Similarly, could being able to pick up the reward impede performance for some reason? We think not, because infants also succeeded when the stimulus and reward could be picked up and manipulated (Velcro condition; Diamond et al., 1999) and they also failed in conditions in which the rewards could not be picked up (Conditions G and H: standard DNMS procedure with jack-in-the-box rewards).

We conclude that infants are able to deduce the abstract nonmatching rule in the DNMS task when the stimulus and reward appear to be physically connected or to be components of a single entity, but that infants are unable to succeed at this task in the 1st year of life if the stimulus and reward do not appear to be physically connected or to be parts of a common entity. These results accord fully with those obtained by Aguiar and Baillargeon (2000) with a means–end A-not-B task with infants and with those obtained by Jarvik (1953, 1956) with a color discrimination task with nonhuman primates (see the introduction). Physical connectedness in the DNMS paradigm is sufficient even when stimuli and rewards are neither spatially nor temporally proximal. Even if the stimulus and reward are some distance apart and the reward does not appear until 5 s after the infant acts on the stimulus, if both the stimulus and the reward are connected to the same piece of apparatus, then infants succeed. In the absence of the perception that the stimulus and reward are components of a single thing, even close spatial and temporal proximity of stimuli and rewards are insufficient for infants to grasp the relation between them in the DNMS task.

Exactly what is crucial for infants to perceive connectedness awaits further study. If components of the large white apparatus had been different colors, perhaps that might have disrupted the perception of connectedness. If stimuli and rewards were attached to the large white apparatus only by Velcro (and hence were detachable), perhaps infants might not have succeeded. Fascinating and potentially critically important questions are why infants can succeed in the connected jack-in-the-box conditions but not in the unconnected ones, why infants can succeed if a toy is attached to the cloth it is on but not if the toy is atop the cloth in the same position but unattached (Aguiar & Baillargeon, 2000), and why nonhuman primates can quickly learn a color discrimination if colored celluloid is attached to the bread reward on which it rests but not if the colored celluloid is still atop the reward but unattached (Jarvik, 1953). Is physical connection so important because then movement of the stimulus and reward are perfectly correlated (Spelke, 1985; Vishton & Badger, 2003), because physical connection makes the stimulus and reward seem more like one thing than two separate ones (DeLoache, 2000), or because of yet a different reason? These are but some of the fascinating questions awaiting future research.

References

Aguiar, A., & Baillargeon, R. (2000). Perseveration and problem solving in infancy. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 27, pp. 135–180). San Diego: Academic Press.

Alberts, E., & Ehrenfreund, D. (1951). Transposition in children as a function of age. *Journal of Experimental Psychology, 41,* 30–38.

Caron, A. J., Caron, R. F., Minichiello, M. D., Weiss, S. J., & Friedman, S. L. (1977). Constraints on the use of the familiarization novelty method in the assessment of infant discrimination. *Child Development, 48,* 747–762.

DeLoache, J. S. (1986). Memory in very young children: Exploitation of cues to the location of a hidden object. *Cognitive Development, 1,* 123–137.

DeLoache, J. S. (1989). Young children's understanding of the correspondence between a scale model and a larger space. *Cognitive Development, 4,* 121–139.

DeLoache, J. S. (1995). Early understanding and use of symbols: The model model. *Current Directions in Psychological Science, 4,* 109–113.

DeLoache, J. S. (2000). Dual representation and young children's use of scale models. *Child Development, 71,* 329–338.

DeLoache, J. S., & Brown, A. L. (1983). Very young children's memory for the location of objects in a large-scale environment. *Child Development, 54,* 888–897.

Diamond, A. (1985). The development of the ability to use recall to guide action, as indicated by infants' performance on A-not-B. *Child Development, 56,* 868–883.

Diamond, A. (1990). Rate of maturation of the hippocampus and the developmental progression of children's performance on the delayed non-matching to sample and visual paired comparison tasks. *Annals of the New York Academy of Sciences, 608,* 394–426.

Diamond, A. (1995). Evidence of robust recognition memory early in life even when assessed by reaching behavior. *Journal of Experimental Child Psychology, 59,* 419–456.

Diamond, A., Churchland, A., Cruess, L., & Kirkham, N. (1999). Early developments in the ability to understand the relation between stimulus and reward. *Developmental Psychology, 35,* 1507–1517.

Diamond, A., Towle, C., & Boyer, K. (1994). Young children's performance on a task sensitive to the memory functions of the medial temporal lobe in adults: The delayed nonmatching-to-sample task reveals problems that are due to non-memory-related task demands. *Behavioral Neuroscience, 108,* 659–680.

Fagan, J. F., III (1973). Infants' delayed recognition memory and forgetting. *Journal of Experimental Child Psychology, 16,* 424–450.

Fagen, J. W., & Rovee-Collier, C. K. (1982). A conditioning analysis of infant memory. In R. L. Isaacson & N. Spear (Eds.), *The expression of knowledge* (pp. 67–111). New York: Plenum Press.

Gellermann, L. W. (1933). Chance orders of alternating stimuli in visual discrimination experiments. *Journal of General Psychology, 42,* 206–208.

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 Development, 45,* 71–77.

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

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York: Wiley.

Hunter, M. A., & Ames, E. W. (1975, April). *Visual habituation and preference for novelty in five-week-old infants.* Paper presented at the meeting of the Society for Research in Child Development, Denver, CO.

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. K. Rovee-Collier & L. P. Lipsett (Eds.), *Advances in infancy research* (pp. 69–95). Norwood, NJ: Ablex.

Iwai, E., Yaginuma, S., & Mishkin, M. (1986). Acquisition of discrimination learning of patterns identical in configuration in macaques (*Macaca mulatta* and *M. fuscata*). *Journal of Comparative Psychology, 100,* 30–36.

Jarvik, M. E. (1953). Discrimination of colored food and food signs by primates. *Journal of Comparative and Physiological Psychology, 46,* 390–392.

Jarvik, M. E. (1956). Simple color discrimination in chimpanzees: Effect of varying contiguity between cue and incentive. *Journal of Comparative and Physiological Psychology, 49,* 492–495.

Jeffrey, W. E., & Cohen, L. B. (1964). Effect of spatial separation of stimulus, response, and reinforcement on selective learning in children. *Journal of Experimental Psychology, 67,* 577–580.

Kuene, M. (1946). Experimental investigation of the relation of language to transposition behavior in young children. *Journal of Experimental Psychology, 36,* 471–486.

Lasky, R. E. (1980). Length of familiarization and preference for novel and familiar stimuli. *Infant Behavior and Development, 3,* 15–28.

McClearn, G. E., & Harlow, H. F. (1954). The effect of spatial contiguity on discrimination learning in rhesus monkeys. *Journal of Comparative and Physiological Psychology, 47,* 391–394.

Mehta, C., & Patel, N. (2000). *StatXact 4 for Windows: Statistical software for exact nonparametric inference user manual.* Cambridge, MA: CYTEL Software Corporation.

Meunier, M., Bachevalier, J., Mishkin, M., & Murray, E. A. (1993). Effects on visual recognition of combined and separate ablations of the entorhinal and perirhinal cortex in rhesus monkeys. *Journal of Neuroscience, 13,* 5418–5432.

Meyer, D. R., Polidora, V. J., & McConnell, D. G. (1961). Effects of spatial S-R contiguity and response delay upon discriminative performances by monkeys. *Journal of Comparative and Physiological Psychology, 54,* 175–177.

Millar, W. S. (1972). A study of operant conditioning under delayed reinforcement in early infancy. *Monographs of the Society for Research in Child Development, 37,* 1–44.

Millar, W. S., & Schaffer, H. R. (1972). The influence of spatially displaced feedback on infant operant conditioning. *Journal of Experimental Child Psychology, 14,* 442–453.

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

Miller, R. E., & Murphy, J. V. (1958). Effect of the spatial relationship between cue, reward, and response in simple discrimination learning. *Journal of Experimental Psychology, 56,* 26–31.

Miller, R. E., & Murphy, J. V. (1964). Influence of the spatial relationships between the cue, reward, and response in discrimination learning. *Journal of Experimental Psychology, 67,* 120–123.

Mishkin, M. (1978). Memory in monkeys severely impaired by combined but not separate removal of amygdala and hippocampus. *Nature, 273,* 297–298.

Murphy, J. V., & Miller, R. E. (1955). The effect of spatial contiguity of cue and reward in the object-quality learning of rhesus monkeys. *Journal of Comparative and Physiological Psychology, 48,* 221–224.

Murphy, J. V., & Miller, R. E. (1959). Spatial contiguity of cue, reward, and response in discrimination learning by children. *Journal of Experimental Psychology, 58,* 485–489.

Murray, E. A., & Mishkin, M. (1998). Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus. *Journal of Neuroscience, 18,* 6568–6582.

Overman, W. H. (1990). Performance on traditional match-to-sample, nonmatch-to-sample, and object discrimination tasks by 12 to 32 month-old children: A developmental progression. *Annals of the New York Academy of Sciences. 608,* 365–393.

Overman, W. H., Bachevalier, J., Turner, M., & Peuster, A. (1992). Object recognition versus object discrimination: Comparison between human infants and infant monkeys. *Behavioral Neuroscience, 106,* 15–29.

Pascalis, O., de Haan, M., Nelson, C. A., & de Schonen, S. (1998). Long-term recognition memory for faces assessed by visual paired comparison in 3- and 6-month-old infants. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 249–260.

Polidora, V. J., Thompson, W. J., & Wayne, J. (1965). Orienting behavior and the S-R spatial discontiguity effect in monkeys. *Journal of Comparative and Physiological Psychology, 59,* 240–245.

Ramey, C. T., & Goulet, L. R. (1971). Effects of spatial separation of stimulus, response, and reward in discrimination learning by children. *Child Development, 42,* 978–982.

Rose, S. A., Gottfried, A. W., Melloy-Carminar, P., & Bridger, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology, 18,* 704–713.

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: Erlbaum.

Rovee-Collier, C. (1990). The "memory system" of prelinguistic infants. *Annals of the New York Academy of Sciences, 608,* 517–542.

Rovee-Collier, C. (1995). Time windows in cognitive development. *Developmental Psychology, 31,* 147–169.

Rovee-Collier, C. (1999). The development of infant memory. *Current Directions in Psychological Science, 8,* 80–85.

Rudel, R. G. (1955). *A re-evaluation of the dichotomy of absolute and relative responses in transposition.* Unpublished doctoral dissertation, New York University, New York, NY.

Rumbaugh, D. M., Richardson, W. K., Washburn, D. A., Savage-Rumbaugh, E. S., & Hopkins, W. D. (1989). Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus–response spatial contiguity. *Journal of Comparative Psychology, 103,* 32–38.

Schuck, J. R., Polidora, V. J., McConnell, D. G., & Meyer, D. R. (1961). Response location as a factor in primate pattern discrimination. *Journal of Comparative and Physiological Psychology, 54,* 543–545.

Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

Spelke, E. S. (1985). Perception of unity, persistence, and identity: Thoughts on infants' conception of objects. In J. Mehler & R. Fox (Eds.), *Neonate cognition: Beyond the blooming buzzing confusion* (pp. 89–113). Hillsdale, NJ: Erlbaum.

Sprent, P. (1993). *Applied nonparametric statistical methods* (2nd ed.). New York: Chapman & Hall.

Squire, L. R., Zola-Morgan, S., & Chen, K. S. (1988). Human amnesia and animal models of amnesia: Performance of amnesic patients on tests designed for the monkey. *Behavioral Neuroscience, 102,* 210–221.

Vishton, P. M., & Badger, A. N. (2003, April). *Infant Gestalt perception and object-directed reaching: Effects of shape, color, and remembered spatiotemporal information.* Paper presented at the biennial meeting of the Society for Research in Child Development, Tampa, FL.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics, 1,* 80–83.

Wunderlich, R. A., & Dorff, J. E. (1965). Contiguity relationships of stimulus, response, and reward as determinants of discrimination difficulty. *Journal of Comparative and Physiological Psychology, 59,* 147–149.

Zola, S. M., Squire, L. R., Teng, E., Stefanacci, L., Buffalo, E. A., & Clark, R. E. (2000). Impaired recognition memory in monkeys after damage limited to the hippocampal region. *Journal of Neuroscience, 20,* 451–463.

Zola-Morgan, S., & Squire, L. R. (1986). Memory impairment in monkeys following lesions limited to the hippocampus. *Behavioral Neuroscience, 100,* 155–160.

Zola-Morgan, S., Squire, L. R., Amaral, D. G., & Suzuki, W. A. (1989). Lesions of the perirhinal and parahippocampal cortex that spare the amygdala and hippocampal formation produce severe memory impairment. *Journal of Neuroscience, 9,* 4355–4370.

Zola-Morgan, S., Squire, L. D., & Mishkin, M. (1982). The neuroanatomy of amnesia: Amygdala-hippocampus versus temporal stem. *Science, 216,* 1337–1339.

Received September 19, 2002
Revision received April 11, 2003
Accepted April 11, 2003 ■