

## Integration of Thought and Action: Arm Weights Facilitate Search Accuracy in 24-Month-Old Children

Martha E. Arterberry  
*Colby College*

Susan J. Hespos  
*Northwestern University*

Rachel A. Herth   
*Colgate University*

Search errors are common in cognitive tasks with infants and toddlers, and these errors reveal important insights to the development of competence and performance. Rivière and Lécuyer (2008, *Journal of Experimental Child Psychology*, 100, 1) demonstrated that 29-month-olds typically make an error during a search task involving invisible displacement. However, performance improves significantly when children wear weighted wrist bands while doing the task. To investigate this phenomenon further, we tested 24-month-old children in an identical search task ( $N = 35$ ). Half the children wore weighted wrist bands, and the rest were in a no-weight condition. To test how far this phenomenon generalizes, we also tested the same children in a second search task where they needed to find a ball that had rolled behind one of four doors. The results showed that children in the no-weight condition replicated previous findings of poor performance on both search tasks. Unlike 29-month-olds, the 24-month-olds in the weighted condition did not immediately show improvement on the search tasks. However, after an initial search attempt, children wearing weights performed significantly better than chance. The findings shed new light on the interplay between thought and action.

An interesting discrepancy has emerged in early cognitive development: Young infants appear to have knowledge that older children fail to use. The first well-documented example of this discrepancy is in the area of object permanence. Despite evidence that very young infants know that hidden objects continue to exist, older children make errors when searching for hidden objects (e.g., Baillargeon, 1987; Baillargeon, Spelke, & Wasserman, 1985; Piaget, 1954; Rivière & Lécuyer, 2008). Keen (2003) pointedly

asked: Why do infants look so smart and toddlers look so dumb? To illustrate this question, she focused on a discrepancy between studies of infants' knowledge of solidity and continuity and toddlers' difficulty in searching for objects in tasks that should rely on the same knowledge. Spelke, Breinlinger, Macomber, and Jacobson (1992) showed that 2-month-old infants looked longer when a ball rolling behind a screen ended up on the far side of a barrier (a physically impossible event) rather than on the near side of the barrier (a physically possible event). Yet, 24-month-olds performed at chance levels when they were asked to choose the correct door (out of four options) to retrieve a ball that had rolled down a ramp behind an occluding panel and stopped at a barrier (the barrier was visible above the panel; Berthier, DeBlois, Poirier, Novak, & Clifton, 2000). Twenty-four-month-olds continued to make errors in simplified versions of the task including: When a transparent panel was used, when the number of doors was reduced to only two, when the children could see the full movement and resting location of the ball before it was occluded, when the children were provided with landmarks, or when the children merely pointed to the location of the ball (Butler, Berthier, & Clifton, 2002; Mangalindan & Schmuckler, 2011; Mash, Keen, & Berthier, 2003; Mash, Novak, Berthier, & Keen, 2006; Shutts, Keen, & Spelke, 2006).

Sometimes a task as simple as retrieving a toy hidden under a cloth is a challenge for young children, and how they integrate goals with the means to achieve them has captivated developmental psychologists as far back as Piaget (1954). For example, in a three-location invisible displacement task, Piaget presented his children with three distinctive objects—a beret, a handkerchief, and a jacket (labeled A, B, and C)—which served as screens to conceal the location of a hidden object. In the task, the child's attention was first drawn to a small pencil in Piaget's hand, after that he placed his closed hand under location A. Then with the hand remaining closed, he passed it under a second cover at location B, and yet again under a third cover at location C. After emerging from location C, he opened his hand to reveal the pencil was gone. The child was asked to find the pencil. In Piaget's version of the task, the hidden toy is always left at location C. Piaget and others report that by 20 months children typically perform a successful search at location C and find the hidden object (Piaget, 1954; Uzgiris & Hunt, 1975).

In a recent variant of the task used by Rivière and his colleagues, a small object is concealed in the experimenter's hand and the hand moves under three identical cloths. In contrast to Piaget's version of the task, the object is always left at location B, and there is a visible bump in the cloth indicating the location of the hidden object (Rivière & Falaise, 2011; Rivière & Lécuyer, 2003, 2008). Even though the visible bump may represent a clue to the correct search location, this task is considerably more difficult, and most 29-month-old children fail the task. What makes this variation more difficult is interesting because it specifies the information children use in their search strategy. Unlike Piaget's invisible displacement task, there are conflicting cues in this version of the task. Attention is drawn to location C when the empty hand is revealed yet there is a visible protrusion of the ball under the cloth at location B. Also, children may have experience with adults hiding objects, and they learn that one does not move to a third hiding location if the object was left at the second location. When 29-month-old children were asked to find the object, 50% of the children searched at location C, 45% searched correctly at location B, and the remaining 5% searched at location A. Hence, Rivière and Lécuyer (2008) named the new variation of the search task the C-not-B error.

Curiously, the C-not-B error persists even when the experimenter opens the hand between each hiding location, indicating the object is still in the hand after moving under A but disappears from the hand after moving under B and again is not there after moving under C (Rivière & Falaise, 2011). Moreover, children understand that the protusion under the cloth indicates the existence of a hidden object in this context. Rivière and Lécuyer (2008) presented children with the three cloths and an object visible next to the cloths, then a screen occluded the display and the experimenter placed the object under the cloth at location B. When the screen was removed, the children were encouraged to retrieve the toy, and 17 of 18 children accurately searched at location B. Rivière and colleagues interpret the majority of searches at location C as a prepotent response incurred by the hand passing under the cloth at location C prior to revealing that it is empty. This action draws attention to location C and triggers an automatic response for the children to search at location C. The argument is that this automatic response must be inhibited such that it does not compete with the knowledge that the bump in location B indicates the location of the toy.

Here, we make a distinction between perception–action responses and representation-guided responses. Perception–action responses are comprised of the visual and motoric information that children use to respond to the task. These include planning and executing of actions in response to the task that were not required in the infant looking-time tasks. Perception–action responses may be guided by any number of variables within the task whether they are relevant or not. For example, a child may have a favorite side or hand position and all responses will be to objects on that side or using that hand position. Another possibility is a child may find movement during a task, such as when an object is being hidden, more attractive than the perceptual information indicating where the object was actually hidden (such as a bump under a cloth). Thus, not all perception–action responses are accurate. Representation-guided responses, in contrast, incorporate knowledge, for example, knowing that an object continues to exist when hidden or that the protrusion under the cloth specifies that location of the object. This knowledge is continuous through development and it is the same understanding of how objects behave and interact that guides infants' sophisticated responses in the studies with younger children. Thus, representation-guided responses rely on an understanding of the task by attending to pertinent information that leads to success and ignoring some perceptual variables that may be more salient but less helpful.

An interesting difference in perception–action and representation-guided responses is illustrated by a second group of children tested by Rivière and Lécuyer (2008) in the C-not-B task. Some of the 29-month-old participants wore weighted wrist bands during the C-not-B task; others wore bright colored bands with no weights. Children in the no-weight (colored band) condition split their choices between B and C as described above. However, children in the weighted condition performed significantly better than children in the no-weight condition: 78% of the children successfully searched at B and only 5% searched at C. Rivière (2014) suggested that the weighted wrist bands interrupted the motor memory allowing children to attend to the protrusion under the cloth that then guided search to the correct B location. In other words, the default response, the perception–action response of searching to the last seen location of the hand, is inhibited by the weights, and thus, children are able to respond to the information provided by the bump in the cloth (a representation-guided response).

This fascinating finding suggests that wrist weights interrupt the searching at location C.

Faced with the evidence that weighted wrists seem to improve performance, it raises the question of why? The benefit of weighted wrists does not appear to be a speed-accuracy trade-off. There were no significant differences in latencies to start acting between conditions, so increased delay is not the explanation for better performance for children in the weighted condition (Rivière & Lécuyer, 2008). Instead, the weights appear to alter the motor planning system. In this sense, action may be more of an embodied experience and the weights alter aspects of the perception-action coupling such that other features of the task can be attended to, like the bump under the cloth at B. Rivière (2014) suggests that when children make the C-not-B error, they rely on automatic motor routines that guide them to the last location of the experimenter's hand prior to revealing that the hand is empty. Further support for this interpretation comes from a task where children's search performance in the C-not-B task improves when their motor response requires a novel adjustment, such as tapping a stick on the correct location or navigating around a transparent barrier to retrieve the hidden object (Rivière & David, 2013). These findings beg the question of whether the competition between a perception-action response versus a representation-guided response generalizes to other search tasks.

Perception-action responses may originate in several ways. One way is likened to a habit that is built up across experiences, such as observing people hiding objects, and learning that the likely location of the hidden object is the last place the experimenter's hand went (Rivière & Lécuyer, 2008). Another way is a perseveration error where a response is reinforced during the course of an interaction. Perseveration errors have been invoked as explanations for the A-not-B task (Marcovitch & Zelazo, 1999; Piaget, 1954; Wellman, Cross, & Bartsch, 1987). The A-not-B task involves only two locations, and after hiding (and finding) an object at A across several trials, the object is hidden in B. Children who successfully retrieve the object at A but continue to search at location A after the object has been hidden at location B have made the A-not-B error. Clearfield, Diedrich, Smith, and Thelen (2006) observed an interesting developmental trajectory. Like many others, they demonstrated the A-not-B error with 7- and 8-month-old infants. The insight relevant to this paper was that Clearfield et al. looked at performance of infants in the months prior to this preservation error. They found a developmental trajectory where early searching was variable, "in the current moment," and yet accurate. Specifically, 5-month-old infants retrieved objects at A and B successfully and did not make the A-not-B error. To account for the younger infants' successes followed by older infants' failures, Clearfield and colleagues demonstrated that 5-month-old infants had correct retrievals at A and B, but their reaches had inconsistent direction and velocity. Between 5 and 8 months, reaching skills became more consistent and controlled. One impact of repetitive and consistent reaches at 7 months was that the reaches also became modulated by the regularities in the just-previous activity. The now-more-skilled reachers made new mistakes; specifically, they continued to search at A when the hiding location was switched to B. More broadly, Clearfield et al. demonstrated that looking at an earlier developmental time point revealed that children progressed along a developmental trajectory that was continuous and becoming more sophisticated over time even though the measured behaviors changed from success at retrieving the hidden object at 5 months and failure to retrieve it at 7 months. The findings from Clearfield et al. (2006) coupled with the performance of

29-month-olds tested by Rivière and Lécuyer (2008) raise the question of how younger children would perform on the C-not-B task. By investigating how children at a slightly younger age perform in this task, we may gain new insights about the proposed competition between a perception–action response, leading to a search error, versus a representation-guided response of looking under the cloth with a bump, and how the result of such competition manifests developmentally.

### The present study

In the present study, we address two questions raised by the research reviewed above. Our first question asked what does performance on the C-not-B task look like at an earlier age? We focused on 24-month-olds because it is the midpoint between the 29-month-olds who show the C-not-B error and Piaget’s children who had passed the invisible displacement task at 20 months when the object was hidden at C. Half of the children were tested with weighted wrist bands, and the other half wore wrist bands with no weight.

Our second question was whether the performance advantage conferred by weighted wrist bands generalizes to a different search task. In selecting a second search task, we capitalized on two attributes of the C-not-B task. First, there was evidence that infants have expectations about the location of the hidden object based on results from looking-time paradigms. For example, Baillargeon (1995) demonstrated that 12-month-old infants expect that a protrusion under a cloth indicates the location of a hidden toy. Second, even though the knowledge is evident in infants, there is something about the search task that makes it challenging to toddlers.

The Door task by Berthier et al. (2000) is a search task that met both these criteria. Spelke et al. (1992) showed that 2-month-old infants expect that the ball should come to rest on the near (not far) side of the barrier. In contrast, 24-month-olds performed at chance levels in a search version of this task (Berthier et al., 2000). In the task, a ball rolls along a ramp behind a panel with four doors, and a wall, the top of which is always visible, stops the ball by one of the doors. The child is asked to open the correct door to find the ball (see Figure 1). It is not until children are 36 months that they are able to consistently open the correct door. Like the C-not-B task, half of the



**Figure 1** Apparatus used in the Door task. The doors are labeled Door 1–4 from left to right.

children participated in the tasks wearing weighted wrist bands and the other half wore wrist bands with no weight.

Our first goal is to see if the performance in the C-not-B for 24-month-olds is similar to that of older children. It is possible that 24-month-old children may show more accurate performance than 29-month-olds, mirroring the findings of Clearfield et al. (2006) in the A-not-B tasks. Twenty-four-month-olds have less experience with invisible displacement tasks than 29-month-olds, and as a result, they might not have developed an expectation that a missing object would be found at the last seen location of the hand and/or that information from a bump under a cloth is important in this context. Alternatively, the behavior might be systematic, but incorrect, like the older children in the work by Rivière and colleagues who make the C-not-B error. Or, the behavior might be at chance levels. Our second goal is to replicate the Door task in the no-weight condition. Berthier et al. (2000) tested 24-month-old children in this task so we expected that our children in the no-weight condition will replicate their chance performance. Our third goal is to evaluate whether weighted wrist bands improved performance. Of particular interest is whether the advantage conferred by weighted wrist bands is similar across the two search tasks. A similar pattern of results would suggest that the weights disrupt perception–action responses allowing the children to make representation-guided responses across differing perceptual and cognitive demands.

## METHOD

### Participants

Thirty-five 24-month-old children ( $M$  age = 23.47,  $SD$  = .57 months; 20 boys) participated. Twenty-nine of the children completed both tasks. Three completed only the C-not-B task, and three completed only the Door task. Children who completed only one task were included in the analyses. The resulting sample size was 16 children in each condition, which is comparable to the sample used by Rivière and Lécuyer (2008) and Berthier et al. (2000). Families were recruited via purchased mailing lists and came from middle-socioeconomic status households (94% of parents had some college or higher). The ethnicity of the participants was 18% Hispanic. The racial composition was 93% White and 7% multiracial (two white and African American and one white and Asian American). Parents received \$20 for their child's participation. The work was conducted under IRB protocols approved by Colby College (# 2014-052, Title: Integrating Perception and Action in Infancy) and Northwestern University (eIRB Protocol # STU00010996, Title: Object Cognition in Infants and Adults). The work was conducted according to guidelines specified in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection.

### Materials and apparatus

#### *Wrist bands*

Two weighted wrist bands, one for each arm, were created by inserting two C-sized batteries into a black sports headband and sewing the ends together. Each wristband



measured 23 cm in circumference, was 5 cm wide, and weighed 136 g. Based on an average body weight of children aged 24 months (11,850.1 g), each band comprised 1.15% of the child's body weight. The amount of weight was comparable to the 200 g weights used with 29-month-olds by Rivière and Lécuyer (2008). No-weight wrist bands were created by inserting two fabric rolls the same size and density of the batteries into a black sports headband and sewing the ends together. The resulting wrist bands were similar in bulk to the weighted wrist bands but only weighed 15 g each. All children wore a long-sleeved garment to hide the wrist bands from view and reduce distraction.

### *C-not-B task*

After Rivière and Lécuyer (2008), three blue cloths, measuring 16 × 16 cm, and an orange ping-pong ball (12 cm in circumference) were used.

### *Door task*

The same apparatus used by Berthier et al. (2000) was used in the present study (Figure 1). The apparatus, constructed of wood and painted white, consisted of a ramp (74.5 cm wide × 14.1 cm high × 14.2 cm deep), a front occluding panel containing four doors (58.7 × 27.8 × 1.8 cm), and a back panel (76.7 × 27.8 × 1.8 cm). The ramp emerged 16.0 cm from the left edge of the occluding panel. One of four green walls could be inserted in grooves in the front and back panels at one of four locations along the ramp. All of the walls were 16.4 cm wide but varied in height (25.5, 27.5, 30, 32.5 cm) such that the top of each wall emerged 7.5 cm above the top of the occluding panel. The wall stopped a blue racquet ball (17.5 cm in circumference) rolling down the ramp to the left of one of the four doors (11.0 × 17.6 cm) in the occluding panel. Each door was painted gray, was hinged at the bottom, and had a wooden knob (3.0 cm in diameter) placed 2.0 cm from the top. Magnets kept the doors closed until the child pulled on the knob. Two dowels (each measuring 75.7 × .5 × 1.3 cm) glued on the ramp allowed the ball to move along a straight trajectory.

### *Video recording*

Two digital cameras provided either an overhead or frontal view of the child. The cameras were attached to an iMac. Digital audio and visual recordings were made using PhotoBooth (© Apple). In the Door task, the child was filmed from overhead, allowing for a clear view of the ball's location, the child's arm movement, and location of door contacted. In C-not-B, the child was filmed from the front so that the child's upper body, including the head and direction of gaze, hands, and objects on the table, was visible.

### *Procedure*

After obtaining consent, parents were asked to place the wrist bands, and if needed, a long-sleeved garment, on their child (some children arrived with long-sleeved shirts). If the child objected to wearing the wrist bands (two intended for the no-weight condition and three intended for the weight condition), the child was tested

without the wrist bands and placed in the no-weight condition. The parents and child were then introduced to the play room, and the child was seated at a child-sized table in his or her own chair. The parent sat behind the child and was asked not to prompt his or her child; however, encouragement or support could be offered as needed. Children participated in up to three tasks (two of which are reported here<sup>1</sup>), and the order of tasks was randomized across participants. Data were still included if children participated in one task (hence the reason why there are 15 children in the no-weight condition for the C-not-B task and 16 in no-weight condition for the Door task).

### *C-not-B task*

The child was first introduced to the ball and the three cloths. The experimenter explained that she was going to hide the ball and that she wanted the child to find it. She laid the three cloths in a row on the table. She emphasized that the child needed to wait until she was done hiding the ball before looking for it. The experimenter then enclosed the ball in her hand, moved it under the first cloth (A), then moved her hand out without opening it and moved it under the second cloth (B). At B, she released the ball, but when she removed her hand her fingers were still curled such that she appeared to still be holding the ball (the ball made a visible bump under B). Then she moved her hand under the third cloth (C). Finally, she removed her hand from C and opened it to show that her hand was empty. The child was asked to find the ball. If the child was not correct on their first attempt, the experimenter removed the ball from B; in other words, children were allowed only one try per trial to find the ball. The procedure was repeated one more time, resulting in a total of two trials. Starting position (with the farthestmost left or right cloth) was randomly varied across participants but remained constant across the two trials. Sixteen children were tested in the no-weight condition, and 16 were tested in the weight condition. Two children in the no-weight condition completed only one trial, one completing only Trial 1 and the second completing only Trial 2, resulting in a total of fifteen children in each trial for the no-weight condition. Both children were only included in the analysis of the trial they completed.

### *Door task*

The apparatus was placed on the table within reach of the child, and the child was introduced to it, particularly the doors and the ball. Modeled after Berthier et al. (2000), children participated in a familiarization session and a test session. The familiarization session comprised two phases. First, the experimenter opened the first door and placed the ball inside it. She then asked the child to retrieve it. When the child was successful, the experimenter cheered and said "good job." Next, the experimenter opened the third door, placed the ball inside it, and asked the child to retrieve it. All children were able to accomplish this part of the task. For the next phase, the trials

---

<sup>1</sup>The third task was a categorization task using the sequential touching procedure (e.g., Mandler, Fivush, & Reznick, 1987; Bornstein & Arterberry, 2010), and children completed the task with or without weights, depending on weight condition. Because this task did not involve search, it was not included in the current manuscript.



began with the apparatus out of reach of the child. First, the panel with the doors was removed and a wall was placed at position 4 (where the fourth door would be) and the ball was rolled down the ramp. The experimenter pointed to the ball and said "look, the ball stopped here." This trial was repeated. Next, the experimenter placed the wall at position 4, tapped it saying "here's the wall," placed the occluding panel with all the doors open in front of the ramp saying "here's the doors," and then "here's the ball," and released the ball to roll down the ramp. Once the ball stopped, the experimenter saying "look, the ball stopped here." This trial was repeated. If the child wanted to retrieve the ball, he or she was allowed to do so, and the experimenter asked for the ball back before beginning the next trial. The next four familiarization trials were similar to the previous trial, but only one door was open (the correct location) when the occluding panel was placed in the apparatus. Again, the experimenter pointed out the wall, the doors, the ball, and where the ball had stopped. Again, if the child desired, he or she was able to retrieve the ball. Trials were conducted for all four wall positions. Order of wall position across trials was determined randomly for each participant.

Test trials were similar to the last familiarization trials except all of the doors were closed when the occluding panel was placed in the apparatus. On each trial, the experimenter inserted the wall saying "here's the wall," placed the occluding panel into place (said "here's the doors"), drew the child's attention to the ball by saying "here's the ball," and then released the ball. The child was asked to find the ball by opening a door when the apparatus was moved within reach. The child was allowed to open up to two doors. When the child successfully found the ball, the experimenter cheered and said "good job." If the child was not successful after opening the second door, the experimenter retrieved the ball from the top of the apparatus, next to the wall, and said "here it is." Children could complete up to 12 trials. The wall location was randomized, without replacement, within each four-trial block. The same wall location was never immediately repeated. If children showed signs of disinterest (e.g., turning away from the apparatus, not attending to the wall or the ball, saying they were done), the session was terminated before 12 trials had been completed.

## Coding

### *C-not-B task*

Our two main dependent variables were coded from the video records. Location of first search was scored to determine correct or incorrect performance. The child did not have to retrieve the ball or remove the cloth; merely touching a cloth was considered a response. One researcher scored all the sessions. A second researcher scored all of the sessions to obtain a measure of reliability for search location; agreement was 100%. The video records were also coded for latency, defined as the time in seconds for the child to begin forward arm movement after the experimenter revealed her empty hand. A second coder coded 36% of the sessions. Agreement was high:  $r = .99$ . In addition, the video records were coded for where the children looked during hiding and after the hand was revealed to be empty but before they touched a cloth by two coders. Agreement again was high:  $r = .95$ .

### *Door task*

Video records were coded for the doors opened by the child on each test trial with Search 1 indicating the first door opened and Search 2 indicating the second door opened (only necessary when a child was incorrect on the first search). One researcher scored all the sessions. A second researcher scored 78% of the sessions to obtain a measure of reliability. Across 301 first and second searches, agreement was 98%. The few disagreements were resolved by the two researchers reviewing the recording together. It was possible for children to touch one door and then fully open another door. This response occurred on only 1.90% of trials, and the analyses did not differ based on whether we coded first door touched versus first door fully opened. In addition, the video records were coded for (a) the time in seconds to initiate a reach toward the first door (i.e., forward movement of the hand) from release of the ball and (b) the time in seconds to fully open the first door from the release of the ball. All of the sessions were coded by the same researcher. In addition, a second observer coded 31% of the sessions to obtain a measure of reliability. Agreement was high,  $r = .96$ .

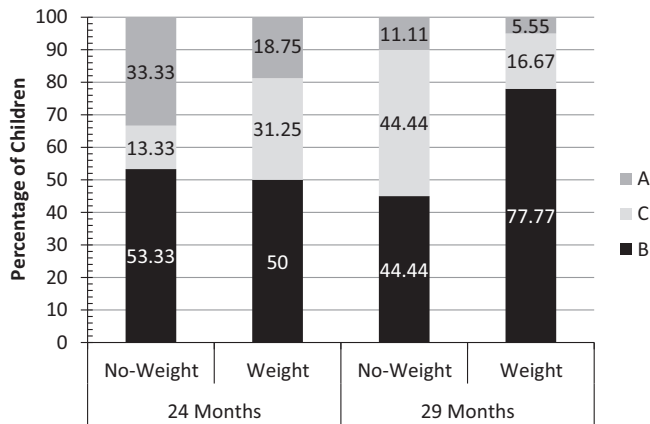
## RESULTS

For each task, the no-weight condition provided the opportunity to test for replication of previous findings. Thus, the analyses first addressed this replication question with the no-weight condition only. Next, analyses focused on differences between the weight and no-weight conditions to directly test the effect of weighted arms on performance on tasks known to be difficult for 24-month-old children. Preliminary analyses explored differences between boys and girls; however, none were found. Thus, all analyses are reported collapsing across sex.

### *C-not-B task*

The first analyses focused on how well 24-month-old children did in the task compared to 29-month-olds. Figure 2 shows the percentage of children who searched in each location across weight conditions on Trial 1 in the present study compared to Rivière and Lécuyer's (2008) study. On Trial 1, searches at location B occurred for 53% of the 24-month-old children in the no-weight condition, which is comparable to the 44% of the no-weight 29-month-olds tested by Rivière and Lécuyer (2008),  $\chi^2(1) = .71$ ,  $p = .400$ . The frequency of searches by 24-month-olds were equally distributed to the three locations and did not differ from what would be expected by chance,  $\chi^2(2) = 3.60$ ,  $p = .165$ . In other words, 24-month-olds were not very successful in finding the object at B, nor did they make the C-not-B error by searching at C. In fact, 24-month-olds in the no-weight condition made significantly *fewer* searches to C compared to the other two locations than the 29-month-olds in Rivière and Lécuyer (2008),  $\chi^2(1) = 4.95$ ,  $p = .026$ ,  $w = .39$ .

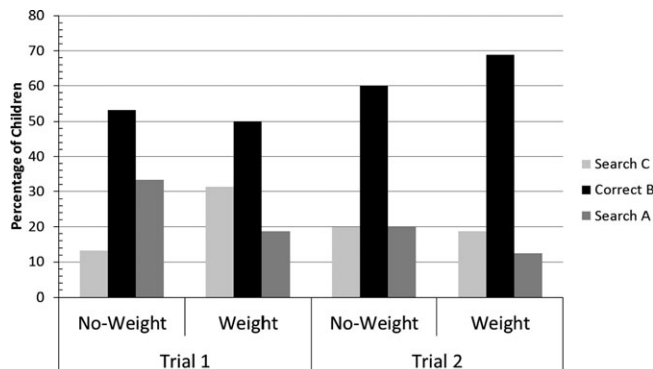
On Trial 1, 24-month-old children in the weight condition also were equally likely to search at the three locations,  $\chi^2(2) = 2.38$ ,  $p = .307$ , thus providing no advantage for weighted wrists and no evidence of the C-not-B error. Additionally, the preference for location B did not differ across the weight conditions for 24- and 29-month-olds,  $\chi^2(1) = 3.03$ ,  $p = .082$ .



**Figure 2** The percentage of children who searched at locations B (the correct location), C, and A on Trial 1 in the weight and no-weight conditions. The 24-month-olds are from the present study; the 29-month-olds' data were reported in Rivière and Lécuyer (2008).

The second analyses focused on whether there was a difference in performance across the weight conditions and across trials. Figure 3 presents the percentage of children who searched at each location across the two trials. As mentioned above, on Trial 1 the percentage of children who searched at locations A, B, and C was not significantly different from chance in either weight condition. On Trial 2, there was a difference between conditions: Children in the no-weight condition again searched equally to the three locations,  $\chi^2(2) = 4.80, p = .091$ ; however, in the weight condition children searched at B significantly more than the other two locations,  $\chi^2(2) = 9.13, p = .011, w = .75$ . Directly comparing the frequency of children who searched at location B across trials 1 and 2 as a function of weight condition revealed no significant difference,  $\chi^2(1) = .09, p = .765$ .

The third analyses focused on whether children changed their strategy from one trial to the next. We compared search locations on Trial 1 and Trial 2 for the children who made an incorrect choice on either the first or second trial. This analysis did not

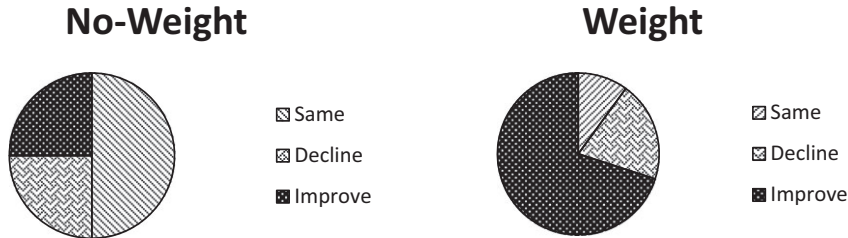


**Figure 3** The percentage of 24-month-old children in the present study who searched at A, B, and C on trials 1 and 2 as function of the no-weight and weight conditions.

include children who correctly searched at B on both trials ( $N = 6$  in both the no-weight and weight conditions) and who did not complete both trials (two participants completed only one trial in the no-weight condition). Each child's performance was coded in one of three ways: The *same* across trials 1 and 2, a *decline* (a correct search on Trial 1 and an incorrect search on Trial 2), or an *improvement* (an incorrect search on Trial 1 and a correct search on Trial 2; see Figure 4). Of the children in the no-weight condition, half searched at the same incorrect location on the first and second trials, thus showing no change from Trial 1 and Trial 2,  $\chi^2(2) = 1.20, p = .273$ . In contrast, only 20% of children tested in the weight condition searched at the same incorrect location on Trial 2, and 70% of the children improved, reflecting a significant change,  $\chi^2(2) = 6.21, p = .013, w = .79$ .

Related to strategy is where children were looking during the hiding phase and after the hand was revealed as empty. From the video recordings, we noted where children were looking during the hiding phases – at the hand/cloth, at the experimenter, or at some other location. As seen in Table 1, the vast majority children looked at the hiding location, with decreasing interest at the third location (some were looking back at earlier hiding locations or the experimenter).

We also coded how often children looked at each location after the hand was revealed to be empty but before they touched the cloth. Across 51 trials for which we had video records, 84% of the children looked at the location where they eventually



**Figure 4** Strategy patterns on Trial 2 in the no-weight and weight conditions for the *C-not-B* task, excluding children who were correct on both trials. Some children searched in the same location on Trial 2 (“Same”), some searched correctly on Trial 1 and were incorrect on Trial 2 (“Decline”), and some were incorrect on Trial 1 and correct on Trial 2 (“Improve”).

**TABLE 1**  
Percentage of Children who Looked at the Hiding Location, Another Location, or the Experimenter During the C-Not-B Task

	Trial 1—Hand Location			Trial 2—Hand Location		
	A	B	C	A	B	C
No-weight	100% to A	100% to B	82% to C 18% to B	92% to A 8% to B	85% to B 15% to Exp	67% to C 25% to B 17% to Exp <sup>a</sup>
Weight	100% to A	88% to B 12% to A	76% to C 24% to B	94% to A 6% to B	81% to B 13% to Exp 6% to A	50% to C 31% to B 19% to Exp

<sup>a</sup>Some children quickly looked to more than one location. Because all locations were coded, these percentages exceed 100.

searched; however, on two trials, children looked at B and then searched at A (one in the weight condition and one in the no-weight); on two other trials, children scanned two or more locations and then searched at A (no-weight) or C (weight); and on one trial, a child (no-weight) looked at A and searched at B.

Finally, latency to initiate the search was analyzed (Table 2). Of the 32 children, 27 video records (12 no-weight and 15 weight) were available for latency coding. A  $2 \times 2$  ANOVA with weight condition (no-weight, weight) as a between-subjects factor and trial (first, second) as a within subjects factor was conducted. The analyses revealed no main effects or interactions, all  $F$ s (1, 24) < 3.31,  $p$ s > .081.

### Door task

Thirty-two children (16 in the no-weight and 16 in the weight condition) completed this task. Recall on each trial, children were asked to open one of four doors to find the ball. Their first attempt was referred to as "Search 1" and chance performance was 25%. If the children were not correct on the first attempt, the incorrect door was left open and the children were allowed to try again within the same trial without the ball or wall location moved. This second attempt was referred to as "Search 2," and chance performance was 33% as there were only three doors left to open. After the second attempt, whether successful or not, a new trial began with a new hiding location. Mean accuracy for Search 1 and Search 2 attempts were calculated across trials. There was no difference in the number of trials completed by children in the no-weight ( $M = 9.19$ ,  $SD = 2.97$ ) and weight ( $M = 7.50$ ,  $SD = 3.01$ ) conditions,  $t(30) = 1.60$ ,  $p = .121$ .

Our first analysis determined whether we replicated the effect found by Berthier et al. (2000) for the 24-month-old age group in the no-weight condition. On Search 1, we found 26% accuracy in our no-weight condition, comparable to the 22% accuracy reported by Berthier et al. (2000). Mean accuracy on a combined score of first and

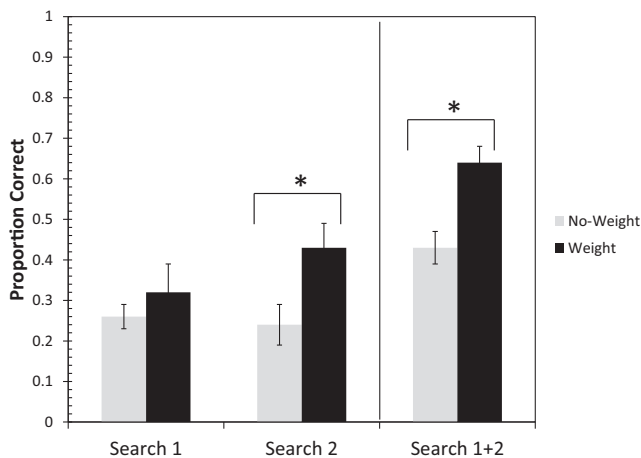
TABLE 2  
Mean Latency (in seconds) to Search by Children in the No-Weight and Weight Conditions by Task

	Condition	
	No-weight	Weight
C-not-B		
Time to initiate reach—Trial 1		
<i>M</i>	2.10	3.48
<i>SD</i>	3.39	4.88
Time to initiate reach—Trial 2		
<i>M</i>	1.36	1.33
<i>SD</i>	1.81	1.89
Door task		
Time to initiate reach		
<i>M</i>	2.20	2.27
<i>SD</i>	1.36	1.00
Time to fully open door		
<i>M</i>	4.40	4.85
<i>SD</i>	1.70	2.06

second searches (“Search 1 + 2”) was 43% for our sample. Although this value was not reported directly by Berthier et al., Figure 2 in their paper (p. 396) suggests that the mean for 24-month-olds was approximately 40%. Compared to Berthier et al., fewer children in the present study completed all 12 trials, likely due to the fact that our task was one of three tasks that children participated in during the session, whereas Berthier et al. only had the single task. When we analyzed our data retaining only children who completed more than eight of the 12 trials (the reported amount in Berthier et al., 2000), the pattern of results was identical to those reported here. Taken together, these analyses suggest that we replicated the phenomena described by Berthier et al. (2000).

Our next analysis compared performance across the weight and no-weight conditions. Figure 5 presents the proportion of correct searches in the no-weight and weight conditions across search attempts. Following Berthier et al. (2000), we tested for the effect of weight on search success on the first, second, and combined 1 + 2 attempts. There was no significant difference on the first search,  $t(30) = .71$ ,  $p = .486$ . There were, however, significant differences for the second search,  $t(30) = 2.41$ ,  $p = .002$ ,  $d = .83$ , and for the combined searches,  $t(30) = 3.28$ ,  $p = .003$ ,  $d = 1.13$ . Children in the weight condition were significantly more successful in finding the ball on the second search and on the two searches combined compared to the children in the no-weight condition.

We adopted a criterion used by Berthier et al. (2000) and Butler et al. (2002) and tallied the number of individual children whose performance exceeded what would be expected by chance, using the binomial distribution with a one-tailed rejection region of .05. In other words, it was a tally of the number of children that performed at a rate of 60% accuracy or better. In the no-weight condition, zero children reached the criterion on the first search and only three children reached the criterion on searches 1 + 2; these results are comparable to the zero and one criterion hits found by Berthier et al. (2000), respectively. In contrast, in the weight condition, two children met the

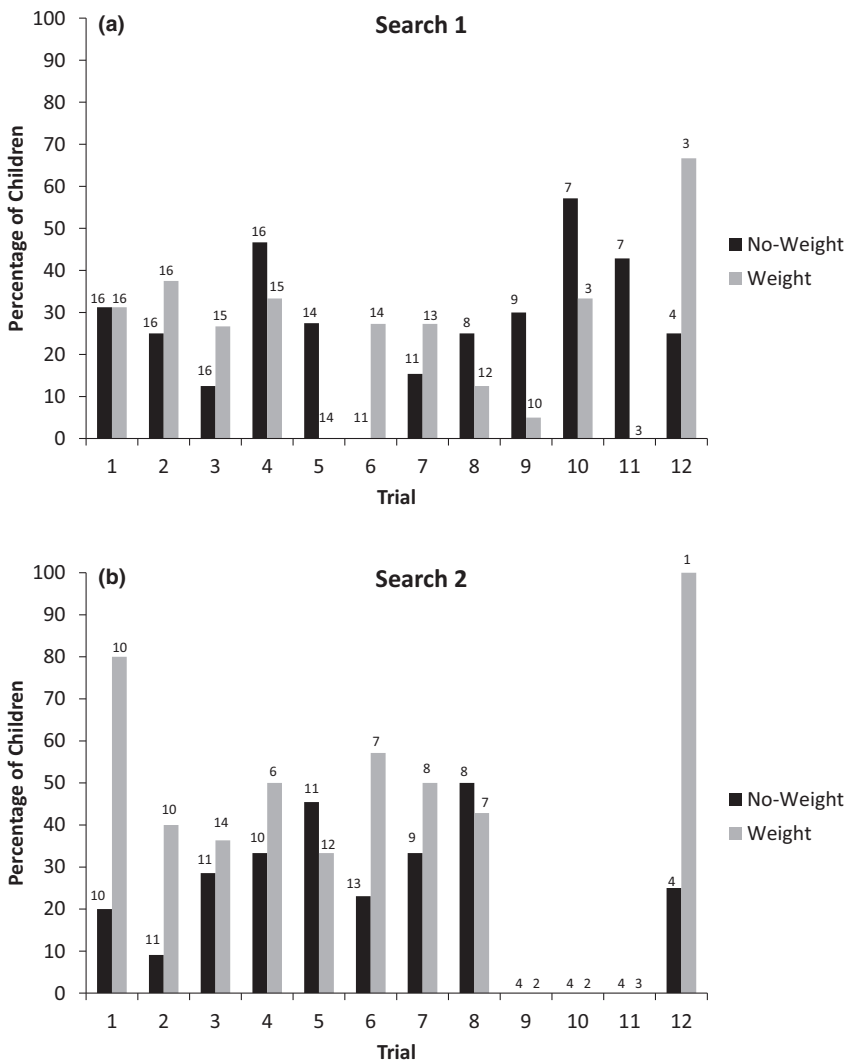


**Figure 5** Mean accuracy for finding the ball on the first attempt (Search 1) and on the second attempt (Search 2) within each trial. Combined performance across the two searches is on the far right. Error bars indicate standard error of the mean. The \* indicates that there is a significant difference between the no-weight and weight conditions on this comparison ( $p < .05$ ).



criterion on the first search and nine children met the criterion on searches 1 + 2. In short, 56.25% of the children in the weight condition exceeded 60% accuracy.

To determine whether children learned where to search across trials, we calculated the percentage of children who were correct on each trial. Figure 6 shows the percentage of children in the weight and no-weight conditions who were correct on the first search (panel A) and, if not correct on the first search, on their second search (panel B) as a function of trial. In addition, we compared mean performance (proportion correct) on the first two trials with that of the last two trials in a  $2 \times 2$  ANOVA with trial (mean of first two, mean of last two) and condition (weight, no-weight) as factors.



**Figure 6** Percentage of children in the weight and no-weight conditions who searched correctly on their first (a) or second (b) search by trial. The values above each bar indicate the number of participants out of 16 who completed each trial.

The analyses revealed no significant main effects or interactions, all  $F$ s  $< .05$ ,  $p > .840$  ( $M$ s ranged from .30 to .33,  $SD = .31$  to .37). These findings demonstrate that children did not improve across the course of the testing session even though they improved from Search 1 to Search 2 within trials.

Following Berthier et al. (2000), we analyzed search patterns to gain insight into which of two possible strategies children used in the task. The first strategy was a perseverative error in which the child might search at the last seen location of the ball, specifically, where it was located on the previous trial. Children in the weight condition showed perseverative errors on 31.02% of trials, and children in the no-weight condition showed perseverative errors on 23.12% of trials. These values are not significantly different from chance (25%),  $t(15) = .42$ ,  $p = .684$  for the no-weight condition and  $t(15) = 1.11$ ,  $p = .286$  for the weight condition. Moreover, there was no significant difference between the two weight conditions,  $t(30) = 1.12$ ,  $p = .273$ . Thus, children were unlikely to be perseverating in this task.

The second strategy involved choosing a door adjacent to the wall, either to its left or right. This strategy implicates attending to the wall and perhaps understanding the wall's physical relation to stopping the ball. To evaluate this strategy, children's first searches were coded as being at a door adjacent to the wall or not. Children in the weight and no-weight conditions searched at an adjacent door on a mean of 83.83% of trials ( $SD = .13$ ) and 64.59% of trials ( $SD = .14$ ), respectively. These percentages are significantly greater than would be expected by chance (43.75%; see Berthier et al., 2000 for the derivation of this value);  $t(15) = 12.51$ ,  $p < .001$ ,  $d = 6.46$  for the weight and  $t(15) = 6.11$ ,  $p < .001$ ,  $d = 3.16$  for the no-weight conditions. Moreover, children in the weight condition were significantly more likely to search at a door adjacent to the wall than children in the no-weight condition,  $t(30) = 4.11$ ,  $p < .001$ ,  $d = 1.50$ . Searching at a door adjacent to the wall does not always lead to success because for three of the four wall locations, there are two possible doors to open. However, searching at an adjacent door is a good strategy for increasing the likelihood of success on the first search, and if incorrect, the child is set up for searching on their second attempt to the other adjacent door which is correct. Similar to Berthier et al., these analyses demonstrate that the adjacent-door strategy is a better description of the data than the perseverative strategy.

Finally, to further evaluate the effect of the weights on children's performance, latency to initiate the reach and latency to fully open the first door were analyzed (Table 1). Of the 32 children, 27 video records (12 no-weight and 15 weight) were available for coding. There were no significant differences between the no-weight and weight children in the latency to initiate a reach,  $t(27) = .17$ ,  $p = .864$ , or to fully open the first door,  $t(27) = .64$ ,  $p = .526$ . Moreover, latency did not differ for correct and incorrect searches;  $t(24) = .61$ ,  $p = .550$  for latency to initiate forward movement of the arm and  $t(24) = .18$ ,  $p = .857$  for latency to open the door fully.

## DISCUSSION

The motivation for this study came from a fascinating finding from Rivière and Lécuyer (2008) demonstrating that weighted wrist bands improved toddlers' performance in a search task. By testing 24-month-olds, our goals were to determine whether the C-not-B error was evident at an earlier age, to replicate chance performance in the

Door task found by Berthier et al. (2000), and to see if the advantage conferred by weighted wrists would generalize to a different search task. In short, we did not find the C-not-B error in younger children, but we did replicate the findings of Berthier et al. (2000) in the Door task. Also, our results confirm that weighted wrist bands led to an improvement in performance on both the C-not-B and Door tasks. Whereas the facilitation of weighted wrists manifested itself in the same way across both tasks, the effects of the weights differed from that of the 29-month-old children in Rivière and colleague's task.

There are three main findings from our results. First, in the no-weight condition, 24-month-old children were at chance performance in both tasks. Thus, we replicated the previous finding from Berthier et al. (2000), and we provide the first report of chance responding for 24-month-olds on the C-not-B task. In the C-not-B task, 24-month-olds' incorrect searches were evenly distributed across the A and C locations, providing no evidence of the C-not-B error. Rivière and colleagues interpret the C-not-B error as a prepotent motor response built up by drawing attention to location C when the experimenter reveals the empty hand after passing under the cloth at location C before asking the child where the toy is hidden. Our data do not refute this interpretation, but instead provide evidence that it takes time for the prepotent motor response to develop because it was not evident at 24 months.

The second main finding from our results was that the weighted wrist bands led to improved performance on two different search tasks. We found that the facilitation of weights emerged on the second, but not the first search attempt. In the C-not-B task, the experimenter passed her hand under each of the three locations and upon revealing that her hand was empty, the child was encouraged to search. The first search attempt was no different from chance. In the second trial after the experimenter again passed her hand under the three locations and asked the child to search, significantly more children chose the B location. This pattern contrasts with the many examples from Rivière and colleagues where 29-month-olds show correct first searches in the weighted condition and consistent behavior across search attempts. On the first trial in the present study, children in both conditions did not appear to know where the object was hidden and searched randomly under one of the three cloths. When they did not meet with success on the first trial, children in the weight condition were able to guide their searches based on the presence of the bump at B, suggesting a switch from a perception-action response to a representation-guided response. In other words, after failing to find the object on the first trial, children wearing weights were able to inhibit the default response of searching to the last seen location of the hand and respond to the information provided by the bump in the cloth.

The variability in responding by the 24-month-olds suggests that 24 months is a transitional period, one in which children of the same age may be using different strategies. Some may be following the linear trajectory of the hiding phase, a strategy identified by Sophian (1986) with this same age group, and thus searching at the first location (A). Others may have attended to the bump, which led them to the correct location (B). And others still may have attended to the last location of the hand, leading them to search at the third location (C). Whereas we concluded that the weights were more likely to enable the children to attend to the bump, the weights certainly did not help all children. Future studies could clarify the nature of the C-not-B error development using a longitudinal design, like Clearfield et al. (2006), to trace the

emergence of the C-not-B error and analyze the consistency of the A and C location responses.

The improvement due to weighted wrists in the C-not-B task is bolstered by the fact that this second search pattern was also found in the Door task. In the Door task, there were four doors to choose from, and children were given two attempts to search within each trial. Performance on the first search was at chance, but performance in the second search was significantly above chance for children in the weighted condition. Thus, the weights appear to have enabled the children to inhibit an incorrect response and attend more carefully to the information indicating the location of the ball, either the bump under the cloth in the C-not-B task or to the ball-wall relation in the Door task. Future studies should investigate whether this behavior is tied to the weighted wrists in particular. For example, would this improvement in performance occur if we put the weights on the children's feet? In other words, do the weights need to be on the limb that is moving (such as the arm while reaching) or is a general disruption of body kinematics sufficient to improve performance?

The third main finding is that these data support Rivière and colleagues' conclusions that there is competition between the perception-action and the representation-guided responding. In the analyses of the Door task, we were able to distinguish a response guided by perseveration and the more cognitively driven error of searching on either side of the barrier. We found that children in either weight condition do not tend to make a perseveration error by searching where they last found the ball. Instead, children searched on either side of the barrier. Searching near the wall, even if incorrect, might lead to a correct search on the next attempt. Children wearing weights were more likely to take advantage of the wall-door relation on their second search than children not wearing weights. It is this successful correction on the second search that we take as evidence for representation-guided responding. All first searches could have been based on the general proximity of the wall to a door based on perceptual information (recall that the wall was visible above the door panel); however, to be accurate children needed to know which door to open based on the ball's trajectory and knowledge of solidity (i.e., balls cannot pass through walls). We see parallels between the representation-guided responses across both tasks: One response is directed to a door adjacent to the wall (in the Door task) and the other is a response to the cloth with the protrusion (in the C-not-B task).

Two additional findings are worth mentioning. First, weights did not slow the children's action. In both tasks, the time to initiate a reach did not differ between the no-weight and weight conditions. Thus, the weights may be encouraging children to integrate thought and action, but the actual action is not delayed. Rivière and Lécuyer (2008) also found no delay between the weight and no-weight conditions, leading them to conclude that the effect of weights is not due to delay in action. It is still possible that the trajectory of the reach is impacted. For example, more time might be spent in planning than controlling the movement, two processes that occur during an ongoing action (e.g., Glover, 2004). In future work, a finer grained analysis of the reach might illuminate differences between no-weight and weight conditions.

A second and unexpected finding is the lack of evidence of learning across trials. In the Door task, we had the opportunity to look at children's performance within the same trial (Search 1 compared to Search 2, in the event that Search 1 was not correct) and to look at performance across trials (recall children completed up to 12 trials). We were stunned to find that there is no evidence of improvement over consecutive trials,

even though there is significant improvement from the first to second search within a trial. This distinction supports the notion that the advantage conferred by weighted wrist bands is influencing perception–action coupling, and not cognition or learning. Supporting this conclusion is the fact that performance in the Door task by children aged 34 months is correlated with measures of inhibitory control (Jenkins & Berthier, 2014). It may very well be the dissociation between these two processes that may help explain why toddlers perform so poorly on search tasks that rely on knowledge that infants possess when tested in looking paradigms (Keen, 2003).

These findings fit nicely within the existing literature on perseveration errors during search tasks. As reviewed in the introduction, Clearfield et al. (2006) found an absence of the A-not-B error at an earlier age, much like we found an absence of the C-not-B error at 24 months of age, whereas Rivière and Lécuyer (2008) find the C-not-B error at 29 months of age. Both of these examples demonstrate younger children do not make the error that is committed by older counterparts, and the deeper explanation is one of developmental continuity and increasing sophistication. In Clearfield et al. (2006), increasing consistency of a repetitive motor skill, or specifically, more skill with reaching, brought on the error. In Rivière and Lécuyer (2008), it could be social sophistication to comprehend the ostensive cue of searching at the last hiding location prior to seeing the empty hand (e.g., location C). Like the younger children tested by Clearfield et al. (2006) who did not make the A-not-B error, the younger 24-month-olds in our task did not reveal evidence for the C-not-B error. A similar ostensive cue may have happened in the Door task when the experimenter pushed the set of doors toward the children and asked them to open one. This action could draw attention to all four doors and the goal of opening any of them, thus distracting the children from the representation-guided response of opening the specific door where the ball was hidden.

Overall, it appears that toddlers' search failures are due to the competition between perception–action responding and representation-guided responding. In the C-not-B task, the competition that led to failure for 29-month-olds was that the experimenter's hand at location C drew attention away from the protrusion at location B. In the Door task, a number of factors might have drawn children's attention away from the relation between the ball and the wall; however, when children wear weights, they are able to more effectively rely on the wall–ball relation such that they can predict behind which door the ball has stopped against the wall. Future research should explore the competition between perception–action responding and representation-guided responding by testing how the ability generalizes to other tasks, including those that do not involve search.

## CONCLUSION

These findings provide new insights into why infants look so smart and toddlers look so dumb. The results demonstrated that 24-month-old children typically fail to search correctly in the C-not-B and Door tasks. However, when children wore weighted wrist bands, their performance improved significantly. The improvement was not evident on the first search attempt, but emerged on the second search attempt. The results confirm and extend previous work by Rivière and colleagues suggesting that the weighted wrist

bands influence competition between perception–action coupling and representation-guided actions.

## ACKNOWLEDGMENTS

This work was supported by a Social Sciences Division grant from Colby College (USA) awarded to Martha Arterberry and a National Science Foundation grant (No. BCS-1423917) awarded to Susan Hespos. Rachel Herth was supported by the 2015 Summer Research Opportunity Program at Northwestern University. We thank the members of the Infant Cognition Lab at Northwestern for their help with data collection and coding, in particular Andrea Azem, Genesis Garcia, Jooshim Kim, and John Smith. The authors declare no conflict of interest with respect to this work.

## REFERENCES

- Baillargeon, R. (1987). Object permanence in 3.5- and 4.5-month-old infants. *Developmental Psychology*, 23, 655–664.
- Baillargeon, R. (1995). A model of physical reasoning in infancy. In C. Rovee-Collier, & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 9, pp. 305–371). Norwood, NJ: Ablex.
- Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in 5-month-old infants. *Cognition*, 20, 191–208.
- Berthier, N. E., DeBlois, S., Poirier, C. R., Novak, M. A., & Clifton, R. K. (2000). Where's the ball? Two and three-year-olds reason about unseen events. *Developmental Psychology*, 36, 394–401.
- Bornstein, M. H., & Arterberry, M. E. (2010). The development of object categorization in young children: Hierarchical inclusiveness, age, perceptual attribute and group versus individual analyses. *Developmental Psychology*, 46, 350–365.
- Butler, S. C., Berthier, N. E., & Clifton, R. K. (2002). Two-year-olds' search strategies and visual tracking in a hidden displacement task. *Developmental Psychology*, 38, 581–590.
- Clearfield, M. W., Diedrich, F. J., Smith, L. B., & Thelen, E. (2006). Young infants reach correctly in A-not-B tasks: On the development of stability and perseveration. *Infant Behavior and Development*, 29, 435–444.
- Glover, S. G. (2004). Separate visual representations in the planning and control of action. *Behavioral and Brain Sciences*, 27, 3–78.
- Jenkins, I. L., & Berthier, N. E. (2014). Working memory and inhibitory control in visually guided manual search in toddlers. *Developmental Psychobiology*, 56, 1252–1262.
- Keen, R. (2003). Representation of objects and events: Why do infants look so smart and toddlers look so dumb? *Current Directions in Psychological Science*, 12, 79–83.
- Mandler, J. M., Fivush, R., & Reznick, J. S. (1987). The development of contextual categories. *Cognitive Development*, 2, 339–354.
- Mangalindan, D. M. J., & Schmuckler, M. A. (2011). What's in a cue? The role of cue orientation in object displacement tasks. *Infant Behavior and Development*, 34, 407–416.
- Marcovitch, S., & Zelazo, P. D. (1999). The A-Not-B error: Results from a logistic meta-analysis. *Child Development*, 70, 1297–1313.
- Mash, C., Keen, R., & Berthier, N. E. (2003). Visual access and attention in two-year-olds' event reasoning and object search. *Infancy*, 4, 371–388.
- Mash, C., Novak, E., Berthier, N. C., & Keen, R. (2006). What do two-year-olds understand about hidden-object events? *Developmental Psychology*, 42, 263–271.
- Piaget, J. (1954). *The construction of reality in the child*. New York, NY: Basic Books.
- Rivière, J. (2014). Embodiment in children's choice. Linking bodily constraints with decisional dynamics. *Current Directions in Psychological Science*, 23, 408–413.
- Rivière, J., & David, E. (2013). Perceptual-motor constraints on decision making: The case of the manual search behavior for hidden objects in toddlers. *Journal of Experimental Child Psychology*, 115, 42–52.



- Rivière, J., & Falaise, A. (2011). What comes first? How selective attentional processes regular the activation of a motor routine in a manual search task. *Developmental Psychology*, 47, 969–975.
- Rivière, J., & Lécuyer, R. (2003). The C-not-B error: A comparative study. *Cognitive Development*, 18, 285–297.
- Rivière, J., & Lécuyer, R. (2008). Effects of arm weight on C-not-B task performance: Implications for the motor inhibitory deficit account of search failures. *Journal of Experimental Child Psychology*, 100, 1–16.
- Shutts, K., Keen, R., & Spelke, E. S. (2006). Object boundaries influence toddlers' performance in a search task. *Developmental Science*, 9, 97–107.
- Sophian, C. (1986). Developments in infants' search for invisibly displaced objects. *Infant Behavior and Development*, 9, 15–25.
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, 99, 605–632.
- Uzgiris, I., & Hunt, J. M. (1975). *Assessment in infancy: Ordinal scales of psychological development*. Urbana, IL: University Press.
- Wellman, H. M., Cross, D., & Bartsch, K. (1987). A meta-analysis of research on Stage 4 object permanence: The A-not-B error. *Monographs of the Society for Research in Child Development*, 5(3), 1–67.