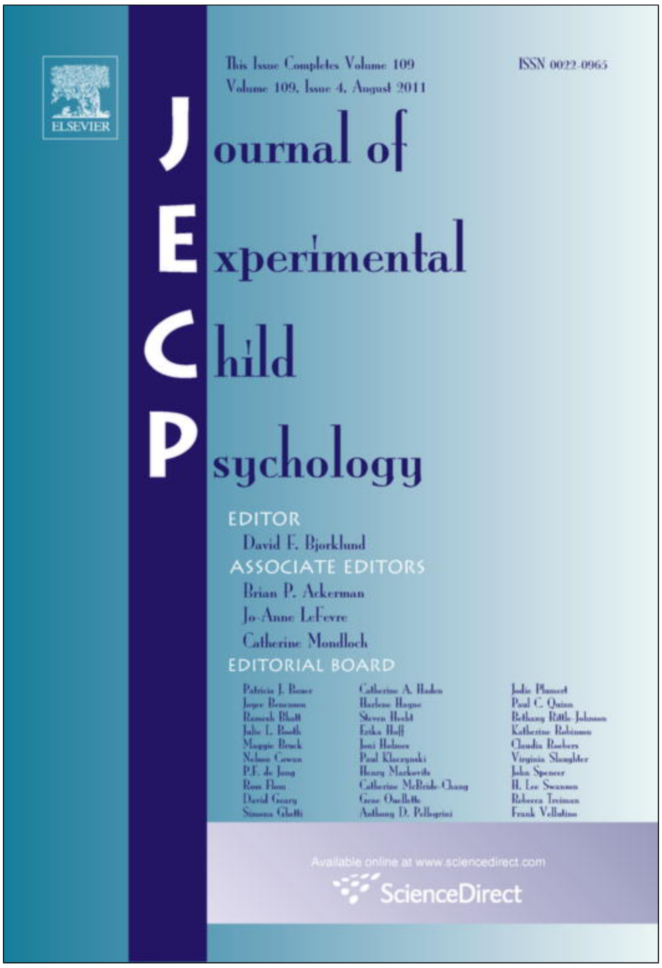
This Issue Completes Volume 109
Volume 109, Issue 4, August 2011
ISSN 0022-0965
ELSEVIER
Journal of
Experimental
Child
Psychology
EDITOR
David F. Bjorklund
ASSOCIATE EDITORS
Brian P. Ackerman
Jo-Anne LeFevre
Catherine Mondloch
EDITORIAL BOARD
Available online at www.sciencedirect.com
ScienceDirect

Journal of Experimental Child Psychology 109 (2011) 497–511

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Why do children lack the flexibility to innovate tools?

Nicola Cutting *, Ian A. Apperly, Sarah R. Beck

School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

ARTICLE INFO

Article history:
Received 9 November 2010
Revised 8 February 2011
Available online 21 March 2011

Keywords:
Tool innovation
Mental flexibility
Ill-structured problems
Problem solving
Causal understanding
Executive function

ABSTRACT

Despite being proficient tool users, young children have surprising difficulty in innovating tools (making novel tools to solve problems). Two experiments found that 4- to 7-year-olds had difficulty on two tool innovation problems and explored reasons for this inflexibility. Experiment 1 ($N = 51$) showed that children's performance was unaffected by the need to switch away from previously correct strategies. Experiment 2 ($N = 92$) suggested that children's difficulty could not easily be explained by task pragmatics or permission issues. Both experiments found evidence that some children perseverated on a single incorrect strategy, but such perseveration was insufficient to explain children's tendency not to innovate tools. We suggest that children's difficulty lies not with switching, task pragmatics, or behavioral perseveration but rather with solving the fundamentally "ill-structured" nature of tool innovation problems.

Introduction

Human life revolves around the use of tools. It is almost impossible to consider life without them. How would we cook without pans and utensils? How would we even catch or dig up our food? Humans are believed to be experts in all tool-related behaviors (Defeyter & German, 2003; Herrmann, Call, Hernàndez-Lloreda, Hare, & Tomasello, 2007). However, despite extensive research indicating children's early competence for tool use, children's tool-making abilities have been neglected in the developmental literature. In this article, we distinguish between two types of tool making: tool manufacture (the ability to make tools after instruction or observation) and tool innovation (independently making a novel tool to solve a problem). The current experiments focused on children's tool innovation and explored whether children's difficulty with tool innovation was due to mental inflexibility.

* Corresponding author.
E-mail address: nxc945@bham.ac.uk (N. Cutting).

0022-0965/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jecp.2011.02.012

Early hominid tool use is thought to have propelled human evolution, making us the advanced social beings we are today (Csibra & Gergely, 2009; Kacelnik, 2009). Tool-related activities are implicated in the development of social behaviors such as imitation, teaching, and language (Csibra & Gergely, 2009; Gibson & Ingold, 1993). The cultural intelligence hypothesis proposes that the advancement of these social capacities has allowed humans to develop cognitive skills not possessed by our nearest primate relatives (Herrmann et al., 2007). Our ability to collaborate and share knowledge permitted massive technological advances in our manufacture and use of an extensive range of tools. Tools quite clearly have played, and continue to play, an integral part in human life.

There is a substantial literature on the development of children's tool-related behaviors. Competent tool use is evident from an early age, as demonstrated by the skillful manipulation of spoons (Connolly & Dalgleish, 1989), hooks and rakes (Brown, 1990), and many more tools during the second year of life. A large literature on social learning shows that young children are also able to learn about novel tools by imitating others from 2 or 3 years of age (McGuigan & Whiten, 2009; Meltzoff, 1995; Want & Harris, 2002). Young children not only can use tools but also show early abilities to infer their intended use (Casler & Kelemen, 2005), their design (Casler, Terziyan, & Greene, 2009), and how they should be categorized (Defeyter, Hearing, & German, 2009).

Furthermore, a recent study showed children to be competent in tool manufacture (making tools after instruction) (Beck, Apperly, Chappell, Guthrie, & Cutting, 2011). Children as young as 3 years of age readily manufactured a simple hook tool when the hook-making action was demonstrated. This is in line with the findings of research investigating infant memory for actions, which shows that at around 30 months of age children readily imitated a model who constructed a nontool object such as a rattle (Barr & Wyss, 2008; Hayne, Herbert, & Simcock, 2003; Herbert & Hayne, 2000). In the current article, we explore a possible limit on children's excellent tool-related capabilities.

Tools were once thought to be a uniquely human phenomenon, but tool-related behaviors are now widely studied comparatively. Recent research has focused on the making of tools. Chimpanzees (*Pan troglodytes*) have demonstrated the ability to manufacture a wide range of tools both in the wild (Boesch & Boesch, 1990) and in captivity (Bania, Harris, Kinsley, & Boysen, 2009; Visalberghi, Fragaszy, & Savage-Rumbaugh, 1995; Povinelli, 2000). However, there is some debate as to whether such behavior, especially when seen in the laboratory, is insightful or merely results from a trial-and-error approach (Povinelli, 2000).

New Caledonian crows (*Corvus moneduloides*) are also well known for their tool manufacturing abilities. Specifically, they manufacture hooks from twigs to retrieve food in the wild (Hunt & Gray, 2002). More recently, impressive tool manufacturing abilities have been seen in the laboratory. To retrieve a bucket from a tall narrow tube, one crow named "Betty" bent a piece of wire into a hook, which she then used to solve the task. What was impressive was that Betty made a tool from a piece of wire, a material that crows would not encounter in the wild. Furthermore, on repeated trials, Betty employed a variety of bending techniques, suggesting that her success was not the result of associative learning (Weir, Chappell, & Kacelnik, 2002). More recently, four rooks, a species that does not use tools in the wild, have also solved this tool manufacturing task (Bird & Emery, 2009).

Tool innovation

Being able to make tools allows individuals to perform a much wider range of acts than they could without tools or with only found tools. However, we should remember that in the corvid (crow family) and child studies described above, when individuals made tools they had already seen an example of the required tools, and in the child study (Beck et al., 2011), when children made tools successfully it was after the experimenter had demonstrated how to make the tools. We term the ability to make tools, after having been instructed or having seen an example, as tool manufacture. This begs the question of where tools come from in the first place. Tool innovation, or making a novel tool to solve a problem, has been largely neglected by the comparative and developmental literatures. This is surprising because tool innovation must be the foundation for all other tool-related behaviors; children's (and adults') evident capacity to make tools and use tools that they see used by others would be of little use if nobody innovated tools in the first place.

There has been only one study of tool innovation to date. Using an apparatus based on that used by Weir and colleagues (2002), Beck and colleagues (2011) investigated children's ability to innovate a simple hook tool needed to retrieve a bucket from a narrow vertical tube. Children were given a straight pipe cleaner, a long piece of string, and some small matchsticks. The most obvious solution was to bend the pipe cleaner into a hook. This is what most adults did when confronted by the task (a few individuals made a functional tool by attaching a matchstick to the pipe cleaner to make an inverted "T"). The critical difference between Beck and colleagues' study and the studies with corvids (Bird & Emery, 2009; Weir et al., 2002) is that the child participants had not seen an example of the appropriate tool within the context of this task. Instead, they needed to imagine the solution themselves; that is, they needed to innovate a novel tool. As mentioned above, when children in the same study saw the experimenter demonstrate making the appropriate tool, they had no difficulties in repeating this tool manufacture.

Children performed remarkably poorly on the tool innovation task. In Beck and colleagues' (2011) study, 3- to 5-year-olds rarely made a hook or any other functional tool, fewer than half of 7-year-olds succeeded, and children did not perform at high levels until 9 or 10 years of age. These findings are even more striking in the context of further evidence presented by the authors. Even 4-year-olds understood that a hook was the best tool for the job and chose it over a straight pipe cleaner significantly more than chance (Beck et al.'s Experiment 1). Children's difficulties persisted even after receiving a warm-up exercise with the materials, which ensured that they knew manipulation of materials was allowed and the pipe cleaner was pliable (Experiment 3). Finally, the fact that tool innovation was the limiting step for children was underscored by the finding that nearly all children successfully completed the task when they received a demonstration of hook bending after their initial failure. This is consistent with literature showing that children are very successful at social learning. This finding raises the question: Why do children find tool innovation to be so difficult?

Because young children clearly have the competence to manufacture and use tools, it seems unlikely that any difficulty with tool innovation would be due to a lack of understanding of what tools are or any difficulty with the practical business of shaping a tool and executing tool-using actions. Instead, we look to a cognitive explanation. Difficulty with tool innovation might be a consequence of the mental inflexibility that is commonly observed during early childhood. One way to characterize this mental inflexibility is to think about children's developing executive control.

Executive control is an umbrella term for psychological processes involved in the conscious control of thought and action (Anderson, 1998; Zelazo & Müller, 2002). Executive control is needed for novel tasks or situations that require concentration, planning, strategy development, coordination, and/or choosing between alternative options (Anderson, 1998; Brocki & Bohlin, 2004; Diamond, 2006). Imagining what kind of tool is needed to solve a problem and how to make it (tool innovation) is likely to tap many of these demands, and to a greater extent, than simply using or manufacturing tools, which relies mainly on imitating actions.

There are different ways in which we might construe the role of mental flexibility in tool innovation. One possibility is that children are able to generate potential tool innovation solutions to the task but find it difficult to move on from unsuccessful ideas and so tend to become "stuck in set." The ability to select and switch between multiple perspectives, tasks, or strategies to determine the optimal option for the current situation is a well-known component of executive function that develops significantly between 3 and 5 years of age (e.g., Chevalier & Blaye, 2009; Diamond, 2006). This is demonstrated in simple card-sorting tasks (Espy, 1997; Frye, Zelazo, & Palfai, 1995), where children begin to demonstrate the ability to shift flexibly between rules. Between 5 and 11 years of age, further improvements in cognitive flexibility occur, with children passing more complex tasks (Luna et al., 2001) and improving in speed and accuracy (Meiran, 1996). It seems plausible that difficulty in switching between alternatives might contribute to children's difficulty with tool innovation.

Thus, in Experiment 1, we investigated the role of switching in tool innovation. We tested children on two tool innovation tasks that required "opposite" solutions (hook making required a pipe cleaner to be bent, and the new task required a pipe cleaner to be unbent). We speculated that if children's difficulty was in switching between strategies, then having succeeded (before or after a demonstration) using one strategy on one task, children might find it to be particularly difficult to adopt a different strategy on the second task. Furthermore, introducing a second tool innovation task also allowed us to generalize Beck and colleagues' (2011) claims, which were based only on a hook-making task.

A second possibility is that children have the capability to innovate the tools required for the tasks, but other features of our tool innovation task created unintended difficulties, making it hard for children to demonstrate this flexible behavior. For example, in the hook innovation task, a straight pipe cleaner is presented along with other distracter items as materials that can solve the task. Children may perseverate with the first material they attempt to solve the task with and fail to switch to another material if the first one proves to be unsuccessful. Alternatively, they may restrict themselves to using only unmodified materials rather than making them into a new tool. We discuss this further in the introduction to Experiment 2, where we adapt the task instructions so as to reduce the chances that children will perseverate with unmodified materials.

A third possibility is that although young children are able to make and use tools, they lack the mental flexibility necessary to innovate tools because tool innovation is an "ill-structured" problem. Executive function researchers distinguish between well-structured and ill-structured problems (Burgess et al., 1996; Goel, 1995). Most commonly used tests of executive function (including those used with children) are well structured insofar as they have a clearly defined set of stimuli (e.g., cards with colored pictures), a clearly defined set of responses (boxes in which to sort the cards), and a clearly defined set of rules (sort according to the color of the picture and then switch to sorting according to shape). In contrast, ill-structured tasks lack information either in their start and goal states or in the transformations needed to get from one state to the other, and so part of the task requirement is for participants to supply this for themselves. The difference between well- and ill-structured executive tasks is underscored by the observation that some brain-injured patients (Shallice & Burgess, 1991) and children with autism (White, Burgess, & Hill, 2009) may pass traditional, well-structured executive function tasks yet show impairment on ill-structured tasks and experience difficulties with mental flexibility in their everyday lives.

Tool innovation is an excellent example of an ill-structured task. Participants have information about the start and goal states but lack information about how to get from one to the other. They must devise and hold in mind a solution to the problem, inhibit irrelevant actions, and plan a sequence of actions to achieve their goal. We return to whether tool innovation might be thought of as an ill-structured task in the General Discussion in the light of our tests of the role of cognitive flexibility.

Experiment 1

The first experiment replicated Beck and colleagues' (2011) hook study with the addition of a second tool innovation task, namely, unbending. In the new task, a pipe cleaner was presented bent in half and needed to be unbent to make it long enough to push a ball from a tube. An unbending task was chosen because it requires the opposite action to the hooks (bending) task. Following Beck and colleagues, children were given a piece of string as a distracter as well as a pipe cleaner (although, unlike Beck et al., we did not include small sticks so as to prevent the making of other functional tools). The distracter material allowed us to see whether the first material children selected was the functionally relevant pipe cleaner. In addition, all children received a warm-up exercise in which they manipulated materials (as in Beck et al.'s Experiment 3) to ensure that they had experience with the materials' properties. Although we did not explicitly check, all children were expected to have had previous experience in working with pipe cleaners in a craft context in school.

Method

Participants

The final sample consisted of 24 4- and 5-year-olds (13 boys and 11 girls, mean age = 4 years 10 months [4;10], range = 4;3–5;3) and 27 6- and 7-year-olds (10 boys and 17 girls, mean age = 6;8, range = 6;3–7;2) from a primary school in South Birmingham, United Kingdom. The ethnic composition of the sample was 91% Caucasian, 7% Black, and 2% other/unknown. A further 5 children were tested but excluded from analysis (3 children from the 6- and 7-year age group who retrieved the sticker without making a functional tool [e.g., by catching the bucket on the folded end of the wire pipe cleaner] and 2 children from the 4- and 5-year age group, 1 who retrieved the sticker without making a tool and 1 who had seen another child perform the task).

Materials

For the warm-up task, a pipe cleaner (length = 29 cm), a pen (length = 14 cm), a piece of string (length = 29 cm), and a template of an "S" shape printed onto an A4 card (height = 12 cm, width = 9 cm) were used. For the "hooks" task, the materials were a transparent plastic tube (height = 22 cm, width of opening = 4 cm) attached vertically to a cardboard base (length = 35 cm, width = 21 cm), a bucket with a wire handle, a pipe cleaner (length = 29 cm), a piece of string (length = 29 cm), and a sticker (see Fig. 1). For the "unbending" task, the materials were a transparent plastic tube (length = 22 cm, width of opening = 4 cm) attached horizontally to a cardboard base (length = 33 cm, width = 15 cm), a pipe cleaner bent in half (unbent length = 22 cm), a piece of string (length = 29 cm), and a small spherical pom-pom (like those used in crafts, diameter = 4 cm) with a sticker attached (see Fig. 2). We used a small clock to time the task.

Procedure

Before testing began, children were instructed by their class teacher not to tell other children how to play the games they would be playing with the experimenter so that they would be a nice surprise for everyone. Participants were tested by a female experimenter in a quiet area just outside the main classroom. Children and the experimenter sat facing each other across a table. First, children completed the warm-up exercise. After this, children received both the hooks and unbending tasks. The order was counterbalanced across participants.

Warm-up task. Children watched as the experimenter demonstrated actions with the string and pipe cleaner (order counterbalanced), which the child then copied. The pipe cleaner was wound around a pen and then removed to demonstrate that it kept its shape. The string was laid over the template to follow the S-shaped pattern.

Hooks task. Children were shown the vertical transparent tube with the bucket containing a sticker already in place in the bottom. They were told that if they could get the sticker out of the tube, they would be allowed to keep it. The experimenter then brought out the string and pipe cleaner and told children that these were things that "may help" to get the sticker out. Children were then given 1 min to try to retrieve the sticker. No feedback was given, but children were given neutral prompts if

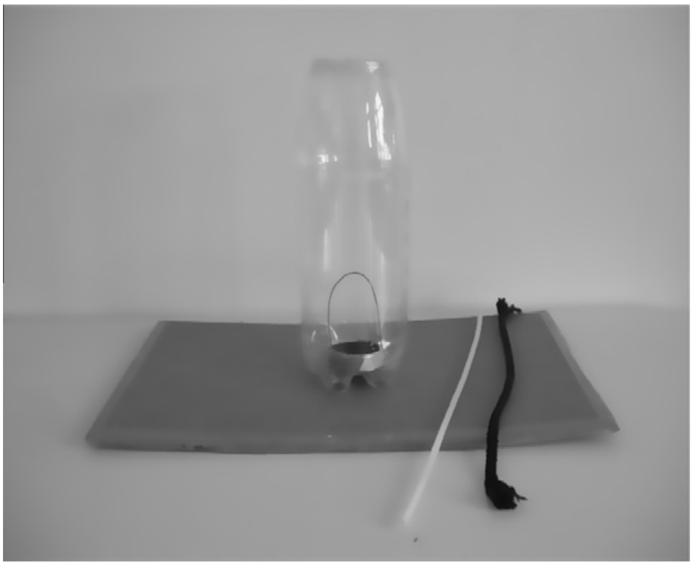

Fig. 1. Apparatus used for hooks task: Tall tube containing bucket (with sticker inside), pipe cleaner, and string.

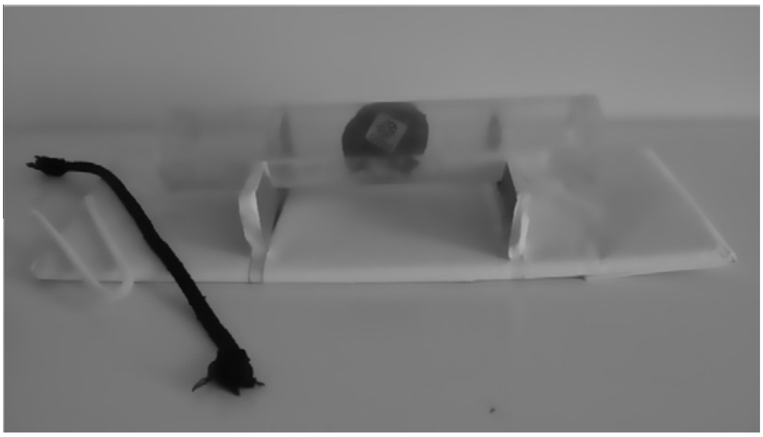

Fig. 2. Apparatus used for unbending task: Narrow horizontal tube containing pom-pom (with sticker attached), bent pipe cleaner, and string.

required. Examples of prompts used include "Can you think how you might be able to get the sticker out?" and "Maybe you could use these things to help you." If, after 1 min had elapsed, children had not retrieved the sticker, they were encouraged by the experimenter to put down the materials they were using. With the materials remaining on view in front of participants, the experimenter then said "Watch this" and, using another pipe cleaner held in the middle, bent one end to form a hook. Children were again encouraged to retrieve the sticker. They were not given the experimenter's hooked pipe cleaner.

Unbending task. Children were shown the horizontal tube with the sticker attached to a pom-pom held in the middle. As in the hooks task, they were told that if they could get the sticker out, they could keep it. The experimenter introduced the string and the bent pipe cleaner as things that "may help" to retrieve the sticker. If, after 1 min had elapsed, children had not retrieved the sticker, they were encouraged to put down the materials they were using. With the materials remaining on view in front of participants, the experimenter then demonstrated "unbending" with another bent pipe cleaner. Children were again encouraged to try to retrieve the sticker (using their own materials).

Measures

Children's behaviors were recorded online using a coding system to differentiate their actions across time. The system coded materials selected (including whether they were touched, picked up, or entered into the tube), whether the material was manipulated and what shape was made, and whether participants were successful before and after the experimenter's demonstration.

Results

There was no difference in success based on gender (hooks task: Fisher's exact test, $p = .739$; unbending task: $\chi^2(1,\ N = 51) = 0.123,\ p = .723$), and so all data were combined for subsequent analyses.

As can be seen in Table 1, children did not assume all materials to be equally useful; there was a strong bias for children to both touch and use the pipe cleaner first in each task. Furthermore, children tended to use the materials as they were presented and very rarely made any attempt to adapt them. However, as is clear in the final column of the table, after tool manufacture was demonstrated, children easily succeeded in these tasks.

First, we focused on the main variable of interest: successful tool innovation before demonstration. Children were coded as successful in the hooks task if they bent the pipe cleaner into a hook, within the 1-min time limit, and used this to retrieve the bucket from the tube. Children were coded as successful in the unbending task if they unbent the pipe cleaner (within the 1-min time limit), making it long enough to push the pom-pom from the tube. It was occasionally unclear whether unbending had

Table 1
Children's behaviors during innovation tasks for Experiment 1.

Age group	Touched first		Used first				Success	
	Pipe cleaner	String	Pipe cleaner as presented[a]	String	Pipe cleaner adapted[b]	Combination[c]	Before demonstration	After demonstration
Hooks task								
4- and 5-year-olds ($n = 24$)	21	3	20	3	1	0	2 (8%)	19 (79%)
6- and 7-year-olds ($n = 27$)	25	2	23	1	2	1	8 (30%)	18 (67%)
Unbending task								
4- and 5-year-olds ($n = 24$)	17	7	17	5	1	1	8 (33%)	14 (58%)
6- and 7-year-olds ($n = 27$)	22	5	20	4	3	0	15 (56%)	11 (41%)

[a] Hooks task: pipe cleaner presented straight; unbending task: pipe cleaner presented bent in half.
[b] Hooks task: pipe cleaner bent into hook; unbending task: pipe cleaner unbent.
[c] Participant combined string and pipe cleaner, usually by tying them together.

been an intentional act because exerting force on the bent pipe cleaner sometimes allowed it to unbend. Because insight is difficult to establish, all cases of unbending were coded as successful.

The low success rates before demonstration for the hooks task are consistent with Beck and colleagues (2011), demonstrating a stable finding that children display difficulties in innovating a simple hook tool. The new unbending task also yielded low success rates, with only one-third of 4- and 5-year-olds and half of 6- and 7-year-olds unbending the pipe cleaner to make it long enough to push the pom-pom from the tube. Because these results may include a small number of children who unbent the pipe cleaner unintentionally, the results for true insightful tool innovation may be lower still.

Comparison of success across age groups reveals a trend that older children successfully innovate more tools than younger children, but unlike Beck and colleagues (2011), we did not find a significant difference between age groups (hooks task: Fisher's exact test, $p = .081$; unbending task: $p = .160$). Therefore, data for the two age groups were combined for subsequent analyses.

Success before demonstration on the unbending task was better than that on the hooks task (McNemar test, $p = .011$). We used chi-square tests to investigate whether task order affected children's performance. Whether the hooks task was presented first or second had no effect on whether children succeeded in making a hook (Fisher's exact test, $p > .999$). Similarly, order had no effect on success in the unbending task, $\chi^2(1, N = 51) = 0.167$, $p = .683$. These results indicate that children's success on one task (whether spontaneous or following demonstration) neither aided nor hindered their spontaneous success on the second task.

Having established there to be no relationship between behaviors between tasks, we decided to look more closely at both unsuccessful and successful (before demonstration) children's behaviors within each task. Although children were not perseverating on techniques across tasks, one possible reason for failure could be that children were perseverating on techniques within a task. We coded unsuccessful children as perseverators if they entered only one "tool" into the tube and persisted in trying to retrieve the sticker with this tool for the whole time period (1 min). As can be seen in Table 2, perseveration was not a common occurrence for either the 4- and 5-year-olds or the 6- and 7-year-olds. Chi-square analyses show there to be no difference in perseveration between the two age groups (hooks task: $\chi^2(1, N = 41) = 0.149$, $p = .699$; unbending task: Fisher's exact test, $p = .613$). Although it is a potential stumbling block to overcome if one first approaches the task with the wrong tool, perseveration cannot explain why children are not successful in innovating tools in this study.

Although unsuccessful children rarely perseverated with one material for the whole time period, few manipulated the materials in any way (e.g., bent the pipe cleaner, combined materials). In the hooks task, only 18% of 4- and 5-year-olds and 26% of 6- and 7-year-olds manipulated materials, and similarly in the unbending task, the figures were only 25% (4- and 5-year-olds) and 17% (6- and 7-year-olds).

Table 2
Frequencies of perseveration in unsuccessful children and number of insertions into tube for successful children for Experiment 1.

Age group	Unsuccessful			Successful			
	n	Perseveration		*n*	Insertion into tube		
		No	Yes		Immediate tool	1 unsuccessful, then tool	2+ unsuccessful, then tool
Hooks task							
4- and 5-year-olds	22	15	7	2	1	1	0
6- and 7-year-olds	19	14	5	8	2	4	2
Unbending task							
4- and 5-year-olds	16	13	3	8	2	4	2
6- and 7-year-olds	12	11	1	15	3	7	5

Next, we examined the actions of successful tool innovators within each task. Successful tool innovators were coded as to the number of different items inserted into the tube before retrieving the sticker. Table 2 shows that the majority of successful tool innovators either entered a successful tool into the tube immediately (i.e., a hook or an unbent pipe cleaner) or entered one unsuccessful tool (always an unmodified material) before making and entering a successful one. These results suggest that tool innovation resulted more from insightful solving of the task than from trial-and-error learning.

Experiment 2

In Experiment 1, children's difficulties with tool innovation were shown to extend beyond hooks to another task, namely, unbending. Investigation of children's success showed that there was no effect of task order, indicating that children's inflexible behavior on one tool innovation task was not modified by prior experience of making a tool on another task. Importantly, children's inflexibility did not appear to be due to perseveration on one unsuccessful strategy. Unsuccessful children in both tasks rarely perseverated with the same material for the whole time period. However, it was also notable that unsuccessful children made few attempts to modify the materials they were given. In our second experiment, we explored the possibility that children may fail to modify the materials because they think they are not allowed to do so.

Although children in Experiment 1 experienced a warm-up task in which they manipulated string and pipe cleaner materials, it remains possible that these children did not realize that they were allowed to alter the materials given in the context of the main task. Alternatively, children may have failed to modify materials due to the pragmatics of the task. Children were presented with the materials as things that "may help" to retrieve the sticker. This may have been interpreted by children as the experimenter proffering the materials as tools that could be used *as presented* as a solution to the task, thereby preventing modification. In Experiment 2, we sought to minimize the likelihood of permission or pragmatics playing a role in children's poor performance on the tool innovation task by telling children that they needed to make something with the materials.

Method

There were two conditions. In the control condition, children received the same instructions as in Experiment 1. In the experimental condition, children received the new instruction to *make something* with the materials. This instruction was used to avoid any assumption children may have had that the materials must be used as they were to solve the task. In addition, we tried to reduce any possibility that children thought the experimenter was giving them premade tools to solve the task by introducing the children to a puppet, "Heinz," who happened to have some materials with him. The aim of the puppet was to draw attention away from the experimenter; making the task appear to be more general rather than one that the experimenter had created and had the answer to. Because of this, we excluded the warm-up phase of the experiment in which children completed an unrelated task that involved manipulating the materials. Previous results (Beck et al., 2011, Experiment 3) indicated that

Table 3
Tool innovation behaviors as a function of age and condition for Experiment 2.

Age group	Condition	*n*	Touch First			Use first					Success	
			Pipe cleaner	String	Matchstick	Pipe cleaner as presented[a]	String	Matchstick	Pipe cleaner adapted[b]	Combination[c]	Before demonstration	After demonstration
Hooks task												
4- and 5-year-olds	Help	22	20	0	2	19	1	1	1	0	1 (5%)	19 (86%)
	Make	22	21	1	0	17	2	0	1	2	2 (9%)	17 (77%)
6- and 7-year-olds	Help	23	21	2	0	19	3	0	1	0	8 (35%)	14 (61%)
	Make	25	23	1	1	17	1	0	4	3	12 (48%)	13 (52%)
Unbending task												
4- and 5-year-olds	Help	22	14	7	1	11	7	1	3	0	10 (45%)	11 (50%)
	Make	22	16	2	4	11	2	4	3	2	8 (36%)	9 (41%)
6- and 7-year-olds	Help	23	14	1	8	7	1	4	9	2	16 (70%)	7 (30%)
	Make[d]	25	19	2	4	9	1	1	8	5	18 (72%)	4 (16%)

[a] Hooks task pipe cleaner presented straight; unbending task pipe cleaner presented bent in half.
[b] Hooks task pipe cleaner bent into hook; unbending task pipe cleaner unbent.
[c] Participant combined string and pipe cleaner, usually by tying them together.
[d] One participant in this group did not attempt to use a "tool" and spent the time trying to make something.

the warm-up exercise had no effect on task success. The materials in the control condition were also presented by a puppet, and the wording was changed to "Here are some things that *can* help you." We used the word "can" rather than "may" (as we had used in Experiment 1) to match the certainty implied by the instructions in the experimental condition. Thus, the only difference between the experimental and control conditions was the instruction to make something.

Participants

The final sample consisted of 44 4- and 5-year-olds (17 boys and 27 girls, mean age = 4;10, range = 4;5–5;5) and 48 6- and 7-year-olds (25 boys and 23 girls, mean age = 6;10, range = 6;5–7;4) from a primary school in South Birmingham, United Kingdom. The ethnic composition of the sample was 48% Caucasian, 27% Black, 10% Asian, and 15% other/unknown.

Materials

The materials for Experiment 2 were the same as in Experiment 1 except for the addition of a short stick (5 cm) presented as an additional distracter (this matched the materials used by Beck et al., 2011), a puppet, and a box ($20 \times 13 \times 5$ cm) in which the puppet carried the materials.

Procedure

Participants were tested in a similar environment to that outlined in Experiment 1. All participants received both the hooks and unbending tasks with the order counterbalanced. Children were alternately assigned to either the control (help) group or the experimental (make) group based on the teacher's class list. Children were introduced to the puppet and were told, "Heinz really likes to play games, so he might come back later to see what we are doing." The procedure for both groups was identical apart from the instructions given by Heinz.

Both the hooks and unbending tasks followed the same procedure as in Experiment 1, but after showing the tube apparatus, the experimenter exclaimed, "Oh look, here's Heinz—let's see what he has to say." The experimenter then listened as Heinz spoke in her ear and then told the children either "Heinz says he has some things here that can help you to get the sticker" (control group) or "Heinz says he has some things here you can make something with to get the sticker" (experimental group). As before, if children had not retrieved the sticker after 1 min, bending or unbending was demonstrated by the experimenter.

Results

Examination of success rates showed there to be no effect of gender (hooks task: $\chi^2(1, N = 92) = 0.058$, $p = .809$; unbending task: $\chi^2(1, N = 92) = 0.097$, $p = .755$), and so all data were combined for subsequent analyses.

Both the hooks and unbending tasks showed the same pattern of behavior seen previously (see Table 3). Children had a strong bias to both touch and use the pipe cleaner first, but very few then went on to manipulate the pipe cleaner and innovate a tool.

This experiment provides further evidence for the stability of the finding that young children do not readily innovate a hook tool to solve a task, with only 3 of 44 4- and 5-year-olds and 20 of 48 6- and 7-year-olds innovating a hook to solve the task. The results for the new unbending task are also consistent with the previous success rates, with 18 of 44 4- and 5-year-olds and 34 of 48 6- and 7-year-olds unbending the pipe cleaner to retrieve the sticker. As in Experiment 1, the unbending task was easier for children to achieve than the hooks task (McNemar test, $p < .001$). Chi-square tests were used to investigate whether task order affected children's performance. Whether each task was presented first or second had no effect on whether children succeeded in making a hook, $\chi^2(1, N = 92) = 0.000$, $p > .999$, or unbending, $\chi^2(1, N = 92) = 0.003$, $p = .956$, indicating the absence of transfer effects.

There was significant improvement in performance with age. Older children were more successful in innovating tools on both the hooks task, $\chi^2(1, N = 92) = 14.869$, $p < .001$, and the unbending task, $\chi^2(1, N = 92) = 8.365$, $p = .004$. Although no age difference was found in Experiment 1, age effects were

Table 4
Frequencies of perseveration in unsuccessful children and number of entries into tube for successful children for Experiment 2.

Age group	Unsuccessful			Successful			
	n	Perseveration		*n*	Entry into tube		
		No	Yes		Immediate tool	1 unsuccessful, then tool	2+ unsuccessful, then tool
Hooks task							
4- and 5-year-olds	41	17	24	3	2	1	0
6- and 7-year-olds	28	24	4	20[a]	7	8	3
Unbending task							
4- and 5-year-olds	26	19	7	18[b]	10	1	6
6- and 7-year-olds	14	13	1	34	20	8	6

[a] Two participants retrieved the sticker without making a hook.
[b] One participant retrieved the sticker without unbending the pipe cleaner.

observed in this age range by Beck and colleagues (2011), and we conclude that the most likely reason for the difference between Experiments 1 and 2 is the larger sample size in Experiment 2.

The main aim of Experiment 2 was to test whether instructing children to make something with the materials helped them to be more flexible at innovating tools. Chi-square analyses revealed no difference between success rates for the experimental and control conditions for either the hooks task, $\chi^2(1, N = 92) = 1.174$, $p = .278$, or the unbending task, $\chi^2(1, N = 92) = 0.057$, $p = .812$. This was also true for the two age groups independently (4- and 5-year-olds—hooks task: Fisher's exact test, $p > .999$; unbending task: $\chi^2(1, N = 44) = 0.376$, $p = .540$; 6- and 7-year-olds—hooks task: $\chi^2(1, N = 48) = 0.861$, $p = .353$; unbending task: $\chi^2(1, N = 48) = 0.034$, $p = .853$). This suggests that it is unlikely that children's difficulty in innovating a tool is due to a misperception that they are not allowed to modify the tool-making materials.

As in Experiment 1, we then examined children's behaviors more closely (see Table 4). For unsuccessful participants, we again coded whether they perseverated on one technique for the whole time period. The performance of 6- and 7-year-olds was consistent with that in Experiment 1. These children did not perseverate with one object. In contrast, 4- and 5-year-olds displayed higher levels of perseverative behavior. Chi-square analysis of the hooks task revealed 4- and 5-year-olds to be significantly more likely than 6- and 7-year-olds to perseverate on one unsuccessful technique for the whole time period, $\chi^2(1, N = 69) = 13.511$, $p < .001$. This same trend was seen for the unbending task but did not reach significance (Fisher's exact test, $p = .222$), most likely due to the lower number of unsuccessful participants.

Examination of the behaviors of successful tool innovators paints a similar picture to Experiment 1. In both tasks, the majority of successful participants succeeded immediately or after just one unsuccessful insertion, suggesting a role for insight rather than trial-and-error learning.

General discussion

The current experiments suggested that young children show striking inflexibility on two tasks that require them to innovate a simple tool, and the two experiments investigated alternative reasons for this inflexibility.

An important first objective was to test whether difficulties previously observed by Beck and colleagues (2011) when children were required to innovate a hook tool would also be apparent on another task. Our novel unbending task was easier to solve than the hooks task, yet overall performance was still poor. Approximately two-thirds of 4- and 5-year-olds and one-third of 6- and 7-year-olds spent their time probing with inadequate materials rather than performing the simple action of unbending the pipe cleaner needed to solve the task. A reason why the unbending task may be easier for children to solve could be because the final shape of the required tool is much simpler to manufacture than the hook. The fact that unbending is easier is consistent with research in the comparative literature showing that chimpanzees have more difficulty in assembling tools than in disassembling them. In Bania and colleagues' (2009) study, chimpanzees were given a tool composed of a long stick with two short sticks that could be added to each end to make an "H" shape. Chimpanzees either needed to assemble a hook to retrieve an object or needed to disassemble the H shape to form the long stick required to probe in a tube. As stated above, chimpanzees found it easier to disassemble the tool, which fits with our finding that children found it easier to unbend, and therefore disassemble, what they had been given than to assemble a hook. Further developmental research is needed to investigate different types of tool manufacture that may have differing levels of complexity.

Having established that tool innovation difficulties are robust across two different tasks, we next considered whether the findings could be explained by children having difficulty in switching between possible task solutions. Experiment 1 revealed that in their second task, children did not perseverate on techniques that had been successful in their first task. For example, children who, before the demonstration, successfully bent the straight pipe cleaner to make a hook on their first task were just as likely to switch to the correct strategy of unbending for their second task compared with children who did not bend the pipe cleaner on their first task. However, it is also noteworthy that children did not demonstrate any positive transfer effects, meaning that succeeding or being shown how to succeed in the first task did not allow children to gain insight and facilitate their tool innovation ability and so did not increase the likelihood of success on the second task. This suggests that tool innovation might not be an all-or-nothing insight that generalizes easily from one task to another.

Experiment 2 investigated whether children's inflexible behavior was due to a misunderstanding that they should not alter the given materials. By telling children that they could make something with the materials, we aimed to overcome any tendency for children to believe that the materials were things that should be used without modification. In fact, children who were prompted to make something were no more likely to make a tool than children who were told only that the materials "could help" with retrieving the sticker. Further evidence against the possibility that children thought they were not permitted to modify materials comes from the absence of transfer effects (in Experiment 1) after the warm-up phase and (in both experiments) after their first task. Given that, in Experiment 1, children modified a pipe cleaner during the warm-up phase, and again when they either solved or were shown the solution to their first task, it seems even less likely that they still believed they were not permitted to modify the materials when they began their second task. Yet we observed no difference in children's levels of success between their first and second tasks. We believe that these considerations make it unlikely that task pragmatics and/or misunderstanding about permission to modify the materials are adequate explanations of children's tool innovation difficulties. Nonetheless, it would be valuable for future work to include yet more explicit indications that the puppet or experimenter no longer needed the materials and that children were allowed to change the materials.

To gain a better understanding of what children were doing within each task, we analyzed the behaviors of both unsuccessful and successful participants. For unsuccessful participants, we focused on perseverative behavior. Although perseveration was rare in Experiment 1, Experiment 2 yielded much higher perseveration rates for the 4- and 5-year-olds, with these younger children perseverating significantly more than the older children. Because the levels of perseveration in each experiment are similar for the 6- and 7-year-olds, we suggest that the difference seen in the younger children is likely to be due to task differences between the two experiments rather than to differences between the two samples. In this regard, it is notable that in Experiment 1 all children received a warm-up exercise in which they manipulated the task materials, whereas this was excluded from Experiment 2 to make the materials appear to be more incidental to the overall task. It is possible that the warm-up exercise in Experiment 1 helped the younger children to avoid perseverative behaviors, perhaps by priming them to manipulate the materials given in the main task. However, despite this finding, it is clear from our results that children's tool innovation difficulties are not due merely to an inability to overcome such perseverative behavior. First, many children did not display perseverative behaviors yet were still not able to innovate tools. Second, for many of the children who succeeded, there was no apparent need to overcome perseveration on an initial unsuccessful solution because they immediately innovated successful tools. Nevertheless, although the current studies suggest that overcoming such perseveration is not the limiting step for tool innovation success, the data do suggest that it may be a necessary condition for success; if children initially use an unsuccessful tool and then fail to stop using it, they can never go on to succeed in innovating a tool.

Taken together, then, Experiments 1 and 2 suggest that children's difficulty with tool innovation might not derive from difficulty with switching between alternative tool innovation solutions or from difficulty in overcoming a bias to view the tool-making materials as having preestablished fixed functions. We can also rule out the possibility that difficulty arises from the need to overcome a tendency to perseverate with incorrect solutions. This raises the question of what other factors might lead to children's apparent lack of flexibility on tool innovation tasks. One possibility is that tool innovation, unlike many tasks examining the development of mental flexibility and executive function in young children (which are well-structured problems), is an intrinsically difficult ill-structured problem (Shallice & Burgess, 1991).

To see why it might be appropriate to view tool innovation as an intrinsically ill-structured task, it is useful to compare the tool innovation tasks with a well-structured tool task. In Beck and colleagues' (2011) Experiment 1, children were given the same goal of retrieving a bucket containing a sticker from a deep narrow container, but with the choice between a straight or hooked pipe cleaner. This is a well-structured task that has clear initial and goal states and clearly defined strategies for how to move between them, and on this task children performed very well from 4 years of age. Together with the evidence of children's success after the experimenter's demonstration of tool making, this clearly demonstrates that children can recognize the solution to the problem when they see it and can execute all of the relevant actions necessary to make and use the tool. What they find to be difficult is generating their own solution when it is not directly supplied.

A requirement to generate a solution that is not directly supplied by the task is the defining feature of ill-structured executive tasks. For example, in the Six Elements Test (Burgess et al., 1996), participants are presented with six tasks to complete and are asked to achieve as many points as possible by completing as many of the tasks as they can within a time limit and while following rules such as needing to attempt every task. Thus, the task explicitly supplies the starting conditions (the games and the rules) and the objective (maximizing points scored on the games), but it is ill-structured because participants must devise their own strategies for tackling the problem. Such problems undoubtedly require multiple executive processes (including memory, inhibition, and switching), but they do not seem to reduce simply to the sum of these components. It is possible to be impaired on ill-structured problems despite showing no impairment on standard well-structured tests of executive function (e.g., Shallice & Burgess, 1991; White et al., 2009). We suggest that children's difficulty with tool innovation may stem from the ill-structured nature of such problems. Although there is little evidence on the development of children's performance on ill-structured executive tasks, it is noteworthy that the ability to solve ill-structured tasks has been specifically associated with regions of medial prefrontal cortex (Brodmann area 10) that show protracted maturation throughout childhood and adolescence (Dumontheil, Burgess, & Blakemore, 2008). If children's difficulty with tool innovation is derived from a broader domain-general difficulty in solving ill-structured problems, it would be expected that individual differences in performance at tool innovation should be correlated with individual differences in performance on other ill-structured problems in nontool contexts, and this relationship should be independent of general intelligence and other executive functions such as inhibition and working memory.

Alternatively, it could be that children's difficulty in generating structured solutions for tool innovation problems lies with a lack of domain-specific knowledge about the mechanical properties of tool-making materials rather than with a domain-general problem with ill-structured tasks. If this were the explanation for children's difficulties with tool innovation, individual differences in successful innovation should correlate with other tasks that require knowledge of the mechanical properties of tools but do not require ill-structured problem solving. Moreover, such a correlation should persist even if children's performance on an ill-structured problem of another kind were partialled out. Future work would be necessary to distinguish between these possibilities.

Finally, whatever the detailed reason for children's difficulties, perhaps the most important conclusion from our experiments is the simple and robust finding that tool innovation is a difficult and late-developing ability. Even when children are excellent tool users and tool manufacturers, they fail to innovate simple tools. It is often noted that children are excellent social learners (e.g., Csibra and Gergely, 2009). Our findings highlight the importance of social learning in children's developing ability to use tools because "reinventing the wheel" for themselves is comparatively difficult. We might speculate that two factors were critical in the historical evolution of humans' tool-rich cultures. The ability to innovate tools is clearly vital for technological advancement, but it is equally important that the valuable products of this effortful process are preserved and passed on through social learning. Either of these abilities has the potential to be the limiting step in the development of tool-rich cultures. However, we venture that the capacity for cognitively demanding tool innovation, rather than tool use or tool manufacture, is what makes human tool culture stand out as uniquely complex.

Acknowledgment

We thank the teachers and children at The Meadows and Woodhouse Primary Schools for their help and participation.

References

Anderson, V. (1998). Assessing executive functions in children: Biological, psychological, and developmental considerations. *Neuropsychological Rehabilitation, 8*, 319–349.

Bania, A. E., Harris, S., Kinsley, H. R., & Boysen, S. T. (2009). Constructive and deconstructive tool modification by chimpanzees (*Pan troglodytes*). *Animal Cognition, 12*, 85–95.

Barr, R., & Wyss, N. (2008). Re-enactment of televised content by 2-year-olds: Toddlers use language learned from television to solve a difficult imitation problem. *Infant Behavior and Development, 31*, 696–703.

Beck, S. R., Apperly, I. A., Chappell, J. M., Guthrie, C., & Cutting, N. (2011). Making tools isn't child's play. *Cognition, 119*, 301–306.

Bird, C. D., & Emery, N. J. (2009). Insightful problem solving and creative tool modification by captive rooks. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 10370–10375.

Boesch, C., & Boesch, H. (1990). Tool use and tool making in wild chimpanzees. *Folia Primatologica, 54*, 86–99.

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology, 26*, 571–593.

Brown, A. L. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science, 14*, 107–133.

Burgess, P. W., Alderman, N., Emslie, H., Evans, J. J., Wilson, B. A., & Shallice, T. (1996). The simplified Six Element Test. In B. A. Wilson, N. Alderman, P. W. Burgess, H. Emslie, & J. J. Evans (Eds.), *Behavioural assessment of the dysexecutive syndrome*. Bury St. Edmunds, UK: Thames Valley Test.

Casler, K., & Kelemen, D. (2005). Young children's rapid learning about artifacts. *Developmental Science, 8*, 472–480.

Casler, K., Terziyan, T., & Greene, K. (2009). Toddlers view artifact function normatively. *Cognitive Development, 24*, 240–247.

Chevalier, N., & Blaye, A. (2009). Setting goals to switch between tasks: Effect of cue transparency on children's cognitive flexibility. *Developmental Psychology, 45*, 782–797.

Connolly, K., & Dalgleish, M. (1989). The emergence of a tool-using skill in infancy. *Developmental Psychology, 25*, 894–912.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*, 148–153.

Defeyter, M. A., & German, T. P. (2003). Acquiring an understanding of design: evidence from children's insight problem solving. *Cognition, 89*, 133–155.

Defeyter, M. A., Hearing, J., & German, T. C. (2009). A developmental dissociation between category and function judgments about novel artifacts. *Cognition, 110*, 260–264.

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 70–95). New York: Oxford University Press.

Dumontheil, I., Burgess, P. W., & Blakemore, S. J. (2008). Development of rostral prefrontal cortex and cognitive and behavioral disorders. *Developmental Medicine and Child Neurology, 50*, 168–181.

Espy, K. (1997). The shape school: Assessing executive function in preschool children. *Developmental Neuropsychology, 13*, 495–499.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*, 483–527.

Gibson, K. R., & Ingold, T. (1993). *Tools, language, and cognition in human evolution*. Cambridge, UK: Cambridge University Press.

Goel, V. (1995). *Sketches of thought*. Cambridge, MA: MIT Press.

Hayne, H., Herbert, H., & Simcock, G. (2003). Imitation from television by 24- and 30-month-olds. *Developmental Science, 6*, 254–261.

Herbert, J., & Hayne, H. (2000). Memory retrieval by 18- to 30-month-olds: Age-related changes in representational flexibility. *Developmental Psychology, 36*, 473–484.

Herrmann, E., Call, J., Hernàndez-Lloreda, M., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science, 317*, 1360–1366.

Hunt, G. R., & Gray, R. D. (2002). Species-wide manufacture of stick type tools by New Caledonian crows. *Emu, 102*, 349–353.

Kacelnik, A. (2009). Tools for thought or thoughts for tools? *Proceedings of the National Academy of Sciences of the United State of America, 106*, 10071–10072.

Luna, B., Thulborn, K. R., Munoz, D. P., Merriam, E. P., Garver, K. E., Minshew, N. J., et al (2001). Maturation of widely distributed brain function subserves cognitive development. *NeuroImage, 13*, 786–793.

McGuigan, N., & Whiten, A. (2009). Emulation and "overemulation" in the social learning of causally opaque versus causally transparent tool use by 23- and 30-month-olds. *Journal of Experimental Child Psychology, 104*, 367–381.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1423–1442.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*, 838–850.

Povinelli, D. J. (2000). *Folk physics for apes: The chimpanzee's theory of how the world works*. Oxford, UK: Oxford University Press.
Shallice, T., & Burgess, P. W. (1991). Higher-order cognitive impairments and frontal lobe lesions in man. In H. S. Levin, H. M. Eisenberg, & A. L. Benton (Eds.), *Frontal lobe function and dysfunction* (pp. 125–138). New York: Oxford University Press.
Visalberghi, E., Fragaszy, D. M., & Savage-Rumbaugh, S. (1995). Performance in a tool-using task by common chimpanzees (*Pan troglodytes*), bonobos (*Pan paniscus*), an orangutan (*Pongo pygmaeus*), and capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology, 109*, 52–60.
Want, S. C., & Harris, P. L. (2002). How do children ape? Applying concepts from the study of non-human primates to the developmental study of "imitation" in children. *Developmental Science, 5*, 1–13.
Weir, A. A. S., Chappell, J., & Kacelnik, A. (2002). Shaping of hooks in new Caledonian crows. *Science, 297*, 981.
White, S. J., Burgess, P. W., & Hill, E. L. (2009). Impairments on "open-ended" executive function tests in autism. *Autism Research, 2*, 138–147.
Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 445–469). Oxford, UK: Blackwell.