Published in final edited form as:
Cognition. 2003 July ; 88(3): 277–316.

Perseverative responding in a violation-of-expectation task in 6.5-month-old infants

Andréa Aguiar[a] and **Renée Baillargeon**[b,*]

[a]University of Waterloo, Waterloo, Ontario, Canada

[b]Department of Psychology, University of Illinois, 603 E. Daniel, Champaign, IL 61820, USA

Abstract

In the present research, 6.5-month-old infants perseverated in a violation-of-expectation task designed to examine their reasoning about width information in containment events. After watching a familiarization event in which a ball was lowered into a wide container, the infants failed to detect the violation in a test event in which the same ball was lowered into a container only half as wide as the ball (narrow-container test event). This negative result (which was replicated in another experiment) was interpreted in terms of a recent problem-solving account of infants' perseverative errors in various means-end tasks (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, CA: Academic Press). It was assumed that the infants in the present experiments (1) did not attend to the relative widths of the ball and container in their initial analysis of the narrow-container test event, (2) categorized the event as similar to the familiarization event shown on the preceding trials, and (3) retrieved the expectation they had formed for that event ("the ball will fit into the container"), resulting in a perseverative error. This interpretation was supported by additional experiments in which different modifications were introduced that led to non-perseverative responding, indicating that 6.5-month-old infants could detect the violation in the narrow-container test event. The present findings are important for several reasons. First, they provide the first demonstration of perseverative responding in a violation-of-expectation task. Second, they make clear the breadth and usefulness of the problem-solving account mentioned above. Finally, they add to the evidence for some degree of continuity between infants' and adults' problem-solving abilities.

Keywords

Infant cognition; Physical reasoning; Violation-of-expectation task; Perseveration

1. Introduction

As adults, we often solve problems by relying on familiar solutions. This ability to retrieve past solutions from memory makes us very efficient as problem solvers: in most cases, less

*Corresponding author. rbaillar@s.psych.uiuc.edu (R. Baillargeon).

time and effort are needed to retrieve as opposed to recompute solutions. To illustrate, consider a subject in an experiment who is asked to solve the problem “11 × 122”. The subject mentally computes the problem’s solution and eventually answers “1342”. On the next trial, the subject is again given the problem “11 × 122”. This time, rather than laboriously recomputing the problem’s solution, the subject swiftly retrieves it from memory – a clear demonstration of efficient problem solving.

Despite its obvious advantages, our ability to retrieve familiar solutions from memory has at least one potential pitfall. Past solutions can be helpful only if they are indeed appropriate for the problems at hand. Consider what would happen if the subject in our hypothetical experiment was given the problem “11 × 132” in a later trial and mistakenly assumed that this was the same problem as before. The solution “1342” would once again be retrieved, resulting in a perseverative error. In the present research, *perseveration* is defined as the retrieval of a familiar solution in a context in which a significant change has been introduced that renders the solution inappropriate; for the problem to be solved, a novel solution must be computed.

Infants, like adults, produce perseverative errors, and developmental researchers have long been interested in understanding the various roots of these errors. The present experiments examined infants’ tendency to perseverate in a task designed to assess their reasoning about width information in containment events. The central perspective of our work, as illustrated in the opening paragraphs, is that perseveration in such tasks is best understood as efficient problem solving gone awry: a familiar solution is retrieved where a novel solution should have been computed, because a critical change in the problem has gone unnoticed.

In the following sections, we first review prior research on infants’ perseverative errors. We then present the problem-solving account we have recently developed to explain some of these errors. Finally, we introduce the present research, which helped confirm and extend this account.

1.1. Prior research on perseveration in infancy

1.1.1. Memory-and-motor tasks—Most of the research on infant perseveration has focused on tasks that require infants (1) to update and remember information and (2) to use this information to select an appropriate motor response. For ease of description, we will refer to tasks with this dual requirement as “memory-and-motor” (ME-MO) tasks.

The first ME-MO task was developed by Piaget (1954) and examined infants’ ability to search for a toy hidden in one of two locations. This task has since been used extensively by developmental researchers (for reviews, see Bremner, 1985; Diamond, 1985; Harris, 1987; Sophian, 1984; Wellman, Cross, & Bartsch, 1986). For example, in a classic longitudinal study, Diamond (1985) tested infants every 2 weeks from 6 to 12 months of age. On each trial, a toy was hidden in one of two identical wells, which were then covered with cloths. Next, a delay was introduced, whose length was slowly increased across test sessions. Following the delay, the infants were allowed to search for the toy. Side of hiding was reversed three to five times in each session. Of particular interest were reversal trials (termed B trials) that followed correct trials (termed A trials): would the infants search at the correct well on the B trials or would they perseverate, returning to the well they had searched successfully on the A trials? Two main results were obtained. First, all of the infants produced perseverative errors on the B trials. Second, longer delays were necessary with age to elicit perseverative errors: thus, whereas 3 s delays produced errors at 8 months, 8 s delays were needed at 10 months, and over 10 s delays at 12 months.

In recent years, a number of researchers have reported perseverative errors in novel ME-MO tasks (e.g. Diedrich, Thelen, Smith, & Corbetta, 2000; Hofstadter & Reznick, 1996; Munakata, 1997; Smith, Thelen, Titzer, & McLin, 1999). For example, Smith et al. (1999) tested 8-month-old infants' ability to reach for one of two visible objects that had been cued by an experimenter. The infants sat in front of two identical wells covered with lids; one lid was labeled the A lid and the other the B lid. In each test trial, an experimenter grasped and waved one of the lids before returning it to its initial position. After a 3 s delay, the infants were allowed to reach. The A lid was used for two trials (A trials) and the B lid for two additional trials (B trials). On the A trials, the infants had a significant tendency to reach for the lid that had been cued by the experimenter, the A lid; on the B trials, the infants again reached for the A lid, even though the experimenter had cued the B lid.

Although superficially different, traditional and novel ME-MO tasks are nonetheless similar in that both require infants (1) to update and remember information (e.g. to remember which of two wells has been used to hide an object or which of two lids has been cued) and (2) to use this information to select one of two alternative motor responses (e.g. to search for the object in the left or right well or to grasp the left or right lid). Not surprisingly, accounts of perseveration in ME-MO tasks (e.g. Ahmed & Ruffman, 1998; Dehaene & Changeux, 1989; Diamond, 1991; Marcovitch & Zelazo, 1999; Munakata, 1998; Smith et al., 1999) typically appeal to (1) limitations in infants' ability to update and maintain information in working memory and (2) limitations in infants' ability to use information in working memory to select a novel motor response over a previously successful but no longer appropriate response. Decreases in perseveration with age are generally attributed to improvements in these two abilities with neurological maturation and motor experience.

1.1.2. Nonmemory-and-motor tasks—Although most of the research on infant perseveration has focussed on ME-MO tasks, perseverative errors have also been obtained in other tasks (e.g. Aguiar & Baillargeon, 2000; Lockman & Pick, 1984; Mackenzie & Bigelow, 1986; Rieser, Doxsey, McCarrell, & Brooks, 1982). These tasks are similar to ME-MO tasks in that they require infants to select one of two alternative motor responses on each trial. The tasks differ from ME-MO tasks, however, in that infants do not need to update and remember information to determine which motor response to select: all of the necessary information remains perceptually available throughout the trial. For ease of description, we will refer to these tasks as "nonmemory-and-motor" (NME-MO) tasks. We next describe two NME-MO tasks, first a detour (Lockman & Pick, 1984) and then a support (Aguiar & Baillargeon, 2000) task.

Lockman and Pick (1984) examined 12- and 18-month-old infants' ability to use the shortest route to reach their mothers. The infants and their mothers were positioned on opposite sides of an 8 foot long barrier, near the left or right end; the two ends were used on alternate trials.

The barrier was too high for the infants to climb over it, but was sufficiently low for them to clearly see their mothers. On the initial trial (A trial), the infants in both age groups almost always went to their mothers by the shortest route (e.g. around the right end of the barrier if they stood near that end). On the next trial (B trial), only the 18-month-old infants changed their response and used the shortest route to reach their mothers; the 12-month-old infants tended to repeat their first response, going to their mothers via the same route across trials. The younger infants thus perseverated even though their mothers were visible above the barrier throughout each trial.

Aguiar and Baillargeon (2000) (see also Aguiar & Menard, 2001, 2002) investigated 9- and 11-month-old infants' ability to retrieve a toy placed on as opposed to behind a support. The infants sat in front of two cloths placed side by side. One cloth had a toy on its far end; the other cloth was folded in half and had an identical toy positioned behind it. After the infants succeeded in pulling the correct cloth and retrieving the toy on two consecutive trials (A trials), they received a final trial (B trial) in which the locations of the two cloths and toys were reversed. Only the 11-month-old infants pulled the correct cloth on the final trial; the 9-month-old infants tended to perseverate, pulling the same cloth they had pulled on the previous trials. The younger infants thus erred even though the two cloths and toys remained visible throughout each trial.

1.2. A problem-solving account of perseveration in NME-MO tasks

To explain the perseverative errors observed in NME-MO tasks (e.g. Aguiar & Baillargeon, 2000; Lockman & Pick, 1984; Mackenzie & Bigelow, 1986; Rieser et al., 1982), we have been developing a new account that appeals to limitations in infants' *problem-solving* abilities (for a detailed application of this account to the support task described earlier and related NME-MO tasks, see Aguiar & Baillargeon, 2000). Before describing this account, we need to make two comments about its domain of application. First, our account focuses on NME-MO tasks of the following format: to start, infants receive one or more trials (A trials) in which they are presented with the same problem; next, infants receive one or more trials (B trials) in which they are given a problem that is similar to the initial problem, except that a critical feature has been changed so that the original solution is no longer valid; for infants to succeed, a novel response must be produced. Second, our account applies only to NME-MO tasks infants are capable of solving: perseveration on the B trials is never due to infants lacking the knowledge necessary to respond correctly to the change introduced.

1.2.1. Three assumptions—Our account of perseverative errors in NME-MO tasks rests on three main assumptions (Aguiar & Baillargeon, 2000). The *first* is that, at the start of each A and B trial in a test session, infants conduct an *initial analysis* of the problem to categorize it as novel or familiar; that is, they judge whether the problem is one they are encountering for the first time in the session or one they have encountered earlier in the session. If infants conclude that the problem is novel, they perform a *further analysis* of the problem and compute its solution. In contrast, if infants conclude that the problem is familiar, they simply retrieve their previous solution (for related ideas, see Logan, 1988; Posner & Raichle, 1994; Suchman, 1987).

The *second* assumption of our account is that infants perseverate on a B trial when their initial analysis is too incomplete to allow them to detect the critical change that has been introduced. As a result, infants mistakenly categorize the problem as familiar and retrieve their previous solution, resulting in a perseverative error.

Finally, the *third* assumption of our account is that whether infants perseverate or respond correctly on a B trial depends in large part on their amount of experience at the task: infants with little experience are more likely to perseverate than are infants with more experience. This last assumption is elaborated in the next section.

1.2.2. Effects of experience

Inexperienced infants: When categorizing a problem, infants with little experience at a task tend to perform a relatively *shallow* initial analysis that bypasses many of the problem's critical features (see Fig. 1). On the first A trial, this analysis is sufficient to yield a correct categorization of the problem as novel (because infants are encountering the problem for the first time in the test session, a correct categorization would be expected on almost any analysis, however superficial). Having categorized the problem as novel, infants then conduct a further analysis of the problem and compute its solution. On subsequent A trials, inexperienced infants' shallow initial analyses are again sufficient to yield a correct categorization of the problem as familiar (because the problem is in fact the same as that on the preceding A trials, a correct categorization would again be expected on almost any analysis). Having categorized the problem as familiar, infants do not conduct any further analysis but simply retrieve their previous solution.[1]

Inexperienced infants' shallow initial analyses are thus sufficient to allow them to perform correctly on all of the A trials. Difficulties arise only on the B trials, when the change introduced involves a critical feature that typically falls outside the set of features inexperienced infants spontaneously attend to in their initial analyses. Failure to notice the change leads infants to miscategorize the problem as familiar and to retrieve their prior solution, resulting in a perseverative error.

Experienced infants: Infants with more experience at a task tend to engage on both A and B trials in a *deeper* initial analysis of the problem that takes into account more of the problem's critical features (see Fig. 1). As a result, infants are more likely to detect the change introduced in the B trials, and hence to respond to it appropriately.

1.3. The present research

According to the account described in the last section, the key to perseveration in NME-MO tasks is problem categorization: infants perseverate when their initial analysis of a novel problem leads them to mistakenly categorize it as familiar, so that they retrieve their previous solution rather than compute a new one. In such an account, there is no reason to expect perseveration to be restricted to tasks that call for motor responses: any task that

[1]For the sake of conceptual clarity, we are drawing here a sharp distinction between the first and subsequent A trials. In practice, however, infants may at times recompute the same solution on two or more A trials before shifting to retrieving the solution from memory.

requires infants to generate solutions or expectations across a series of A and B trials could conceivably lead to perseveration.

In the adult problem-solving literature, it has long been known that adults can perseverate in tasks that require solving similar A and B problems in succession. In the classic experiments of A.S. and E.H. Luchins (e.g. Luchins, 1942, 1946; Luchins & Luchins, 1950), adults solved two sets of pencil-and-paper problems in which they had to work out how to obtain a specific volume of liquid from a large tank using any or all of three empty jars (X, Y, and Z). The volume of liquid to be obtained and the capacity (measured in quarts) of the three jars changed from problem to problem. The problems in the first set (set A) were solved most efficiently by the formula $Y - X - 2Z$. All but one of the problems in the second set (set B) could be solved by this formula or more efficiently by simpler formulas such as $X - Z$ or $X + Z$; the remaining problem could be solved only by the $X - Z$ formula. Most adults continued to use the formula they had generated for the first set of problems when solving the second set. Indeed, the subjects' perseverative tendency was so robust that many failed to notice that the original formula led to an incorrect answer in the $X - Z$ problem. These findings make clear that, in "routinized situations" (Medin, Ross, & Markman, 2001), even adults may retrieve familiar solutions instead of computing more efficient or even appropriate ones.

The present research explored the possibility that infants might, under certain conditions, respond perseveratively in a *violation-of-expectation* (VOE) task. In such a task, infants typically see two test events, one consistent (expected event) and one inconsistent (unexpected event) with the expectation being examined in the experiment. With appropriate controls, evidence that infants look reliably longer at the unexpected than at the expected event is taken to indicate that they (1) possess the expectation under examination, (2) detect the violation in the unexpected event, and (3) are surprised by this violation. The term "surprise" is used here simply as a shorthand descriptor, to denote a state of heightened attention or interest induced by an expectation violation.

The possibility that perseverative errors could be found in a VOE task was first suggested by the results of an experiment on 6.5-month-old infants' reasoning about width information in containment events. This experiment built on prior results with 8.5-month-old infants (Aguiar & Baillargeon, 1998, 2000). In the next section, we first describe these prior results and then introduce our experiment with 6.5-month-old infants, which is reported in Experiment 1.

1.3.1. Prior experiments with 8.5-month-old infants—In a series of experiments (Aguiar & Baillargeon, 1998, 2000), we examined 8.5-month-old infants' ability to reason quantitatively and qualitatively about width information in containment events. The distinction between quantitative and qualitative reasoning strategies is derived from computational models of everyday physical reasoning (e.g. Forbus, 1984). A strategy is said to be quantitative if it requires one to encode and use information about absolute quantities (e.g. object O is "this" wide, where "this" stands for some absolute measure of O's width). In contrast, a strategy is said to be qualitative if it requires one to encode and use information about only relative quantities (e.g. object O is wider than object P).

The infants were assigned to a quantitative or a qualitative condition. The infants in both conditions first received familiarization trials in which they saw a large ball being lowered into a wide, shallow container. Next, the infants received test trials in which they saw the same ball being lowered into a tall container that was either much narrower (narrow-container event) or slightly wider (wide-container event) than the ball. In the *quantitative* condition, the container and ball were shown successively at the start of each familiarization and test trial (see Fig. 2); hence, the infants had to remember the absolute width of the ball to predict whether it could be lowered into the container. In the *qualitative* condition, the ball was held directly above the container at the start of each familiarization and test trial (see Fig. 3); the infants could thus judge whether the ball would fit into the container simply by visually comparing their widths.

The infants in both the quantitative and qualitative conditions looked reliably longer at the narrow- than at the wide-container test event. These and control results (Aguiar & Baillargeon, 1998) indicated that the infants (1) realized that the width of the ball relative to that of the container determined whether the ball could be lowered into the container, (2) succeeded in comparing the widths of the ball and container whether these were shown successively or simultaneously, (3) judged that the ball could be lowered into the wide but not the narrow container, and hence (4) were surprised in the narrow-container event when this last expectation was violated.

In Experiment 1, younger, 6.5-month-old infants were tested in the same quantitative and qualitative conditions. Prior research (e.g. Baillargeon, 1991, 1994a, 1995) indicates that, when infants first identify a continuous variable as relevant to an event category, they can reason about the variable qualitatively but not quantitatively: they are not able at first to remember absolute information about the variable. Thus, we predicted that, at some point prior to 8.5 months of age, infants would succeed (i.e. would look reliably longer at the narrow- than at the wide-container test event) in the qualitative but not the quantitative condition. The results of Experiment 1 ran counter to this prediction, and provided the impetus for the present research.

2. Experiment 1

2.1. Method

2.1.1. Participants—The participants were 32 healthy term infants, 16 male and 16 female, ranging in age from 6 months, 3 days to 6 months, 21 days (*M* = 6 months, 13 days). Sixteen infants, eight male and eight female, were randomly assigned to the quantitative (*M* = 6 months, 14 days) and qualitative (*M* = 6 months, 13 days) conditions. Another five infants failed to complete four valid test trials and were eliminated, two because of fussiness, two because of inattentiveness, and one because he looked for the maximum amount of time allowed (60 s) on all four test trials. Infants' names in this and in the following experiments were obtained from birth announcements in the local newspaper. Parents were contacted by letters and follow-up phone calls; they were offered reimbursement for their transportation expenses but were not compensated for their participation.

2.1.2. Apparatus—The apparatus was similar to that in Aguiar and Baillargeon (1998) and consisted of a wooden display box 124 cm high, 102 cm wide, and 57 cm deep that was mounted 76 cm above the room floor. The infant faced an opening 56 cm high and 95 cm wide in the front of the apparatus; between trials, a muslin-covered frame 60 cm high and 101 cm wide was lowered in front of this opening. The floor of the apparatus was covered with contact paper decorated with pastel flowers, and the side walls were painted white. Positioned against each side wall, 19 cm from and parallel to the front of the apparatus, was a large panel constructed of foam-core board and covered with the same contact paper as the floor. The two panels served to restrict the infant's view to the center portion of the apparatus; the right panel (from the infant's perspective) was 74 cm high and 41 cm wide and the left panel was 74 cm high and 29 cm wide. The back wall of the apparatus was constructed of red foam-core board and had a T-shaped opening, filled with a red cloth fringe, centered 10 cm above the floor between the two panels; the top portion of the opening was 30 cm high and 30 cm wide, and the lower portion was 33 cm high and 13 cm wide.

A blue cardboard screen, 33 cm high and 41 cm wide, was used to occlude the area between the two panels. The screen stood in a wooden track at the back of the right panel; a wooden rod attached to the right edge of the screen allowed an experimenter to slide the screen along the track.

The ball used in the events was 15 cm in diameter, made of Styrofoam, and attached to a wooden rod 1 cm in diameter and 10 cm long; the ball and rod were painted green.

Five different containers were used in the events: one in the familiarization event, two in the narrow-container test event, and two in the wide-container test event. All five containers were made of aluminum cans, and all had their top and bottom rims lined with black tape.

The container used in the familiarization event was 21 cm in diameter and 8 cm in height. It was covered on the inside with white paper and on the outside with blue paper decorated with small red dots.

The two containers used in the narrow-container test event were 7.5 cm in diameter and 19 cm in height. Both were covered on the inside with red paper and on the outside with yellow paper decorated with small rainbows. One of the containers was pre-filled with a small ball 5 cm in diameter attached to a rod 20 cm long. The container was sufficiently tall that the infants could not see the ball; only the top 6 cm of the ball's rod protruded above the container. During the narrow-container event, the pre-filled container was surreptitiously substituted for the empty container, to create the illusion that the large ball had been inserted in the container.

The two containers used in the wide-container test event were 15.5 cm in diameter and 19 cm in height. Both were covered on the inside with red paper and on the outside with orange paper decorated with multi-colored balloons. As with the narrow containers, one of the wide containers was pre-filled with a small ball identical to the one described above (this ball was again not visible to the infants; only the top 6 cm of the ball's rod protruded above the container). Although the large ball could simply have been lowered into the empty wide container during the wide-container event, the same substitution procedure was used with the narrow and wide containers to make the test events as similar as possible.

During the events, an experimenter's left hand reached through the opening in the back wall of the apparatus to raise and lower the ball. The experimenter wore a long black glove as well as a bib made of the same red cloth as the fringe in the opening; the bib covered the experimenter's torso and made him or her less noticeable when standing behind the opening. To help the experimenter move the ball at a constant pace, a column of numbered marks was placed on the back wall of the apparatus, next to the opening.

The infants were tested in a brightly lit room. Three 20 W fluorescent light bulbs were attached to the front and back walls of the apparatus to provide additional light. Two frames, each 182.5 cm high and 71 cm wide and covered with blue cloth, stood at an angle on either side of the apparatus; these frames served to isolate the infants from the experimental room.

2.1.3. Events—Three experimenters worked in concert to produce the events: the first raised and lowered the ball, the second manipulated the screen and the pre-filled containers (from an opening in the right wall of the apparatus behind the right panel), and the third manipulated the ball and the empty containers (from an opening in the left wall of the apparatus behind the left panel). To help the experimenters adhere to the events' scripts, a metronome beat softly once per second. The numbers in parentheses indicate the number of seconds taken to perform the actions described.

2.1.3.1. Quantitative condition

Familiarization event: At the start of the familiarization event, the wide, shallow container rested on the apparatus floor, centered between and 5 cm behind the two panels. After the computer signaled that the infant had accumulated 3 s of looking at the container, the first event cycle began. The second experimenter slid the screen to the left until it filled the area between the two panels and hid the container (1 s). After a 1 s pause, the first experimenter's hand introduced the ball into the apparatus through the opening in the back wall; the hand held the ball by the tip of its rod and positioned it about 3 cm above the screen, centered above the hidden container (2 s). Next, the hand lowered the ball behind the screen until it lay centered inside the container (2 s). In this position, the first experimenter's hand and forearm, the ball, and the ball's rod were completely hidden from the infant's view. After a 1 s pause, the second experimenter slid the screen open (2 s) to reveal the hand holding the ball upright in the container (2 s); about half of the ball was visible above the container's rim. The first event cycle ended at this point and was immediately followed by the second event cycle. The second experimenter slid the screen shut to hide the ball and container (1 s). After a 1 s pause, the hand raised the ball to its starting position above the screen (2 s), and then lowered it again inside the container (2 s). After another 1 s pause, the second experimenter slid the screen open (2 s) to reveal the hand holding the ball upright in the container (2 s). The first and second event cycle thus each lasted about 11 s; the second cycle was repeated until the computer signaled that the trial had ended (see below).

The infants saw the familiarization event just described on the second, third, and fourth familiarization trials; on the first familiarization trial, the infants saw a simplified version of the event that involved no screen. At the start of the event, the hand held the ball about 28 cm above the container. After a 1 s pause, the hand lowered the ball inside the container (2 s), paused for 1 s, and then raised the ball back to its starting position (2 s). Each event cycle thus lasted about 6 s; cycles were repeated until the trial ended.

Narrow-container test event: At the start of the narrow-container test event, the empty narrow container stood on the apparatus floor centered between and 12 cm behind the two panels. After the computer signaled that the infant had accumulated 3 s of looking at the container, the first event cycle began. The second experimenter slid the screen shut, thus hiding the container (1 s); the third experimenter then removed the container from the apparatus (1 s). The first experimenter's hand introduced the ball into the apparatus and positioned it above the screen, as in the familiarization event (2 s). Next, the hand lowered the ball behind the screen until it rested on the apparatus floor (2 s). The third experimenter swiftly removed the ball and the second experimenter replaced it with the pre-filled narrow container (1 s). Next, the second experimenter slid the screen open (2 s) to reveal the hand holding the tip of the rod protruding from the container's rim (2 s). The first event cycle ended at this point and was immediately followed by the second event cycle. The second experimenter slid the screen shut to hide the pre-filled container (1 s). Next, the second experimenter removed the pre-filled container and the third experimenter replaced it with the ball (1 s). The hand raised the ball to its starting position above the screen (2 s), and then lowered it again to the apparatus floor (2 s). The ball was exchanged with the pre-filled container (1 s), and the second experimenter opened the screen (2 s) to reveal the hand holding the tip of the rod protruding from the container (2 s). As in the familiarization event, the first and second event cycle thus each lasted about 11 s, and the second cycle was repeated until the trial ended.

Wide-container test event: The wide-container test event was identical to the narrow-container test event, except that the empty and pre-filled wide containers were used; they were positioned between and 8 cm behind the two panels.

2.1.3.2. Qualitative condition

Familiarization event: At the start of the familiarization event, the first experimenter's hand held the ball about 5 cm above the wide, shallow container. After the computer signaled that the infant had accumulated 3 s of looking at the ball and container, the first event cycle began. The second experimenter slid the screen shut, thus hiding the container and ball (1 s); after a 1 s pause, the hand raised the ball about 3 cm above the screen (2 s). From this point on, the trial proceeded exactly as in the familiarization event of the quantitative condition. Each event cycle lasted about 11 s, as in the quantitative condition.

The infants saw the familiarization event just described on all but the first familiarization trial; on that trial, the infants saw a simplified version of the event identical to that shown in the quantitative condition.

Narrow- and wide-container test events: To start, the first experimenter's hand held the ball about 5 cm above the empty narrow (narrow-container event) or wide (wide-container event) container. After the computer signaled that the infant had accumulated 3 s of looking at the ball and container, the first event cycle began. The second experimenter slid the screen shut, thus hiding the container and ball (1 s); the third experimenter then removed the container (1 s). Next, the hand raised the ball above the screen (2 s). From this point on, the trial proceeded exactly as in the narrow- and wide-container test events of the quantitative condition. Each event cycle lasted about 11 s, as in the quantitative condition.

2.1.4. Procedure—Prior to the experiment, each infant was allowed to manipulate, one at a time and for a few seconds, the ball, the familiarization container, and the (empty) test containers. During the experiment, the infant sat on a parent's lap in front of the apparatus. The infant's head was approximately 60 cm from the panels and 98 cm from the back wall of the apparatus. Parents were asked to avoid interacting with their infant during the experiment, to keep the infant seated at all times, and to close their eyes during the test trials.

The infant's looking behavior was monitored by two observers who viewed the infant through peepholes in the cloth-covered frames on either side of the apparatus. The observers could not see the events from their viewpoints and they did not know the order in which the test events were presented. Each observer depressed a button linked to a computer when the infant attended to the events. The looking times recorded by the primary observer were used to determine when a trial had ended (see below).

The infants in the quantitative and qualitative conditions were tested using a two-phase procedure that consisted of a familiarization and a test phase.[2] During the *familiarization* phase, the infants received four trials; they saw the simplified version of the familiarization event on the first trial, and the full version appropriate for their condition on the next three trials. Each trial ended when the infant either (1) looked away from the event for 2 consecutive seconds after having looked for at least 6 (simplified version) or 11 (full version) cumulative seconds or (2) looked at the event for 60 cumulative seconds.

During the *test* phase, the infants again received four trials; they saw the narrow- and wide-container test events appropriate for their condition on alternate trials. In each condition, half of the infants saw the narrow-container event first, and half saw the wide-container event first. Each trial ended when the infant (1) looked away from the event for 2 consecutive seconds after having looked for at least 11 cumulative seconds (one event cycle) or (2) looked at the event for 60 cumulative seconds. The 11 s minimum value was chosen to ensure that the infants had the opportunity to observe the ball being lowered into the narrow or wide container.

To measure interobserver agreement during the test trials, each trial was divided into 100 ms intervals and the computer determined in each interval whether the two observers agreed on whether the infant was looking at the event; the number of intervals in which the observers

[2]The computer program (Barrett-Baillargeon Baby Program) used to conduct the present and other VOE experiments is available free of charge at ftp://www.psych.uiuc.edu/SHARE/babyPrograms/.

agreed was divided by the total number of intervals in the trial. Agreement in this and in the following experiments averaged 93% or more per trial per infant.

Preliminary analyses of the test data in this and in the following experiments revealed no significant interaction involving sex and event; the data were therefore collapsed across sex in subsequent analyses.

2.2. Results

Fig. 4 presents the mean looking times at the test events of the infants in Experiment 1. It can be seen that the infants in the quantitative condition looked longer at the narrow-than at the wide-container event, but that those in the qualitative condition looked about equally at the two events.

The infants' looking times at the test events were averaged and compared by means of a 2 × 2 × 2 mixed-model analysis of variance (ANOVA), with condition (quantitative or qualitative) and order (narrow- or wide-container event first) as between-subjects factors and with event (narrow- or wide-container) as a within-subject factor. Planned comparisons revealed that, whereas the infants in the quantitative condition looked reliably longer at the narrow- (M = 40.5, SD = 15.5) than at the wide-container (M = 31.5, SD = 11.9) event ($F(1, 28) = 4.46$, $P < 0.05$), those in the qualitative condition did not ($F(1, 28) = 0.24$) (narrow-container event: M = 32.3, SD = 11.9; wide-container event: M = 34.4, SD = 13.7). Inspection of the individual infants' mean looking times yielded similar findings: whereas 12 of the 16 infants in the quantitative condition looked longer at the narrow-than at the wide-container event (Wilcoxon T = 30, P = 0.05), only eight of the 16 infants in the qualitative condition did so (Wilcoxon T = 60, $P > 0.05$).

The ANOVA indicated that the main effects of condition ($F(1, 28) = 0.66$) and event ($F(1, 28) = 1.31$, $P > 0.05$) were not significant, nor was the condition × event interaction ($F(1, 28) = 3.39$, $P > 0.05$). Finally, there was a significant event × order interaction ($F(1, 28) = 7.37$, $P < 0.025$). Post-hoc comparisons revealed that the infants who saw the narrow-container event first looked reliably longer at this event (M = 42.0, SD = 13.7) than at the wide-container event (M = 30.4, SD = 13.8) ($F(1, 28) = 7.45$, $P < 0.025$), whereas those who saw the wide-container event first tended to look equally at the two events ($F(1, 28) = 1.23$,$P > 0.05$) (narrow-container event: M = 30.8,SD = 12.8; wide-container event: M = 35.5, SD = 11.3). Because this effect did not interact with condition, it does not bear on the present hypotheses and will not be discussed further.

2.3. Discussion

Like the 8.5-month-old infants in our previous research (Aguiar & Baillargeon, 1998), the 6.5-month-old infants in the quantitative condition looked reliably longer at the narrow- than at the wide-container test event. This result suggested that the infants (1) realized that the width of the ball relative to that of each test container determined whether the ball could be lowered into the container, (2) remembered the absolute width of the ball, (3) judged that the ball could be lowered into the wide but not the narrow container, and (4) were surprised when this last expectation was violated.

In contrast to the 6.5-month-old infants in the quantitative condition, those in the qualitative condition tended to look equally at the narrow- and wide-container test events, suggesting that they failed to notice the violation in the narrow-container event. This negative result differed from our previous results with 8.5-month-old infants, who succeeded in *both* the quantitative and qualitative conditions (Aguiar & Baillargeon, 1998, 2000).

The results of Experiment 1 gave rise to three questions. First, why did the infants in the qualitative condition fail to detect the violation in the narrow-container test event? Second, what development between 6.5 and 8.5 months of age made it possible for the older infants in our previous research to succeed in the qualitative condition? Finally, why did the 6.5-month-old infants in the quantitative condition of Experiment 1 succeed in detecting the violation in the narrow-container event? Each question is addressed in turn.

2.3.1. 6.5-month-old infants in the qualitative condition—Given the successful performance of the 6.5-month-old infants in the quantitative condition, it was unlikely that the infants in the qualitative condition failed to detect the violation in the narrow-container test event because they lacked the physical knowledge necessary to do so. It was also unlikely that the events shown in the qualitative condition required more complex reasoning than did those in the quantitative condition: even for adults, one would predict that comparing the widths of a ball and container would be easier when the two are shown simultaneously as opposed to successively. Furthermore, as mentioned earlier, prior findings indicate that after identifying a continuous variable as relevant to an event category, infants typically succeed at reasoning about the variable first qualitatively and only after some time quantitatively (e.g. Baillargeon, 1991, 1994a, 1995). Finally, it seemed unlikely that the infants in the qualitative condition were more bored or less engaged by their test events than were the infants in the quantitative condition; recall that the main effect of condition in Experiment 1 was not significant, suggesting that there was no reliable difference in the overall looking times of the infants in the two conditions during the test trials.

Why, then, did the infants in the qualitative condition fail to detect the violation in the narrow-container test event? The hypothesis we eventually came to was that the infants were responding to the event *perseveratively*. In line with the problem-solving account we had developed to explain perseverative errors in NME-MO tasks (Aguiar & Baillargeon, 2000), we speculated that during the familiarization trials (A trials), the infants compared the width of the ball to that of the wide, shallow container, and generated an expectation that the ball would fit into the container. During the test trials (B trials), the infants performed a shallow initial analysis of each test event that did not include information about the relative widths of the ball and container, and thus did not specify that a critical change had been introduced in the width of the container. As a result, the infants miscategorized the test events as similar to the familiarization event. This miscategorization in turn led the infants to retrieve the expectation they had formed when watching the familiarization event ("the ball will fit into the container") and to apply it to each test event. Because this expectation was correct for the wide- but not the narrow-container event, the infants failed to detect the violation in the latter event.

2.3.2. 8.5-month-old infants in the qualitative condition—Why did the 8.5-month-old infants in our previous research (Aguiar & Baillargeon, 2000), unlike the 6.5-month-old infants in Experiment 1, succeed in the qualitative condition? One hypothesis, derived from our problem-solving account, was that these older and more experienced infants performed a more complete initial analysis of each test event that *did* include information about the relative widths of the ball and container. As a result, the infants realized, when shown the narrow-container test event, that they faced a novel problem distinct from that shown on the preceding, familiarization trials. The infants therefore computed a novel expectation (e.g. "the ball will not fit into the narrow container"), and responded with increased attention when it was violated.

What factors could explain the development that takes place between 6.5 and 8.5 months of age? We suspect that, with more experience at reasoning about the relative widths of objects and containers, infants come to routinely encode this information in their initial analyses of containment events. Thus, although both 6.5- and 8.5-month-old infants would realize that the width of an object relative to that of a container determines whether the object can be lowered into the container, only the older, more experienced infants would routinely attend to such information in their initial analyses of containment events.

2.3.3. 6.5-month-olds in the quantitative condition—Why did the 6.5-month-old infants in the quantitative condition *not* perseverate in their responses to the test events? During the familiarization trials, the infants had to remember the width of the ball to later compare it to that of the container. This process no doubt required more attention and effort than did, in the qualitative condition, visually comparing the widths of the ball and container as one stood above the other. We speculated that the fact that the width comparison process was more effortful for the infants in the quantitative condition rendered this process more salient for them. As a result, the infants were more likely to include width information in their initial analyses of the test events.

2.3.4. Overall plan of the experiments—Experiments 2, 3, and 4 were designed to test our hypothesis that the 6.5-month-old infants in the qualitative condition of Experiment 1 failed to detect the violation in the narrow-container test event because they responded to this event perseveratively. Experiment 5 examined our hypothesis that the 6.5-month-old infants in the quantitative condition of Experiment 1 did not perseverate because the width comparison they had to perform during the familiarization trials was more challenging and thus more salient for them, so that they were more likely to attend to the relative widths of the ball and container in their initial analyses of the test events. The subjects in all four experiments were 6.5-month-old infants, and all were tested with modifications of the procedure used in the *qualitative* condition of Experiment 1. In these modifications, only the *familiarization* event was changed; the test events remained the same as before.

We reasoned that our interpretation of the results of Experiment 1, if valid, would be important for at least two reasons. First, it would provide the first demonstration that infants can perseverate in VOE tasks – tasks that do not require selecting one of two alternative motor responses. Second, such a demonstration would add to the evidence that there exists some degree of continuity between infant and adult problem solving in routinized situations

(Aguiar & Baillargeon, 2000). From the present perspective, there is clearly a tantalizing similarity between the infant who applies a familiar (but no longer appropriate) expectation to a novel containment event, and the adult who applies a familiar (but no longer efficient or appropriate) solution to a novel volume problem (e.g. Luchins, 1942, 1946; Luchins & Luchins, 1950). We return to this issue in the general discussion in Section 9.

3. Experiment 2

Our interpretation of the negative results obtained in the qualitative condition of Experiment 1 was that, in their initial analysis of the narrow-container test event, the infants did not detect the critical change introduced because they did not attend to the relative widths of the ball and container; this feature was not one of those the infants spontaneously attended to when judging whether the event was similar or different from that shown on the preceding, familiarization trials. This interpretation led to the following prediction: if a change was introduced in a feature the infants *did* attend to in their initial analysis of the event, then they should detect this change and respond non-perseveratively. Experiment 2 was designed to examine this prediction.

We speculated that a feature the infants were very likely to encode when watching each test event was its *category*. Recent research suggests that infants form distinct event categories, such as containment, collision, and support events, and reason and learn in terms of these categories (e.g. Casasola, Cohen, & Chiarello, in press; Hespos & Baillargeon, 2001a; Luo, 2001; McDonough, Choi, & Mandler, in press; Wang, Baillargeon, & Paterson, in press; Wilcox & Baillargeon, 1998; Wilcox & Chapa, 2002; for reviews, see Baillargeon, 2002; Baillargeon & Wang, 2002). In Experiment 1, the familiarization and test events all involved containment events. In Experiment 2, the familiarization event was modified so that containment events were now shown *only* in the test events.

The infants were tested with the same procedure as in the qualitative condition of Experiment 1, except that no container was present in the familiarization trials (see Fig. 5): the ball was lowered to the apparatus floor in a simple display event (a display event is an event in which one or more objects remain stationary or move, without interacting in any way; e.g. Baillargeon & Luo, 2002). The infants thus saw a *display* event during the familiarization trials, and *containment* events during the test trials.[3] We reasoned that the infants should (1) notice that the test events belonged to a different event category than the familiarization event, (2) conclude that the test events were novel, (3) perform a further analysis of each test event to compute an appropriate expectation for it, and (4) be surprised when their expectation for the narrow-container event was violated.

One aspect of the above reasoning may be worth emphasizing: it assumed that providing the infants with *less* information about the test situation during the familiarization trials would

[3]In this discussion, we ignore the segment of each event in which the screen was slid between the two panels to hide the apparatus; instead, we focus on those segments in which the ball was lowered to or lifted from the apparatus floor or container. Because the screen hid a large portion of the apparatus, we suspect that the screen segment of each event (like the lowering and raising of the curtain in front of the apparatus between trials) played relatively little role in the infants' categorization of the primary event the ball was involved in.

result in a better test performance. Such a prediction obviously runs counter to what developmental researchers expect for most test situations.

3.1. Method

The participants in Experiment 2 were 16 healthy term infants, eight male and eight female, ranging in age from 6 months, 4 days to 6 months, 21 days (M = 6 months, 15 days). Another three infants were eliminated, two because of fussiness and one because of inattentiveness.

The apparatus, events, and procedure used in Experiment 2 were identical to those in the qualitative condition of Experiment 1, except that no container was present during the familiarization trials; the ball was simply lowered to the apparatus floor.

3.2. Results

Fig. 6 presents the mean looking times at the test events of the infants in Experiment 2; for comparison purposes, the looking times of the infants in the qualitative condition of Experiment 1 are also shown. It can be seen that the infants in Experiment 2, unlike those in the qualitative condition of Experiment 1, looked longer at the narrow- than at the wide-container event.

The looking times of the infants in Experiment 2 at the test events were averaged and compared by means of a 2 × 2 mixed-model ANOVA, with order (narrow-container event first or wide-container event first) as a between-subjects factor and with event (narrow- or wide-container) as a within-subject factor. The main effect of event was significant ($F(1, 14) = 5.13$, $P < 0.05$), indicating that the infants looked reliably longer at the narrow- (M = 38.1, SD = 11.8) than at the wide-container (M = 29.1, SD = 12.1) event. Thirteen of the 16 infants looked longer at the narrow- than at the wide-container event (Wilcoxon $T = 30$, $P = 0.05$).

The ANOVA also revealed a significant main effect of order ($F(1, 14) = 5.25$, $P < 0.05$), indicating that the infants who saw the narrow-container event first (M = 38.1, SD = 15.1) looked reliably longer overall than did those who saw the wide-container event first (M = 29.1, SD = 7.5).

3.3. Discussion

Unlike the infants in the qualitative condition of Experiment 1, those in Experiment 2 looked reliably longer at the narrow- than at the wide-container test event. This result suggested that the infants (1) qualitatively compared the width of the ball to that of each test container, (2) judged that the ball could be lowered into the wide but not the narrow container, and (3) were surprised when this last expectation was violated.

Why did the infants in Experiment 2 detect the violation in the narrow-container test event, but those in the qualitative condition of Experiment 1 did not? Our interpretation was that when performing their initial analysis of each test event, the infants in Experiment 2 noticed that it belonged to a different event category than the familiarization event. The infants concluded that the test event was novel, and they then proceeded to analyze it further to compute an appropriate expectation for it. In contrast, the infants in the qualitative condition of Experiment 1 noted in their initial analysis of each test event that it again involved a containment event; because they did not attend to another critical feature of the event, the relative widths of the ball and container, the infants concluded that the event was similar to the familiarization event, and they simply retrieved the expectation they had formed for that event.

However, there was another, simpler explanation for the results of Experiment 2. Perhaps the change that the infants detected in their initial analyses of the test events was not a change in event category, as suggested above, but rather a change in the number of objects present. After all, the infants saw only a ball in the familiarization event, and both a ball and a container in the test events. Perhaps the addition of a second object in the test events was sufficient to induce the infants to categorize them as novel. Experiment 3 was designed to examine this alternative explanation.

Recent research (e.g. Casasola et al., in press; Hespos & Baillargeon, 2001a; Luo, 2001; for reviews, see Baillargeon, 2002; Baillargeon & Wang, 2002) suggests that infants view *occlusion* and *containment* events as distinct event categories. Experiment 3 built on these findings. The infants were tested with the same procedure as in the qualitative condition of Experiment 1 with one exception: the back half and bottom of the familiarization container were removed, leaving only its front to serve as a rounded occluder (see Fig. 5; this modification was adapted from Hespos & Baillargeon, 2001a). The ball was lowered behind this rounded occluder to the apparatus floor.

We reasoned as follows. If the infants in Experiment 2 succeeded because they noticed, in their initial analysis of each test event, that the ball was no longer the only object present, then the infants in Experiment 3 might respond perseveratively, because the familiarization and test events always involved the ball and some other object (an occluder or a container). On the other hand, if the infants in Experiment 2 succeeded because they detected, in their initial analysis of each test event, that it belonged to a different event category than the familiarization event, then the infants in Experiment 3 should also succeed, because the familiarization and test events again belonged to different event categories.

4. Experiment 3

4.1. Method

The participants in Experiment 3 were 16 healthy term infants, eight male and eight female, ranging in age from 6 months, 7 days to 6 months, 19 days (M = 6 months, 13 days). Another four infants were eliminated, three because of fussiness and one because he looked for the maximum amount of time allowed on all four test trials.

The apparatus, events, and procedure used in Experiment 3 were identical to those in the qualitative condition of Experiment 1, except that the familiarization container was replaced with an occluder identical to the container with its back half and bottom removed; the ball was lowered to the apparatus floor behind this rounded occluder.

4.2. Results

Fig. 6 presents the mean looking times at the test events of the infants in Experiment 3. It can be seen that, unlike the infants in the qualitative condition of Experiment 1 but like those in Experiment 2, the infants in Experiment 3 looked longer at the narrow- than at the wide-container event.

The looking times at the test events of the infants in Experiment 3 were averaged and analyzed as in Experiment 2. The only significant effect was that of event ($F(1, 14) = 6.68$, $P < 0.025$), indicating that the infants looked reliably longer at the narrow- ($M = 36.3$, SD = 12.4) than at the wide-container ($M = 29.2$, SD = 9.8) event. Thirteen of the 16 infants looked longer at the narrow- than at the wide-container event (Wilcoxon $T = 24.5$, $P < 0.025$).

4.3. Discussion

The infants in Experiment 3 detected the violation in the narrow-container test event. This result suggested that the infants (1) noticed in their initial analyses of the test events that these belonged to a different event category than the familiarization event, (2) categorized the test events as novel relative to the familiarization event, (3) analyzed the test events further to compute novel expectations, (4) determined that the ball could be lowered into the wide but not the narrow container, and finally (5) were surprised when this last expectation was violated.

The results of Experiment 3 also supported our interpretation of the results of Experiment 2. Specifically, they suggested that the infants in Experiment 2 responded non-perseveratively, not simply (or not only) because they detected a change in the number of objects present in the apparatus, but because they realized, in their initial analyses of the test events, that these belonged to a different event category than the familiarization event.

5. Experiment 4

We have argued that the infants in the qualitative condition of Experiment 1 responded *perseveratively*. When performing their initial analysis of each test event, the infants attended to one critical feature of the event, its event category, but they ignored another critical feature, the relative widths of the ball and container. As a result, the infants mistakenly concluded that the event was similar to that shown during the familiarization trials and they retrieved their prior expectation, leading – in the case of the narrow-container event – to a perseverative error.

We have also suggested that the infants in Experiments 2 and 3 responded *non-perseveratively*. When performing their initial analysis of each test event, they detected that the event did not belong to the same event category as the familiarization event. They concluded that the event was novel and proceeded to compute an appropriate expectation for it.

Because the preceding interpretations rested crucially on the negative result obtained in the qualitative condition of Experiment 1, it was important that it be confirmed. Experiment 4 was carried out for this purpose.

The infants were tested using the same procedure as in the qualitative condition of Experiment 1 with one exception: the familiarization container was replaced with a wide rectangular wicker basket (see Fig. 5).We suspected that increasing the perceptual dissimilarity between the familiarization and test containers would have little effect on the infants' performance. Recall that the infants in the qualitative condition of Experiment 1 saw test containers that differed from the familiarization container not only in width, but also in height, pattern, and color; nevertheless, the infants responded perseveratively, suggesting that they paid little attention to these featural differences in their initial analyses of the test events. It thus seemed likely that the infants in Experiment 4 would also respond perseveratively, despite the fact that the familiarization and test containers differed in shape, material, and texture, as well as in width, height, pattern, and color.

5.1. Method

The participants in Experiment 4 were 16 healthy term infants, eight male and eight female, ranging in age from 6 months, 4 days to 6 months, 20 days (M = 6 months, 10 days). One additional infant was eliminated, because he looked for the maximum amount of time allowed on all four test trials.

The apparatus, events, and procedure used in Experiment 4 were identical to those in the qualitative condition of Experiment 1, except that the familiarization container was replaced with a cream-colored wicker basket 23.2 cm wide, 17.3 cm long, and 8.6 cm tall.

5.2. Results

Fig. 6 presents the mean looking times at the test events of the infants in Experiment 4. It can be seen that, like the infants in the qualitative condition of Experiment 1 and unlike those in Experiments 2 and 3, the infants in Experiment 4 looked about equally at the narrow- and wide-container events.

The looking times at the test events of the infants in Experiment 4 were averaged and analyzed as in Experiment 2. The main effect of event was not significant ($F(1, 14) = 0.43$), indicating that the infants tended to look equally at the narrow- ($M = 36.4$, SD = 15.6) and wide-container ($M = 34.0$, SD = 12.7) events. Only seven of the 16 infants looked longer at the narrow- than at the wide-container event (Wilcoxon $T = 59$, $P > 0.05$).

5.3. Discussion

Like the infants in the qualitative condition of Experiment 1, and unlike those in Experiments 2 and 3, the infants in Experiment 4 failed to detect the violation in the narrow-container test event. Together, these results provide strong support for the hypothesis that, when performing their initial analysis of each test event, the infants in the present research attended to the category of the event, but not to another critical feature, the relative widths of the ball and container.

The results obtained in the qualitative condition of Experiment 1 and in Experiment 4 also suggested that the infants tended to pay little attention, in their initial analyses of the test events, to the featural properties of the containers. In Experiment 4, the test containers differed from the familiarization container not only in width but also in shape, material, texture, height, pattern, and color; nevertheless, the infants responded perseveratively, retrieving the same expectation they had generated for the familiarization event.

In all our discussions of perseveration in NME-MO and VOE tasks so far, we have focused primarily on infants' ability to detect *critical* changes. However, the ability to ignore *non-critical* changes is undoubtedly an important part of efficient problem solving. Our remarkable capacity for re-using familiar solutions would be severely limited if it could be invoked only with problems in all respects identical to ones encountered before. The better we are at ignoring non-critical changes when categorizing problems as familiar or novel, the more efficient we can be as problem solvers.

The results of the qualitative condition in Experiment 1 and of Experiment 4 raise interesting questions for future research about infants' responses to non-critical changes. Why did the infants respond perseveratively, despite the many featural differences between the familiarization and test containers? At least two possibilities come to mind. The first is that these differences were actually not very salient to the infants, perhaps because they involved the stationary containers rather than the moving ball. On this view, the infants might have responded non-perseveratively had they been shown, for example, a ball with a different pattern and color during familiarization and test. A second possibility is that the infants readily detected the marked featural differences between the familiarization and test containers, but tended to ignore or dismiss these differences in their initial analyses of the test events because they viewed them as largely irrelevant. To draw an analogy with an adult problem-solving situation, consider once again the experiments by Luchins and Luchins (e.g. Luchins, 1942, 1946; Luchins & Luchins, 1950) described in the introduction in Section 1. The subjects obviously realized that the volume of liquid to be obtained and the capacity of the three jars changed from trial to trial; nevertheless, they tended to retrieve the same formula across trials because they viewed the problem as remaining essentially the same despite these superficial changes.

Did the infants in the qualitative condition of Experiment 1 and in Experiment 4 simply fail to notice the changes in the featural properties of the test containers? Or did they register these changes, but dismiss them as unimportant (as indeed most of them were: the material, texture, height, pattern, and color of a container should not affect whether a ball of a given diameter can be lowered into it)? Additional research is needed to answer these questions.

6. Comparison of Experiments 1–4

In an overall analysis of Experiments 1–4, we compared the test responses of two groups of infants: those in the qualitative condition of Experiment 1 and in Experiment 4, who saw containment events during familiarization and test (same-event group, n = 32), and those in Experiments 2 and 3, who saw containment events during test but not familiarization (different-event group, n = 32). The infants' looking times were averaged and compared by means of a 2 × 2 × 2 mixed-model ANOVA, with group (same- or different-event) and order (narrow-container event first or wide-container event first) as between-subjects factors and with event (narrow- or wide-container) as a within-subject factor. Both the main effect of event ($F(1, 60) = 4.62$, $P < 0.05$) and the group × event interaction ($F(1, 60) = 4.30$, $P < 0.05$) were significant. Planned comparisons confirmed that the infants in the different-event group looked reliably longer at the narrow- ($M = 37.2$, SD = 11.9) than at the wide-container ($M = 29.1$, SD = 10.8) event ($F(1, 60) = 8.94$, $P < 0.005$), but that those in the same-event group tended to look equally at the two events ($F(1, 60) = 0.00$) (narrow-container: $M = 34.3$, SD = 13.8; wide-container: $M = 34.2$, SD = 13.0).

6.1. A physical-reasoning account?

Could the negative results obtained in the same-event group be explained in terms of standard physical-reasoning limitations? Is it truly necessary to appeal to a perseveration account to make sense of these data?

To date, negative results in VOE tasks designed to examine infants' physical reasoning have typically been attributed to one of four causes: (1) infants cannot detect the violation in the unexpected event because they have not yet identified the variable being tested in the experiment (e.g. Aguiar & Baillargeon, 1999; Baillargeon & DeVos, 1991; Baillargeon, Needham, & DeVos, 1992; Hespos & Baillargeon, 2001a; Kotovsky & Baillargeon, 1998; Wang et al., in press; Wilcox, 1999); (2) infants have identified the relevant variable, but this variable is a continuous one and infants have not yet developed the ability to reason about it quantitatively; as a result, when tested under quantitative conditions, infants fail to detect the violation in the unexpected event (e.g. Baillargeon, 1991, 1994a, 1995; Luo & Baillargeon, 2003); (3) infants know the relevant variable and can detect the violation in the unexpected event, but they are able to generate an explanation for this violation and so do not respond to it with increased attention overall (e.g. Aguiar & Baillargeon, 2002; Baillargeon, 1994b); and finally (4) infants know the relevant variable and could, under better testing circumstances, detect the violation in the unexpected event, but they are confused or distracted by aspects of the experimental procedure and process the test events only superficially (e.g. Baillargeon, 1995, 2000; Baillargeon, Kotovsky, & Needham, 1995).

None of the preceding explanations could plausibly account for the discrepant results of the same- and different-event groups. First, it is clear from the results of the different-event group that 6.5-month-old infants have identified width as a containment variable – that they realize that the relatives widths of an object and container determine whether the one can be lowered into the other. Second, the infants in the same- and different-event groups were all tested under qualitative, not quantitative conditions; moreover, recall that the infants in the quantitative condition of Experiment 1, like those in the different-event group, did succeed in detecting the violation in the narrow-container test event (we return to this finding in Experiment 5). Third, it is difficult to see why the infants in the same- but not the different-event group would have been able to generate an explanation for the narrow-container event (e.g. "a small ball was surreptitiously substituted for the large ball"), as both groups saw exactly the same test events. Finally, there is no evidence that the infants in the same-event group were confused or distracted by the test procedure; such confusion typically manifests itself by a low level of looking overall, which was not the case. In the ANOVA reported above, the main effect of group was not significant ($F(1, 60) = 0.19$), suggesting that the overall mean looking times of the infants in the same- ($M = 34.3$, SD = 13.3) and different-event ($M = 33.2$, SD = 12.0) groups did not differ reliably.

There is thus no reason to suppose that a physical-reasoning account could explain the negative results of the same-event group better or more parsimoniously than the perseveration account we have offered.

7. Experiment 5

The results of Experiment 1 indicated that the infants in the quantitative condition, unlike those in the qualitative condition, succeeded in detecting the violation in the narrow-container test event. As mentioned earlier, our interpretation of this positive finding was that having to remember the width of the ball at the start of the familiarization event to later compare it to that of the container (quantitative condition) required more attention and effort than visually comparing the widths of the ball and container as one stood above the other (qualitative condition). Because the width comparison process was more effortful for the infants in the quantitative condition, it became more salient for them. As a result, the infants were more likely to include width information in their initial analyses of the test events, and hence to respond non-perseveratively.

These speculations predicted that infants might respond non-perseveratively if tested with a procedure similar to that of the qualitative condition in Experiment 1, but with a familiarization event requiring a *more difficult* width comparison. Experiment 5 was designed to examine this prediction: the opening of the familiarization container was greatly reduced, so that the ball now fit tightly into the container (see Fig. 5). In Experiment 1, the familiarization container was 6 cm wider than the ball, so that even a cursory visual comparison was sufficient for the infants to ascertain that the ball could be lowered into the container. We reasoned that making the opening of the container smaller would require the infants to pay closer attention to (or to engage in a more careful scrutiny of) the relative widths of the ball and container. Such a process might have the effect of making the width comparison process more salient for the infants, and thus of inducing them to attend to the relative widths of the ball and container in their initial analyses of the test events, thereby ensuring a non-perseverative performance.

7.1. Method

The participants in Experiment 5 were 16 healthy term infants, eight male and eight female, ranging in age from 6 months, 5 days to 6 months, 18 days (M = 6 months, 11 days). Another two infants were eliminated because of drowsiness.

The apparatus, events, and procedure used in Experiment 5 were identical to those in the qualitative condition of Experiment 1, except that a black cover with a central circular opening 15.3 cm in diameter was placed on top of the familiarization container; with this cover added, the ball (which was 15 cm in diameter) now fit closely into the container.

7.2. Results

Fig. 7 presents the mean looking times at the test events of the infants in Experiment 5; for comparison purposes, the looking times of the infants in the quantitative and qualitative conditions of Experiment 1 are also shown. It can be seen that the infants in Experiment 5, unlike those in the qualitative condition of Experiment 1 but like those in the quantitative condition, looked longer at the narrow- than at the wide-container event.

The looking times at the test events of the infants in Experiment 5 were averaged and analyzed as in Experiment 2. The main effect of event was significant ($F(1, 14) = 6.88$, $P < 0.025$), indicating that the infants looked reliably longer at the narrow- ($M = 38.6$, SD = 11.9) than at the wide-container ($M = 30.9$, SD = 9.8) event. Twelve of the 16 infants looked longer at the narrow- than at the wide-container event (Wilcoxon $T = 24$, $P < 0.025$).

7.3. Discussion

The infants in Experiment 5, unlike those in the qualitative condition of Experiment 1, looked reliably longer at the narrow- than at the wide-container test event. Why did reducing the opening of the familiarization container help the infants detect the violation in the narrow-container event? We believe that this manipulation made it more difficult for the infants to judge whether the ball could be lowered into the container. Because the width comparison process required more attention and effort from the infants, it became more salient for them. As a result, the infants were more likely to consider the relative widths of the ball and container when performing their initial analyses of the test events. The infants were thus more likely to correctly categorize the test events as novel, to compute appropriate expectations for them, and to respond with increased attention when their expectation for the narrow-container event was violated.

The results of Experiment 5 also supported the hypothesis that the infants in the quantitative condition of Experiment 1 succeeded because the task of remembering the absolute width of the ball and comparing it to that of the familiarization container required additional attention and effort; this apparently had the effect of highlighting the width comparison process for the infants, making it more likely that they would include width information in their initial analyses of the test events.

Together, the results of the quantitative condition in Experiment 1 and of Experiment 5 thus suggest that, although 6.5-month-old infants do not routinely attend to width information in their initial analyses of containment events, it is possible, through various contextual manipulations, to induce them to do so. The manipulations employed here consisted of initial containment trials that required a more challenging width comparison – either because it could be performed only quantitatively (Experiment 1), or because it involved the qualitative comparison of a ball and container very similar in width (Experiment 5). These trials highlighted the width comparison process for the infants, and as such induced them to again attend to the relative heights of the ball and container in their initial analyses of the test events.

The present findings are generally consistent with evidence from the literature on adults' skill acquisition that initial training trials influence the kinds of strategies and computations adults use on later trials when processing novel stimuli (e.g. Doane, Alderton, Sohn, & Pellegrino, 1996; Kerr & Booth, 1978; Medin & Bettiger, 1994; Pellegrino, Doane, Fisher, & Alderton, 1991; Schmidt & Bjork, 1992). For example, in a visual discrimination task (e.g. Doane et al., 1996), adults were found to be significantly better at judging whether two random polygon stimuli were similar or different if they were initially trained on pairs of polygons that were highly similar and thus required a precise comparison strategy, as opposed to pairs that were highly dissimilar and thus could be judged by means of a more global comparison strategy.

Before leaving this section, we would like to return briefly to the positive results of Experiments 2 (in which the infants saw no container during the familiarization trials) and 3 (in which the infants saw an occluder during these trials). We have argued that these infants succeeded because they detected the change in event category between familiarization and test and this change led them to categorize the test events as novel. This reasoning can explain why the infants responded non-perseveratively in the *first* test trial, but cannot explain why they did so *across* all four test trials. Why didn't the infants apply the containment expectation they had generated for the first test trial to all subsequent trials? It is perhaps not surprising that the infants who saw the narrow-container event first did not perseverate; the violation in the event would have caused them to pay special attention in subsequent trials. However, what of the infants who saw the wide-container event first? Why did they not respond perseveratively on the remaining test trials?

At least two possibilities exist. One is that infants are unlikely to perseverate when a change in container is introduced after a single trial: recall that the infants in Experiments 2 and 3 would have seen the wide container in the first test trial and the narrow container in the second trial. In the qualitative condition of Experiment 1 and in Experiment 4, where perseveration was observed, the infants saw the same ball and container on four familiarization trials prior to the test trials.

Another possible explanation for why the infants in Experiments 2 and 3 who saw the wide-container test event first did not respond perseveratively in the subsequent test trials is that, because the ball and wide test container were very similar in width (recall that the ball was 15 cm in diameter and the wide container 15.5 cm), the infants had to compare them carefully to determine whether containment was possible – as did the infants in Experiment 5, who saw the wide, shallow container with a reduced opening during the familiarization trials. This careful comparison would have rendered the infants more likely to include width information in their initial analyses of subsequent events, thereby ensuring a successful performance.

8. Additional analyses

We have argued that the infants in the quantitative condition of Experiment 1 and in Experiment 5 did not perseverate during the test trials because they were forced to perform a more difficult width comparison during the familiarization trials. Analyses of the familiarization data support the notion that the infants who faced a more difficult width comparison responded differently than did those who faced an easier comparison.

In these analyses, we compared the familiarization responses of two groups of infants: (1) a difficult-comparison group (*N* = 32), which involved the infants in the quantitative condition of Experiment 1 and in Experiment 5; and (2) an easy-comparison group (*N* = 32), which involved the infants in the qualitative condition of Experiment 1 and in Experiment 3. The infants in both groups saw at least a portion of the wide, shallow container during familiarization (those in Experiment 1 saw the whole container, those in Experiment 3 saw the front of the container serving as a rounded occluder, and those in Experiment 5 saw the whole container with a reduced opening; recall that, in contrast, the infants in Experiment 2 saw no container and those in Experiment 4 saw a very different container, a wicker basket). The analyses focused on the infants' responses during the last three familiarization trials, when they were shown the full as opposed to the simplified version of the familiarization event.

Because many infants in the difficult- and easy-comparison groups looked for the maximum amount of time allowed, 60 s, during the familiarization trials, the first analysis compared, on each trial, the numbers of infants in the two groups who looked for 60 s. No reliable difference was found on the first trial (25/32 infants in the difficult-comparison group and 20/32 infants in the easy-comparison group, $\chi_1^2=1.88$, $P > 0.05$), or on the second trial (20/32 in each group, $\chi_1^2=0.0$), but one was found on the third trial: whereas 16 of the 32 infants in the difficult-comparison group looked for 60 s, only eight of the 32 infants in the easy-comparison group did so ($\chi_1^2=4.26$, $P < 0.05$).

A second analysis compared the mean looking times of the infants in the difficult- and easy-comparison groups on each trial. Once again, no reliable difference emerged on the first trial (difficult-comparison: $M = 54.2$, SD = 13.4; easy-comparison: $M = 48.1$, SD = 17.7; $F(1, 124) = 3.39$, $P > 0.05$), or on the second trial (difficult-comparison: $M = 51.2$, SD = 15.8; easy-comparison: $M = 48.9$, SD = 17.8; $F(1, 124) = 0.45$), but one was found on the third trial ($F(1, 124) = 5.35$, $P < 0.025$): the infants in the difficult- comparison group ($M = 46.8$, SD = 15.1) looked reliably longer than did those in the easy-comparison group ($M = 39.1$, SD = 17.8).

The data obtained in the third familiarization trial thus support the notion that the infants in the difficult-comparison group faced a more challenging task than did those in the easy-comparison group. During the first two familiarization trials, the infants in both groups tended to be at or near ceiling in their looking times; on the third trial, however, as the infants began to habituate to the familiarization event, differences consistent with our account began to emerge. The present design, involving as it did only three full-version familiarization trials, was not ideally suited for uncovering such differences, but it should be kept in mind that the discovery of perseverative responding in Experiment 1 was a fortuitous one, and that our initial design was never intended to track differences such as those pursued here. Future designs will involve multiple trials in order to track relevant differences more effectively.

9. General discussion

The present experiments revealed three main findings. First, the infants in the qualitative condition of Experiment 1 and in Experiment 4 responded perseveratively to the narrow- and wide-container test events. When performing their initial analysis of each test event, the infants noticed that it again involved a containment event, but failed to attend to the relative widths of the ball and container and hence did not realize that a change had been introduced in this critical feature. As a result, the infants miscategorized the event as similar to that shown on the familiarization trials, and they retrieved the expectation they had formed during these trials ("the ball will fit into the container"). Because this expectation was correct for the wide- but not the narrow-container event, the infants failed to detect the violation in the latter event.

Second, the infants in Experiments 2 and 3 responded non-perseveratively to the narrow- and wide-container test events. In their initial analyses of the events, the infants noted that these belonged to a different event category (containment) than the familiarization event (display in Experiment 2, occlusion in Experiment 3). The infants therefore categorized the events as novel and analyzed them further to compute appropriate expectations for them. When their expectation for the narrow-container test event was violated, the infants responded with increased attention.

Third, the infants in the quantitative condition of Experiment 1 and in Experiment 5 also responded non-perseveratively to the narrow- and wide-container test events, albeit for a different reason than the infants in Experiments 2 and 3. When computing an expectation for the outcome of the familiarization event, the infants had to devote more attention and effort to compare the widths of the ball and container, either because (1) this comparison could be carried out only quantitatively (Experiment 1) or (2) the comparison could be carried out qualitatively but the opening of the container was greatly reduced so that a careful inspection was necessary to determine whether the ball could be lowered into the container (Experiment 5). The increased attention and effort required by the width comparison process rendered it more salient, making it more likely that the infants would continue attending to the widths of the ball and container in subsequent trials.

The present findings are important for several reasons. First, they provide the first experimental demonstration that infants can respond perseveratively in VOE tasks. Until now, demonstrations of perseverative responding in infancy have involved only ME-MO and NME-MO tasks (e.g. Aguiar & Baillargeon, 2000; Diamond, 1985; Lockman & Pick, 1984; Smith et al., 1999), which require infants to select one of two alternative motor responses on each trial (e.g. to search for an object in a left or right well, or to go to a parent around the left or right end of a barrier). Second, the present results are readily explained in terms of the problem-solving account we had originally developed to explain infants' perseverative errors in NME-MO tasks (Aguiar & Baillargeon, 2000), and as such attest to the breadth and usefulness of this account. We do not wish to claim that our account can explain *all* forms of perseveration in infants – clearly, it cannot. In particular, there is no mechanism in the account to explain the delay effects typical of ME-MO tasks: infants in these tasks typically succeed under no-delay conditions, which suggests that (1) infants readily detect the critical change introduced in the B trials and (2) some additional mechanism is needed to explain why infants, having initially registered this critical change, cease attending to it over time (e.g. Diamond, 1985). Nevertheless, the fact that our account can explain perseveration in both NME-MO and VOE tasks, and can distinguish between manipulations likely to produce (e.g. Experiment 4) and to prevent (e.g. Experiments 2, 3, and 5) perseverative responding, encourages attempts to develop it further.

Finally, a third reason why the present research is important is that it emphasizes the continuity between the problem-solving abilities of infants and adults. Like adults, infants attempt to be efficient problem solvers who re-use previous solutions when possible, and like adults, infants sometimes err in re-using solutions in modified contexts that call for novel solutions. Although to our knowledge no researcher has yet examined whether, in routinized problem-solving situations, adults with less experience at a task are more likely to perseverate than are adults with more experience, it would not be surprising if this turned out to be the case. After all, there already exists a vast literature indicating, for example, that adult experts and novices differ in the information they attend to and use when categorizing and solving problems (e.g. Chi, Feltovich, & Glaser, 1981; Murphy & Wright, 1984; Reitman, 1976; Tanaka & Taylor, 1991).

The sense of continuity in problem solving one gains from the present research is especially important in that it comes at a time when the issue of how much continuity there exists between infant and adult cognition is hotly debated (e.g. Bogartz, Shinskey, & Speaker, 1997; Haith & Benson, 1998; Thelen & Smith, 1994). It would be difficult to explain the present results without assuming that (1) infants use their physical knowledge to compute expectations about the likely outcomes of events and (2) in situations where similar events recur, infants tend to retrieve rather than recompute their expectations. The general picture that emerges from the present research is thus one in which infants, far from being cognitively destitute, possess problem-solving abilities that are certainly rudimentary but still on a continuum with those of adults.

9.1. Links to other investigations

How do the present results relate to other findings in the infant physical reasoning literature? In this section, we first examine the results of Sitskoorn and Smitsman (1995), which appear at first glance to be inconsistent with our own. Next, we consider VOE tasks generally and ask why perseverative responding is not a common phenomenon. Finally, we discuss how the present results might be integrated with recent findings on the development of infants' physical reasoning.

9.1.1. Prior results with containment events—Sitskoorn and Smitsman (1995) obtained results with 6-month-old infants that appear inconsistent with those of the present research. Their infants were habituated to two events presented on alternate trials. In both events, a block was repeatedly lowered into a box; one event involved a small block and a box with a small opening, and the other event involved a large block and a box with a large opening. Following habituation, the block and box pairs were rearranged, and the infants saw two test events. In one (large-opening event), the *small* block was lowered into the box with the *large* opening. In the other event (small-opening event), the *large* block was lowered into the box with the *small* opening (the box's side rims were partly collapsible). The results indicated that the infants looked reliably longer at the small- than at the large-opening event.

Why did the infants not perseverate during the test trials and look equally at the test events they were shown, like the infants in the qualitative condition of Experiment 1 and in Experiment 4? One possible explanation has to do with the fact that Sitskoorn and Smitsman (1995) changed *both* the block and the container on alternate habituation as well as test trials. In our experiments, the ball remained the same throughout the experimental session; only the container was changed. As was mentioned earlier, it may be that when observing an event in which an object is lowered into a container, infants have a natural bias to attend more to the moving object than to the stationary container. Such a bias could lead infants to include information about the featural properties of the moving object in their initial analysis of the event on each trial. A change in the width of the moving object would thus be likely to be detected, leading to a correct categorization of the event as novel.

The preceding speculations suggest that 6.5-month-old infants tested with the present procedure might be less likely to respond perseveratively if the width of the ball rather than that of the container was changed between familiarization and test. For example, infants might see events involving a medium-size container and a small ball (familiarization event), a medium ball (medium-ball test event), or a large ball (large-ball test event). If the infants did include information about the width of the ball in their initial analyses of the events, then they should succeed in detecting the violation in the large-ball test event. In other words, the infants should respond non-perseveratively, like those tested by Sitskoorn and Smitsman (1995) and unlike those tested in the qualitative condition of Experiment 1 and in Experiment 4.

9.1.2. Perseveration in other VOE tasks?—One concern raised by the present research might be the following: why has evidence of perseverative responding in a VOE task not been reported until now? If infants can respond perseveratively to repeated events, and thus fail to detect violations they would otherwise easily detect, why do such errors not occur more frequently?

We believe that at least two factors may have contributed to the rarity of perseverative errors in VOE tasks to date. First, many VOE tasks have presented infants with events from different event categories during familiarization and test (e.g. Baillargeon, 1986; Baillargeon, Spelke, & Wasserman, 1985; Hespos & Baillargeon, 2001a, b; Kotovsky & Baillargeon, 2000; Spelke, Kestenbaum, Simons, & Wien, 1995). As shown in the present research, infants typically attend to and detect these category changes, making them less likely to respond perseveratively. Second, VOE tasks presenting infants with events from the same event category in familiarization and test have often involved changes across or within trials in the featural properties or trajectories of moving objects (e.g. Baillargeon & DeVos, 1991; Kotovsky & Baillargeon, 1998; Sitskoorn & Smitsman, 1995; Spelke et al., 1992; Wilcox, 1999; Wilcox & Baillargeon, 1998). One example of such a VOE task is that of Sitskoorn and Smitsman (1995), which was described in the last section. Another example comes from the work of Spelke, Breinlinger, Macomber, & Jacobson (1992). In one experiment, 4-month-old infants were habituated to the following event sequence. To start, a 10 cm ball was held above a large horizontal surface positioned 15 cm above the apparatus floor; the ball was centered above an 11 cm gap in the surface. Next, a screen was raised to hide the gap, and the ball was dropped behind the screen. Finally, the screen was removed to reveal the ball resting on the apparatus floor below the gap. Following habituation, the infants saw two test events similar to the habituation event, except that a smaller (6 cm) or a larger (17 cm) ball was used. The infants looked reliably longer at the large- than at the small-ball event, suggesting that they responded non-perseveratively to the events: they judged that the small but not the large ball could pass through the gap in the horizontal surface, and they were surprised when this last expectation was violated.[4]

The preceding speculations suggest that a constellation of factors may be needed to elicit perseveration in a VOE task: (1) infants should be shown events from the same event category in familiarization and test; (2) changes across trials should involve stationary as opposed to moving objects; and (3) infants should be relatively inexperienced still at reasoning about the variable being manipulated in the experiment (recall that older, 8.5-month-old infants do not respond perseveratively when tested with the procedure used in the qualitative condition of Experiment 1; Aguiar & Baillargeon, 2000).

Keeping in mind these various factors, Scott and Aguiar (2002) recently conducted an experiment designed to elicit additional evidence of perseverative responding in a VOE task. This experiment built on results by Hespos and Baillargeon (2001a) which indicate that, at about 7.5 months of age, infants identify height as a containment variable: they begin to appreciate that the height of an object relative to that of a container determines whether the object can be fully or only partly hidden inside the container. Scott and Aguiar first showed 9-month-old infants the following familiarization event. To start, an experimenter’s hand held a short rod attached to the top of a tall cylindrical object; the object stood next to a very tall container. After a pause, a screen hid the object and container. Next, the hand lifted the object above the screen and then lowered it into the container. Finally, the screen was removed to reveal the container with the experimenter’s arm protruding from it. Following familiarization, the infants saw two test events similar to the familiarization event, except that the container was either as tall as the object minus the rod (tall-container event) or much shorter than the object (short-container event). At the end of each event, the screen was removed to reveal the hand holding the rod protruding from the container. Infants in a display condition saw events similar to those shown in the first, containment condition, except that the very tall container was absent in the familiarization event (as in the present Experiment 2).

The results indicated that the infants in the display condition detected the violation in the short-container test event, but that those in the containment condition did not, presumably

[4]A few experiments present an exception to the preceding discussion (e.g. Aguiar & Baillargeon, 1999, 2002). In these experiments, infants aged 2.5–3.5 months saw occlusion events in both familiarization and test, the same small object moved back and forth across the apparatus in familiarization and test, and the changes introduced in test involved the very large, stationary occluder. The infants in these experiments did not respond perseveratively, perhaps because very young infants rarely respond perseveratively, or perhaps because changes in a stationary object are more salient when this object is very large.

because they were responding to the test events perseveratively. In a second experiment, older, 10.5-month-old infants tested with the containment condition procedure responded non-perseveratively. Together, these results not only confirm the present findings, but also make clear that a maturational account of the discrepant responses of the 6.5- and 8.5-month-old infants in the present research and in Aguiar and Baillargeon (2000) is unlikely to be correct. Evidence that 8.5- or 9-month-old infants respond perseveratively when reasoning about the containment variable *height*, which they only recently identified (e.g. Hespos & Baillargeon, 2001a), but not about the containment variable *width*, which they identified several months earlier (e.g. Wang, Baillargeon, & Brueckner, 2003), underscores the role of experience in perseveration in VOE tasks (for evidence concerning the role of experience in NME-MO tasks, see Aguiar & Baillargeon, 2000).

9.1.3. The development of infants' physical reasoning—Prior research has revealed two distinct steps in the development of infants' reasoning about a variable in an event category; the present findings suggest that a third step may also be involved. In the *first* step, infants identify the variable (for a discussion of this process, see Aguiar & Baillargeon, 2002; Baillargeon, 2002) and come to routinely include information about it when representing events from the category. We have suggested that, when watching a physical event, infants build a specialized physical representation that is used to predict and interpret the event's outcome. Early in the representation process, infants categorize the event and then access their knowledge of the event category selected. This knowledge specifies the variables that have been identified as relevant to the category, and hence that should be included in the physical representation. Thus, as infants identify more and more of the variables relevant to an event category, their physical representations become more and more complete. In the initial stages, however, when few variables have been identified, infants' physical representations tend to be sparse and impoverished.

To illustrate these points, consider the following situation. Adults watching an object being lowered toward a container might attend to: (1) the relative widths of the object and container (to judge whether the object could be inserted into the container); (2) the relative heights of the object and container (to determine whether the object could be fully hidden inside the container); and (3) the transparency of the container (to judge whether the object could be seen through the walls of the container). In contrast, infants aged 4–6 months might include only the width information in their physical representation of the event: recent research suggests that, although width is identified at about 4 months of age as a containment variable (Wang et al., 2003), height is not identified until about 7.5 months (e.g. Hespos & Baillargeon, 2001a), and transparency until about 10 months (e.g. Luo, 2001).

After infants have identified a variable as being relevant to an event category, a *second* development takes place in the case of continuous variables. As was discussed earlier, infants who have just identified a continuous variable can typically reason about it qualitatively, but not quantitatively (e.g. Baillargeon, 1991, 1994a, 1995; Luo & Baillargeon, 2003). Perhaps infants learn with experience that they cannot rely on qualitative information about a continuous variable always being provided as an event unfolds, and that including absolute information about the variable ensures that such information will be available if needed to predict the event’s outcome.

Infants who have identified a variable as being relevant to an event category, and can reason both qualitatively and quantitatively about the variable, may still fail to reveal these achievements if tested, as in the present research, with repeated events from the category. Under these conditions, infants do not engage in the careful monitoring of variables described above, but limit themselves to an initial analysis aimed at determining whether the event before them is similar to that presented before (in which case they retrieve their prior expectation) or different from it (in which case they compute a novel expectation). The *third* development uncovered in the present research is that infants who are less experienced at reasoning about a variable in an event category are less likely to spontaneously include information about the variable in their initial analyses, and hence are more likely to perseverate when critical changes in the variable go undetected. The development of infants’ reasoning about a variable in an event category thus appears to be relatively complex and protracted.

9.2. Future directions

Throughout the paper, we have alluded to various ways in which we plan to build, or are already building, on the present results. Some of these new experiments are intended to provide converging evidence of perseverative responding in VOE tasks. Others are intended to explore the conditions necessary to elicit such responses. Yet others examine the contents of the initial analyses performed by infants with more or less experience at a task, and attempt to specify what factors (in and out of the laboratory) can bring about changes in these contents.

Beyond these more immediate investigations, we hope to continue expanding our problem-solving account by considering perseveration in tasks other than NME-MO and VOE tasks. One interesting challenge will be to understand how our research relates to tasks that require the retrieval of familiar solutions in non-routinized settings (i.e. subjects do not receive multiple trials involving similar problems); what might be perseverative errors sometimes occur when modest task distortions – which render familiar solutions invalid – go undetected.

To illustrate, consider tasks in which adults are asked questions with a distorted term, such as “How many animals of each kind did *Moses* take on the ark?”. Most adult subjects answer such questions without noticing that Noah has been replaced with Moses (e.g. Erickson & Mattson, 1981; Reder & Kusbit, 1991). Reder and Kusbit (1991) argued that subjects fail to notice the distortion because they make an incomplete match between their representation of the question and their previously stored answer. However, adults readily detect the distortion if, for instance, Nixon replaces Noah (Erickson & Mattson, 1981). According to Reder and Kusbit, this is because the subjects’ partial match process is sensitive to the level of conceptual similarity between the words in the distorted question and the subject’s previously stored answer. Subjects overlook substitutions such as Moses for Noah because the two words invoke similar concepts (i.e. biblical characters); however, they detect substitutions such as Nixon for Noah because the two words invoke different concepts (i.e. a presidential figure and a biblical character). We are struck by the parallels between these findings and those of Experiments 2 and 3, where infants shown events from different event categories during familiarization and test responded non-perseveratively.

Much younger subjects may also fall prey to the same sort of retrieval error under mild distortion conditions. An anecdote involving Baillargeon's (RB) son Antoine (AD), aged 28 months, illustrates this point well. One morning, RB asked AD to play a guessing game with her; she would describe various objects and he would guess what they were (Baillargeon, 1993). RB said she was thinking of an animal with a very long neck, and AD correctly guessed a giraffe. RB then said she was thinking of something AD put on his feet when he went outside to keep his toes warm (it was then winter), and AD correctly guessed boots. Later that day, RB asked AD to play their guessing game a second time (showing off for Daddy, of course). She first said she was thinking of an animal with a very long neck, and AD again correctly guessed a giraffe. RB's next question was as follows: "I am thinking of something you put on your head when you go outside to keep your ears warm", and AD quickly responded "boots"! Because AD well knew the difference between boots and hats and was familiar with both words, it seemed likely that he had been lulled by the partial contextual similarity to retrieve a previously correct but now inappropriate solution.

Examples of this type are no doubt relatively common. From our perspective, they are especially interesting because they suggest a way of taking our research beyond the routinized and somewhat artificial situations of the laboratory, into the everyday world of problem solving, where familiar solutions are efficiently re-used, often for the better but occasionally for the worse.

Acknowledgments

This research was supported by grants to the first author from the Natural Sciences and Engineering Research Council of Canada, the University of Waterloo, and CAPES-Brasilia/ Brasil (BEX-2688), and by a grant to the second author from the National Institute of Child Health and Human Development (HD-21104). We thank Dov Cohen, Cindy Fisher, Greg Murphy, Kris Onishi, and Brian Ross for helpful comments and suggestions, Yuyan Luo and Karen Menard for their help with the data analyses, and Deepa Block, Laura Brueckner, Beth Cullum, Laura Glaser, Sue Hespos, Gavin Huntley-Fenner, Lisa Kaufman, Laura Kotovsky, Melsie Minna, Helen Raschke, and Teresa Wilcox for their help with the data collection. We also thank the parents who kindly agreed to have their infants participate in the research.

References

Aguiar A, Baillargeon R. 8.5-month-old infants' reasoning about containment events. Child Development. 1998; 69:636–653. [PubMed: 9680677]

Aguiar A, Baillargeon R. 2.5-month-old infants' reasoning about when objects should and should not be occluded. Cognitive Psychology. 1999; 39:116–157. [PubMed: 10462457]

Aguiar, A.; Baillargeon, R. Perseveration and problem solving in infancy. In: Reese, HW., editor. Advances in child development and behavior. Vol. 27. San Diego, CA: Academic Press; 2000. p. 135-180.

Aguiar A, Baillargeon R. Developments in young infants' reasoning about occluded objects. Cognitive Psychology. 2002; 45:267–336. [PubMed: 12528903]

Aguiar, A.; Menard, KR. Perseverative errors in means-end support tasks: further findings; Paper presented at the Biennial Meeting of the Society for Research in Child Development; Minneapolis, MN. 2001 Apr.

Aguiar, A.; Menard, KR. Perseverative responding in means-end support tasks; Paper presented at the Biennial International Conference on Infant Studies; Toronto, Canada. 2002 Apr.

Ahmed A, Ruffman T. Why do infants make A-not-B errors in a search task, yet show memory for the location of hidden objects in a non-search task? Developmental Psychology. 1998; 34:441–453. [PubMed: 9597355]

Baillargeon R. Representing the existence and the location of hidden objects: object permanence in 6- and 8-month-old infants. Cognition. 1986; 23:21–41. [PubMed: 3742989]

Baillargeon R. Reasoning about the height and location of a hidden object in 4.5- and 6.5-month-old infants. Cognition. 1991; 38:13–42. [PubMed: 2015755]

Baillargeon, R. The object concept revisited: new directions in the investigation of infants' physical knowledge. In: Granrud, CE., editor. Visual perception and cognition in infancy. Hillsdale, NJ: Erlbaum; 1993. p. 265-315.

Baillargeon R. How do infants learn about the physical world? Current Directions in Psychological Science. 1994a; 3:133–140.

Baillargeon R. Physical reasoning in young infants: seeking explanations for unexpected events. British Journal of Developmental Psychology. 1994b; 12:9–33.

Baillargeon, R. A model of physical reasoning in infancy. In: Rovee-Collier, C.; Lipsitt, LP., editors. Advances in infancy research. Vol. 9. Norwood, NJ: Ablex; 1995. p. 305-371.

Baillargeon R. Reply to Bogartz, Shinskey, and Schilling; Schilling; and Cashon and Cohen. Infancy. 2000; 1:447–462.

Baillargeon, R. The acquisition of physical knowledge in infancy: a summary in eight lessons. In: Goswami, U., editor. Blackwell handbook of childhood cognitive development. Oxford: Blackwell; 2002. p. 46-83.

Baillargeon R, DeVos J. Object permanence in 3.5- and 4.5-month-old infants: further evidence. Child Development. 1991; 62:1227–1246. [PubMed: 1786712]

Baillargeon, R.; Kotovsky, L.; Needham, A. The acquisition of physical knowledge in infancy. In: Sperber, D.; Premack, D.; Premack, AJ., editors. Causal cognition: a multidisciplinary debate. Oxford: Clarendon Press; 1995. p. 79-116.

Baillargeon, R.; Luo, Y. Encyclopedia of cognitive science. Vol. 3. London: Nature Publishing Group; 2002. Development of the object concept; p. 387-391.

Baillargeon R, Needham A, DeVos J. The development of young infants' intuitions about support. Early Development and Parenting. 1992; 1:69–78.

Baillargeon R, Spelke ES, Wasserman S. Object permanence in 5-month-old infants. Cognition. 1985; 20:191–208. [PubMed: 4064606]

Baillargeon R, Wang S. Event categorization in infancy. Trends in Cognitive Sciences. 2002; 6:85–93. [PubMed: 15866192]

Bogartz RS, Shinskey JL, Speaker CJ. Interpreting infant looking: the event set £ event set design. Developmental Psychology. 1997; 33:408–422. [PubMed: 9149920]

Bremner JG. Object tracking and search in infancy: a review of data and a theoretical evaluation. Developmental Review. 1985; 5:371–396.

Casasola M, Cohen LB, Chiarello E. Six-month-old infants' categorization of containment spatial relations. Child Development, in press. (in press).

Chi MTH, Feltovich PJ, Glaser R. Categorization and representation of physics problems by experts and novices. Cognitive Science. 1981; 5:121–152.

Dehaene S, Changeux JP. A simple model of prefrontal cortex function in delayed-response tasks. Journal of Cognitive Neuroscience. 1989; 1:244–261. [PubMed: 23968508]

Diamond A. Development of the ability to use recall to guide action, as indicated by infants' performance in AB. Child Development. 1985; 56:868–883. [PubMed: 4042750]

Diamond, A. Neuropsychological insights into the meaning of object concept development. In: Carey, S.; Gelman, R., editors. The epigenesis of mind. Hillsdale, NJ: Erlbaum; 1991. p. 67-110.

Diedrich FJ, Thelen E, Smith LB, Corbetta D. Motor memory is a factor in infant perseverative errors. Developmental Science. 2000; 3:479–494.

Doane SM, Alderton DL, Sohn YW, Pellegrino JW. Acquisition and transfer of skilled performance: are visual discrimination skills stimulus specific? Journal of Experimental Psychology: Human Perception and Performance. 1996; 22:1218–1248.

Erickson TA, Mattson ME. From words to meaning: a semantic illusion. Journal of Verbal Learning and Verbal Behavior. 1981; 20:540–552.

Forbus KD. Qualitative process theory. Artificial Intelligence. 1984; 24:85–168.

Haith, MM.; Benson, JB. Infant cognition. In: Damon, W.; Kuhn, D.; Siegler, R., editors. Handbook of child psychology. Vol. 2. New York: Wiley; 1998. p. 199-254.(Series Ed.) & (Vol. Eds.),

Harris, PL. The development of search. In: Salapatek, P.; Cohen, LB., editors. Handbook of infant perception. Vol. 2. New York: Academic Press; 1987. p. 155-207.

Hespos SJ, Baillargeon R. Infants' knowledge about occlusion and containment events: a surprising discrepancy. Psychological Science. 2001a; 12:141–147. [PubMed: 11340923]

Hespos SJ, Baillargeon R. Knowledge about containment events in very young infants. Cognition. 2001b; 78:204–245.

Hofstadter M, Reznick JS. Response modality affects human infant delayed-response performance. Child Development. 1996; 67:646–658. [PubMed: 8625732]

Kerr R, Booth B. Specific and varied practice of a motor skill. Perceptual and Motor Skills. 1978; 46:395–401. [PubMed: 662537]

Kotovsky L, Baillargeon R. The development of calibration-based reasoning about collision events in young infants. Cognition. 1998; 67:311–351. [PubMed: 9775513]

Kotovsky L, Baillargeon R. Reasoning about collision events involving inert objects in 7.5-month-old infants. Developmental Science. 2000; 3:344–359.

Lockman, JJ.; Pick, HL. Problems of scale in spatial development. In: Sophian, C., editor. Origins of cognitive skills. Hillsdale, NJ: Erlbaum; 1984. p. 3-26.

Logan GD. Toward an instance theory of automatization. Psychological Review. 1988; 95:492–527.

Luchins AS. Mechanization in problem solving. Psychological Monographs. 1942; 54(248):1–95.

Luchins AS. Classroom experiments on mental set. American Psychological Journal. 1946; 59:295–298.

Luchins AS, Luchins EH. New experimental attempts at preventing mechanization in problem solving. Journal of General Psychology. 1950; 42:279–297.

Luo, Y. Infants' knowledge about transparency in occlusion and containment events; Paper presented at the Biennial Meeting of the Society for Research in Child Development; Minneapolis, MN. 2001 Apr.

Luo Y, Baillargeon R. Qualitative but not quantitative reasoning about height information in occlusion events in 5-month-old infants. Manuscript in preparation. 2003

Mackenzie BE, Bigelow E. Detour behavior in young human infants. British Journal of Developmental Psychology. 1986; 4:139–148.

Marcovitch S, Zelazo PD. The A-not-B error: results from a logistic meta-analysis. Child Development. 1999; 70:1297–1313.

McDonough L, Choi S, Mandler JM. Understanding spatial relations: flexible infants, lexical adults. Cognitive Psychology. (in press).

Medin DL, Bettiger JG. Presentation order and recognition of categorically related examples. Psychonomic Bulletin and Review. 1994; 95:492–527.

Medin, DL.; Ross, BH.; Markman, AB. Cognitive psychology. 3rd ed.. Forth Worth, TX: Harcourt College; 2001.

Munakata Y. Perseverative reaching in infancy: the roles of hidden toys and motor history in the AB task. Infant Behavior and Development. 1997; 20:405–416.

Munakata Y. Infant perseveration and implications for object permanence theories: a PDP model of the AB task. Developmental Science. 1998; 2:161–184.

Murphy GL, Wright JC. Changes in conceptual structure with expertise: differences between real-world experts and novices. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1984; 10:144–155.

Pellegrino JW, Doane SM, Fisher SC, Alderton D. Stimulus complexity effects in visual comparisons: the effects of practice and learning context. Journal of Experimental Psychology: Human Perception and Performance. 1991; 17:781–791. [PubMed: 1834790]

Piaget, J. The construction of reality in the child. New York: Basic Books; 1954.

Posner, MI.; Raichle, ME. Images of mind. New York: Scientific American Library; 1994.

Reder LM, Kusbit GW. Locus of the Moses illusion: imperfect encoding, retrieval, or match? Journal of Memory and Language. 1991; 30:385–406.

Reitman JS. Skilled perception in Go: deducing memory structures from inter-response times. Cognitive Psychology. 1976; 8:336–356.

Rieser JJ, Doxsey PA, McCarrell NJ, Brooks PH. Wayfinding and toddlers' use of information from an aerial view of a maze. Developmental Psychology. 1982; 18:714–720.

Schmidt RA, Bjork RA. New conceptualizations of practice: common principles in three paradigms suggest new concepts for training. Psychological Science. 1992; 3:207–217.

Scott, J.; Aguiar, A. Reasoning about height relations in a containment task: a developmental trend in perseverative responding; Paper presented at the Biennial International Conference on Infant Studies; Toronto, Canada. 2001 Apr.

Sitskoorn SM, Smitsman AW. Infants' perception of dynamic relations between objects: passing through or support? Developmental Psychology. 1995; 31:437–447.

Smith LB, Thelen E, Titzer B, McLin D. Knowing in the context of action: the task dynamics of the A-not-B error. Psychological Review. 1999; 106:235–260. [PubMed: 10378013]

Sophian, C. Developing search skills in infancy and early childhood. In: Sophian, C., editor. Origins of cognitive skills. Hillsdale, NJ: Erlbaum; 1984. p. 27-56.

Spelke ES, Breinlinger K, Macomber J, Jacobson K. Origins of knowledge. Psychological Review. 1992; 99:605–632. [PubMed: 1454901]

Spelke ES, Kestenbaum R, Simons DJ, Wein D. Spatiotemporal continuity, smoothness of motion, and object identity in infancy. British Journal of Development Psychology. 1995; 13:113–142.

Suchman, LA. Plans and situated actions: the problem of human–machine interaction. Cambridge: Cambridge University Press; 1987.

Tanaka JA, Taylor M. Object categories and expertise: is the basic level in the eye of the beholder? Cognitive Psychology. 1991; 23:457–482.

Thelen, E.; Smith, LB. A dynamic systems approach to the development of cognition and action. Cambridge, MA: MIT Press; 1994.

Wang S, Baillargeon R, Brueckner L. Young infants' reasoning about hidden objects: evidence from violation-of-expectation tasks with test trials only. Manuscript submitted for publication. 2003

Wang S, Baillargeon R, Paterson S. Detecting continuity and solidity violations in infancy: a new account and new evidence from covering events. Cognition. (in press).

Wellman HM, Cross D, Bartsch K. Infant search and object permanence: a meta-analysis of the A-not-B error. Monographs of the Society for Research in Child Development. 1986; 51(3) Serial No. 214).

Wilcox T. Object individuation: infants' use of shape, size, pattern, and color. Cognition. 1999; 72:125–166. [PubMed: 10553669]

Wilcox T, Baillargeon R. Object individuation in infancy: the use of featural information in reasoning about occlusion events. Cognitive Psychology. 1998; 17:97–155. [PubMed: 9878104]

Wilcox T, Chapa C. Infants' reasoning about opaque and transparent occluders in an object individuation task. Cognition. 2002; 85:B1–B10. [PubMed: 12086715]

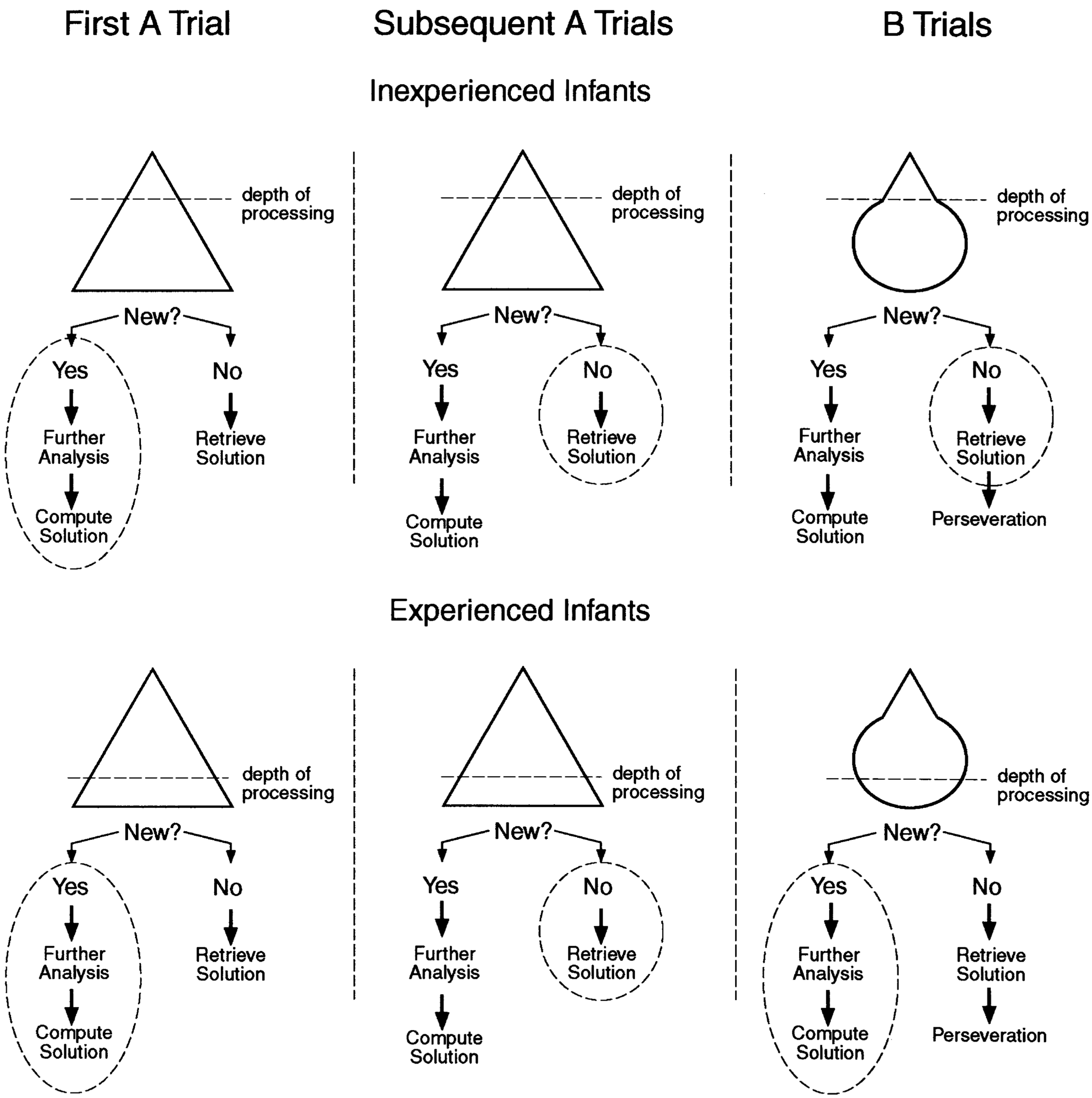

Fig. 1.

Schematic description of the third assumption of our problem-solving account of infant perseveration in NME-MO and related tasks.

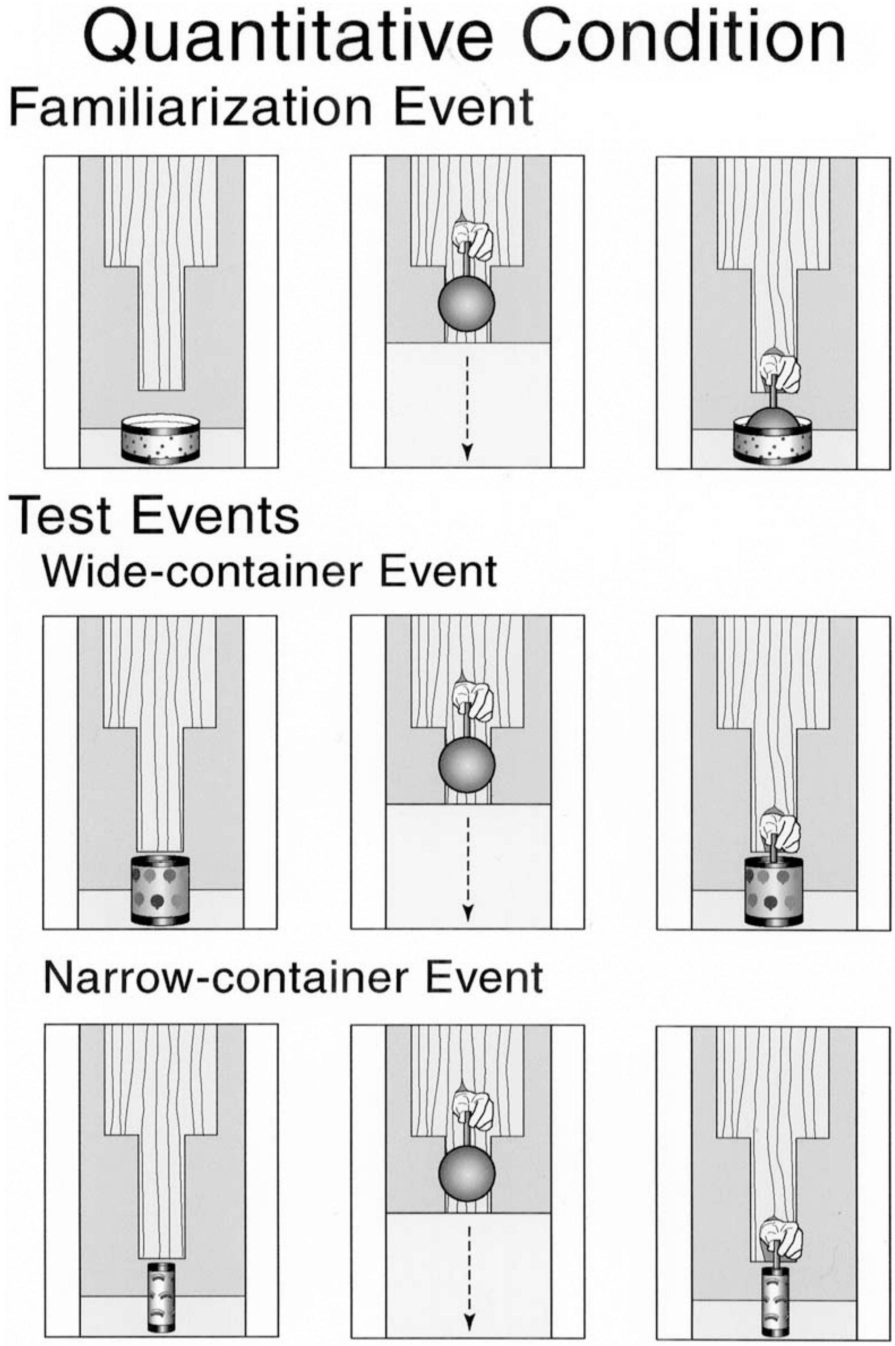

Fig. 2.

Schematic drawing of the familiarization and test events shown in the quantitative condition of Experiment 1.

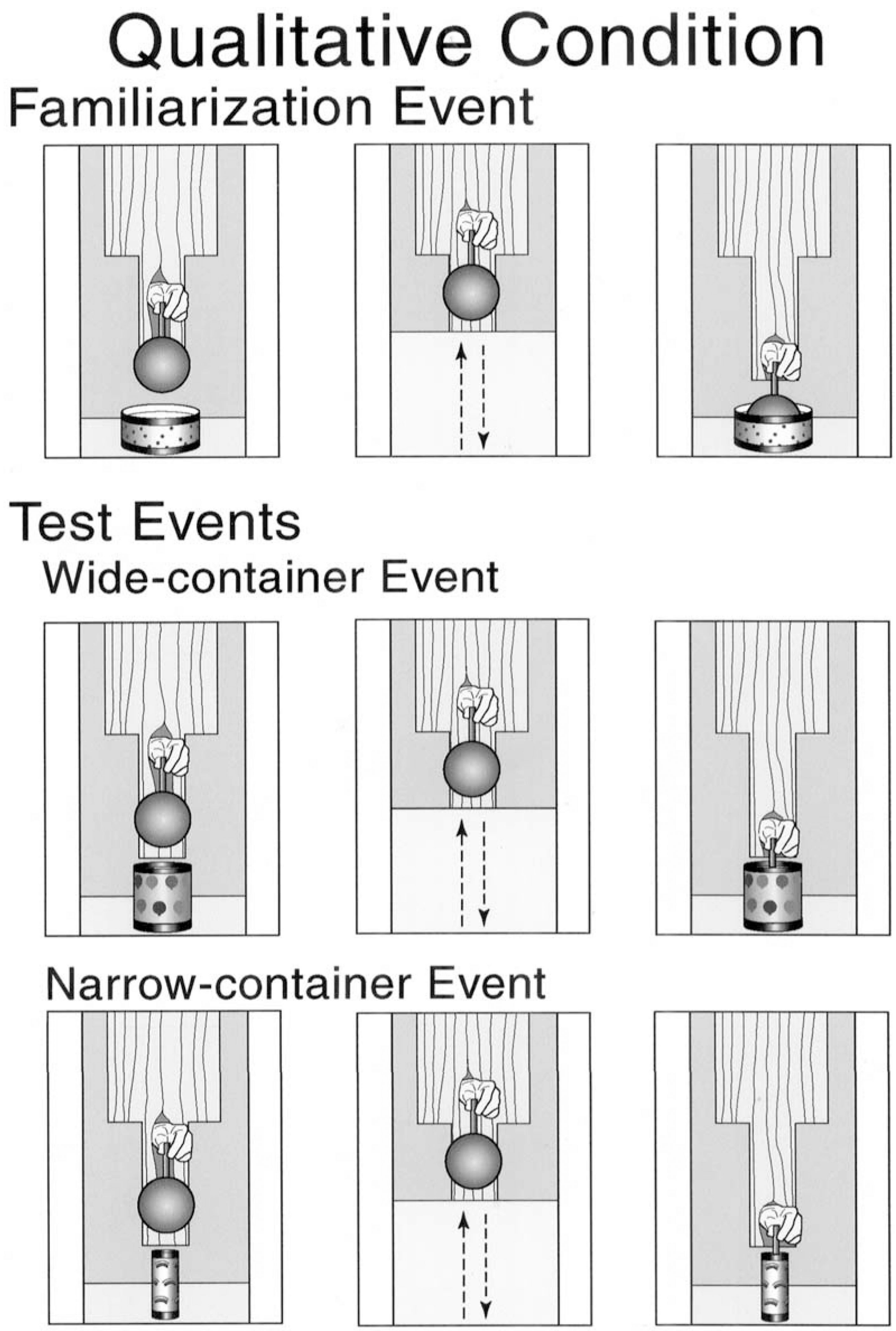

Fig. 3.
Schematic drawing of the familiarization and test events shown in the qualitative condition of Experiment 1.

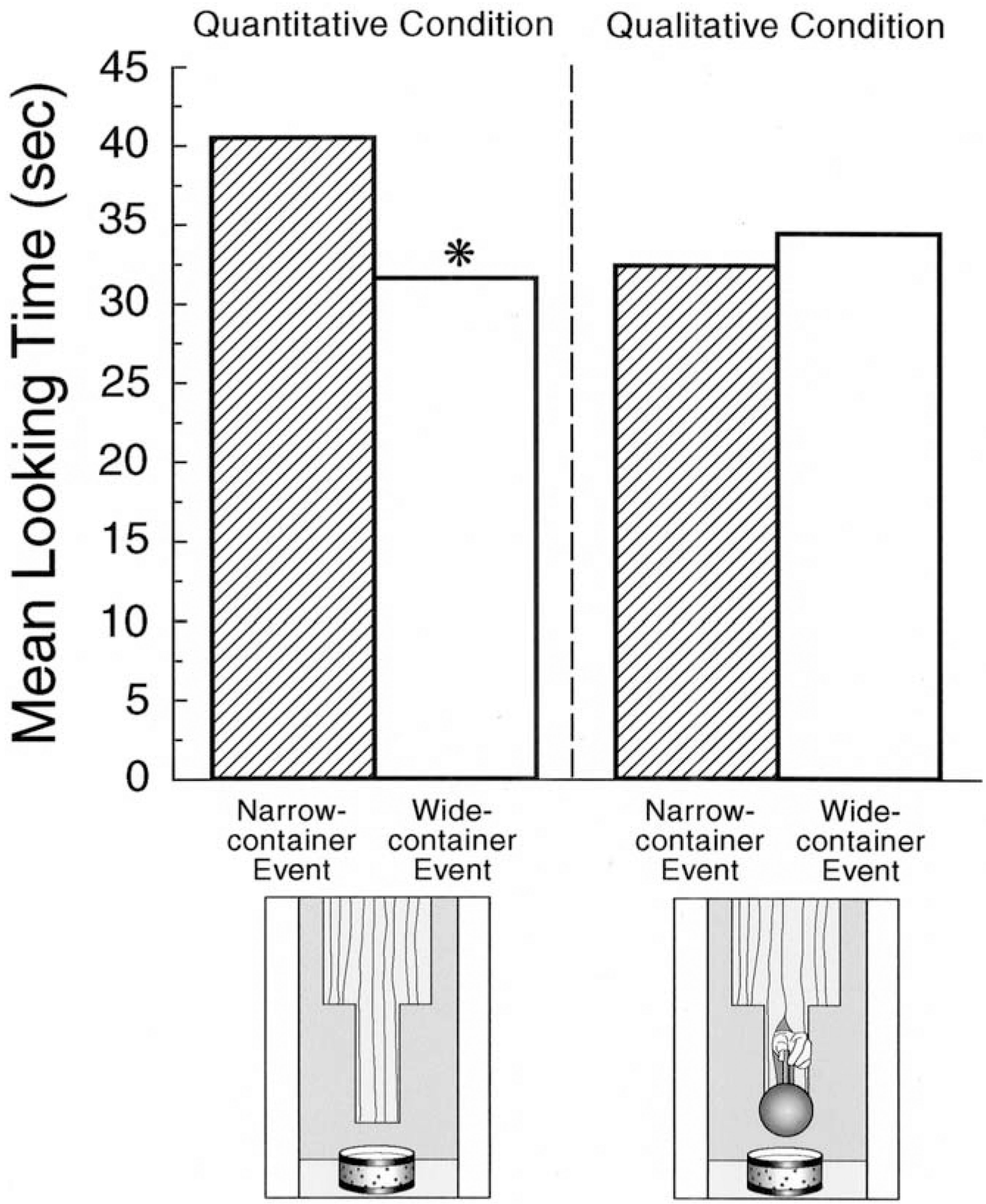

Fig. 4.

Mean looking times at the narrow- and wide-container test events of the infants in the quantitative and qualitative conditions of Experiment 1.

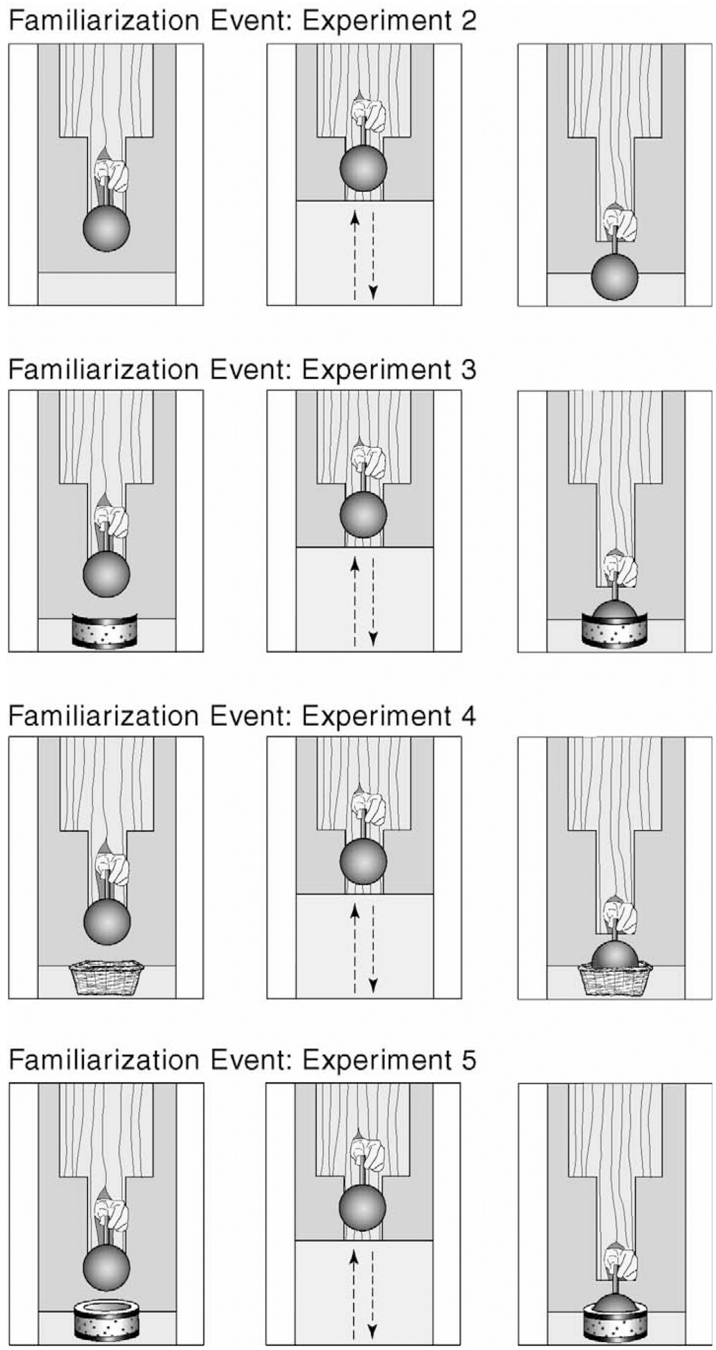

Fig. 5.

Schematic drawing of the familiarization events shown in Experiments 2–5.

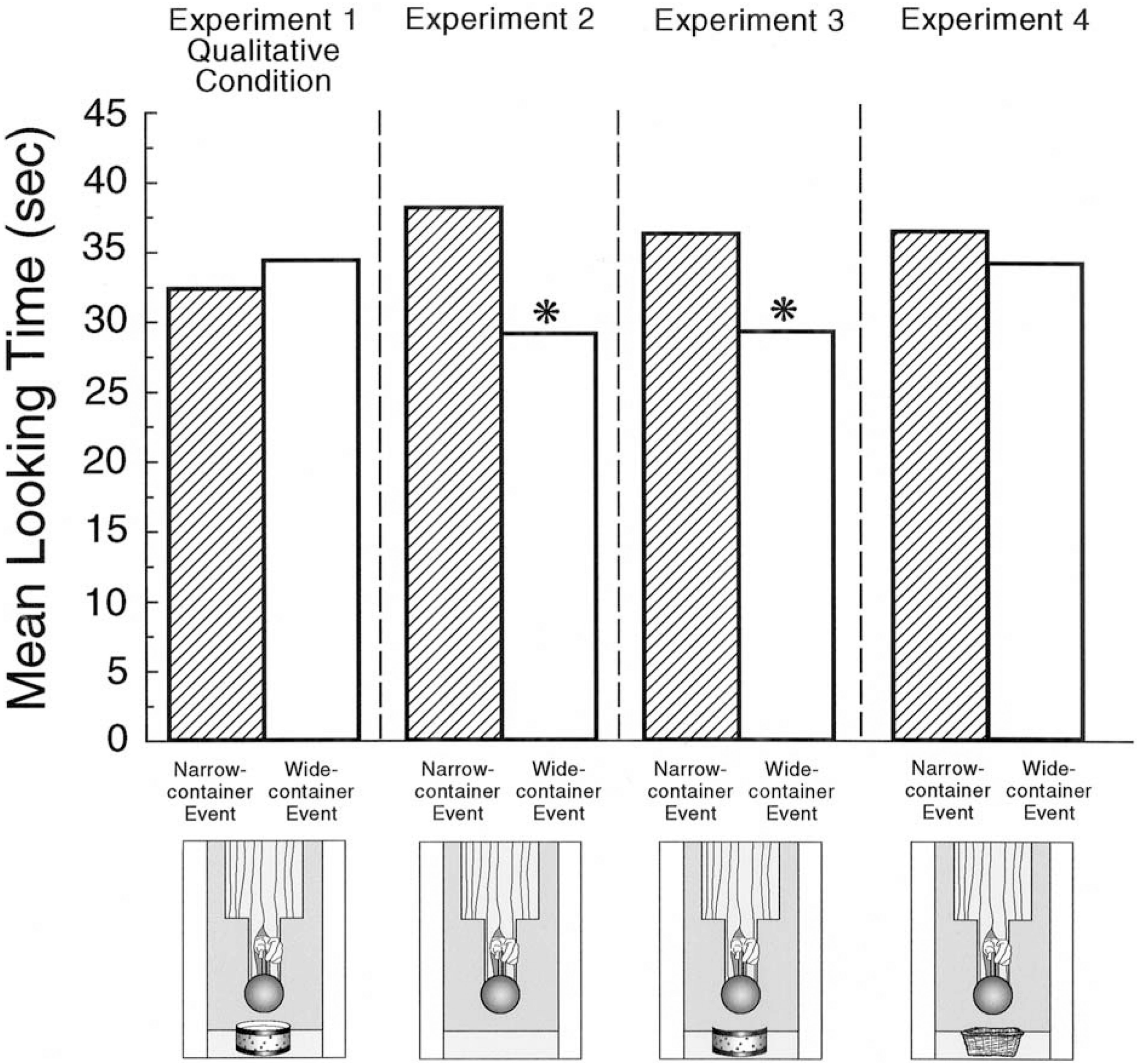

Fig. 6.
Mean looking times at the narrow- and wide-container test events of the infants in the qualitative condition of Experiment 1 and in Experiments 2–4.

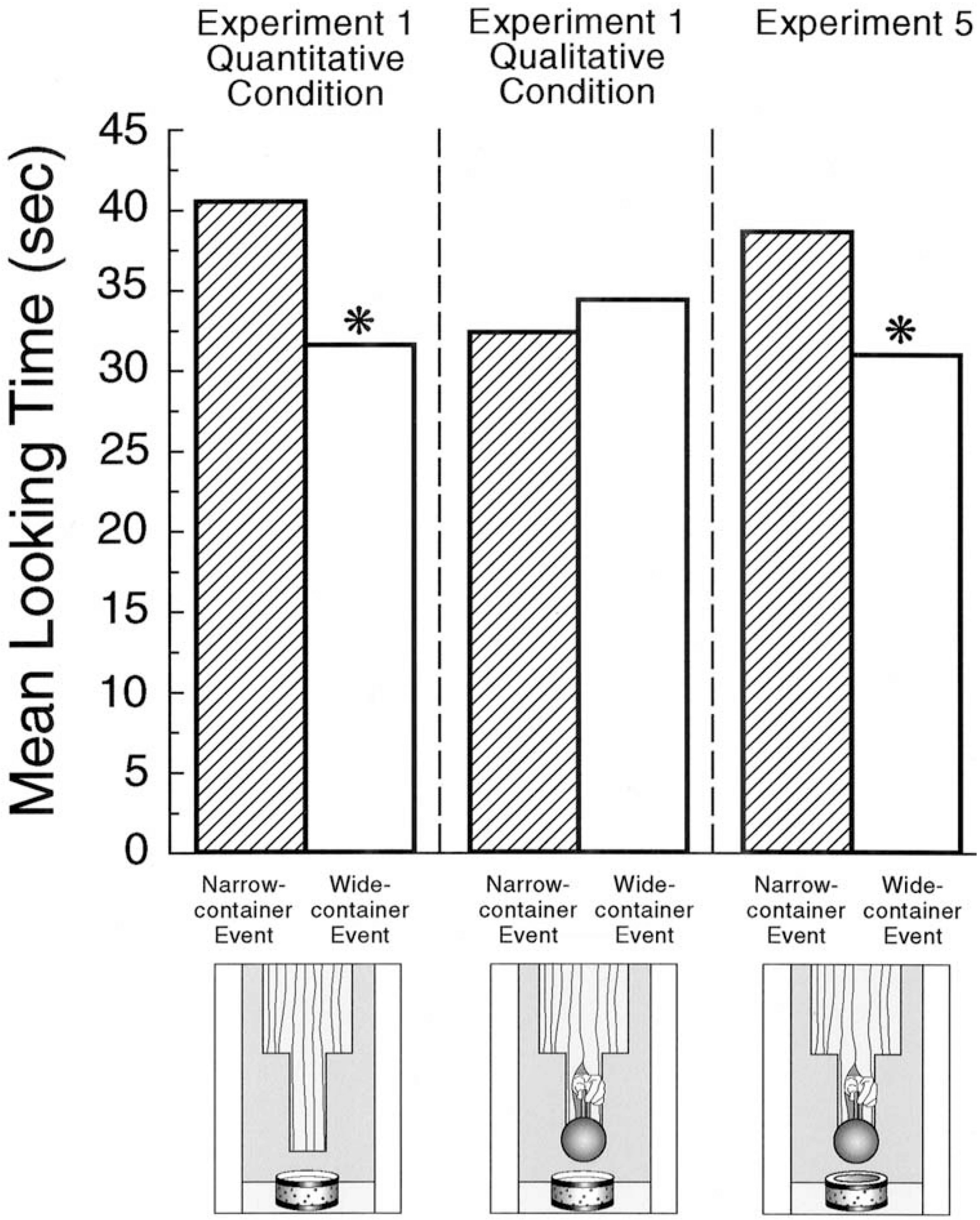

Fig. 7.

Mean looking times at the narrow- and wide-container test events of the infants in the quantitative and qualitative conditions of Experiment 1 and in Experiment 5.