The Influence of Number of A Trials on 2-Year-Olds' Behavior in Two A-Not-B-Type Search Tasks: A Test of the Hierarchical Competing Systems Model

By: Stuart Marcovitch and Philip David Zelazo

Marcovitch, S., & Zelazo, P. D. (2006). The influence of number of A trials on 2-years-old's behavior in Two A-Not-B-Type search tasks: A test of the hierarchical competing systems model. *Journal of Cognition and Development*, *7*, 477-501.

Abstract:
Age-appropriate modifications of the A-not-B task were used to examine 2-year-olds' search behavior. Several theories predict that A-not-B errors will in-crease as a function of number of A trials. However, the hierarchical competing systems model (Marcovitch & Zelazo, 1999) predicts that although the ratio of perseverative to nonperseverative errors will increase, the likelihood of perseveration will be an inverted U-shaped function. In Experiment 1, children received 1, 6, or 11 A trials in a multistep multilocation search task. In Experiment 2, children received 3, 7, 11, or 15 A trials in a sandbox task. Results from Experiment 1 showed that the ratio of perseverative to nonperseverative errors increased with number of A trials, and results from both Experiments 1 and 2 indicated that the likelihood of perseverative errors was highest following a moderate number of A trials. These findings are consistent with the suggestion that search behavior provides an index of the development of conscious control.

Article:
As children develop, they are increasingly able to represent multiple aspects of a problem, plan a future course of action, keep that plan in mind and act on it, and detect and use information about errors. This growing ability to engage in deliberate, goal-directed thought and action is often studied under the rubric of executive function, and it is now clear that executive function emerges early in development, probably around the end of the first year of life, and develops across a wide range of ages (for review, see Zelazo & Müller, 2002). However, executive function is a functional construct, and as such, it is defined solely in terms of its behavioral outcome. The actual mechanisms that make executive function possible are likely varied, and they remain a matter of considerable debate. This study was designed to examine these mechanisms in the context of children's search for hidden objects.

Executive function is often most conspicuous in its absence. Because measures of executive function typically involve conflict among different possible responses, one of which is correct and one of which is prepotent, failures of executive function at all ages are often manifested as perseveration, which has been de-fined as the repeated production of an action or thought in the absence of an appropriate stimulus (e.g., Hauser, 1999; Sandson & Albert, 1984, 1987; Werner, 1946). However, the specific situations in which perseveration occurs and, ultimately, the specific psychological processes that are likely to be involved vary as a function of age.

One widely used measure of perseveration in infants is Piaget's (1954) A-not-B search task. The A-not-B task has been employed successfully as an assessment of cognitive abilities toward the end of the first year of life. In a typical A-not-B task, an object is hidden conspicuously at one location (A) for a number of trials before it is then hidden in the same manner at a second location (B). The majority of 8- to 12-month-old infants search perseveratively at the A location on B trials, committing what is termed the A-not-B error (for reviews see Marcovitch & Zelazo, 1999; Thelen, Schöner, Scheier, & Smith, 2001; Wellman, Cross, & Bartsch, 1986).

Based on their meta-analysis of the A-not-B error, Marcovitch and Zelazo (1999) proposed that two hierarchically organized dissociable systems influence search behavior: a response-based system that corresponds to motor memory and learns incrementally according to associative principles, and a conscious representational system that potentially has the ability to control the response-based system. The conscious representational system is capable of reflecting on (or reprocessing) its own representations, allowing for the integration of information about the structure of the task and about the organism's own behavior vis-à-vis that structure. A simple schematic of the hierarchical competing systems model is displayed in Figure 1. As can be seen in Figure 1a, a stimulus elicits activity in both the response-based and the representational systems. In the absence of conscious reflection, behavior is determined jointly by the two systems but primarily by the response-based system. Figure 1b illustrates the effects of conscious control. Here, the infant is able to reflect on the contents of the representational system, as indicated by the reciprocal arrow. When reflection occurs, the influence of the representational system on the response is potentiated and conscious control can override the influence of previous behavior.

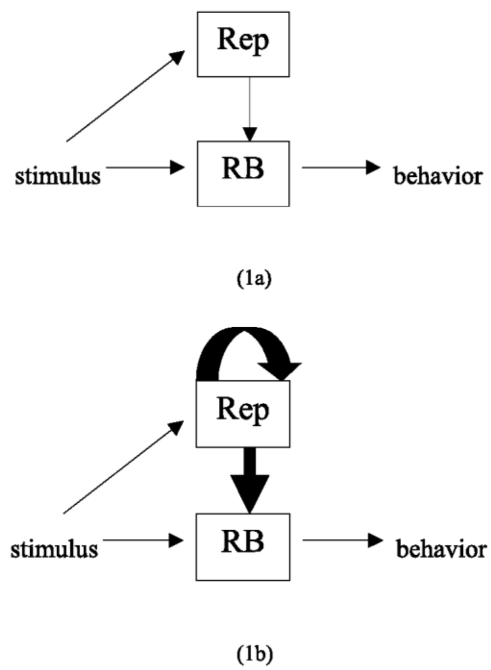

FIGURE 1 Schematic representation of the hierarchical competing systems model in the absence (1a) or presence (1b) of reflection on a representation of the task. Line thickness indicates the relative strength of the influence. RB = response based system, Rep = representational system.

The hierarchical competing systems model has been quantified to account for behavior in a specific task at a specific age—namely, the standard A-not-B task at 9 months of age (see Marcovitch & Zelazo, 1999). According to this quantification, the likelihood of search at any particular location can be calculated as the sum of a normal distribution (reflecting the influence of the representational system) and an exponential distribution (reflecting the influence of the response-based system). This quantification generates a set of novel predictions regarding infant search. It predicts that as the number of A trials increases, infants on B trials will be more likely to search at A and less likely to search at other incorrect locations (i.e., at locations between A and B; see Marcovitch & Zelazo, 1999, for discussion). This prediction follows from the fact that as the number of A trials increases, habit strength (a parameter in the exponential distribution) increases toward an asymptote at which the likelihood of searching at location A versus other incorrect locations is maximal. This prediction also makes sense intuitively, because repeated reaching to the A location should strengthen the tendency to search at A but not at other incorrect locations. However, this prediction has never been tested directly.

In addition, the hierarchical competing systems model predicts a non-monotonic, inverted-U-shaped relation between the number of A trials and performance on B trials (see Marcovitch, Zelazo, & Schmuckler, 2002, for discussion). According to the hierarchical competing systems model, the contribution of the response-based system to eventual search increases incrementally with the amount of experience at the A location, but this contribution reaches an asymptote, consistent with basic learning theory (e.g., Hull, Felsinger, Gladstone, & Yamaguchi, 1947; Rescorla & Wagner, 1972; Thorndike, 1921). However, as infants obtain experience with the task, they are also increasingly likely to reflect on the experimental context, including the task constraints and affordances (e.g., the fact that objects can be hidden in any location). The likelihood of reflection on the experimental context corresponds to the cumulative probability that reflection will occur on any given trial, and it may be influenced by the availability of additional memory resources as the task becomes automatized (cf. Grant & Berg, 1948; Hasher & Zacks, 1979; Logan, 1988). Taking into account the combined contributions of both the representational and the response-based systems, incorrect search on B trials would be most likely to occur following a moderate number of A trials because at this point the response-based system has reached a level of maximum influence, but the representational system has not. In contrast, search tends to be more accurate after either fewer or more A trials. This surprising prediction has recently been confirmed with 9-month-old infants. Using a standard version of the A-not-B task, Marcovitch et al. (2002) reported that 9-month-old infants were more likely to commit the A-not-B error after six A trials than after one or eleven A trials.

Compared to 8- to 12-month old infants, 2-year-old children display an impressive range of cognitive and linguistic skills—leading Piaget (1936/1952, 1954) to propose that it is around 2 years of age that children first acquire the ability to generate mental representations of reality. For example, unlike younger children, 2-year-olds succeed in invisible displacement tasks (Call, 2001), display mirror self-recognition and other signs of self-awareness (e.g., Amsterdam, 1972; Kagan, 198 1; Lewis & Brooks-Gunn, 1979), imitate the action an actor intended instead of the action performed (Meltzoff, 1995), and show the beginnings of pretend play (Leslie, 1987). In addition, most children experience a “naming explosion” between about 1.5 and 2 years of age (Reznick & Goldfield, 1992), which provides them with a relatively large corpus of words that potentially can be used to mediate their cognitive skills. Indeed, in light of these acquisitions, Zelazo (2004; Zelazo & Zelazo, 1998) has argued that there are considerable changes between infancy and 2 years of age in the function of the conscious representational system.

Despite their cognitive and linguistic advances, however, 2-year-old children are still prone to perseverative errors in simple but age-appropriate A-not-B tasks (e.g., Schutte & Spencer, 2002; Sophian & Wellman, 1983; Spencer, Smith, & Thelen, 2001; Zelazo, Reznick, & Spinazzola, 1998), and some researchers have suggested that the fundamental processes responsible for the A-not-B error in in-fancy are present across the lifespan (e.g., Marcovitch & Zelazo, 1999; Thelen et al., 2001). The hierarchical competing systems model, for example, postulates that the response-based system is already well developed by late infancy, and then re-mains developmentally invariant. In contrast, the conscious representational sys-tem develops considerably over the course of childhood (Zelazo, 2004). Given that 2-year-old children perseverate in age-appropriate A-not-B tasks, the hierarchical competing systems model predicts that at this age and in these tasks, there will be (a) an increase in the ratio of perseverative to non-perseverate errors during B trials as a function of number of A trials, and (b) an inverted-U-shaped relation between the number of A trials and the likelihood of perseveration on B trials, such that a moderate number of A trials is most likely to elicit errors.

Experiment 1 was designed to assess the effect of A trial experience on the multistep multilocation search task, which has been found to elicit perseverative errors in 2-year-old children (Zelazo et al., 1998). Similar to the standard A-not-B task used with infants, the multistep multilocation search task consists of discrete locations (i.e., a heterogeneous search space). Therefore, search must occur at one of the designated locations, and cannot occur between two locations; performance on this task is scored categorically (e.g., the proportion of children who searched perseveratively on the first B trial).

In Experiment 2, however, the role of A-trial experience was assessed in a different context, namely a sandbox. In the sandbox paradigm (Spencer et al., 2001), children search for objects hidden in sand, which is a homogeneous search space. Spencer et al. emphasized that the homogeneity of the sand allows researchers to measure graded responses such as the distance between the search location and the actual hiding location. This measure may reflect the degree of perseveration on a given trial, which may in turn reveal the strength of children's prepotent response tendencies.

A homogeneous search space may elicit different patterns of performance than a heterogeneous space. For example, Spencer et al. (200 1) noted that in the typical A-not-B task, infants commit more errors when the lids that cover the hiding wells are the same color, minimizing perceptual cues. In the sandbox task, the lack of perceptual cues is taken even further; the potential hiding locations are perceptually indistinguishable from each other and the rest of the search space. Indeed, successful search is dependent on place-learning cues such as the object's location relative to the sides of the sandbox or landmarks in the room, both of which are external to the search space (cf. Hermer & Spelke, 1994, 1996; Newcombe, Huttenlocher, Drummey, & Wiley, 1998).

In research using the sandbox paradigm, Spencer et al. (2001) found that on B trials, 18- to 31-month-old children typically initiated contact with the sand between the A and B locations (partially perseverative) as opposed to exactly at the A location (strictly perseverative), exactly at the B location (strictly correct), or at any other location (incorrect but nonperseverative). Although children demonstrated an overall search bias toward the midline, Spencer et al. reported that perseverative search occurred whether the A and B locations were on different sides of the midline or on the same side of the midline.

In Experiment 2, 2-year-old children received different numbers of A trials on the sandbox task. Because the dependent variables were quantitative, they were suitable for linear and quadratic regression analyses, and the presence of linear and quadratic trends could be assessed. The results from the sandbox task may offer an interesting comparison to the results from the multistep multilocation search task, given that search patterns in a homogeneous search space may differ from those in a heterogeneous one and given that the two tasks have different response requirements (see the following).

EXPERIMENT 1

Experiment 1 examined the effects of A trial experience on 2-year-olds' performance on a five-location multistep multilocation search task (Zelazo et al., 1998). In this task, candy is conspicuously hidden at one location (the A location) and then conspicuously hidden at another location (B). Children are required to execute a four-step retrieval process to find the hidden candy: (a) remove a foam barrier, (b) pull a tray, (c) choose a stimulus, and (d) pull the stimulus to reveal the candy.

Children received 1, 6, or 11 A trials. Most competing systems accounts of perseveration (e.g., Munakata, 1998; Thelen et al., 2001) postulate a positive relation between the number of A trials and the likelihood of perseveration on B trials. In contrast, the hierarchical competing systems model predicts a nonmonotonic trend, with poorest performance arising after a moderate number of A trials. This model also predicts that as the number of A trials increases, there will be an in-crease in the ratio of perseverative to nonperseverative errors (i.e., errors at locations other than the A location).

Method

Design and overview. Children were assigned randomly to receive 1, 6, or 11 A trials. In each condition, half of the children received A trials to the left of midline (Location 1 or 2) and B trials to the right of midline (Location 4 or 5). For the remaining children, this contingency was reversed. Within each condition, children were equally likely to receive any of the eight possible A–B pairings (1–4, 1–5, 2–4, 2–5, 4–1, 4–2, 5–1, 5–2). The middle location (Location 3) was used for training only.

For approximately half of the children in each condition, the hiding locations were demarcated by cards depicting familiar objects (e.g., truck, duck, hat). The remaining children were presented with cards with abstract line drawings. This manipulation, which was included as part of a separate line of research, was intended to be a preliminary assessment of whether familiarity with the stimuli affected performance at this age. However, because there was no effect of stimulus familiarity, and this variable did not interact with number of A trials, it will not be considered further in this article.

Participants. Seventy-one 2-year-old children (*M* = 26.4 months; range = 24.2–28.5 months) participated in this experiment and received 1 (*n* = 24), 6 (*n* = 24), or 11 (*n* = 23) A trials. Children were recruited from a database containing names of parents who had expressed interest in participating in research. Demographic information was not recorded systematically.

Data from an additional 32 children (31%) were excluded from the analyses because of failure to finish the training phase (*n* = 1) or failure to reach the A trial criterion (*n* = 3 1).1 Chi-square analyses did not reveal a relation between the probability of exclusion and the assigned number of A trials, X*2* (2, *N* = 103) = 2.43, *ns*.

Materials. The hiding apparatus was a 55 cm × 60 cm × 30 cm wooden box (see Figure 2). There was a side door through which the experimenter hid a candy. An opaque door at the front of the box could be removed to reveal the contents of the box; when in place (resting 3 cm above the base of the box), it concealed them. To prevent children from reaching for the candy during the hiding phase, a trans-parent Plexiglas window (which could be lowered completely) was used. A sliding tray on the base of the box could be extended as far as 20 cm from the door. Velcro swatches, on which cardboard symbols could be placed, were affixed to the tray. The swatches were placed 10 cm apart from each other. Each symbol was about 4 cm2 and was connected to a transparent plastic bag by a 45-cm string. The bags were used to contain the candies and were kept at the back of the box. A 50 cm × 10 cm × 7 cm foam barrier was placed in front of the tray during hiding. In addition, several small toys were available for hiding if the parent was uncomfortable with the children searching for candy.

Procedure. The child entered the test room and was seated at a small table on which the apparatus was already in place. The experimenter sat to the left of the child; a trained assistant sat across from the experimenter (on the child's right) and lowered and raised the Plexiglas window and the opaque door at the appropriate times. The parent sat behind the child so as to observe the experiment yet not distract the child. If a child was uncooperative or refused to stay in the seat, then the parent was allowed to sit at the table with the child on his or her lap. All parents were instructed not to reveal any information about the correct location or to label any of the symbols at any point.

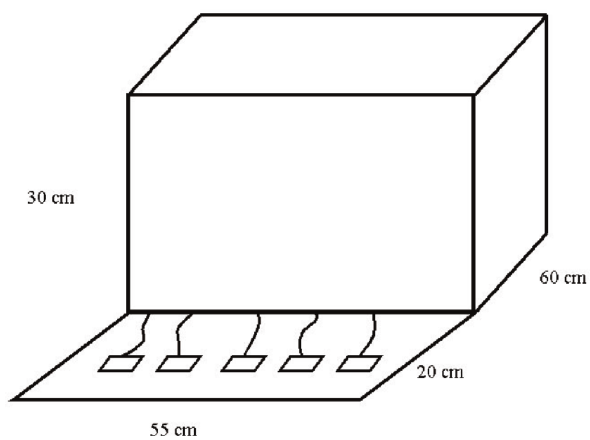

FIGURE 2 Apparatus for the multistep multilocation search task.

A training phase was required to familiarize the child with the apparatus and the task. The child was taught each of the following steps in succession: (a) pull the symbol (i.e., a yellow circle, which was used only in the training phase) to obtain the candy, (b) pull the tray (revealing the yellow circle symbol at the middle position and its accompanying string), and (c) remove the foam barrier (revealing the tray). After modeling each step, the experimenter asked the child, "Can you find the candy?" For the successful execution of each step, the child had to perform all previously learned steps (i.e., to find the candy, a child had to remove the foam barrier, pull the tray, and then pull the symbol). The child was given the candy on successful execution. The training phase was considered complete only if the child was able to execute the entire retrieval sequence without assistance.

After training, the Plexiglas window was lowered and five distinct symbols were placed on the Velcro swatches. For all trials (A and B), the transparent Plexiglas window was lowered, and the candy was placed conspicuously at the proper location. The experimenter held up the bag containing the candy, pointed to the appropriate symbol, and said, “Here’s the candy. Right over here. This is what you pull to get the candy. Okay, let’s go get the candy.” At no time did the experimenter label the symbols. The Plexiglas was removed, the opaque screen was lowered, and the foam barrier was placed in front of the apparatus. The child was then encouraged to look for the candy immediately and had to remove the barrier, pull the tray, and then pull the appropriate symbol attached to a string. Note that no systematic delay was imposed; rather the delay depended on how long it took the child to perform the multistep procedure. The trial was considered complete when the string was pulled so that the bag became visible. Children who chose correctly were given the candy. If the child chose incorrectly or did not respond for 30 sec, the experimenter pushed in the tray, lowered the Plexiglas window, showed the child the location of the candy, and then repeated the trial. For A trials, the testing continued until the child searched correctly the prescribed number of times. B trials were administered only if the child completed all the A trials and were repeated until the child found the reward three times.[2]

Results

The mean number (and *SD*) of A trial errors was 1.4 (2.2) for the 1 A trial condition, 1.8 (2.0) for the 6 A trial condition, and 2.2 (3.0) for the 11 A trial condition. The number of A trial errors did not differ by condition, $F(2, 68) = 0.73$, *ns*.

The primary variable of interest was performance on the first B trial. Across all conditions, the majority of errors (69%) were perseverative (i.e., at the A location) as opposed to nonperseverative (i.e., search at one of the other three incorrect locations). Analyses were conducted to address two questions. First, did the number of A trials affect the ratio of children making perseverative errors to those making nonperseverative errors? A chi-square analysis revealed that the proportion of children making perseverative errors (out of all children making errors) increased with the number of A trials (1 A trial: 44%; 6 A trials: 79%; 11 A trials: 90%), $X^2(2, N = 45) = 7.69$, $p < .05$, $\Phi c = 0.41$. See Table 1 for a breakdown of how children searched on the first B trial.

Second, was the effect of number of A trials on the probability of perseveration U-shaped? Quadratic logistic regression was conducted on the likelihood of committing a perseverative error on the first B trial with number of A trials as the quantitative independent variable (see Figure 3). The best-fitting logit function was $Y = \text{logit}(-.51 - .21X + 0.79X^2)$. The second-order coefficient differed significantly from 0, $Wald = 5.05$, $p < .05$, which is indicative of a U-shaped function. In contrast, the first-order coefficient did not differ from 0, $Wald = 0.68$, *ns*, which con-firms the absence of a monotonic relation.

TABLE 1

Breakdown of Search Behavior on the First B Trial in Experiment 1

	Number of A Trials		
	1	*6*	*11*
Number of children correct	8	5	13
Number of children perseverative	7	15	9
Number of children nonperseverative	9	4	1

It is possible that the U-shaped pattern was an artifact of grouping non-perseverative errors with correct responding. To control for differences in the likelihood of making nonperseverative errors, an additional regression analysis was conducted after omitting those children in all conditions who made nonperseverative errors (N = 14). The results did not change: The best-fitting logit function was Y = logit (–1.10 +.06X + 0.91X^2). The second-order coefficient differed significantly from 0, *Wald* = 4.78, $p < .05$, whereas the first-order coefficient did not, *Wald* = 0.05, *ns*.

Discussion

The number of A trials influenced 2-year-olds' search behavior on the multistep multilocation search task. Specifically, the ratio of children making perseverative errors to children making nonperseverative errors increased monotonically with increases in the number of A trials. In addition, more children made perseverative errors after six A trials than after one or eleven A trials. This nonmonotonic, U-shaped relation between the number of A trials and B trial performance was not an artifact of differences in the likelihood of making nonperseverative errors (e.g., due to fatigue); the same results were obtained after omitting children who made nonperseverative errors (and indeed, only one child [4%] in the eleven A trial condition committed a nonperseverative error).

Whereas several accounts of perseverative search (e.g., Munakata, 1998; Thelen et al., 2001) might predict that the ratio of perseverative to nonperseverative errors will increase with number of A trials, only the hierarchical competing systems model predicts a nonmonotonic relation between the number of A trials and the likelihood of perseveration. According to this model, increases in number of A trials not only produce increases in habit strength, but also provide opportunities for reflecting on the task structure.

One noteworthy limitation to this design with the multistep multilocation search task might obscure the true nature of the relation between the number of A trials and performance: The fact that the dependent variable is discrete (i.e., search is either correct or incorrect) precludes quantification of the degree of perseveration on any given trial. The possibility remains that true differences in performance might only be revealed using a task that yields a more sensitive mea-sure of perseveration.

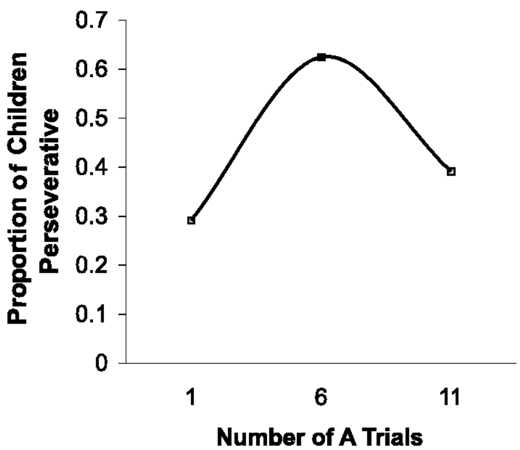

FIGURE 3 Proportion of children in Experiment 1 who searched perseveratively by the number of A trials.

EXPERIMENT 2

Experiment 2 was designed to assess the potential nonmonotonic relation between the number of A trials and B trial search performance on the sandbox task, which has previously been employed to provide quantitative measures of perseveration in 2-year-old children (Spencer et al., 2001). Because the sandbox consists of a homogeneous testing space, there are no discrete locations at which to search. Perhaps for this reason, Spencer et al. found that children often searched near, but not at, the correct location. Specifically, on B trials, children tended to search away from the B location and toward the A location. The distance from where the search commenced (i.e., from where the child's hand initially touched the sand) to the actual location of the target was used as a quantitative measure of perseveration. More precisely, the closer that B trial search was to the A location, the greater the degree of perseverative behavior.

Spencer et al. (2001, Experiment 4) varied the number of A trials and concluded that more A trials produced stronger attractors to the A location. However, there are a number of reasons this finding should be tested for replication and extended. First, children either received one or three A trials, whereas most previous research has only found strong effects of number of A trials with a wider range of A trials (see Marcovitch & Zelazo, 1999; Marcovitch et al., 2002). In Experiment 2 in this article, children received 3, 7, 11, or 15 A trials to assess the effects of A trials over a wider range.

Second, the children tested by Spencer et al. (2001) received their training trials at the A location. This procedure biases children toward the A location before the first A trial is administered (see Smith, Thelen, Titzer, & McLin, 1999). In the experiment discussed here, children received their training trials at the midline of the sandbox, which was never used as an A or B location and never situated between A and B. The method of using the middle location exclusively for training is consistent with the training procedures used in Experiment 1.

Third, Spencer et al. (2001) used no delay between hiding and searching in the A trials, but used a 10-sec delay on the B trials. It is possible that the sudden introduction of a delay (after receiving A trials without a delay) may have alerted children to a change in procedure. In this experiment, the more traditional approach of administering A and B trials in an identical manner was taken.

Finally, Spencer et al. (2001) did not measure response time. Unlike in traditional A-not-B paradigms, in the sandbox task the object is always retrieved by the participant. Even if search is perseverative in nature (i.e., away from the B location and toward the A location), children are permitted to search until the object is found. Thus, the time it takes to find the object should reflect perseverative tendencies and in certain situations may even be a more sensitive measure than distance based on where the search commenced. If children are continuously (and perhaps recurrently) attracted to the A location, it should take longer to find the object than if they are not. One advantage of using response time is that it provides a measure of performance even for those children who inadvertently touch the sand or use search strategies that involve purposely touching the sand at an inappropriate area (e.g., starting at one side and scooping the object or using both hands and converging on the object). Obviously, in these cases, the location where the child's hand initially touches the sand yields an inaccurate index of the level of perseverative activity, and response time may provide a better measure. Thus, in this experiment, response time was measured and used as a dependent variable along with the standard measure of distance.

Method

Design and overview. Children were assigned randomly to one of four conditions: 3, 7, 11, or 15 A trials. Within each condition, there were four possible A locations and two possible B locations. Both the A and B locations were on the same side of midline (i.e., they were both to the child's left or both to the child's right). Spencer et al. (2001) employed this procedure to reduce the effect of children's natural tendency to search at midline (Newcombe et al., 1998). If the midline is defined as 0, then the possible A locations to the left of midline were –5 (i.e., 5 cm to the left) and –45. The B location for these two A locations was –25. Likewise, the A locations to the right of midline were +5 and +45. The B location for these two A locations was +25. For each child, objects were first hidden at one of the A locations (on each of the designated number of A trials) and then hidden at the corresponding B location. The A location was counterbalanced perfectly across children in each condition (i.e., one fourth of the children received A trials at -45, one fourth at –5, one fourth at +5, and one fourth at +45).

Participants. Sixty-four 2-year-old children (M = 25.6 months; range = 24.0–27.9) participated in this experiment. Children were recruited in the same manner as the participants in Experiment 1. None of the children in Experiment 2 participated in Experiment 1.

An additional 14 children (18%) were tested, but their results were not included in the analyses. Ten children, distributed evenly across all four conditions, were eliminated because they refused to complete all of the A trials and at least one B trial. In addition, three children, all in the 3 A trial condition, were replaced due to equipment malfunction, and one child in the 15 A trial condition was replaced due to parental interference.

Materials. The sandbox apparatus was 150 cm × 30 cm × 45 cm (see Figure 4). It was constructed from plywood approximately 1.25 cm thick and painted green. The sandbox had a false bottom that consisted of a platform 30 cm from the ground. Thus, the receptacle for the sand was 15 cm deep. The sandbox was filled with sterile, nontoxic play sand. The distance between the top of the sandbox and the sand was 2.5 cm. Strips of measuring tape were affixed to the walls of the sandbox facing the scorer and a video camera (i.e., these strips were hidden from the child's view). Two purple footprints were situated in front of the sandbox in the middle for the child to stand on during the experiment. A rectangular block of wood 25 cm × 6 cm was used to smooth over the sand before the child was permit-ted to search. Also, a skunk puppet was used to divert the child's attention away from the sandbox during the delay.

The majority of children searched for candies that were placed in semi-transparent film canisters. However, if either the parents or the children were not comfortable searching for candies, a variety of small toys was available to be used.

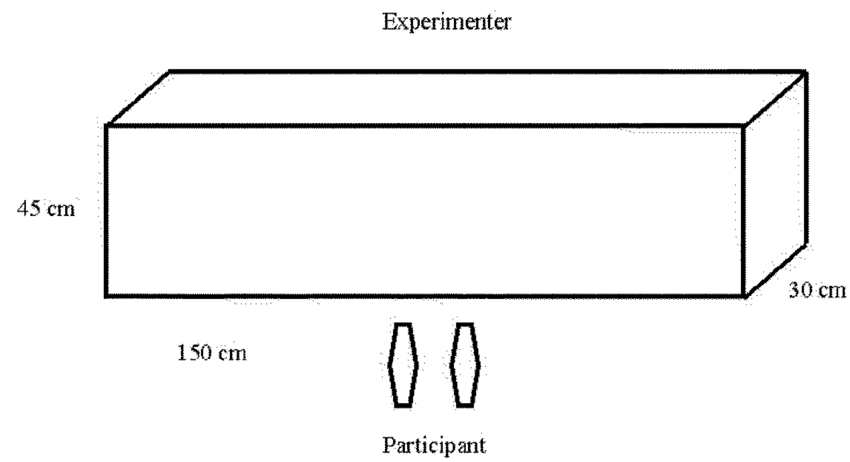

FIGURE 4 **Specifications of the sandbox apparatus.**

Procedure. The procedure was similar to that used by Spencer et al. (2001). The parent, child, and experimenter touched and played with the sand until the child was comfortable with the sandbox. The parent sat on a chair located behind the child, and the experiment began.

Prior to the A trials there were 3 training trials at the midline location. The child was positioned on the footprints and the parent hugged the child while the experimenter placed the container in the appropriate place. On the first training trial, the experimenter placed the container on top of the sand, and the child was allowed to retrieve it immediately. On the second training trial, the experimenter buried the container so that only the lid was visible. Again, the child was allowed to retrieve it with no delay. On the final training trial, the container was completely covered by sand, but the sand was not smoothed over (i.e., there were indentations in the sand that could cue the child to the location of the canister). Immediately thereafter, the child was allowed to search for the container. On all training trials, when the child searched successfully, the parent opened the container and gave the candy to the child. If on any of the training trials the child was unable to retrieve the container, then the experimenter helped the child by pointing to the area where the container was hidden. On these trials, the child was not given the candy, and the trial was re-administered until the child performed the action correctly without assistance.

The A trials were identical to the last training trial except that the container was hidden at the A location and a delay was imposed between hiding and search. During this time, the parent continued to hug the child while the experimenter smoothed over the sand with the block of wood. Then, the experimenter used a puppet to divert the child's gaze away from the sand and toward the experimenter's face. Once the child was no longer looking at the sandbox, the experimenter backed away from the sandbox while counting aloud to 3. Then, the parent was instructed to release the child as the experimenter encouraged the child to find the candy (the total time of the delay was approximately 10 sec). The trial ended after the container was found. If the container was not found within 30 sec, the experimenter pointed out the hiding location to the child (and the child was assigned a time of 30 sec). As long as the child touched the sand at any point during the trial, he or she was given the candy after the trial concluded. B trials were identical to the A trials, except that the candy was now hidden at the B location.

An online scorer was present during the experiment and seated behind the experimenter. However, the entire procedure was videotaped and scoring was completed from videotapes as well as online. In both cases, scorers measured the time elapsed from when the parent released the child to when the child retrieved the container. Also, the scorers determined the position of the hand's first contact with the sand to the nearest centimeter. Following Spencer et al. (2001), in those situations where the child touched the sand with both hands simultaneously, the hand closest to the actual location was used. Because of the precision that could be achieved by frame-by-frame analyses of the videotape, scores derived from videotape were used unless the location of contact was impossible to discern, in which case the result of the online scorer was used. The online scorer's results were used in only 57 of 793 trials (7%).

Results

Distance measure—original scoring. Following Spencer et al. (2001), the magnitude of the perseverative error on the B trial was measured as the distance (in cm) between the B location and the point where the child made initial contact with the sand. Spencer et al. assigned positive error values to reaches that were toward the A location, and negative values to reaches that were away from the A location (e.g., if the A location was to the left of B, but the child searched to the right of B).

Performance on the A trials was analyzed with errors toward the B location as positive errors and errors away from the B location as negative errors. There were no differences across conditions whether A trial performance was measured as the error score on the first A trial, $F(3, 60) = 0.\ 14$, *ns*; the error score on the last A trial, $F(3, 60) = 0.69$, *ns*; or the average error across all A trials, $F(3, 60) = 1.50$, *ns* (see Table 2, Original distance scoring—all A trials). To determine overall accuracy during the A trials, we examined whether the error scores of the first A trial and the last A trial differed significantly from zero, which occurred in three situations: the first A trial in the 7 A trial condition, $t(15) = 1.8, .05 < p <. 1$; the last A trial in the 3 A trial condition, $t(15) = 2.5$, $p < .05$; and the last A trial in the 15 A trial condition, $t(15) = 3.7$, $p < .01$. Clearly, not all children searched accurately during A trials (see the following).

The variable of primary interest was performance on the first B trial, which was initially analyzed including all participants and using the scoring method adopted from Spencer et al. (2001; see Table 2, Original distance scoring—first B trial). A one-way analysis of variance on the error score for the B trial failed to reveal an effect of the number of A trials, $F(3,60) = 0.47$, *ns*, Cohen's $f = .\ 14$. Although there were no differences across conditions, there was an overall bias to search toward the A location ($M = +5.70$ cm), $t(63) = 3.40$, $p < 0.01$, replicating Spencer at al.'s main finding that B trial search is biased toward the A location.

Distance measure–modified scoring. Two modifications to the scoring method were instituted. First, as Diamond (2001) noted, performance on the B trials is difficult to interpret if children do not perform adequately on the A trials. As noted earlier, a number of children were not searching accurately even on the A trials. Thus, an arbitrary criterion was established so that data from the children who did not search within 10 cm of the target averaged across the last two A trials were eliminated from the analyses ($N = 17$, evenly distributed across conditions).

TABLE 2
Mean (and Standard Errors) of Different Measures of Search Behavior in Experiment 2

	Number of A Trials											
	3			7			11			15		
	M	*SE*	*n*	*M*	*SE*	*n*	*M*	*SE*	*n*	*M*	*SE*	*n*
Original distance scoring—all A trials	6.2	(2.9)	16	5.0	(2.0)	16	1.5	(1.2)	1	8.4	(2.9)	16
Original distance scoring—1st B trial	5.6	(3.9)	16	2.6	(3.4)	16	8.3	(3.0)	16	6.3	(3.2)	16
Modified distance scoring—1st B trial	8.9	(2.2)	12	6.4	(2.2)	12	10.1	(2.4)	1	10.0	(2.4)	11
Mean response time—1st B trial	8.6	(1.3)	11	10.7	(2.6)	12	20.8	(3.3)	12	9.2	(2.3)	12

Second, the scoring system was modified so that all scores were constrained between 0 and 20. In the original scoring system (Spencer et al., 2001), search that occurred beyond the A location was given a larger value than if the child searched exactly at A (i.e., a pure perseverative error). We modified the scoring so that all search at the perseverative location or beyond was assigned a score of 20, signifying maximum perseverative behavior. Similarly, according to the original scoring system, incorrect search away from the A location was given a negative score, effectively counting as more correct than if search occurred exactly at the B location, which would have received a score of 0. Thus, all search that occurred away from the A location was assigned a score of 0, indicating no perseverative error.[3]

The modified distance scores are displayed in Table 2 (Modified distance scoring—first B trial). Despite these modifications, no effect of the number of A trials was revealed, $F(3, 43) = 0.57$, *ns*, Cohen's $f = .13$.

Response time measure. Another measure of performance was the time needed to retrieve the container. The maximum time allowed was 30 sec, after which the child was shown the location of the object and assigned a time of 30 sec. Following the logic from the modified distance measure described previously, data from children who did not find the object within an arbitrarily chosen time of 10 sec averaged across the last 2 A trials were eliminated from all analyses ($N = 17$, evenly distributed across conditions). Note that although the overall pattern of results did not change by removing these children, we believed that it was necessary from a theoretical perspective (Diamond, 2001). There were significant correlations between the time taken to find the container and both the original distance score, $r = 0.38$, $p < .05$, and the modified distance score, $r = 0.55$, $p < .05$.

Table 2 (Mean response time—first B trial) shows the mean response time for each condition. A one-way ANOVA revealed an effect of the number of A trials, $F(3,43) = 5.08$, $p <.01$, *Cohen's* $f =.33$. Post hoc Tukey tests at the.05 level indicated that the mean response time in the 11 A trials condition was significantly longer than it was in the other three conditions, which did not differ from one another.

One goal of this experiment was to assess whether the relation between the number of A trials and B trial performance was U-shaped (i.e., quadratic), as op-posed to linear. To that end, planned contrasts on the response times for the B trial were used to assess the presence of linear (–3 –1 1 3) and quadratic (1 –1 –1 1) trends. Results revealed a significant quadratic trend, $F(1, 43) = 7.22$, $p < .05$, but no linear trend, $F(1, 43) = 1.07$, *ns*.

The nature of the relation between the number of A trials and B trial performance was further assessed by conducting linear and quadratic regression on the response times for the B trial. Linear regression produced the equation: $Y = 0.3X + 9.8$, but the analysis did not reveal that linear regression was appropriate, $F(1, 45) =.77$, *ns*.

Further, the model accounted for very little variance, R^2 =.02, and the linear coefficient did not differ significantly from 0, $t(45) = .88$, *ns*. In contrast, quadratic regression produced the equation $Y = –.2X^2 + 4.2X – 3.5$. The analysis revealed that quadratic regression was appropriate, $F(2, 44) = 3.68$, $p < .05$, and the model accounted for a moderate amount of variability, $R^2 = .14$. Finally, both the quadratic and linear coefficients differed significantly from 0, $t(44) = 2.55$, $p < .05$ and $t(44) = 2.68$, $p < .05$, respectively.

The U-shaped pattern did not appear to be an artifact of outliers in the 11 A trial condition. First, a genuine majority of children in this condition (7 out of 12) failed to find the toy within 30 sec. In contrast, only one other child (in the 15 A trial condition) failed to find the toy within 30 sec. Second, to guard against the possibility that the nonmonotonic relation between response times and the number of A trials was an artifact of extreme response times rather than a shift in the distribution response times, a nonparametric Kruskal–Wallis test was conducted on median response times (10. 1, 7.2, 30.0, and 5.0 for the 3, 7, 11, and 15 A trial conditions, respectively). This analysis, which is resistant to outliers, revealed that the median response times differed by condition $X^2(3, N = 47) = 9.26$, $p < .05$. Furthermore, contrast-coded variables can be used on rank transformations to assess the existence of a quadratic trend. This procedure is advantageous because it controls for outliers and non-normality (Judd & McClelland, 1989). Indeed, a significant quadratic trend was found using this nonparametric analysis, $F(1, 43) = 4.29$, $p <.05$.

Discussion

A primary advantage of the sandbox search task over other search tasks is that it elicits quantitative measures of perseverative behavior. Consistent with the results from Experiment 1, a relation was revealed between the number of A trials and performance on the B trial in the sandbox task when response time was used. Because there were four levels of the independent variable, both linear and quadratic trend analyses were conducted, and the quantitative nature of the dependent variable permitted quadratic regression analysis. Both techniques provided support for the existence of a U-shaped pattern of behavior across a wide range of A trials.

Spencer et al. (2001) found an effect of the number of A trials on performance using the distance between initial contact with the sand and the B location. However, in the experiment described here, no effect of the number of A trials was revealed using that measure, although in all conditions strong biases toward the A location were found, replicating Spencer et al.'s main findings. Several procedural differences between this experiment and Spencer et al.'s Experiment 4 might account for the different pattern of results.

First, Spencer et al. (2001) included all children in their analyses, as opposed to using a criterion for acceptable A trial performance. Although the absence of such a criterion could, in principle, have influenced Spencer et al.'s results, it is unlikely to have done so because children in that experiment were reported to have searched accurately on the A trials.

Second, this experiment employed a 10 sec delay on both A and B trials, whereas Spencer et al. (2001) used different delays on the A and B trials. Recent evidence suggests that delays as long as 10 sec lead to larger midline biases (Schutte, Spencer, & Schöner, 2003). Even though this experiment was counter-balanced so that the effects of midline biases would be minimal, it is possible that the 10 sec delay may have obscured an effect of A trials.

Finally, Spencer at al. (2001) found a difference in performance after 1 A trial as compared to performance after 3 A trials. In this experiment, all children received a minimum of 3 A trials. It is possible that effect of number of A trials on the distance measures only occurs within a small range (i.e., between 1 and 3 A trials). Although this hypothesis needs to be tested directly within the context of a single experiment, current results using the response time measure suggests that performance does in fact change across a wide range of A trials.

The failure to find an effect of the number of A trials using the distance measures might be due to inadvertent touches of the sand and search strategies (e.g., using two hands and converging on the object) that do not conform to the assumptions inherent in our scoring measure. An advantage of measuring response times is that it does not matter if the child inadvertently touches the sand or intentionally touches the sand in the wrong location. For this reason, we argue that response time provides another useful measure of perseverative behavior. Generally, response times should increase proportionally to the level of difficulty required to find the object. Specifically, perseverative search should elicit longer response times than correct search, although longer response times might also be due to other factors such as distractibility or forgetting.

Clearly, response times and distance are related, as indicated by the significant correlation between the two measures in this experiment. Previous work (Schutte et al., 2003; Spencer et al., 2001), together with this experiment (in which there was an overall bias to search toward the A location), suggests that distance is a good index of perseverative behavior. This experiment shows that response time also provides a sensitive measure of performance in some contexts. Further work is needed to explore more fully the properties of the two measures, as well as the nature of the relation between them.

GENERAL DISCUSSION

Two experiments were conducted to examine the effects of number of A trials on behavior in two A-not-B-type search tasks: the multistep multilocation search task (Experiment 1) and the sandbox task (Experiment 2). Experiment 1 confirmed the prediction from the hierarchical competing systems model that the ratio of perseverative to nonperseverative errors would increase monotonically as a function of number of A trials. In addition, both Experiments 1 and 2 provided evidence of a nonmonotonic relation between the number of A trials and B trial performance. In Experiment 1, 2-year-old children were more likely to commit perseverative errors after 6 A trials, as opposed to 1 or 11 A trials. In Experiment 2, 2-year-old children's response times were maximal following 11 A trials (vs. 3, 7, or 15 A trials), and there was clear evidence of a quadratic trend.

In both experiments, children were required to remember the location of a hid-den object over a brief delay, and then to execute a response to find the object. However, there are at least two fundamental differences between the tasks employed in the two experiments. First, in the multistep multilocation search task (Experiment 1), the search locations were discrete and clearly demarcated with distinct symbols, whereas the sandbox task (Experiment 2) consisted of a homogeneous search space with no local cues to demarcate the location of the hidden object. Second, in the multistep multilocation search task, a multistep response was required before reaching toward the object, whereas reaching toward the object was the initial response in the sandbox task. The overall nonmonotonic relation between number of A trials and B trial performance was revealed despite these methodological differences, although these differences may help account for subtle variations in the details of the results. For example, because there were no discrete hiding locations in the sandbox task, it was probably relatively difficult to encode the location of the object. The relative level of difficulty in the tasks may explain why perseveration (indexed by response time) was maximal after 11 A trials in Experiment 2, whereas it was maximal after 6 A trials in Experiment 1.

Most competing systems accounts of search behavior predict that the effect of the number of A trials should be monotonically increasing, perhaps to an asymptote (e.g., Diamond, Cruttenden, & Neiderman, 1994; Munakata, 1998; Smith et al., 1999; Thelen et al., 2001). For this reason, it is likely that many extant theories can account for the finding (from Experiment 1) that as the number of A trials increases, children are more likely to search at A and less likely to search at other incorrect locations. However, only Marcovitch and Zelazo's (1999) hierarchical competing systems model predicts that the effect of number of A trials on the over-all likelihood of perseveration is not monotonic, but U-shaped.

According to this model, the representational system begins to influence search behavior around the end of the first year of life (i.e., 8–12 months) as shown by the pattern of results from multilocation A-not-B tasks with 9-month-old infants (Bjork & Cummings, 1984; Cummings & Bjork, 1983; Diamond et al., 1994). In these tasks, children do not consistently err at the A location, but rather commit a considerable number of errors between the A and B locations, as if they are pulled away partially from the perseverative location toward the correct location. This pattern of responding can be interpreted as a partial, incomplete influence of the representational system on search behavior.

With development, the influence of the representational system increases as a function of the growth of reflective processing through a series of levels of consciousness (Zelazo, 2004; Zelazo & Zelazo, 1998). In particular, it is argued that during the second half of the second year, children acquire the ability to reflect on the contents of consciousness through a functional process of recursion or re-processing of information via prefrontal cortical circuits. As a result, 2-year-olds potentially can increase the influence of the representational system on their behavior.

By 2 years of age, children typically pass the standard A-not-B task. However, if the task is made more complicated and demands more cognitive resources, the contribution of the representational system may be disrupted. Under these conditions, children may instead be influenced by the response-based system and perseverate, which is consistent with what was found in this set of experiments. In Experiment 1, a multistep procedure was required to find the candy. The only difference between A and B trials was the execution of the final step, which involved selecting the appropriate stimulus card. Under these conditions, and after a moderate number of A trials (6), the influence of the representational system was disrupted and perseveration was likely to occur. Only after a larger number of A trials (11) were most children arguably able to reflect on the task and override the tendency to search at the A location. Experiment 2 yielded similar results when response time was used as an index of perseveration, although in this experiment, the influence of the response-based system was most pronounced after 11 A trials, and it was not until 15 A trials that children's reaction times improved.

The nonmonotonic influence of number of A trials on the likelihood of perseveration has been observed in 9-month-old infants using a standard version of the A-not-B task (Marcovitch et al., 2002). In that experiment, infants were more likely to commit the A-not-B error after 6 A trials than after 1 or 11 A trials. Demonstrating that 2-year-old children are prone to the same nonmonotonic influences previously observed in 9-month-olds is important for several reasons. First, 2-year-olds are very different from 9-month-olds, and it is by no means obvious that similar results would be obtained across these ages and despite the major developments that take place during the first two years of life. Given that the U-shaped effect has now been found in 2-year-olds, however, it is possible to ex-amine the way in which this effect may interact with other aspects of development occurring during the second year, such as improvements in working memory (e.g., Wiebe & Bauer, 2005) and self-awareness (e.g., Kagan, 1981). Second, the fact that similar results were obtained in 2-year-olds demonstrates that the U-shaped effect is robust across at least two age groups and three tasks (although not, apparently, across all measures on all tasks) and this fact informs and constrains our interpretation of the phenomenon. Rather than being peculiar to a particular period in infancy (and task used with infants), whatever processes give rise to this phenomenon must be operative at least throughout the toddler years, consistent with models such as the hierarchical competing systems model.

The hierarchical competing systems model further postulates that the processes that govern search behavior are enacted even when the objects are visible (i.e., not hidden or behind a transparent barrier). The influence of the response-based system will be unaffected by the visibility of the object, and will continue to provide interference on B trials. In contrast, the continued presence of the object will affect the influence of the representational system. Specifically, the visible object will serve the dual role as an attention-grabbing object (cf. Thelen et al., 2001) as well as a mnemonic cue that increases the likelihood of reflection. Thus, compared to hidden objects, correct search performance with visible objects should occur more often (although not always; see Butterworth, 1977, and Harris, 1974).

Relative to 2-year-olds and, a fortiori, to 9-month-olds, adults are capable of considerable reflection; as a result, the contribution of the representational system is relatively strong. Indeed, it is difficult to imagine that an attentive adult would ever commit a perseverative error on a simple A-not-B task regardless of the number of A trials experienced (Marcovitch & Zelazo, 2001). However, adults do perseverate in other situations. Several researchers have catalogued a host of common perseverative tendencies (i.e., action slips; Heckhausen & Beckmann, 1990; Nor-man, 198 1). The standard explanation for action slips, which is consistent with the hierarchical competing systems model, is that perseverative behavior occurs when attention is focused elsewhere, utilizing the cognitive resources necessary to maintain a goal in working memory. The result is an inappropriate action based on a prepotent tendency (i.e., perseverative behavior).

Thus, the hierarchical competing systems model provides the foundation of a developmental account of perseveration across the life span. In this respect, the model is similar to dynamic field theory (Spencer et al., 2001; Thelen et al., 2001), which shares the position that the processes that elicit perseverative behavior in infants are the same processes responsible for perseverative behavior throughout development. Indeed, both theories are similar insofar as (a) they both capture the interactive contributions of multiple determinants of behavior, (b) they both model effects of repeated experience within a particular context, and (c) they both view perseverative behavior as a systematic response indicative of psychological organization. Currently, however, dynamic field theory cannot account for the non-monotonic influence of number of A trials on the likelihood of perseveration.

Alternative interpretations of the findings from the experiments discussed here are possible, however. For example, the finding that the ratio of perseverative to nonperseverative errors increases monotonically with number of A trials is consistent with the notion that B trial errors result from two processes: a perseverative tendency to search at A and a random error process that may result in either perseverative or nonperseverative errors (cf. Diedrich, Highlands, Spahr, Thelen, & Smith, 2001; Thelen et al., 2001). From this perspective, it is reasonable to assume that as the number of A trials increases, the perseverative tendency strengthens but random errors become less likely because children become more skilled with the task.[4] One difference between this possibility and the account offered by the hierarchical competing systems model is that the latter does not postulate a random error process. Whether a random error process is required to account for children's search behavior is an important question that deserves to be explored further.

It is more difficult to imagine alternative interpretations of the nonmonotonic effect of number of A trials on the likelihood of perseveration. One possibility is that children in the 11 A trial condition habituated to the repeated response of reaching to the A location, and thus were more likely to respond to a novel event (i.e., the reward hidden at the B location). Although this interpretation differs from that offered by the hierarchical competing systems model insofar as it does not (necessarily) implicate the process of reflection, it is not inconsistent with this model. Indeed, according to the hierarchical competing systems model, automatization of A trial performance may increase the likelihood of reflection by freeing up attentional resources; the process of habituation may be closely related to the process of automatization. Clearly, further research should investigate the possible roles of habituation, automatization, and reflection in producing the U-shaped function.

Further research might also examine more closely the circumstances in which a nonmonotonic effect of A trials is observed. Although the U-shaped effect has now been found at two ages and with three tasks, the results of Experiment 2 suggest that the effect may be dependent on particular combinations of tasks and dependent measures—at least within the range of number of A trials tested. In this experiment, using the sandbox task, we failed to find an effect of number of A trials when distance was used as a dependent measure, although we did find an effect for response time. This failure cannot currently be accounted for by the hierarchical competing systems model, and further research on the limits of the phenomenon may require that the model be revised or rejected.

In summary, the results of these experiments revealed that the ratio of perseverative to nonperseverative errors in A-not-B-type search tasks increases monotonically as a function of number of A trials, whereas the likelihood of perseveration varies nonmonotonically as a function of number of A trials across a wide range of trials. These results, which were obtained with 2-year-olds, extend previous findings from 9-month-old infants and suggest that the U-shaped effect is not limited to a particular age or task. Although further work is required to explore more fully the limits on this phenomenon, the presence of the phenomenon in children's search behavior was predicted by the hierarchical competing systems model. According to this model, a U-shaped function emerges as a result of the interaction between a conscious representational system and a response-based sys-tem. Age-related changes in search are attributed to the development of the ability to reflect on the task structure and to use that information to guide behavior consciously. From this perspective, search behavior in young children potentially provides a useful index of the development of conscious control.

Notes:

1The percentage of excluded participants is in line with other A-not-B studies in which participants were required to pass both training and the A trial phase (e.g., Marcovitch et al., 2002).

2B trials were administered until children searched correctly three times, but for ease of explanation and to remain consistent with other A-not-B studies, we only report performance on the first B trial. The pattern of results was similar when all B trials were considered, but the effects were not significant be-cause many children quickly corrected their perseverative behavior after being incorrect on the first B trial.

3By instituting the modified scoring system, we may be artificially increasing the degree of perseveration because errors that were clearly nonperseverative (away from the A location) did not con-tribute to the average error score (i.e., they were assigned a score of 0). However, for the purpose of this analysis, we are not concerned about the absolute strength of the perseverative error, but rather the relative strength of the perseverative error across conditions. We are assuming that any artificial bias toward the A location will be equal across conditions. Recall that the original scoring system, which allowed for negative scores, also did not reveal an effect of A trials.

4We would like to thank John Spencer for suggesting this interpretation.

ACKNOWLEDGMENTS

Funding for this research was provided by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada to P. D. Zelazo.

We would like to thank Brett Hendrie and Janet MacNeil for their assistance in conducting these studies and Dr. John Spencer for helpful comments on an earlier draft of the manuscript.

REFERENCES

Amsterdam, B. (1972). Mirror self-image reactions before age two. *Developmental Psychobiology, 5,* 297–305.

Bjork, E. L., & Cummings, E. M. (1984). Infant search errors: Stage of concept development or stage of memory development. *Memory & Cognition, 12,* 1–19.

Butterworth, G. E. (1977). Object disappearance and error in Piaget's stage IV task. *Journal of Experimental Child Psychology, 23,* 391–401.

Call, J. (2001). Object permanence in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*), and children (*Homo sapiens*). *Journal of Comparative Psychology, 115,* 159–17 1.

Cummings, E. M., & Bjork, E. L. (1983). Perseveration and search on a five-choice visible displacement hiding task. *Journal of Genetic Psychology, 142,* 283–291.

Diamond, A. (2001). Looking closely at infants' performance, and experimental procedures, in the A-not-B task [Commentary on Thelen et al.]. *Behavioral and Brain Sciences, 24,* 38–4 1.

Diamond, A., Cruttenden, L., & Neiderman, D. (1994). AB with multiple wells: 1. Why are multiple wells sometimes easier than two wells? 2. Memory or memory + inhibition? *Developmental Psychology, 30,* 192–205.

Diedrich, F. J., Highlands, T. M., Spahr, K. A., Thelen, E., & Smith, L. B. (200 1). The role of target distinctiveness in infant perseverative reaching. *Journal of Experimental Child Psychology, 78,* 263–290.

Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology, 38,* 404–411.

Harris, P. L. (1974). Perseverative search at a visibly empty space by young infants. *Journal of Experimental Child Psychology, 18,* 535–542.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388.

Hauser, M. D. (1999). Perseveration, inhibition and the prefrontal cortex: A new look. *Current Opinion in Neurobiology, 9,* 214–222.
Heckhausen, H., & Beckmann, J. (1990). Intentional action and action slips. *Psychological Review, 97,* 36–48.
Hermer, L., & Spelke, E. S. (1994, July 7). A geometric process for spatial reorientation in young chil-dren. *Nature, 370,* 57–59.
Hermer, L., & Spelke, E. S. (1996). Modularity and development: The case of spatial reorientation. *Cognition, 61,* 195–232.
Hull, C. L., Felsinger, J., Gladstone, I., & Yamaguchi, H. (1947). A proposed quantification of habit strength. *Psychological Review, 54,* 237–254.
Judd, C. M., & McClelland, G. H. (1989). *Data analysis: A model-comparison approach.* San Diego, CA: Harcourt Brace Jovanovich.
Kagan, J. (198 1). *The second year.* Cambridge, MA: Harvard University Press.
Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Re-view, 94,* 412–426.
Lewis, M., & Brooks-Gunn, J. (1979). *Social cognition and the acquisition of the self.* New York: Plenum.
Logan, G. D. (1988). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors, 30,* 583–598.
Marcovitch, S., & Zelazo, P. D. (1999). The A-not-B error: Results from a logistic meta-analysis. *Child Development, 70,* 1297–1313.
Marcovitch, S., & Zelazo, P. D. (2001). On the need for conscious control and conceptual understanding [Commentary on Thelen et al.]. *Behavioral and Brain Sciences, 24*, 48–49.
Marcovitch, S., Zelazo, P. D., & Schmuckler, M. A. (2002). The effect of number of A trials on performance on the A-not-B task. *Infancy, 3,* 519–529.
Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31,* 838–850.
Munakata, Y. (1998). Infant perseveration and implications for object permanence theories: A PDP model of the AB task. *Developmental Science, 2,* 161–184.
Newcombe, N., Huttenlocher, J., Drummey, A. B., & Wiley, J. G. (1998). The development of spatial location coding: Place learning and dead reckoning in the second and third years. *Cognitive Development, 13,* 185–200.
Norman, D. A. (198 1). Categorization of action slips. *Psychological Review, 88,* 1–15.
Piaget, J. (1936/1952). *The origins of intelligence in children.* New York: International Universities Press.
Piaget, J. (1954). *The construction of reality in the child.* New York: Basic Books.
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
Reznick, J. S., & Goldfield, B. A. (1992). Rapid change in lexical development in comprehension and production. *Developmental Psychology, 28,* 406–413.
Sandson, J., & Albert, M. L. (1984). Varieties of perseveration. *Neuropsychologia, 6,* 715–732. Sandson, J., & Albert, M. L. (1987). Perseveration in behavioral neurology. *Neurology, 37,*1736–1741. Schutte, A. R., & Spencer, J. P. (2002). Generalizing the dynamic field theory of the A-not-B error be-
yond infancy: Three-year-olds' delay- and experience-dependent location memory biases. *Child Development, 73,* 377–404.
Schutte, A. R., Spencer, J. P., & Schöner, G (2003). Testing the dynamic field theory: Working memory for locations becomes more spatially precise over development. *Child Development, 74,* 1393–1417. Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychological Review, 106,* 235–260.
Sophian, C., & Wellman, H. M. (1983). Selective information use and perseveration in the search behavior of infants and young children. *Journal of Experimental Child Psychology, 35,* 369–390.
Spencer, J. P., Smith, L. B., & Thelen, E. (200 1). Tests of a dynamic systems account of the A-not-B error: The influence of prior experience on the spatial memory abilities of 2-year-olds. *Child Development, 72,* 1327–1346.

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). A dynamic field theory of infant perseverative errors. *Behavioral and Brain Sciences, 24*, 1–18.
Thorndike, E. L. (192 1). *Educational psychology: Briefer course.* New York: Teachers College Press.
Wellman, H. M., Cross, D., & Bartsch, K. (1986). Infant search and object permanence: A meta-analysis of the A-not-B error. *Monographs of the Society for Research in Child Development, 51* (3, Serial No. 214).
Werner, H. (1946). Abnormal and subnormal rigidity. *The Journal of Abnormal and Social Psychology, 41,*15–24.
Wiebe, S. A., & Bauer, P. J. (2005). Interference from additional props in an elicited imitation task: When in sight, firmly in mind. *Journal of Cognition and Development, 6,* 325–363.
Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences, 8*,12–17.
Zelazo, P. D., & Müller, U. (2002). Executive functions in typical and atypical development. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 445–469). Oxford, England: Blackwell.
Zelazo, P. D., Reznick, J. S., & Spinazzola, J. (1998). Representational flexibility and response control in a multistep multilocation search task. *Developmental Psychology, 34,* 203–214.
Zelazo, P. R., & Zelazo, P. D. (1998). The emergence of consciousness. *Advances in Neurology, 77,* 149–165.