Journal of Experimental Child Psychology 108 (2011) 580–592

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Toddlers benefit from labeling on an executive function search task

Stephanie E. Miller *, Stuart Marcovitch

Department of Psychology, University of North Carolina at Greensboro, Greensboro, NC 27402, USA

ARTICLE INFO

Article history:
Available online 26 November 2010

Keywords:
Executive function
Labeling
Representation
Reflection
Cognitive development
Toddler search

ABSTRACT

Although labeling improves executive function (EF) performance in children older than 3 years, the results from studies with younger children have been equivocal. In the current study, we assessed performance in a computerized multistep multilocation search task with older 2-year-olds. The correct search location was either (a) not marked by a familiar picture or given a distinct label, (b) marked by a familiar picture but not given a distinct label, (c) marked by a familiar picture and labeled by the experimenter, or (d) marked by a familiar picture and labeled by the participant. The results revealed that accuracy improved across conditions such that children made the fewest errors when they generated the label for the hiding location. These findings support the hierarchical competing systems model, which postulates that improved performance can be explained by more powerful representations that guide search behavior.

Introduction

Executive function (EF) refers to the cognitive processes that play a role in conscious control over thought and action (Zelazo, Carter, Reznick, & Frye, 1997). The emergence and development of EF has become a major concentration in cognitive development (for a review, see Garon, Bryson, & Smith, 2008), with many theories of EF focusing on children's ability to represent information and use representations to guide behavior (e.g., Marcovitch & Zelazo, 2006, 2009; Munakata, 1998; Zelazo, 2004; Zelazo et al., 1997). This representational ability transforms during the preschool period as children begin to use language and symbols to form internal representations of their external environment (Piaget, 1959; Vygotsky, 1934/1986). Indeed, most studies that assess the benefit of labeling stimuli are conducted with children 3 years of age or older (e.g., Kirkham, Cruess, & Diamond, 2003). In contrast, few studies have examined linguistic representation and EF in children younger than 3 years of

* Corresponding author. Fax: +1 336 334 5066.
E-mail address: semille3@uncg.edu (S.E. Miller).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2010.10.008

age, and the study of this younger age group is important because considerable improvements in language occur early in life and the foundational abilities of EF may be affected by verbal mediation (e.g., Luria, 1979). The goal of the current study was to determine whether toddlers under 3 years of age benefit from the labeling of visual cues on an age-appropriate EF task.

The acquisition of language is perhaps one of the most important developments during childhood. Piaget (1959) suggested that the development of language (within a more broad symbolic function) was so powerful that it changed the way in which children thought (i.e., the shift from the sensorimotor stage to the preoperational stage). Vygotsky (1934/1986) further argued that early in life children use and understand speech in relation to the external social context (i.e., external speech or speech that must be spoken aloud to others to have meaning). Eventually, children learn to speak to themselves (i.e., private speech), which ultimately transitions to inner speech (i.e., covert speech to oneself) used to regulate thought and behavior. Vygotsky (1934/1986) and Luria (1979) hypothesized that both private and inner speech influence the control that individuals have over actions. Furthermore, Luria demonstrated that language initiates the control of behavior in 3.5- to 4-year-olds as they transition from impulsive responses to responses controlled by the meaning of an utterance.

Language and symbols have influenced EF performance in preschoolers as well. For example, Kirkham and colleagues (2003) demonstrated that 3-year-olds who were asked to label the relevant sorting dimension on the Dimensional Change Card Sort (DCCS) performed better than children who did not label themselves but heard the experimenter label (but see Müller, Zelazo, Lurye, & Liebermann, 2008). Furthermore, Rattermann and Gentner (1998) reported that in a relational search task where children needed to find a sticker based on the relationship between two objects, 3-year-olds were more likely to retrieve the sticker correctly based on object relationships (e.g., two objects of the same size) when the experimenter labeled the relationship (e.g., told children that the sticker was hidden under the baby/small animal). In addition, Müller, Zelazo, Hood, Leone, and Rohrer (2004) found that children labeling the nondominant correct response improved 3-year-olds' performance in an interference control task. Labeling may be beneficial because it provides distance from the immediate context, encourages abstract representation of the problem (Carlson, Davis, & Leach, 2005; DeLoache, 2000; Homer & Nelson, 2005, 2009; Jacques & Zelazo, 2005) and redirects attention to the relevant aspects of the task (Kirkham et al., 2003).

The limited linguistic ability of children younger than 3 years presents challenges to studying the development of EF because of the complexity of the verbal instructions (e.g., Gerstadt, Hong, & Diamond, 1994; Zelazo, Müller, Frye, & Marcovitch, 2003). One method for assessing EF in children younger than 3 years is through age-appropriate variations of Piaget's (1954) A-not-B task. Participants watch as a desirable object is hidden at one hiding location (Location A) and are subsequently allowed to retrieve it. After the object is hidden at Location A for a number of times, the object is then hidden conspicuously at a new hiding location (Location B). A perseverative error occurs when participants search incorrectly at Location A on the B trials. Marcovitch and Zelazo (2009) argued that success on the A-not-B task requires all of the components of EF—updating, inhibiting, and shifting (see also Garon et al., 2008; Miyake et al., 2000). To succeed in the A-not-B task, children must inhibit a dominant incorrect response (searching at Location A) and update the representation of the hiding location (from Location A to Location B) to shift search behavior successfully (switch from searching at Location A to searching at Location B). Perseverative errors may be evidence of a lapse in EF, specifically with difficulty in initiating or maintaining conscious control over actions.

The hierarchical competing systems model (HCSM) (Marcovitch & Zelazo, 2006, 2009) proposes that two systems interact to produce search behavior: a habit-based response system and a representational system. According to this framework, habits are formed through repeated actions (e.g., the habit of driving home after work is based on repeated actions). The representational system influences behavior through conscious thought and representation (e.g., to deviate from the habitual drive home for a trip to the grocery store, one needs to plan the change of direction). Furthermore, the representational system can be strengthened by external (e.g., seeing a billboard for the grocery store) and internal (e.g., verbally planning the route to the grocery store) sources. In the A-not-B task, the habit and representational systems work in concert on A trials but work in opposition on the first B trial; perseveration occurs when the representational system fails to override the habit system. According to the HCSM, labeling encourages reflection on the conscious representation, which strengthens the representational system, increasing the likelihood that it will determine behavior.

Unlike studies conducted with older children, studies examining EF in children younger than 3 years typically do not reveal a consistent benefit from a labeling condition. For example, Sophian and Wellman (1983) found that when children were provided with visual hiding information on B trials, 2-year-olds performed the same as older 2.5- and 4-year-olds. However, when children were provided with only a verbal statement of where the object was hidden on the B trial, 2-year-olds did considerably worse compared with 2.5- and 4-year-olds. Similarly, Marcovitch and Zelazo (2006) found that 2-year-olds did not benefit from easily labeled pictures marking hiding locations. In contrast, Homer and Nelson (2009) reported that 2.5- and 3-year-olds who labeled the hiding location in a scale model search task (i.e., a task where children used a smaller scale model to inform search in an identical larger room) (DeLoache, 1989) were more likely to search correctly than those who did not, but only when they had previous experience with the task. It may be that children younger than 3 years can benefit from labeling only when they label the stimuli actively, as opposed to listening passively to adults labeling the hiding location for them (e.g., Sophian & Wellman, 1983) or merely providing a task environment conducive to eliciting a linguistic label (e.g., Marcovitch & Zelazo, 2006). In other words, 2.5-year-olds might need to generate language in a social context (e.g., Vygotsky, 1934/1986) for linguistic representations to aid in problem representation and provide cognitive distance in an EF task.

This is consistent with Zelazo's (2004) levels of consciousness (LoC) model, which postulates that labeling one's experiences is necessary for young children to develop higher levels of consciousness. Toward the end of the first year of life, children begin to point, allowing them to represent an object and link a semantic memory to current experience. Zelazo and Zelazo (1998) suggested that developments in language (i.e., pointing and labeling) might provide the impetus for the emergence of the first level of consciousness (i.e., recursive consciousness). Labeling enables an additional representation to be held in mind (e.g., the word *ball* along with the image of a round object), allowing the contents of consciousness to be reprocessed and reflected on. This ability is at the core of the LoC model and is necessary for higher levels of consciousness (e.g., creation of rules and reflection on rule structure) to emerge. The self-generation of a label may be the optimal method to ensure that young children label their current experiences and reprocess them at a higher level of consciousness. Furthermore, generation should be especially important to young children who are just beginning to appreciate dual representations through self-initiated pointing and labeling.

The goal of the current study was to assess whether older 2-year-olds can benefit from the labeling of visual cues by examining performance in an age-appropriate computerized version of the A-not-B task (i.e., a multistep multilocation search task) (Carlson, 2005; Marcovitch & Zelazo, 2006; Zelazo, Reznick, & Spinazzola, 1998). In this task, children watched on a computer screen as a star entered one of five boxes. Following the hiding event, there was a 10-s delay during which three gray blocks occluded the hiding location. Next, children removed the blocks by pressing them in the correct sequence on the screen. Once the last block was removed, the hiding locations appeared and children searched for the star by pressing one of the boxes. Children needed to find the star correctly six times in Location A before the star was moved to Location B.

Children were assigned to one of four conditions, ordered by the availability of visual and labeling cues (see Table 1 and Fig. 1). In the *no picture* condition, the hiding locations were identical gray boxes with no distinct markers. The experimenter labeled the location unhelpfully by calling it "this" box. In the *no verbal label* condition, a familiar picture (e.g., flower, dog) denoted each hiding location but was not labeled by the experimenter. In the *experimenter verbal label* condition, on all trials the experimenter labeled the picture on the hiding location explicitly. In the *child verbal label* condition, on all

Table 1
Characteristics of experimental conditions.

Condition	Visual cue	Verbal label present	Verbal label generated
No picture			
No verbal label	X		
Experimenter verbal label	X	X	
Child verbal label	X	X	X

trials children were required to label the picture denoting the hiding location. The HCSM predicts improvement incrementally across all conditions, with the best performance occurring in the child verbal label condition. Children who are asked to generate a label are more likely to become aware of its benefit and the potential to label a different hiding location and, thus, search correctly at Location B. Generation of the label should be the most helpful because young children may need to speak aloud in a social context for speech to have meaning (e.g., Vygotsky, 1934/1986).

Method

Participants

A total of 153 older 2-year-olds participated in this study (M = 33.06 months, SD = 1.80, range = 29.4–36.6 months, 77 girls and 76 boys). Participants were recruited from child-care centers, preschools, and a database of parents interested in participating in research on cognitive development. The sample was approximately 79% Caucasian, 19% African American, 1.5% American Indian, and 0.5% Asian, although the parents of 26 children did not provide demographic information. Of the sample that reported income, the majority (71%) reported an average annual income of $60,000 or higher. Age and sex of the participants by condition are presented in Table 2.

Materials

The computerized version of the multistep multilocation search task was programmed using the SuperLab Pro (Version 4.0) software program. Stimuli were presented on a Dell laptop computer (Latitude D600) with a 14-inch monitor, and children responded on a Magic Touch touch screen placed directly on the laptop monitor. Children were seated in front of the computer and saw stimuli presented in full screen view centered on a white background (28.5 × 21.6 cm). A yellow star (2.7 × 2.5 cm) entered one of five boxes (4.4 × 4.0 cm) used for hiding locations. In some conditions, easily identified pictures (dog, flower, car, apple, and pencil) were centered on the front of each box (approximate size = 2.5 × 2.0 cm) (see Fig. 1). Children touched the hiding location on the screen to retrieve the star.[1]

Three blocks were required for the multistep sequence (see Fig. 2): one red block in the bottom left-hand corner of the screen (12.6 × 5.0 cm), one yellow block in the bottom right-hand corner of the screen (13.5 × 5.0 cm), and one green block centered above the red and yellow blocks (24.6 × 8.0 cm). A pleasant noise (Microsoft Windows Operating System tada.wav) sounded for correct responses, and an unpleasant noise (Microsoft Windows Operating System Windows XP Battery critical.wav) sounded for incorrect responses. Small stickers were used as rewards for correct searches.

Procedure

Training

The experimenter introduced the multistep procedure to children by presenting the sequence in a backward training fashion (i.e., presenting the final step first) (see Marcovitch & Zelazo, 2006). Pilot work indicated that a delay of 10 s and a three-part multistep sequence were necessary to make the task difficult enough to elicit perseveration in this age range. Children needed to complete the entire training procedure independently (i.e., without experimenter help) before moving on to the testing phase. The training phase began with one unmarked gray box centered on the computer screen (i.e., the eventual middle location) against a white background. The lid to the box opened, a yellow star entered the box, and the lid closed. Following the hiding of the star, the experimenter instructed children to touch the box to find the star. After children touched the box, the star appeared and children heard the reinforcing sound. In the next step, the star entered the same gray box, but after the hiding event a green block appeared on the screen over the gray box. Children touched the green block on the

[1] The majority of children used a finger to respond. Three children, whose responses were not picked up by the touch screen, used a stylus on some trials.

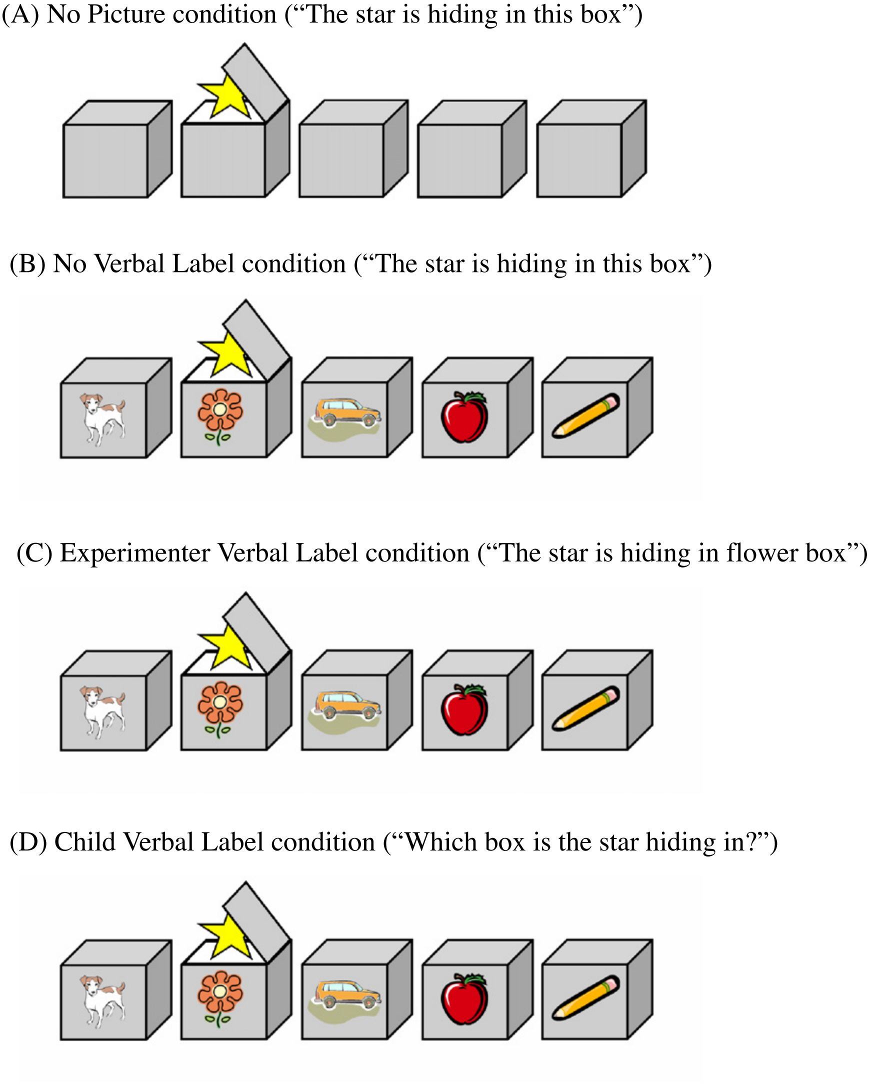

Fig. 1. Hiding event by condition (and experimenter prompt).

Table 2
Mean age by condition and sex.

Condition		Age (months)
No picture	Female (*n* = 19)	33.14 (1.59)
	Male (*n* = 19)	33.46 (2.13)
No verbal label	Female (*n* = 19)	32.58 (2.12)
	Male (*n* = 19)	33.40 (1.69)
Experimenter verbal label	Female (*n* = 18)	33.17 (1.66)
	Male (*n* = 20)	32.60 (1.72)
Child verbal label	Female (*n* = 21)	33.77 (1.69)
	Male (*n* = 18)	33.21 (1.79)

Note. Standard deviations are in parentheses.

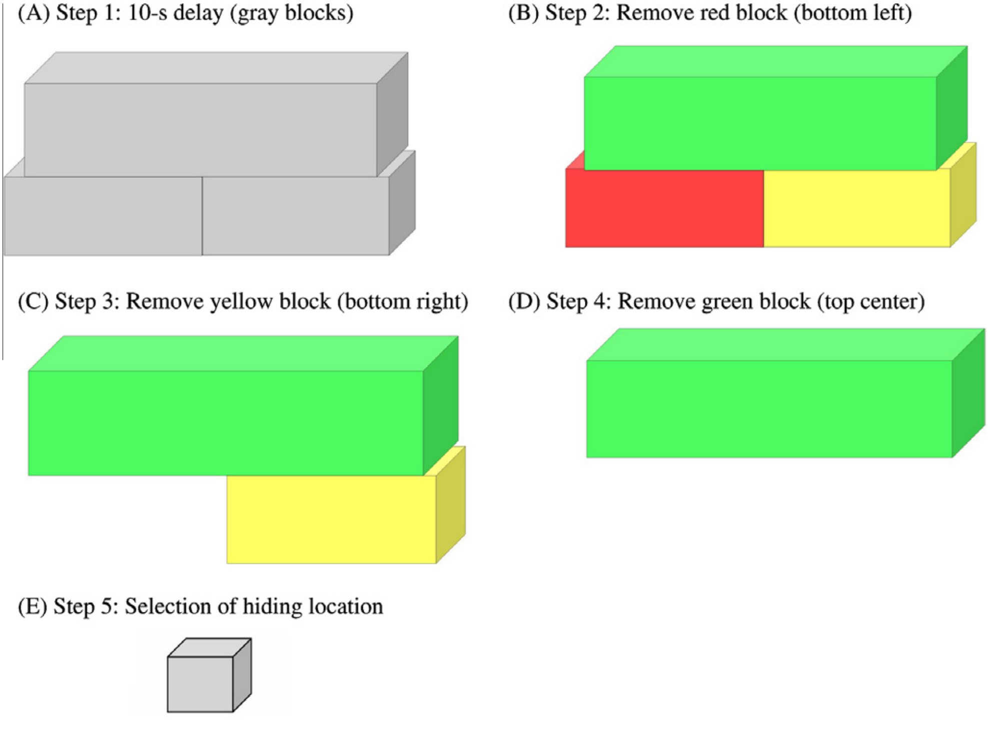

Fig. 2. The multistep procedure. (For interpretation of the references to color in this figure, the reader is referred to the Web version of this article.)

screen to make it disappear and then pressed the gray box to retrieve the star. After the next hiding event, both the green block and the yellow block appeared on the screen. Children first touched the yellow block, then touched the green block, and finally touched the box to get the star out. After the next hiding event, the green block, the yellow block, and the new red block appeared on the screen. Children pressed the three blocks in the correct sequence (red block, yellow block, and green block) to make each block disappear and then pressed the gray box to retrieve the star. In the final component of the training phase, the complete multistep procedure, including the delay, was revealed to children. After the hiding event, three gray blocks appeared on the screen. There was a computer-timed 10-s delay during which the experimenter encouraged children to count aloud to 5. After the delay, the blocks changed colors, indicating that children could begin the multistep procedure and search for the star. Children were required to touch the blocks in the correct sequence before they retrieved the star (see Fig. 2).

Testing

The A and B trials were similar to training trials except that five gray boxes were now presented on the screen during the hiding event. Locations A and B were counterbalanced with the stipulations that Location B was on the opposite side of the midline as Location A and that the middle box was never used as a hiding location to minimize interference from the training trials (see Marcovitch & Zelazo, 2006).

Children were randomly assigned to one of four conditions. In the *no picture* condition, the hiding locations were unmarked. When children were ready to begin, the star entered the box at Location A (see Fig. 1A). The experimenter noted where the star was by pointing and saying, "The star is hiding in this box, and you are going to find it right here." There was a 10-s delay during which the experimenter encouraged children to count aloud to 5, after which the multistep sequence was performed to find the star. Note that the last block pressed (the green block) covered all five boxes so that the boxes were revealed to the children at the same time (see Diamond, Cruttenden, & Neiderman, 1994). Children were given a sticker if they searched correctly. The experimenter administered A trials until children correctly found the star on six occasions (perseveration was most likely after six A trials with 2-year-olds on a noncomputerized version of this task) (Marcovitch & Zelazo, 2006). After children retrieved the star correctly on the sixth A trial, the hiding locations reappeared and children observed the star as it entered a new location (Location B). B trials were administered until one correct B response was obtained or until children stopped responding.

The procedure for the *no verbal label* condition was identical to that for the no picture condition (i.e., the experimenter pointed to the box and said, "The star is hiding in this box, and you are going to find it right here") except that the boxes presented to children had pictures of the easily identified objects on them (see Fig. 1B).

In the *experimenter verbal label* condition, the procedure was identical except that the experimenter labeled the correct location (e.g., "The star is hiding in the flower box, and you are going to find it right here").

In the *child verbal label* condition, the procedure was the same as in the experimenter verbal label condition except that children were asked to label the location (e.g., "Which box is the star in? [Child answers.] That's right, and you are going to find it right here"). If children did not answer immediately, they were prompted further (e.g., "What is that a picture of? Can you tell me what that is?"). If children continued to not respond or named the wrong box, they were told the correct location of the star and were instructed to name the box (this occurred infrequently [$n = 4$] and never on the critical first B trial).[2]

Accuracy and response time were measured on each trial. To encourage precise responding (i.e., location of touch on the touch screen), on training and A trials, the location of the touch response needed to be exactly within the hiding location boundaries to be counted as correct. On B trials, however, the search area was enlarged to accept touch responses 4.0 cm above or below the search location. The measuring of response time began at the end of the multistep procedure, as soon as the five boxes became available for search, and ended when one of the hiding locations was touched.

Results

Outlier analysis

Response time data were analyzed to determine whether there were any outliers on the first B trial response because extremely low response times may indicate unintentional responses on the touch screen (e.g., if children tapped the last block twice, it would be scored as a very fast response time when selecting the boxes), whereas extremely high response times may indicate that children were not focused on the task. Because response times were positively skewed, they were subjected to a logarithmic transformation (Sheskin, 2004). A box plot analysis on the log response times on the first B trial indicated that seven negative outliers fell below the log response time of 2.645 ($Q1 - 1.5 * \text{interquartile range}$) and one positive outlier fell above the log response time of 3.965 ($Q3 + 1.5 * \text{interquartile range}$). These eight cases, which were evenly distributed across conditions, were removed. Measures of skewness and kurtosis on the first B trial were 2.11 and 5.32 before the logarithmic transformation, but they improved to 0.588 and 0.118 after the transformation and outlier removal. The average age and sex of the participants by condition did not change significantly as a result of the outlier removal.

A trial accuracy

Because preliminary analyses did not reveal any main effects or any interaction with the sex of participants, the variable was not included in the reporting of any of the analyses.

We analyzed A trial performance because of the possibility that incorrect searches during A trials may weaken the cumulative habit and make perseveration less likely (e.g., Diedrich, Thelen, Smith, & Corbetta, 2000; Thelen, Schöner, Scheier, & Smith, 2001). Table 3 displays the number of errors across

[2] Removal of these four children from the analyses did not reveal any differences in the pattern of performance.

Table 3
Mean number of errors on A trials by condition.

Condition	Number of A trial errors
No picture	1.57 (0.25)
No verbal label	1.00 (0.20)
Experimenter verbal label	1.57 (0.27)
Child verbal label	1.46 (0.24)

Note. Standard errors are in parentheses.

A trials. The majority of children (63%) had little difficulty with A trials, as demonstrated by their scores at the lower limit (i.e., scores of 0 or 1 error). A Kruskal–Wallis test (necessary because of the non-normal distribution of errors) did not reveal a difference in total number of A trial errors by condition, $\chi^2(3) = 2.71$, $p > .05$.

B trial accuracy

The primary dependent variable was accuracy on the first B trial (see Fig. 3). Notably, 89% of the errors made on the first B trial could be classified as perseverative (defined as search at Location A or between Location A and Location B) (see Diamond et al., 1994; Marcovitch & Zelazo, 1999; Spencer, Smith, & Thelen, 2001). First, we examined the effect of the visual cue by comparing the no picture condition with all other conditions (i.e., no verbal label, experimenter verbal label, and child verbal label). Children who received visual cues performed marginally better (71% correct) than children who received indistinct hiding locations (57% correct), $\chi^2(1) = 2.66$, $p \leqslant .10$, Cramer's $\Phi = .14$. The second analysis examined the effect of labeling by comparing the no picture and no verbal label conditions with the experimenter verbal label and child verbal label conditions. Children who generated or heard labels performed marginally better (75% correct) than children who did not receive the label (60% correct), $\chi^2(1) = 3.59$, $p < .10$, Cramer's $\Phi = .16$.[3]

Finally, we examined whether improvement increased across conditions by using Bartholomew's test for order, intended to assess gradients in proportions across qualitatively ordered samples (Fleiss, Levin, Cho Paik, & Fleiss, 2003). The results support the hypothesis that the proportion of children answering correctly on the B trials follows a prescribed order, namely that the proportion correct increases as the condition provides increases in visual and labeling cues, $\chi^2(c_1 = .50,\ c_2 = .49) = 4.78$, $p < .05$. In addition, we conducted a test for a linear relationship between the ordered conditions and accuracy with the Mantel–Haenszel chi-square. The significant result provided evidence of this linear relationship, that is, that increases across condition (or increases in visual and label cues) are associated with increases in accuracy, $\chi^2(1) = 4.71$, $p < .05$.

B trial error run

We defined error run as the number of B trial errors before children searched correctly or withdrew from the task ($n = 7$). Table 4 displays the average error run by condition. A Kruskal–Wallis test, justified by the non-normal distribution on the error run, approached significance, $\chi^2(3) = 7.39$, $p < .10$. Follow-up Mann–Whitney *U* tests comparing error run in the no picture condition to other conditions revealed a marginal reduction in the experimenter verbal label condition, $U(N = 74) = 548.0$, $p < .10$, and a significant reduction in the child verbal label condition, $U(N = 72) = 462.5$, $p < .05$.

Exclusion of poor A trial performers

As mentioned earlier, incorrect A trial performance may be indicative of off-task behavior or may weaken the habit toward Location A and make B trial responses difficult to interpret (e.g., Diedrich

[3] It was possible for children in the no verbal label condition to generate a verbal label spontaneously during the hiding event (e.g., say the star is hiding in the dog box). However, this was observed infrequently ($n = 5$), and only one child generated the label on the critical B trial. Removal of these five children did not alter the pattern of performance.

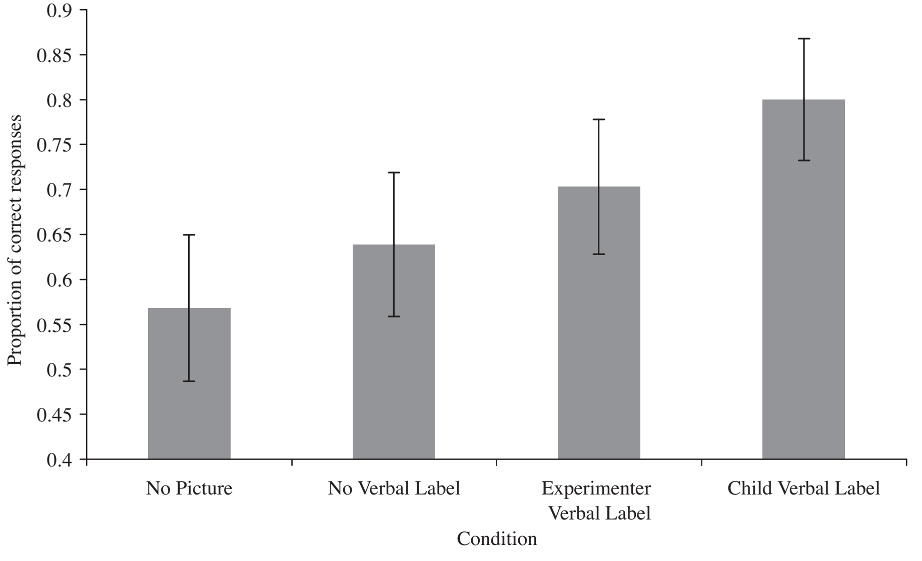

Fig. 3. Proportions of children answering correctly by condition.

Table 4
Error run on B trials by condition.

Condition	Error run
No picture	2.27 (0.61)
No verbal label	1.08 (0.36)
Experimenter verbal label	0.68 (0.26)
Child verbal label	0.37 (0.18)

Note. Standard errors are in parentheses.

et al., 2000; Thelen et al., 2001). We reanalyzed data including only children who committed 0 or 1 A trial error ($n = 91$). Overall, the analyses revealed similar patterns to the complete sample. First, we examined the effect of the visual cue. The difference in accuracy approached significance, with children who received visual cues performing better (77% correct) than children who searched in the no picture condition (59% correct), $\chi^2(1) = 2.76$, $p \leqslant .10$, Cramer's $\Phi = .17$. The second analysis examined the effect of labeling, with children who had labels present performing better (86% correct) than children who did not receive labeling cues (62% correct), $\chi^2(1) = 6.49$, $p < .05$, Cramer's $\Phi = .27$.

Bartholomew's test for order was also conducted. The results support the hypothesis that the proportion of children answering correctly increases as the condition provides increases in visual and labeling cues, $\chi^2(c_1 = .44,\ c_2 = .52) = 7.13$, $p < .05$. In addition, we conducted a test for a linear relationship between the ordered conditions and accuracy with the Mantel–Haenszel chi-square. The significant result provides evidence of this linear relationship, that is, that increases across condition (or increases in visual and label cues) are associated with increases in accuracy, $\chi^2(1) = 6.61$, $p < .05$.

Finally, error run was also analyzed. The error runs were 1.27, 1.11, 0.18, and 0.35 for the no picture, no verbal label, experimenter verbal label, and child verbal label conditions, respectively. A Kruskal–Wallis test revealed a difference in error run by condition, $\chi^2(3) = 7.98$, $p < .05$. Mann–Whitney U tests revealed that children in the verbal label conditions (both child and experimenter) had smaller error runs compared with the no picture condition, both Us(Ns = 44 and 42) < 174.0, both ps < .05.

Discussion

The current research was designed to assess the influence of differential visual and labeling cues on performance of older 2-year-olds in an EF task. The results indicated improvement as the level of visual and linguistic support increased. Importantly, young children benefited most when the labeling was self-generated.

This ordered improvement across conditions was predicted by the HCSM (Marcovitch & Zelazo, 2006, 2009) that postulated that the addition of visual and verbal label cues increases the likelihood that children will reflect on the hiding event, which they can later use to guide behavior. It is important to note that improvement in performance cannot be explained by visual and labeling cues encouraging better overall performance on this task because these additional cues did not influence performance on the A trials. This suggests that there is something specific about labeling that influences the executive component of the task (switching behavior). In addition, it must also be noted that the presence of pictures improved performance (i.e., the no picture condition compared with all other conditions). This finding is not surprising and is in line with research suggesting that distinctive hiding locations improve A-not-B performance (e.g., Bremner, 1978; Bremner & Bryant, 2001; Smith, Thelen, Titzer, & McLin, 1999). However, the mere presence of distinct visual cues cannot account for the ordered improvement across conditions.

The superior performance by children who generated a label suggests that simply providing children with a label might not be enough for them to elicit reflection of their current experiences. This finding is consistent with the Vygotskian perspective that young children use and produce language in a social context. Specifically, children may need to generate language to speak to others for it to have representational meaning necessary to guide behavior. Although older children are probably more adept at private and inner speech, younger children may need a more supportive social context (i.e., experimenter prompting) to encourage them to form a linguistic representation of the hiding location. According to the LoC theory (Zelazo & Zelazo, 1998), once a linguistic representation is active, children are able to hold two representations simultaneously (i.e., the *word* and the *object* that the word is to represent), permitting further reflection on the representation. Active generation is the first instance that allows children to link their current experience to memory, and at 2.5 years of age children may need this active experience to form a linguistic representation.

The benefit from self-generated material is also consistent with the generation effect (Slamecka & Graf, 1978), a phenomenon (usually found in adults) that generated material is better remembered compared with presented material. The effort hypothesis (Jacoby, 1978) maintains that generation increases interest, which requires more cognitive processing resources. This theory further postulates that the benefits of generation may come into play at the level of the central executive (responsible for the allocation of attention) in the working memory system (Baddeley, 1996). Generated material may recruit more attention, leading to better processing of the material in working memory. The current study supports this hypothesis given that children who generated a label for the hiding location on the B trial may have allocated more attention to the hiding location, and because it is actively processed in working memory, this information may be maintained in mind longer and be available during the time of search. Following a similar line of reasoning, Kirkham and colleagues (2003) suggested in the context of the DCCS that generation redirects attention such that a label may assist children in disengaging from the attentional inertia by allowing the shift of attention to the new relevant information.

However, theories that promote reflection (e.g., HCSM, LoC) differ from attention-based theories in the role that labeling plays in the control of behavior. Attention theories do not specifically hypothesize that there is a labeling generation effect per se; in fact, the increase in attention could occur in any modality (e.g., visual, motoric). In contrast, reflection-based theories give special credence to linguistic generation, especially early in life when dual representation emerges (Zelazo & Zelazo, 1998).

The benefit of linguistic labels in this task also supports other representational accounts such as Munakata's (1998) active-latent account. From this perspective, flexible thinkers (i.e., those who correctly switch behavior in EF tasks) rely on active representations when representing task-relevant information, whereas perseverators rely on latent representations.

Kharitonova, Chien, Colunga, and Munakata (2009) found that switchers' active representations of task-relevant information were more abstract (i.e., they were related to higher order rules and generalized to new stimuli). Therefore, the active-latent account would also predict that the presence of verbal labels should encourage children to use strong active representations (i.e., an abstract verbal label that can be represented separately from the task context) to control behavior.

Although the active-latent account and HCSM both focus on the importance of representations in EF, the HCSM emphasizes the importance of generating a label. In fact, the HCSM speculates that even in conditions where children are not required to explicitly generate a label, it is likely that successful children are generating a label spontaneously (and/or covertly). One possible experimental manipulation that may distinguish the HCSM from the active-latent account is the co-occurrence of an experimenter label with a child-generated label.[4] The HCSM predicts that a child-generated label would influence behavior equivalently whether or not there was an additional experimenter label. Conversely, the active-latent account would suggest that children are more likely to form a stronger active representation to guide behavior when both the experimenter and the child label the hiding location, and this condition would lead to increased task switching.

Finally, the results from the current study are broadly consistent with the dynamic systems perspective (Spencer & Schutte, 2004; Thelen et al., 2001), which focuses on how multiple systems (e.g., memory and attention processes) interact with external cues in the task environment to influence the motor plan and behavior of children. Consistent with the trends found in the current study, this theory would hypothesize that visual and labeling cues in the environment improve performance because they increase the distinctiveness between Location A and Location B (Bremner & Bryant, 2001; Smith et al., 1999). However, the pattern of improvement across conditions and the benefit of generating labels in particular suggest that the presence of visual and labeling cues alone cannot explain performance. From the dynamic systems perspective, labeling an item may serve to differentiate it further from other similar items, but there is no stated reason why self-generated labels should be more useful than experimenter-generated labels. In contrast, the HCSM and the LoC model do specify how self-generated labels elicit the strongest from of representation.

In conclusion, the current research provides compelling evidence that children younger than 3 years can benefit from labeling on EF tasks and that performance improves with the addition of more visual information and gradations of labeling. From a developmental standpoint, it will be important to determine how children come to generate labels spontaneously (i.e., without external prompting) and use that information to guide themselves seamlessly from one event to another.

Acknowledgments

This research was supported by Grant 5R03HD054609-02A1 awarded to the second author from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development or the National Institutes of Health. The authors express their thanks to Jared Cook, Lilly Lawson, Natalie Hellman, Dani Ryan, and Lindsey Waldron for help with data collection and to the families who participated in the study.

References

Baddeley, A. (1996). The concept of working memory. In S. Gathercole (Ed.), *Models of short term memory* (pp. 1–27). Hove, UK: Psychology Press.

Bremner, A., & Bryant, P. (2001). The effect of spatial cues on infants' responses in the AB task, with and without a hidden object. *Developmental Science, 4*, 408–415.

Bremner, J. G. (1978). Spatial errors made by infants: Inadequate spatial cues or evidence of egocentrism? *British Journal of Psychology, 69*, 77–84.

[4] We thank an anonymous reviewer for suggesting the experimental manipulation distinguishing the HCSM from the active-latent account.

Carlson, S. M. (2005). Developmentally sensitive measures of executive functioning in preschool children. *Developmental Neuropsychology, 28*, 595–616.
Carlson, S. M., Davis, A. C., & Leach, J. G. (2005). Less is more: Executive function and symbolic representation in preschool children. *Psychological Science, 16*, 609–616.
DeLoache, J. S. (1989). Young children's understanding of the correspondence between a scale model and a larger space. *Cognitive Development, 4*, 121–139.
DeLoache, J. S. (2000). Dual representation and young children's use of scale models. *Child Development, 71*, 329–338.
Diamond, A., Cruttenden, L., & Neiderman, D. (1994). AB with multiple wells: Why are multiple wells sometimes easier than two wells? Memory or memory + inhibition? *Developmental Psychology, 30*, 192–205.
Diedrich, F. J., Thelen, E., Smith, L. B., & Corbetta, D. (2000). Motor memory is a factor in infant perseverative errors. *Developmental Science, 3*, 479–494.
Fleiss, J. L., Levin, B., Cho Paik, M., & Fleiss, J. (2003). *Statistical methods for rates and proportions* (3rd ed.). Hoboken, NJ: Wiley-Interscience.
Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*, 31–60.
Gerstadt, C. G., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3½–7 year olds on a Stroop-like day night test. *Cognition, 53*, 129–153.
Homer, B. D., & Nelson, K. (2005). Seeing objects as symbols and symbols as objects: Language and the development of dual representation. In B. D. Homer & C. S. Tamis-LeMonda (Eds.), *The development of social cognition and communication* (pp. 29–52). Mahwah, NJ: Lawrence Erlbaum.
Homer, B. D., & Nelson, K. (2009). Naming facilitates young children's understanding of scale models: Language and the development of symbolic understanding. *Journal of Cognition and Development, 10*, 115–134.
Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior, 17*, 649–667.
Jacques, S., & Zelazo, P. D. (2005). Language and the development of cognitive flexibility: Implications for theory of mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 144–162). Oxford, UK: Oxford University Press.
Kharitonova, M., Chien, S., Colunga, E., & Munakata, Y. (2009). More than a matter of getting "unstuck": Flexible thinkers use more abstract representations than perseverators. *Developmental Science, 12*, 662–669.
Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science, 6*, 449–476.
Luria, A. R. (1979). *The making of mind: A personal account of Soviet psychology*. Cambridge, MA: Harvard University Press.
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. (2006). 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. *Journal of Cognition and Development, 7*, 477–501.
Marcovitch, S., & Zelazo, P. D. (2009). A hierarchical competing systems model of the emergence and early development of executive function. *Developmental Science, 12*, 1–25.
Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.
Müller, U., Zelazo, P. D., Hood, S., Leone, T., & Rohrer, L. (2004). Interference control in a new rule use task: Age related changes, labeling, and attention. *Child Development, 75*, 1594–1609.
Müller, U., Zelazo, P. D., Lurye, L. E., & Liebermann, D. P. (2008). The effects of labeling on preschool children's performance in the Dimensional Change Card Sort task. *Cognitive Development, 23*, 395–408.
Munakata, Y. (1998). Infant perseveration and implications for object permanence theories: A PDP model of the A-not-B task. *Developmental Science, 1*, 161–211.
Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.
Piaget, J. (1959). *Language and thought of the child*. New York: Humanities Press.
Rattermann, M. J., & Gentner, D. (1998). The effect of language and similarity: The use of relational labels improves young children's performance in a mapping task. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from cognitive, computational, and neural sciences* (pp. 274–282). Sofia, Bulgaria: New Bulgarian University.
Sheskin, D. J. (2004). *Handbook of parametric and nonparametric statistical procedures* (3rd ed.). Boca Raton, FL: CRC Press.
Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 592–604.
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., & Schutte, A. R. (2004). Unifying representations and responses: Perseverative biases arise from a single behavioral system. *Psychological Science, 15*, 187–193.
Spencer, J. P., Smith, L. B., & Thelen, E. (2001). Test of a dynamic systems account of the A-not-B error: The influence of prior experience on the spatial memory abilities of two-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.
Vygotsky, L. S. (1986). *Thought and language (A. Kozulin, Trans.)*. Cambridge, MA: MIT Press (Original work published 1934).
Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences, 8*, 12–17.
Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology, 1*, 198–226.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development, 68* (3, Serial No. 274).

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. In H. H. Jasper, L. Descarries, V. F. Castellucci, & S. Rossignol (Eds.), *Consciousness: At the frontiers of neuroscience* (Advances in Neurology, Vol. 77, pp. 149–165). New York: Lippincott-Raven.