Developmental Psychology
2014, Vol. 50, No. 7, 1951–1962

0012-1649/14/$12.00 http://dx.doi.org/10.1037/a0036695

Space or Physics? Children Use Physical Reasoning to Solve the Trap Problem From 2.5 Years of Age

Amanda M. Seed
University of St. Andrews

Josep Call
University of St. Andrews and Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

By 3 years of age, children can solve tasks involving physical principles such as locating a ball that rolled down a ramp behind an occluder by the position of a partially visible solid wall (Berthier, DeBlois, Poirer, Novak, & Clifton, 2000; Hood, Carey, & Prasada, 2000). However, the extent to which children use physical information (the properties of the wall) remains unclear because spatial information would suffice (the location of the wall in relation to the ball). We confronted 2- to 6-year-old children with a ball resting on a shelf inside a clear plastic-fronted box. To retrieve the ball, children had to roll it away from a trap or barrier using their fingers. Crucially, a single object acted as a barrier or supporting surface in different conditions, thus requiring a flexible response. Preschoolers solved the task and the critical transfers from 2.5 years of age (Study 1). Interestingly, 2.5-year-olds required to use a tool to displace the ball performed significantly worse than those who could use their fingers (Study 2). In contrast, 2.5- to 4.5-year-olds failed a *covered* trap box that provided only 2-dimensional predictive cues without any visible physical information, and even 6.5-year-olds performed significantly worse on the covered task compared to the uncovered one (Studies 3 and 4). Our results suggest that children from around 2.5 years of age integrate spatial and physical information when solving problems like the trap box task, rather than simply exploit spatial relationships between features.

Keywords: physical cognition, core knowledge, tool use, causality, symbols

Objects in the environment have physical properties that constrain the ways in which they can interact with one another (e.g., solid objects cannot pass through one another). These properties, such as solidity, continuity, weight, and support, can be directly sensed, but the principles themselves can also be represented at a deeper level of abstraction. “Abstract” representations are not equivalent or reducible to concrete, analogue sensory input; instead the information has undergone further processing in which meaning is extracted, leading to conceptual, as opposed to perceptual, knowledge (Carey, 1985; Karmiloff-Smith, 1992; Mandler, 2004).

The last 3 decades of research have shown that young infants possess considerable knowledge about object interactions (for reviews, see Baillargeon, 1999, 2004, 2008; Spelke, Breinlinger, Macomber, & Jacobson, 1992). However, the evidence is mixed with regard to whether that knowledge, uncovered with perception-based measures, is also recruited for solving manual search tasks during early childhood. Whereas some studies do show a good concordance between looking time and manual search (e.g., Hespos & Baillargeon, 2006, 2008), other studies show a stark dissociation between the two (Berthier, DeBlois, Poirer, Novak, & Clifton, 2000; Hood, Carey, & Prasada, 2000). For instance, in the shelf task, 2.5-year-olds, but not 2-year-olds, were able to locate an object that was dropped onto a shelf behind an occluder by opening a door in the occluder at the level of the shelf, instead of one below it (Hood et al., 2000; see Figure 1A). Similarly, 3-year-olds, but not 2- or 2.5-year-olds, could locate a ball that was rolled down a ramp behind an occluder by using the location of a wall that stopped the ball’s motion, which was visible protruding above the occluder (Berthier et al., 2000; see Figure 1B). The failure of 2- to 2.5-year-olds to solve these tasks is surprising given that infants possess expectations about object solidity and arrested motion (Keen, 2003; Spelke et al., 1992). One possible reason for this discrepancy is that coordinating perception, attention, knowledge, prediction, and action within the context of a multifaceted task may overload immature executive resources and may explain the failure of the younger children on these tasks (Keen, 2003; Keen et al., 2008; Keen & Shutts, 2007). Another possibility is that the looking-time research does not reveal physical concepts in very young infants, but rather expectations based on familiarity, an issue that has been

This article was published Online First April 28, 2014.

Amanda M. Seed, Centre for Social Learning and Cognitive Evolution and Scottish Primate Research Group, School of Psychology and Neuroscience, University of St. Andrews; Josep Call, Centre for Social Learning and Cognitive Evolution and Scottish Primate Research Group, School of Psychology and Neuroscience, University of St. Andrews, and Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.

We thank the Max Planck Society for funding. Amanda M. Seed was funded by a fellowship from the Royal Commission for the Exhibition of 1851. We would like to thank Jana Jurkat for her help and guidance, Sarah Girlich for her assistance with testing, and Henriette Beranek for assistance with testing and for performing the reliability coding.

Correspondence concerning this article should be addressed to Amanda M. Seed, School of Psychology and Neuroscience, University of St. Andrews, St. Mary’s Quad, St. Andrews, KY16 9JP, Scotland. E-mail: ams18@st-andrews.ac.uk

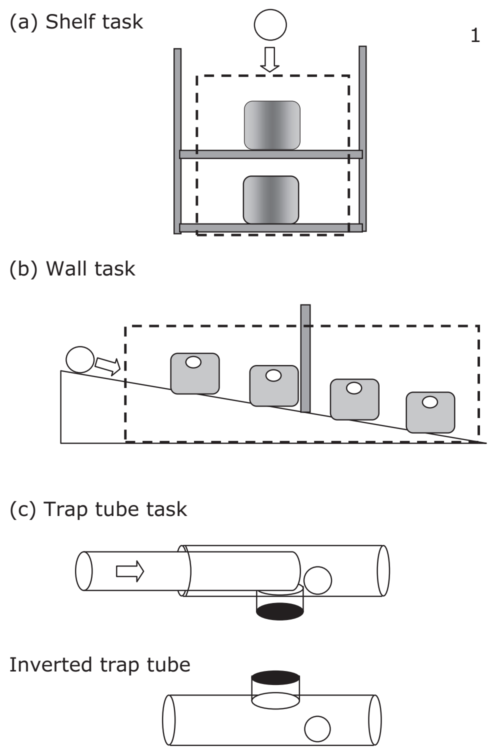

Figure 1. Tasks used in previous studies of physical reasoning. The locations of opaque occluders are indicated with dotted lines but shown as transparent for clarity.

the topic of fierce and repeated debate (see Baillargeon, 1999, 2004, and commentaries).

A central issue that needs to be considered is the cognitive abilities underpinning children's responses. Are they using physical principles such as solidity and continuity to infer the ball's final location? Or could their success be based on spatial-relational information learned during familiarization or generalized from past experience, such as searching for the ball in the location closest to its point of disappearance, without the physical properties of the objects involved (solidity, support) playing a role in their decision-making process? This is a crucial distinction, because to understand when the ability to reason about objects and their physical properties develops, experiments need to show not only that subjects can predict or produce object interactions, but also that their computations are based on representations of physical properties (e.g., a solid object cannot move through a barrier).

To better understand the nature of 3-year-olds' computations, Keen et al. (2008) switched the orientation of the ramp in the wall task midway through trials such that the trajectory of the ball was changed. The 3-year-olds successfully switched from opening the door to the left of the wall to opening the one to the right, and were therefore shown not to be using an inflexible associative rule of the form "open the door to the left of the wall." Nevertheless, the children's reasoning might still have been grounded in space rather than physics, provided that they integrated information about the trajectory of the ball into their decisions. In fact, neither the shelf task nor the wall task allow reasoning based on physical laws to be completely disambiguated from reasoning based on spatial-relational features, because the spatial and physical features of the object (its location in relation to the ball and the role played by its solidity) are always confounded in the events used.

Another limitation of the wall task is that it only involves one type of physical event. An adult-like theory of solidity can be used flexibly, for example, to predict that a solid structure can stop a moving object regardless of its trajectory, and that an object will fall if not supported by a solid surface. If the 3-year-olds are indeed using physical rather than spatial reasoning to solve the wall task, does this mean that they have a fully fledged theory of solidity? Two findings from infant research suggest that using the principle of support to predict the fate of moving objects may be developmentally dissociated from the principles of solidity and continuity. First, Spelke et al. (1992) found that 4-month-old babies did not look longer when a partially occluded vertical trajectory was revealed to end "in midair" than when it ended on a solid shelf. Nor did they look longer at a display in which an object's horizontal trajectory was revealed to have ended at a point at which it must have passed over a gap in a continuous surface than one in which the object must have fallen through the gap. In these experiments, babies seemingly did not expect moving objects to fall through a gap, despite the fact that the same study did find longer looking in 2-month-olds when objects appeared to have violated the principle of solidity (passed through a solid barrier) both when their occluded trajectories were vertical (falling, as in the shelf task) and when they were horizontal (traveling along a continuous surface, as in the wall task). Not until between the ages of 4.5 and 6 months do children reveal expectations about events involving progressively more complex support relations in both looking and action measures; for example, from 4.5 months they look longer if an object is pushed from a supporting platform but remains in midair (Hespos & Baillargeon, 2008; Needham & Baillargeon, 1993).

One task that has been used to investigate physical reasoning in nonhuman animals and children is the trap tube task (Visalberghi & Limongelli, 1994), which requires subjects to use a tool to push a reward out of a horizontal tube with a trap along its length through which the reward would drop if pushed over it (see Figure 1C). Horner and Whiten (2007) tested children on this task and found that 3- and 4-year-old children failed to avoid the trap above chance levels within the 10 trials given. Furthermore, demonstrations of successful or unsuccessful trials did not improve their performance. Want and Harris (2001) also found that 2- and 3-year-old children could not solve the task without demonstration. However, 3-year-olds, but not 2-year-olds, performed above chance if they were first shown a demonstration in which the demonstrator inserted the stick from the wrong side, and then corrected themselves to move the reward away from the trap. This was not the case if they were simply shown a correct demonstration. Finally, E. Limongelli (personal communication, 1994) tested 2- to 3-year-old children over more trials (up to 40). None of the 2.5-year-old children and only about half of the 3.5-year-olds consistently solved the task. Here then we have a further inconsistency: Children from the age of 2.5 years solve manual search tasks involving the principles of solidity and continuity of surface in the shelf and wall paradigms described above and yet fail the trap tube task. This could suggest that success on the search tasks does not in fact reflect robust knowledge of these physical principles and their causal role in object interactions: This develops later, at the age when they can pass the trap tube task too (Want & Harris, 2001).

The trap tube task possesses two key features that make it an ideal paradigm to further investigate physical reasoning in children. First, since the relation between their elements (i.e., traps, barriers, rewards) can be construed along both spatial and physical dimensions, it is possible to disentangle spatial and physical reasoning by using a series of task configurations. Faced with different configurations where the same solid object can play the role of a supporting surface or a barrier requires subjects to use information about physical properties flexibly. More specifically, the trap tube task involves three physical principles: that solid objects do not pass through one another, that they travel along continuous paths, and that they fall when unsupported. To date, trap tasks that investigate the use of the latter principle have found negative results at the age of 3–4 years (in the absence of demonstration), but by reducing task demands, we may find earlier success with this paradigm, and provide evidence for flexible physical reasoning.

Second, the tool use component potentially increases the task's complexity in particularly interesting ways. Although several authors have invoked limitations in physical or causal reasoning abilities to explain children's age-related changes in performance (e.g., Horner & Whiten, 2007; Want & Harris, 2001), the high cognitive demands associated with tool use may have significantly contributed to the observed outcomes. Tool use has been suggested to increase the load on the executive brain, as three items need to be related: the tool, the object, and the obstacle (Fragaszy & Cummins-Sebree, 2005). In fact, when the tool use component in a trap tube task was removed, chimpanzees' performance improved substantially. This is important because it highlights that even the simplest form of tool use (i.e., move the reward to the left or to the right) may take executive resources away from the task of reasoning about which way to move the reward (Byrne, 2004). Seed, Call, Emery, and Clayton (2009) adapted the trap tube task to reduce peripheral task demands by removing the tool component and found that all of the eight chimpanzees successfully avoided the trap. Moreover, in a follow-up task, chimpanzees that used their hands to move the reward directly were found to perform significantly better than a group tested on the same task with a tool.

The primary goal of this study was to investigate (using the trap box) the development of physical reasoning in children, paying special attention to the contribution of physical and spatial reasoning. Our secondary goal was to investigate the effect that using tools had on performance. Since the kind and range of movement that could be executed with the tool once it was inserted into the trap box were quite restricted (i.e., insert from the left vs. right side), it was not a foregone conclusion that performance would deteriorate, unless tool use per se, despite its apparent simplicity, drained some of the cognitive resources needed for physical reasoning. Finally, successfully avoiding the trap in the first phase of the experiment would show a competence in young children to predict the outcome of object interactions before performing an action to make an object move. This would meaningfully extend the findings from the previous literature on physical reasoning, which has been based largely on manual search, after an event has taken place. Because of the well-documented inhibitory control demands associated with manual search, using a different kind of action might provide important information for triangulating on what develops between the ages of 2 and 3 as children become more competent at solving a variety of physical problems.

Study 1

The trap box task is a box with a transparent front, with a reward (a sticker inside a plastic ball) sitting in the middle of a shelf inside it (for a labeled diagram, see Figure 2). Subjects can move the reward along the length of the shelf with their fingers, through a gap cut into the transparent front of the box. There are exits on each side of the box aligned with the shelf, but the shelf falls short of these exits on either side, so that before each exit there is a vertical channel where rewards can fall down. There is another pair of exits in the bottom of the box at the end of these channels. Before each trial the box is configured by the experimenter using removable pieces (see Figure 2A). One piece (the shelf piece) extends the length of the shelf in one direction so that rewards can be recovered from the exit on that side. Another piece (the trap piece) forms an obstacle midway down one of the vertical channels so that balls that fall down it become trapped inside the box.

In the initial phase of the experiment, the experimenter configures the box so that one channel is obstructed by the trap piece (see Figure 2C, Tasks A and B). For half of the participants, the shelf piece is added on the other side of the box, so that the reward can be pushed along the shelf to the side exit (A). For the other half of participants the other channel is unobstructed, so that the reward can be pushed off the shelf, down the channel, and out of the bottom exit on successful trials (B). During testing, the position of the trap (to the left or the right of the shelf) varies at random, with the constraint that it does not appear on the same side more than twice in a row. If participants solved the first task (e.g., A), they were then confronted with the other one (B). To solve both A and B configurations, children must move the reward away from the trap piece. Both can therefore be solved using the location of the trap piece without encoding its functional significance. The rationale for having the two tasks is that they give children experience with different positive features: the uninterrupted channel and the shelf piece. To disentangle the spatial and physical strategies, these two positive features are pitted against one another in the transfer phase. In this phase, children who learned to solve Tasks A and B receive two transfer tests (see Figure 2C, Tasks C and D). These both feature the two previously rewarded cues (the shelf piece and the unobstructed channel), but require opposite responses to them. In Task C, the side exits are blocked so that subjects need to push the reward away from the shelf piece down the unobstructed channel. In Task D, the bottom exits are blocked so that they need to push the reward toward the shelf piece to the side exit instead. We reasoned that a solution based solely on the perceptual/spatial features of the objects would not allow subjects to solve both of these configurations. Simply using a rule based on the spatial location of the shelf piece in relation to the reward without integrating information about its role as a solid supporting surface and the locations of the openings would lead to failure on at least one configuration in the transfer phase.

Method

Participants. Eight 2-, sixteen 2.5-, and eight 3.5-year-olds were tested from kindergartens across Leipzig. In each age group

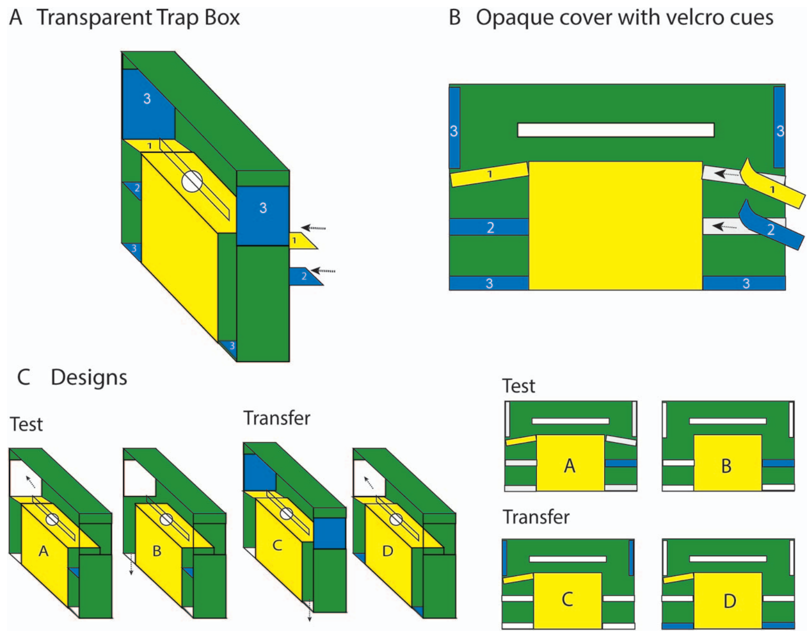

Figure 2. The trap box task. (A) The transparent version of the task, with all removable pieces inserted in their respective locations. 1 = shelf piece; 2 = trap piece; 3 = blocker pieces. (B) The cover that is placed over the Perspex front in the covered version (Studies 2 and 3), showing all of the two-dimensional cues in their respective locations. (C) The configuration of the four designs (A–D) in the transparent and covered conditions. The color version of this figure appears in the online article only.

we tested equal numbers of boys and girls. We used a recruitment window of 2 months above and below the target age. See Table 1 for the mean and range of the age in months of each group. A further eight children were dropped from the study due to experimental error (two) or because they did not complete all phases of testing, either because of fussiness (four) or illness (two).

Materials. The trap box was made from wood, painted green, and was 42 cm wide, 30 cm high, and 8.5 cm deep (see Figure 2).

Table 1
Performance of Participants on Each Task Administered

	Age				Task 1 (A or B)				Task 2 (B or A)				Task 3 (C or D)				Task 4 (D or C)			
		Months				Fail				Fail				Fail				Fail		
Condition	Years	*M*	Range	*n*	Pass	SB	BC	No	Pass	SB	BC	No	Pass	SB	BC	No	Pass	SB	BC	No
Study 1																				
Transparent	2	24.1	23–26	8	1	6	0	1	1	0	0	0	0	0	0	1	0	0	0	1
Transparent	2.5	29.5	29–31	16	8	8	0	0	8	0	0	0	8	0	0	0	7	0	1	0
Transparent	3.5	41.8	41–43	8	8	0	0	0	8	0	0	0	8	0	0	0	8	0	0	0
Study 2																				
Transparent–tool	2.5	30	28–32	16	3	11	1	1												
Study 3																				
Covered	2.5	29.8	29–31	8	0	8	0	0												
Covered	3.5	40.4	40–42	8	0	8	0	0												
Covered	4.5	54.3	53–55	8	0	7	1	0												
Covered	5.5	66.3	65–67	8	3	4	1	0	3	0	0	0	3	0	0	0	2	0	0	1
Covered	6.5	77.8	76–79	8	3	4	1	0	2	1	0	0	2	0	0	0	2	0	0	0
Study 4																				
Transparent	6.5	77.4	76–79	8	8	0	0	0	8	0	0	0	8	0	0	0	8	0	0	0
Covered	6.5	77.4	76–78	8	5	0	1	2	5	0	0	0	5	0	0	0	5	0	0	0

Note. "Pass" shows the number of children who reached criterion. Those who failed, where applicable, are shown as the numbers that employed the following strategies: SB = side bias; BC = below chance; No = no discernible strategy. Only children who passed the first box (A or B, counterbalanced) proceeded to the second. Children who passed both A and B received both transfer tasks (the third and fourth tasks administered). The order of administrating A and B in Phase 1 and C and D in Phase 2 was counterbalanced. Note that children tested with a tool only received the first task to solve.

The shelf inside the box was 25 cm wide and 20 cm high, painted yellow. The reward was a sticker inside a transparent ball (4-cm radius), which was positioned in the middle of the shelf from the back of the box by the second experimenter. The transparent front of the box was a clear plastic sheet, with a gap cut into it, wide enough to allow children to insert their fingers (1.5 cm) and move the reward along the 23-cm length of the shelf, but not wide enough for the ball to pass through. The channels on either side of the shelf were 7.5 cm wide. The box had four exits cut into it that the reward could pass through: two side exits just below the shelf and two bottom exits in line with the channels. The configuration of the box was adjusted from the back by the second experimenter by adding and removing wooden pieces (10 cm $\times$ 7.3 cm $\times$ 0.8 cm).

The blue trap piece was inserted halfway down either of the channels to obstruct the passage of the reward. The yellow shelf piece was inserted in line with the shelf at the top of either of the vertical channels to extend its length to the side exits (the shelf sloped downward so that the ball rolled out of the side exit once pushed onto the shelf; this was important in Task C, when pushing the ball over the shelf was incorrect because the side exits were blocked—the ball would roll beyond the reach of the children's fingers, and so a mistake could not be corrected). Two blue blocker pieces were used to block either the side or the bottom exits. The experimenter added the reward in the middle of the shelf, and removed trapped rewards, through doors in the back of the box. An orange screen was placed around the box to prevent children from trying to look behind it. Figure 2 shows the configuration of the four designs.

Procedure. During a testing session, children were first encouraged by Experimenter 1 to take the ball from all four exits in the box (Experimenter 2 held the ball in place at each exit). Then trials began. A sticker was shown to the child by Experimenter 1, placed in the ball, and given to Experimenter 2. Experimenter 2 configured the box with the removable pieces, adding pieces from left to right in cases where more than one piece needed to be added, and checked that the child had looked at the configuration. If not, Experimenter 1 and/or Experimenter 2 encouraged them to do so by saying "Look." Experimenter 2 put the ball in the middle of the shelf. Experimenter 2 then looked down to avoid cuing the child. In the first trial, and afterward if reticent, children were encouraged by Experimenter 1 to get the ball out of the box and told that they could use their fingers to move the ball from side to side. Experimenter 1 sat out of the child's eyeline when they were working on the box to avoid giving unintentional cues to the correct answer. If the child looked around for help, Experimenter 1 looked directly at the child, shrugged and responded with "Just try and get the ball out of the box." If children moved the ball in the correct direction, they put the sticker in the sticker book that they were given at the beginning with help from Experimenter 1 if needed. If they moved it toward the trap, Experimenter 2 took the ball out of the box from the back, took the sticker out, and put it away in another box. Then Experimenter 2 removed the pieces, and the next trial began. Trials continued until criterion was reached (16 or more trials out of 20 correct), until the child declined to participate, or until the maximum of 20 trials for the day was reached. The order of trials (left or right side correct) was randomized, with the constraint that there were no more than two in a row on the same side. Children were tested on consecutive days (or as near as possible, with a maximum of 7 days between sessions).

Phase 1: Testing (Boxes A and B). Half of each age group (equal numbers of boys and girls) received Configuration A first, and half Configuration B. Testing continued until criterion or a maximum of 40 trials (across two sessions over 2 days) was reached. If unsuccessful, testing ended. If successful, children were given the other design the next day (B if they had received A, and vice versa) with the same criteria as in the first box.

Phase 2: Transfer (Boxes C and D). Half of the successful children from Phase 1 (balanced numbers of boys and girls) received Configuration C first, and half Configuration D. They received a maximum of 20 trials on each box. If they scored 8 or more out of 10 on their first 10 trials, they proceeded directly to the other task on the same day.

Scoring and analysis. All testing was recorded on MiniDV tape. Trials were scored live by Experimenter 2 as correct or incorrect. The data satisfied the conditions of normality, and so parametric statistical tests were conducted with SPSS (Version 19). Significance was set at $p < .05$. Children received a maximum of 20 trials per day; however, performance was analyzed in terms of proportion correct over blocks of 10 trials, as this was the minimum number a child could receive per condition.

Error analyses for children who failed to reach criterion was conducted. A side-bias strategy was reported if children responded more to one side than the other over the 40 trials received according to a two-tailed binomial test. A win–stay, lose–shift strategy was reported if children stuck with the same direction after correct trials and shifted to the other side after incorrect trials for more trials than expected by chance according to a binomial test. However, such a strategy was never observed, and is not referred to again. "Below chance" was reported if children were significantly below chance according to a binomial test across two-successive sessions (the inverse of the criteria required to pass). If none of the above applied, "no strategy" was recorded.

Twenty percent of all tapes (Studies 1–4) were selected at random and coded for reliability. Across the resulting 41 sessions coded, interobserver reliability was 99.9%, rising to 100% after the tape was reviewed.

Results

Phase 1: Testing (Boxes A and B). One of the 2-year-olds, half of the 2.5-year-olds, and all 3.5-year-olds reached criterion on both boxes (see Table 1). Two of the 2.5-year-olds and three of the 3.5-year-olds were significantly above chance in their first 10 trials (scoring 8 or more correct out of 10). In contrast, the 2-year-old girl only reached criterion after 30 trials on Task A and 40 trials on Task B. One-sample t tests showed that the 2.5-year-olds and the 3.5-year-olds were above chance but the 2-year-olds were not: 2-year-olds ($M = 4.97$, $SD = 1.02$), $t(7) = -0.09$, $p > .05$; 2.5-year-olds ($M = 6.52$, $SD = 1.73$), $t(15) = 3.52$, $p < .01$; 3.5-year-olds ($M = 8.19$, $SD = 0.84$), $t(7) = 10.7$, $p < .01$. There was no significant difference between performance on Box A and Box B (all comparisons, $p > .05$). Children who failed the task predominantly responded with a side-biased strategy; this was the case for six of the seven 2-year-olds (the other child had no discernible strategy) and all eight 2.5-year-olds (Table 1). Children very rarely corrected an error; once they started moving the ball in one direction, they almost always continued to do so. All of the successful children solved the second box they received (Table 1).

Phase 2: Transfer (Boxes C and D). The successful 2-year-old was not above chance on either of transfer tasks C and D. All but one of the successful 2.5-year-olds tested passed both transfer tasks; one boy failed Task C, scoring 0/10 in his first 10 trials, but passed Task D. All of the 3.5-year-olds passed both transfers. Table 1 provides a summary of the individual data.

Discussion

Half of the 2.5-year-olds and all of the 3.5-year-olds successfully anticipated the effect of a physical obstacle and moved the ball along an unobstructed path, but only one 2-year-old was able to do so. This developmental pattern is similar to that found from manual search paradigms, suggesting that there may be developmental changes during the second year of life in cognitive processes that allow for successful prediction of object interactions both when searching after an event has taken place and when producing object movement in a problem-solving task. The poor performance at 2 years of age might reflect conceptual or executive functioning limitations. All but one of the 2- and 2.5-year-olds who failed the task had a significant side bias, the sort of strategy that could point toward perseverative responding due to inhibitory control limitations, which has been observed in the children of this age who fail the wall task (Berthier et al., 2000). However, it should be noted that such a strategy is also a good way to get half of the rewards if you do not know how to solve the task, and so the pattern of errors alone cannot give a decisive answer to this question.

A looking-time version of the task could help to tease these alternatives apart. Mash, Novak, Berthier, and Keen (2006) found that 2-year-olds looked longer when the outcome of a puppet's search in the wall task was inconsistent with reality (i.e., when it opened the right door and found no ball), compared to consistent with reality (i.e., when it opened the wrong door and found no ball). A similar approach could be used in this task to investigate the expectations of 2-year-olds. Success on manual search tasks positively correlates with performance on inhibitory control tasks (Baker, Gjersoe, Sibielska-Woch, Leslie, & Hood, 2011), and it seems likely that inhibitory control also plays a role here. A correlationary or even a training study could be a useful direction for future research.

The lack of a tool-using component may explain the much improved performance of preschoolers on the trap box as compared to the trap tube task. Three 2.5-year-olds and two 3.5-year-olds were significantly above chance within the first 10 trials, in comparison to the 3- and 4-year-olds tested on the traditional paradigm by Horner and Whiten (2007) and the 2- and 3-year-olds tested by Want and Harris (2001), all of whom failed to reach significance within the 10 trials given if they had not seen a demonstration. Similarly, children in the current study outperformed those tested by E. Limongelli (personal communication, 1994). To further investigate whether the difference between the tasks is ascribable to the tool-using component, in Study 2 we tested 2.5-year-olds on a tool-using version of the trap box problem to provide a direct comparison.

Almost all of the successful children solved both transfer tasks (although interestingly not the 2-year-old), suggesting that by 2.5 years, children used the physical properties of objects, and not just their spatial locations in relation to the reward, to solve the task. They successfully moved the ball away from the shelf piece when the side exits were blocked and the bottom exits were open, and yet toward that same feature when the opposite was true. This reinforces the case for physical reasoning by 2- to 3-year-old children, as it has been proposed for successful performance in the wall and shelf tasks (Keen et al. 2008). However, the exact nature of children's representations is still difficult to ascertain. Whether children "perceive" or "reason" about the continuity of surfaces, or the inability of objects to pass through one another, is, we argue, still an open question. Can object properties such as solidity and continuity be perceived as "affordances" without children having abstract knowledge that they could use to make "offline" predictions in the absence of visual feedback at this age? We return to this issue in the General Discussion.

Bluff, Weir, Rutz, Wimpenny, and Kacelnik (2007), commenting on the use of this task with rooks and chimpanzees (Seed et al., 2009; Seed, Tebbich, Emery, & Clayton, 2006), argued that the whole set of tasks could be solved by learning to respond to any asymmetry in surface-level visual features without any recourse to physical cognition, be it perceptually or conceptually based. Several of the successful children made a number of errors before finding the solution. This raises the possibility that rapid task-specific learning based on arbitrary features could account for their performance, as opposed to knowledge of the physical principles involved. The transfer tasks C and D were designed to rule out heuristics based on arbitrary cues, whether by avoiding the trap or preferring the uninterrupted channel or the shelf piece. However, given that some visible discriminative cue must almost always be available for the solution of a task, it is difficult to falsify this kind of argument through transfers with different perceptual characteristics based on the same physical principles. To counter this suggestion, we decided to look at performance on a task that only provides asymmetric featural cues, without a functional context. This follows in Studies 3 and 4.

Study 2

In Study 1 we found that children began to pass the trap box task at a younger age than previously found for trap tasks (2.5 years compared to 3.5 years or older). In this study we wanted to compare performance with and without a tool on the same apparatus. With some simple modifications, we constructed a tool-using version of the trap box. We tested sixteen 2.5-year-old children, since the ability to solve the task without a tool seems to be emerging at this age. We predicted that performance would be poorer in this group compared to that of the 16 children tested without a tool in Study 1. We only tested children on one configuration of the box (half received A, and half B), as in this study we were solely interested in children's ability to solve the trap box task when they had to use a tool as compared to when they could use their hands. Although it might have been interesting to investigate how they performed with a tool on the transfer tasks, the combination of the critical transfer tasks C and D was not possible because we could not block the side exits in the tool-using version of the task.

Method

Participants. Sixteen 2.5-year-old children from kindergartens across Leipzig were included in the study. We tested equal numbers of boys and girls, and used a window of 2 months above and below the target age. See Table 1 for the mean and range of the age in months of each group. A further two children were dropped from the study because of fussiness.

Materials. The trap box from Study 1 was used, but the transparent Perspex front with a gap in it was switched for one that was intact, so that the reward could only be moved by inserting a tool from the side exits. The tool was 50 cm long and 2 cm in diameter, made of plastic, and red in color.

Procedure. The procedure was the same as in Study 1, except that in the first trial, and afterward if reticent, children were told by Experimenter 1 that they could get the ball out of the box using the tool, saying "This is the stick. You can use it this way [demonstrates], and this way [demonstrates]. Try to get the ball out of the box, using the stick." Experimenter 1 demonstrated that the tool could be inserted from either exit by pushing the tool inside the box until it was adjacent to, but not in contact with, the ball in the center of the shelf. Half of the children (equal numbers of boys and girls) received Configuration A, and half Configuration B. Testing continued until criterion was reached as in Study 1.

Scoring and analysis. Scoring and analysis were the same as in Study 1.

Results

Three of the 16 children tested solved the task (two received Box A, and one Box B). A one-sample *t* test showed that as a group, in contrast to the group tested without a tool, mean performance out of 10 ($M = 5.34$, $SD = 1.51$) was not significantly different from chance, $t(15) = 0.91$, $p > .05$. An independent-samples *t* test revealed that performance of the group tested on the non-tool-using version in Study 1 ($M = 6.52$, $SD = 1.73$) was significantly better than the group tested with a tool in this study, $t(15) = 2.05$, $p < .05$. Children who failed the task largely responded with a significant side bias (11 children); one child was significantly below chance, and another had no discernible strategy. Children never corrected an error midway through responding; once they had inserted the tool from one side, they proceeded to push the ball in that direction until it was either trapped or out of the box.

Discussion

The trap box task was more difficult for 2.5-year-olds when they were required to use a tool as compared to when they could use their hands, as revealed by the significant difference in their average performance. This supports our suggestion that the challenges associated with tool use may have been a significant factor in the chance performance of younger children on tool-using versions of the trap problem: 2- and 3-year-olds in Want and Harris (2001); and even 3- and 4-year-olds in Horner and Whiten (2007). We suggest that from the age of 2.5, children have the representational resources needed to solve the trap problem, in contrast to the conclusions of these authors that causal or physical knowledge may be the limiting factor.

Although the performance of the group tested with a tool did not differ from chance, three of the 16 children were able to solve the task, even using a tool. One important factor contributing to this difference may have been the amount of information provided about how to use the tool (although not about how to solve the task). As in the previous studies, children were shown that the tool was relevant for solving the task, and the action for moving the tool was demonstrated. However, we showed the children that the tool needed to be inserted into the box, while in both previous studies the tool was moved over the top of the apparatus in the no-demonstration conditions (in which, as in our study, no information about the consequences of moving the reward in either direction was given). Some of the children copied the demonstration and waved the tool over the top of the apparatus when it was their turn, which may have posed an additional hurdle to finding the solution.

Studies 1 and 2 have demonstrated that children from the age of 2.5 can solve the trap task, suggesting that in contrast to conclusions from previous research, the conceptual resources needed to anticipate the effect of a trap or barrier are beginning to emerge from this age, in line with research from other paradigms. (e.g., Berthier et al., 2000; Hood et al., 2000; Keen et al., 2008). We have argued that the transfer tests from Study 1 reveal that successful children's representations integrate information about the spatial location of a feature and the physical role it plays in the task (as a supporting surface, or barrier). However, a low-level explanation based on rapid learning based solely on surface-level perceptual features remains a possible alternative explanation. In any set of transfer tasks, both visual cues and physical information are available, and so to rule out this explanation we needed to take a different approach. In Study 3 we challenged children with a version of the trap box task that provided only arbitrary visual cues without any physical information visible.

Study 3

We aimed to reinforce the case for the use of physical information by 2.5- and 3.5-year-olds by comparing the performance of subjects in Study 1 with a group tested on a covered version of the task with perceptually similar cues stuck to its surface (see Figure 2B). In the covered condition, the front of the trap box was permanently occluded by a green cardboard screen, so that the mechanics of the task were invisible at all times. The discriminatory cues were plastic strips, the same size and color as the trap, shelf, and blocker pieces of the real puzzle box, attached to the cover with Velcro in the same location as their real counterparts behind the cover (Figure 2B). These cues were 100% predictive of the solution, although they provided no information about their functional relevance (as traps or barriers). If performance is similar on the two tasks, a task-specific learning account based on arbitrary perceptual features remains a feasible explanation for the performance of the 2.5- and 3.5-year-olds on the transparent task.

Method

Participants. We tested eight 2.5-, 3.5-, 4.5-, 5.5-, and 6.5-year-olds on the covered task. There were equal numbers of boys and girls in each group. We used a recruitment window of

2 months above and below the target age. See Table 1 for the mean and range of the age in months of each group. An additional 10 children were excluded from the study, due to experimental error (four) or not completing all phases of testing due to illness (three), family vacation (one), or unwillingness to continue (two).

Materials. The trap box from Study 1 was used, but the front face was covered with cardboard. This was colored green with a yellow rectangle in the center in the same position as the shelf in the transparent task. A slit was cut in the cardboard in line with the gap in the Perspex front of the box, to allow children to see and move the reward. Importantly, it was impossible for children to see the real shelf pieces or traps at any time during the experiment. White Velcro was stuck to the screen in the locations corresponding to the gaps in the box where the shelf, trap, and blocker pieces can be inserted. The corresponding Velcro was stuck to pieces of plastic that were the same color, width, and height as the removable pieces. To solve the task, children had to learn to use the pattern of cues in order to move the ball in the correct direction. The cues always predicted the right answer in the same way the real trap and shelf pieces did in Experiment 1 (e.g., in Task A, move the ball toward the yellow line in line with the ball and away from the blue line halfway down the side of the box).

Procedure. The procedure was the same as in Experiment 1, except that after configuring the box (the location of the removable pieces was obscured from the child by the cover), Experimenter 2 then attached the appropriate plastic cues to the box before checking that the child had looked at the configuration, and encouraging the child to do so if he or she had not. Scoring and analysis was identical to Study 1.

Results

Phase 1: Testing (Boxes A and B). None of the 2.5- or 3.5- or 4.5-year-old children tested solved the first box they were given within 40 trials. Three of the 5.5-year-olds and three of the 6.5-year-olds did pass the covered task. However, none of the groups performed significantly above chance level according to one-sample t tests: 2.5-year-olds ($M = 5.03$, $SD = 0.53$), $t(7) = 0.17$, $p > .05$; 3.5-year-olds ($M = 5.03$, $SD = 0.816$), $t(7) = 0.55$, $p > .05$; 4.5-year-olds ($M = 4.59$, $SD = 0.97$), $t(7) = -1.18$, $p > .05$; 5.5-year-olds ($M = 5.31$, $SD = 2.84$), $t(7) = 0.312$, $p > .05$; 6.5-year-olds ($M = 5.28$, $SD = 2.71$), $t(7) = 0.29$, $p > .05$. A univariate analysis of variance revealed a significant effect of condition between the performance of the 2.5- and 3.5-year-old children and those tested in Experiment 1, $F(1, 28) = 39.5$, $p < .01$; no significant effect of age, $F(1, 28) = 2.512$, $p > .05$; and no significant interaction, $F(1, 28) = 2.512$, $p > .05$.

Most of the children who failed the task had a significant side bias, although one 4.5-year-old, one 5.5-year-old, and one 6.5-year-old responded significantly below chance. These children attended to thediscriminative cue but used it consistently incorrectly—something that we never saw in first box performance on the transparent task in the younger children (Study 1). Performance is summarized in Table 1.

Phase 2: Transfer (Boxes C and D). All of the successful children from Phase 1 except one 5.5-year-old passed both transfer tasks C and D (see Table 1). The use of one set of arbitrary discriminative cues was therefore generalized to new cues, despite the fact that a familiar cue (the yellow shelf cue) had to be used in opposite ways in Tasks C and D.

Discussion

Strikingly, 2.5- and 3.5-year-old children completely failed to solve the covered version of the task, although children at the same age solved the transparent task in Experiment 1. This difference in performance implies that learning rules of action based on arbitrary perceptual cues over the course of the experiment, in the way that one might learn to walk on a green light and stop at a red one, cannot explain the performance of children at this age on the transparent version, as such a strategy would result in success on the covered task too. Children failed to use the spatial location of features in relation to a reward to make a decision about which way to move a ball when these features had no physical relevance to the outcome, and when they had no visual information about the path the ball could take. Combined with the fact that in Study 1, children at this age were able to use the same cue flexibly when the mere location of a given feature was not sufficient to solve the task (to solve transfer tasks C and D), we think there is a strong case for the use of physical information to solve tasks involving the principles of solidity and support by 2.5 years of age. This rules out Bluff et al.'s (2007) explanation based on stimulus response learning in the absence of any object-related cognition. However, this task alone does not speak to the nature of the physical representations employed in the transparent task.

The covered task removes all spatial and physical information relevant to solving the trap task—not only the location and physical properties of the obstacles, but also the possible paths the ball could take (and does take over successive trials). Previous research on the wall task has shown that manipulating the degree of occlusion plays a critical role in determining the ability of young children to anticipate the effect of a physical obstacle. When provided with visual access to all elements of the wall task (ramp, ball, wall), 2- and 2.5-year-olds could place a doll correctly on the ramp where they would "catch" the ball (i.e., immediately beside the wall toward the direction in which the ball would be rolled; Kloos & Keen, 2005). When an opaque screen with doors was introduced, covering the ramp and the lower part of the wall, the same children were at chance. Removing all visual support of the elements involved appeared to prevent the children from reasoning about the ball's future path. It would be interesting to examine the role of visual support to problem solving by occluding parts of the trap box task.

Despite its lack of functional information, the covered task is a problem with a solution: Some 5.5- and 6.5-year-olds were able to solve it. In some respects this finding might be considered puzzling, as the task was designed to rule out the use of simple rules based on surface-level perceptual features. However, the failure of more than half of even 5- to 6-year-olds to avoid the side marked with the "trap" sticker, especially in Task B, where this was the only mark added to the box in full view of the child, suggests that the opaque task is not being solved in the same way as the younger children were solving the transparent task. Future studies will be needed to determine at which age most children can solve this task. In Study 4 we compared the performance of older children on the covered and transparent versions of the task directly, predicting that they would perform better on the transparent task than the covered one.

Study 4

Method

Participants. We tested a further sixteen 6.5-year-olds—eight on the covered task and eight on the transparent task—for direct comparison. There were equal numbers of boys and girls in each group. We used a recruitment window of 2 months above and below the target age. See Table 1 for the mean and range of the age in months of each group. No children were excluded.

Materials and data analysis. Materials and data analysis were the same as in Study 1.

Results

Phase 1: Testing (Boxes A and B). Performance of the 6.5-year-olds on the transparent task was near perfect, and significantly above chance according to a one-sample t test ($M = 9.69$, $SD = 0.26$), $t(7) = 51.24$, $p < .001$. As in Study 3, performance of the 6.5-year-olds on the covered task was much more variable, and although five children passed the task, their performance as a group was not different from chance ($M = 6.88$, $SD = 3.45$), $t(7) = 1.54$, $p > .05$. An independent-samples t test revealed a significant difference between 6.5-year-olds tested in the two conditions, $t(14) = 2.30$, $p < .05$.

Phase 2: Transfer (Boxes C and D). As in Study 3, in the covered condition, all of the successful children from Phase 1 passed both transfer tasks C and D (see Table 1). However, the eight 6.5-year-olds tested on the transparent condition ($M = 9.94$, $SD = 0.18$) performed significantly better on the transfer tests than the seven who proceeded to the transfer tasks in the covered condition ($M = 9.07$, $SD = 0.35$) across Studies 3 and 4, $t(13) = 6.25$, $p < .001$.

Discussion

Though 6.5-year-old children were able to find the solution to the covered task, performance was significantly poorer than on the transparent version. This makes it unlikely that both tasks are solved with the same strategy. Instead, we argue that like the younger children, 6.5-year-olds use physical information to solve the transparent version of the task: Removing all spatial and physical information impedes performance precisely because this information is processed in the transparent task.

It was not the purpose of this study to unpack the mechanisms employed to solve the covered task beyond differentiating it from the transparent one, and several different psychological mechanisms might be recruited. With hindsight it would have been very useful to ask the successful children to explain how they solved the task at the end of the experiment. Though we predicted a difference between the two tasks, the fairly late emergence of the ability to solve the covered task was somewhat surprising. For instance, children as young as 2 years of age are capable of using landmarks to locate hidden targets in an spatial array (DeLoache & Brown, 1983), and 4-year-olds are even capable of learning (after repeated trials) that a target is hidden between two landmarks (Spetch & Parent, 2006; Uttal, Sandstrom, & Newcombe, 2006). Additionally, children are able to pick up on statistical regularities between visual features and responses in other paradigms from a young age. For example, Gopnik et al. (2004) found that 2- to 3-year-old children learned a causal relationship between an arbitrary cue (blue, not red block) and an outcome (activates a sound box) in very few trials. We are not sure why around half of the 6.5-year-olds did not either readily employ a landmark-based strategy or recognize that moving the ball toward the trap sticker would cause the reward to become trapped. It might be harder to relate an arbitrary cue to an outcome if the mechanics of a problem are within the grasp of the participant, as it may cause them to discount certain types of cue as being causally relevant. While the mechanics were quite opaque in Gopnik et al. (in fact the experimenter activates the box surreptitiously), they might have been more obvious in our study (push the ball to an exit). In the covered task, the blue cue is a thin piece of plastic that cannot possibly impede the passage of a ball. That it should have relevance to the correct solution is at odds with its physical properties. Indeed, new learning has been shown to interact with previous domain specific knowledge, making relationships that run counter to previous knowledge harder to learn (e.g., talking to a machine to make it go; Schulz & Gopnik, 2004). Of course there are a great many differences between the two paradigms, but this task might prove an interesting paradigm for further examination of the interaction between previous knowledge and inferences based on patterns of dependence in the future.

Another possible mechanism that older children might recruit is the ability to interpret a cue as relevant by virtue of the experimenter's communicative intent. Children as young as 3 years of age can use the hiding location of a miniature toy in a scale model of room to locate a large toy in the full-size room (see DeLoache, 2004, for a review). However, these children were given detailed instructions about the model–room relation, and with less complete instruction DeLoache (2004) reported that no 3- to 4-year-olds and only 67% of 5- to 7-year-old children were successful, a result that fits well with the results that we have from the covered task. Future experiments could manipulate variables such as familiarity with the object involved, ostension, and intentionality to reinforce the case for this possible explanation. This type of action task might be an interesting new way to look at the development of symbolic reasoning, to uncover a developmental trajectory that is not specific to certain task constraints.

Another possibility is that subjects could use the spatial configuration of the cues to solve the task by analogy (the trap cue, for example, intersects the last point at which the ball can be seen as it is moved in one direction and the location where it might exit the box, and so might be identified as analogous to a barrier). This notion could be tested by manipulating the degree of iconicity, by using symbols that do not have the same spatial correspondence with the real task, which should reduce performance if spatial analogies play a significant role.

Although uncovering the specific nature of the cognitive challenges posed by the covered box will require further study, what is clear from Experiments 3 and 4 is that without visual access to information about the physical structures involved in the task, simply learning to associate a perceptual feature with a certain response (move the ball to the left or right) is not at all trivial for children, and is not a good candidate explanation for their success on the transparent task.

General Discussion

There were three main findings in this study. First, children from the age of 2.5 years can solve a task involving solidity, the continuity of surfaces, and the fate of unsupported objects to make a decision about which way to move an object in order to obtain it. Second, using a tool negatively affected the performance of 2.5-year-old children even though the required movements were relatively simple. Third, blocking visual access to the physical properties of the trap box and replacing them with nonfunctional cues that possessed the same predictive power as the physical properties massively reduced children's success so that only half of the 5.5- and 6.5-year-old children (and none of the younger ones) were able to consistently solve the task. We discuss each of these findings in turn.

Our findings resolve some previous disparity in the literature concerning children's difficulty with the trap problem as compared to other physical tasks involving similar physical principles but different action demands: the shelf and wall tasks. The developmental trajectory seen in the non-tool-using version of the trap task is in line with performance on the action-based shelf task (Hood et al., 2000), where 2.5-year-olds are also above chance at the group level. Though in the original version of the wall task (Berthier et al., 2000), children did not pass the task at this age at the group level, three of the 16 2.5-year-olds tested did pass at the individual level, and as discussed above, by 2.5 years of age children can begin to use the wall's location to predict where a ball will stop if the path to be taken by the ball is never visually occluded. Together, this work suggests that by 2.5 years of age, object representations are becoming robust enough to support decision making about where to move a reward as well as reaching for a reward after an event has taken place, while this ability might be much more fragile in younger 2-year-olds.

A novel feature of this paradigm was the use of a dual-functioning platform. Unlike previous tasks, simply noting the location of an obstacle (such as a wall or shelf) was not sufficient for children to solve both transfer tasks C and D. The task was designed to separate spatial from physical reasoning because children must use physical information about an object in different contexts—a piece of wood could either support and facilitate getting a reward or become the surface of a trap into which the reward would fall and be unattainable. Almost all of the children who correctly avoided the trap in the first phase of the experiment responded flexibly to the same feature based on the role it played (as a support or a barrier). This provides support for the notion that children do utilize physical information (e.g., that objects will travel along continuous paths and the one object cannot pass through another). However, the form that these representations take (e.g., would they better characterized as perceptually or conceptually based? to what degree can they be considered abstract and rule governed?) is still a matter for debate (Penn & Povinelli, 2007; Povinelli & Penn, 2012; Seed, Hanus, & Call, 2012).

A previous study (Seed et al., 2009) found that although 100% of chimpanzees passed the non-tool-using version of the trap box (in between 30 and 100 trials), just two of six chimpanzees that proceeded to the transfer phase were able to solve both Tasks C and D. In contrast, 15 out of 16 2.5- and 3.5-year-olds solved both C and D. This difference is difficult to interpret, and could indicate either that object knowledge is not ubiquitous in chimpanzees as it is in humans or that chimpanzees are more likely to employ a distance-based heuristic, despite having object knowledge, than 2- to 3-year-olds are (see Seed et al., 2012, for further discussion). Nevertheless, the implication is that the behavior of chimpanzees is more heavily influenced by the perceptual features of objects than that of even 2.5-year-old children, to the extent that when it conflicts with functional knowledge, erroneous behavior can result. This finding could implicate changes in our recent evolutionary past.

The results from the tool-using version of the task highlights that, as for the wall task, seemingly peripheral task demands play a very important role in determining successful problem solving (reviewed in Keen et al., 2008 and Seed et al., 2012; see also Baker et al., 2011). When tested with a tool, 2.5-year-olds were no longer above chance at the group level on the trap box task (though three out of 16 children did solve it). This study therefore suggests that using a tool imposes a cognitive load that can mask children's physical reasoning, just as we found to be the case for chimpanzees (Seed et al., 2009). Coordinating even simple tool use with existing physical knowledge is quite challenging, as evidenced by the fact that it takes children years to achieve a seamless integration. Moreover, integrating physical knowledge with motor responses is not the only challenge faced by children, since they also need to coordinate their actions, holding in mind all of the relevant information while inhibiting any prepotent tendencies. Just as unsuccessful children's pattern of errors in the wall task were nonrandom (most children searched behind a favorite door; Berthier et al., 2000), most children who failed the trap box task tended to have a favorite side, sometimes moving the ball in this direction on 100% of trials. However, it should be noted that this pattern of errors continued into older childhood, since most of the older children failing the covered task had a significant side bias.

Finally, we introduced another approach to disentangling the use of object properties from distance-based heuristics: comparing performance on a covered version of the task that only provided two-dimensional discriminative features. Perhaps unsurprisingly, given the wealth of evidence that children are building a representational model of the physical world over the first years of life, removing visual access to the internal structure of the problem made it much harder for children to succeed; even by the age of 6.5 years, only half of the children succeeded. Learning to associate a perceptual feature with a certain response (move the ball to the left or right) is not at all trivial for children, at least not in this setup, and is not a good candidate explanation for their success on the transparent task.

This result is apparently at odds with the use of symbolic information (i.e., an arbitrary cue that represents a physical object or property) in children as young as 3 years of age (DeLoache, 2004). Two things, however, need to be considered. First, symbolically representing physical objects or properties in a problem-solving task might be harder than in other contexts (e.g., pretend play) in which the laws of physics do not play such a prominent role. Future studies are needed to investigate whether this is the case. Second, studies that have shown 3-year-olds performing well in dual-representation or analogy-based tasks have provided a considerable amount of verbal scaffolding (DeLoache, 2004; Loewenstein & Gentner, 2005). Without such verbal scaffolding, 3-year-olds typically fail, whereas 5-year-olds succeed—precisely the age when we start to see some children succeed in the covered version of the trap box task.

Future work is needed to unpack the specific demands posed by this task. In particular, removing visual access to the path the ball can take, removing information about the mechanics of the task, and occluding different aspects of the problem at different stages of the task will be important directions for future work. Additionally, it will be of particular interest to isolate the specific extra challenges posed by tool use, given the central importance of this ability to our species' evolutionary success. Finding out what changes occur over child development that allow older children to overcome the cognitive challenges associated with using a tool to quickly solve a novel problem with greater facility than younger children or adult chimpanzees will be an interesting question for future research. Recently there have been some prominent calls for more research into development of tool-using skill in childhood (Keen, 2011), and we echo that sentiment.

References

Baillargeon, R. (1999). Young infants' expectations about hidden objects: A reply to three challenges. *Developmental Science, 2,* 115–132. doi:10.1111/1467-7687.00061

Baillargeon, R. (2004). Infants' reasoning about hidden objects: Evidence for event-general and event-specific expectations. *Developmental Science, 7,* 391–414. doi:10.1111/j.1467-7687.2004.00357.x

Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science, 3,* 2–13. doi:10.1111/j.1745-6916.2008.00056.x

Baker, S. T., Gjersoe, N. L., Sibielska-Woch, K., Leslie, A. M., & Hood, B. M. (2011). Inhibitory control interacts with core knowledge in toddlers' manual search for an occluded object. *Developmental Science, 14,* 270–279. doi:10.1111/j.1467-7687.2010.00972.x

Berthier, N. E., DeBlois, S., Poirer, 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. doi:10.1037/0012-1649.36.3.394

Bluff, L. A., Weir, A. A. S., Rutz, C., Wimpenny, J. H., & Kacelnik, A. (2007). Tool-related cognition in New Caledonian crows. *Comparative Cognition & Behavior Reviews, 2,* 1–25. doi:10.3819/ccbr.2008.20001

Byrne, R. W. (2004). The manual skills behind hominid tool use. In A. E. Russon & D. R. Begun (Eds.), *Evolutionary origins of great ape intelligence* (pp. 31–44). Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511542299.005

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

DeLoache, J. S. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences, 8*(2), 66–70. doi:10.1016/j.tics.2003.12.004

DeLoache, J. S., & Brown, A. L. (1983). Very young children's memory for the location of objects in a large-scale environment. *Child Development, 54,* 888–897. doi:10.2307/1129893

Fragaszy, D. M., & Cummins-Sebree, S. E. (2005). Relational spatial reasoning by a nonhuman: The example of capuchin monkeys. *Behavioral and Cognitive Neuroscience Reviews, 4,* 282–306. doi:10.1177/1534582306286573

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111,* 3–32. doi:10.1037/0033-295X.111.1.3

Hespos, S. J., & Baillargeon, R. (2006). Décalage in infants' knowledge about occlusion and containment events: Converging evidence from action tasks. *Cognition, 99,* B31–B41. doi:10.1016/j.cognition.2005.01.010

Hespos, S. J., & Baillargeon, R. (2008). Young infants' actions reveal their developing knowledge of support variables: Converging evidence for violation-of-expectation findings. *Cognition, 107,* 304–316. doi:10.1016/j.cognition.2007.07.009

Hood, B., Carey, S., & Prasada, S. (2000). Predicting the outcomes of physical events: Two-year-olds fail to reveal knowledge of solidity and support. *Child Development, 71,* 1540–1554. doi:10.1111/1467-8624.00247

Horner, V., & Whiten, A. (2007). Learning from others' mistakes? Limits on understanding a trap-tube task by young chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Journal of Comparative Psychology, 121,* 12–21. doi:10.1037/0735-7036.121.1.12

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

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. doi:10.1111/1467-8721.01234

Keen, R. (2011). The development of problem solving in young children: A critical cognitive skill. *Annual Review of Psychology, 62,* 1–21. doi:10.1146/annurev.psych.031809.130730

Keen, R., Berthier, N., Sylvia, M. R., Butler, S., Prunty, P. K., & Baker, R. K. (2008). Toddlers' use of cues in a search task. *Infant and Child Development, 17,* 249–267. doi:10.1002/icd.550

Keen, R., & Shutts, K. (2007). Object and event representation in toddlers. In C. von Hofsten & K. Rosander (Eds.), *Progress in brain research* (Vol. 164, pp. 227–235). Amsterdam, the Netherlands: Elsevier. doi:10.1016/S0079-6123(07)64012-6

Kloos, H., & Keen, R. (2005). An exploration of toddlers' problems in a search task. *Infancy, 7,* 7–34. doi:10.1207/s15327078in0701_3

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology, 50,* 315–353. doi:10.1016/j.cogpsych.2004.09.004

Mandler, J. M. (2004). *The foundations of mind: Origins of conceptual thought*. New York, NY: Oxford University Press.

Mash, C., Novak, E., Berthier, N. E., & Keen, R. (2006). What Do two-year-olds understand about hidden-object events? *Developmental Psychology, 42,* 263–271. doi:10.1037/0012-1649.42.2.263

Needham, A., & Baillargeon, R. (1993). Intuitions about support in 4.5-month-old infants. *Cognition, 47,* 121–148. doi:10.1016/0010-0277(93)90002-d

Penn, D. C., & Povinelli, D. J. (2007). Causal cognition in human and nonhuman animals: A comparative, critical review. *Annual Review of Psychology, 58,* 97–118. doi:10.1146/annurev.psych.58.110405.085555

Povinelli, D. J., & Penn, D. C. (2012). Through a floppy tool darkly: Toward a conceptual overthrow of animal alchemy. In T. McCormack, C. Hoerl, & S. Butterfill (Eds.), *Tool use and causal cognition* (pp. 69–88). Oxford, England: Oxford University Press.

Schulz, L. E., & Gopnik, A. (2004). Causal learning across domains. *Developmental Psychology, 40,* 162–176. doi:10.1037/0012-1649.40.2.162

Seed, A. M., Call, J., Emery, N. J., & Clayton, N. S. (2009). Chimpanzees solve the trap problem when the confound of tool use is removed. *Journal of Experimental Psychology: Animal Behavior Processes, 35,* 23–34. doi:10.1037/a0012925

Seed, A. M., Hanus, D., & Call, J. (2012). Causal knowledge in corvids, primates, and children: More than meets the eye? In T. McCormack, C. Hoerl, & S. Butterfill (Eds.), *Tool use and causal cognition* (pp. 89–110). Oxford, England: Oxford University Press.

Seed, A. M., Tebbich, S., Emery, N. J., & Clayton, N. S. (2006). Investigating physical cognition in rooks, *Corvus frugilegus*. *Current Biology, 16,* 697–701. doi:10.1016/j.cub.2006.02.066

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99,* 605–632. doi:10.1037/0033-295X.99.4.605

Spetch, M. L., & Parent, M. B. (2006). Age and sex differences in children's spatial search strategies. *Psychonomic Bulletin & Review, 13,* 807–812. doi:10.3758/BF03194001

Uttal, D. H., Sandstrom, L. B., & Newcombe, N. S. (2006). One hidden object, two spatial codes: Young children's use of relational and vector coding. *Journal of Cognition and Development, 7,* 503–525. doi:10.1207/s15327647jcd0704_4

Visalberghi, E., & Limongelli, L. (1994). Lack of comprehension of cause–effect relations in tool-using capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology, 108,* 15–22. doi:10.1037/0735-7036.108.1.15

Want, S. C., & Harris, P. L. (2001). Learning from other people's mistakes: Causal understanding in learning to use a tool. *Child Development, 72,* 431–443. doi:10.1111/1467-8624.00288

Received February 22, 2011
Revision received February 13, 2014
Accepted February 21, 2014 ■