



Contents lists available at ScienceDirect

# Journal of Experimental Child Psychology

journal homepage: [www.elsevier.com/locate/jecp](http://www.elsevier.com/locate/jecp)



## Using a bad tool with good intention: Young children's imitation of adults' questionable choices

Cara DiYanni<sup>a</sup>, Deborah Kelemen<sup>b,\*</sup>

<sup>a</sup>Department of Psychology, Rider University, Lawrenceville, NJ 08648, USA

<sup>b</sup>Department of Psychology, Boston University, Boston, MA 02215, USA

### ARTICLE INFO

#### Article history:

Received 30 June 2007

Revised 23 April 2008

Available online 30 June 2008

#### Keywords:

Imitation

Artifacts

Tool use

Intention

Function

Design

Trust

Social influence

### ABSTRACT

We present three studies exploring 2- to 4-year-olds' imitation on witnessing a model whose questionable tool use choices suggested her untrustworthiness. In Study 1, children observed the model accidentally select a physically optimal tool for a task and then intentionally reject it for one that was functionally nonaffordant. When asked to perform the task for her, children at all ages ignored the model's intentional cues and selected the optimal tool. Study 2 found that when the model's nonaffordant tool choice was emphasized by claims about its design, 3-year-olds increased imitation. They also imitated, as did 2-year-olds, when the model selected a suboptimal rather than nonaffordant tool. The 4-year-olds consistently avoided imitation. Study 3 replicated these findings with new tools and participants. Additional measures indicated that knowledge about artifact design predicted children's tendency to ignore the model. These results shed light on developmental trends in the social and cognitive functions of imitation.

© 2008 Elsevier Inc. All rights reserved.

### Introduction

One of the most significant insights to emerge from studies of young children's object-directed imitation is that, from early on, children's imitative actions on objects are guided by the intentionality of modeled behaviors. For example, Meltzoff (1995) found that 18-month-olds were as likely to produce the intended goal of a model's failed intentional acts as the successful attempts, for example, failing or succeeding to pull apart the two ends of a dumbbell (see also Bellagamba & Tomasello, 1999; Johnson, Booth, & O'Hearn, 2001). Also, Carpenter, Akhtar, and Tomasello (1998) found that, given the

\* Corresponding author. Fax: +1 617 353 6933.

E-mail address: [dkelemen@bu.edu](mailto:dkelemen@bu.edu) (D. Kelemen).

opportunity to imitate either an intentional or accidental object-directed act, 14- and 18-month-olds privileged the intentional behavior. Thus, after witnessing an effect (e.g., a toy lighting up) that was potentially caused by either an accidental (“Whoops!”) or deliberate (“There!”) manipulation, children selectively reproduced the experimenter’s intentional action (see also Carpenter, Call, & Tomasello, 2002). Recent work has also shown that, in addition to influencing procedural imitation (how to operate an object), children’s imitative attention to intention also influences their tool choice and artifact categorization. For example, after seeing a model briefly intentionally use one of two distinctive but functionally equivalent tools to operate a light box, 3-year-olds enduringly returned to that tool as being exclusively *for* the box task but made no such rapid, exclusive function mappings if the demonstrated act was accidental (Phillips, Kelemen, Seston, & Casler, 2008).

In short, children’s imitative focus on others’ intentional actions is highly adaptive. In terms of its cognitive function, it yields a variety of generalizable, culturally pertinent knowledge that is particularly informative in relation to social objects such as artifacts. In terms of its social function, it facilitates identification, thereby heightening a sense of interpersonal connectedness and mutual understanding between child and model (Hobson & Meyer, 2006). The resulting social alignment also acts as a social lubricant (Užgiris, 1981; Užgiris, 1984; see also Chartrand & Bargh, 1999; Dijksterhuis & Bargh, 2001; Meltzoff & Moore, 1995; Nielsen, 2006; Tomasello, Carpenter, Call, Behne, & Moll, 2005).

However, although there are obvious cognitive and social benefits to children’s imitative attention to intention, it is also in children’s interest to exercise some caution and discernment when attending to others’ deliberate actions as a basis for guiding their own behavior. One reason is that certain features of agents’ intentional actions can be unworthy of imitative generalization because they are consequences of situational constraints and, therefore, are irrelevant to efficient goal achievement in other contexts. Indeed, findings already exist that young children are sensitive to such situational influences from as early as infancy. For example, Gergely, Bekkering, and Király (2002) found that 14-month-olds witnessing a model inefficiently operating a light box with her forehead rather than her hand were less likely to replicate the odd forehead action themselves if there was an obvious situational reason to discount the unusual means–end action (e.g., the model’s hands were occupied rather than free). Gergely and colleagues have argued that underlying this selective means–ends imitation behavior is a “naive theory of rational action” (Csibra & Gergely, 2003; Gergely & Csibra, 2005); by default, infants presume that agents always act in the most rational (e.g., efficient) way. Therefore, on seeing inexplicable, inefficient means–ends actions, particularly when accompanied by ostensive–communicative (“pedagogical”) cues, they imitate them on the basis that they must yield relevant information and must have been chosen for good reason (Gergely & Csibra, 2006; see also related work on children’s “underimitation” of causally irrelevant acts, e.g., Nielsen, 2006; Schulz, Hooppell, & Jenkins, 2008; Want & Harris, 2001; but see also Lyons, Young, & Keil, 2007; Nagell, Olguin, & Tomasello, 1993).

These results importantly demonstrate that imitative behavior involves an interpretive process and that children regulate their imitative tendencies by considering the situational constraints on a rational model’s choice of action. But there is also another reason why children should be discriminating before imitating somebody else’s behavior. Although it might be valid to assume that a social informant’s intentional actions generally occur for good reason, it might not be as advisable to assume that those actions always yield reliable cultural information. That is, although some inefficient actions are reasonable responses to contexts influenced by cultural conventions and, therefore, are worthy of imitative generalization (e.g., the now physically redundant act of bowing or shaking hands in addition to verbal greeting), other physically inefficient intentional actions, while also rational, might not be normative because they result from an individual agent’s ignorance, incompetence, tomfoolery, or even trickery. If there is any evidence to suggest that a social informant’s information might be fallible for the latter reasons, it would be beneficial for children to weigh that evidence before engaging in imitative generalization.

Recently, a line of work has emerged that indicates that, at least in relation to word learning, young children do indeed use various cues to evaluate the reliability of an agent’s information before incorporating it into their own repertoire. For example, Sabbagh and Baldwin (2001) found that both 3- and 4-year-olds are less likely to learn a word from a hesitant, behaviorally uncertain labeler than from

someone expressing no uncertainty and that 4-year-olds can also use explicit statements about ignorance as a basis for discounting someone's labeling information (see also Koenig & Harris, 2005). Jaswal (2006) found that 3- and 4-year-olds will treat artifact labeling information from someone with privileged insights into an object's identity (i.e., the designer) as more reliable than that from someone with less knowledge (i.e., the discoverer). Finally, Koenig, Clement, and Harris (2004) recently showed that 3- and 4-year-olds will track agents' prior records of accuracy in naming familiar objects when deciding which of two individuals to trust as a source of information about novel labels (see also Clement, Koenig, & Harris, 2004; Jaswal & Neely, 2006; Koenig & Harris, 2005).

But although this body of research importantly indicates that children can be discerning about the actions of linguistic informants, particularly when given direct statements about those individuals' states of knowledge or unambiguous evidence of inaccuracy, what remains unknown is how children's imitative behavior might be influenced by more implicit behavioral cues as to the potential unreliability of a model's actions. Specifically, what do children do when learning about unfamiliar objects if presented with a model who actively favors a dubious and costly choice of action? The current three experiments address this question. The degree of imprudence displayed by the model varied, but in each study preschoolers were given reason to question the reliability of a model's actions after witnessing her state a desired goal and then threaten efficient achievement of that goal by favoring the use of an inefficient rather than efficient tool for the intended task. More specifically, drawing on aspects of Carpenter and colleagues' (1998) method, preschoolers observed a model with a stated aim (e.g., cookie crushing) accidentally select and reject a physically "good" tool for the stated goal ("Oops!") and then intentionally select a physically "bad" nonoptimal or suboptimal tool ("There!") before getting interrupted prior to task completion and asking children to select between the tools to perform the task on her behalf.

In each trial of the studies, the issue of interest was which kind of information would guide children's own behavior. Would children's attention to intention, social motivations, and monitoring for potentially significant cultural-conventional information lead them to follow the model's intentional choice of a relatively bad tool? Alternatively, would they set aside interpersonal motivations, evaluate the model's social cues as unreliable based on their own physical knowledge, and select the good tool for the task? Importantly, then, these scenarios differ from most imitation studies, including those on "rational imitation" (Gergely et al., 2002; Schwier, van Maanen, Carpenter, & Tomasello, 2006; Tomasello & Carpenter, 2005; see also Killen & Uzgiris, 1981), in which the costliness of the modeled action never brings the reliability of the agent's information into question. For example, in Gergely and colleagues' (2002) study, children observed a model act to achieve a goal and, in fact, clearly successfully attain that goal, albeit via an unusual manner of action. In contrast, in the current studies, children witnessed someone who actively rejected an optimal choice of action in favor of one that, in addition to being more inefficient, was sufficiently threatening to her stated (unfulfilled) aim that it clearly raised questions about the trustworthiness of her social information.

Because one important view of imitation emphasizes its social interpersonal function, particularly in children from late toddlerhood onward (Nielsen, 2006; Uzgiris, 1981; Uzgiris, 1984), Study 1 was conducted to get a baseline sense of 2-, 3-, and 4-year-olds' tendency to align themselves with an adult's highly questionable and costly tool choice. In Study 1, the model rejected an extremely functionally affordant good tool in favor of an extremely nonaffordant bad one. In Study 2, we then presented children with two trials that manipulated the social or physical context of the model's actions. In the *design cue* trial, children once again saw the extremely contrasted Study 1 tools except that this time the model gave greater general social authority and emphasis to her inefficient choice by stating that the tool had been designed for the task. In a second *suboptimal tool* trial, the model gave the same social cues as in Study 1 but with new tools that made the imprudence of her choice of action less extreme; she rejected a physically good tool in favor of one that was physically suboptimal or simply "less adequate" rather than actively "bad." Finally, Study 3 explored the generalizability of the findings of the first two studies with a new set of participants and new tools. It also included additional independent measures of children's understanding of artifact design and sensitivity to agent reliability (a modified version of Koenig et al., 2004) to explore possible contributions of these capacities to children's decisions to imitate or not imitate.

## Study 1

### Method

#### Participants

The participants were 68 preschoolers recruited from local day care centers and play groups in the Boston area: 20 2-year-olds (11 girls and 9 boys, mean age = 30 months, range = 24–35), 23 3-year-olds (12 girls and 11 boys, mean age = 41 months, range = 36–47), and 25 4-year-olds (14 girls and 11 boys, mean age = 53 months, range = 48–59). Children were tested in a quiet place in their home or preschool or at our laboratory and were given a small gift for their participation.

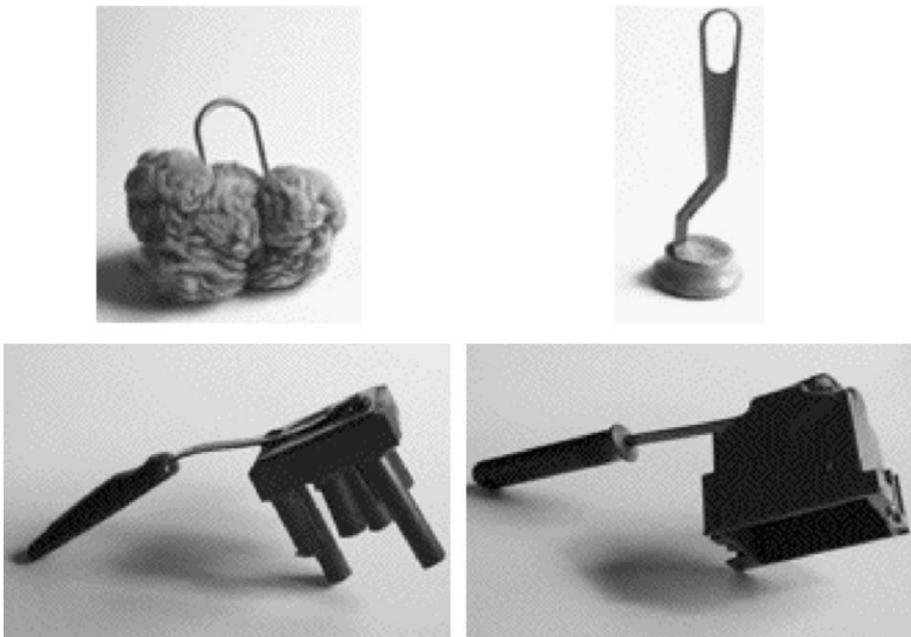
#### Materials

Two novel tools were created (see top panels of Fig. 1). The “good” functionally affordant tool was highly appropriate for the task of crushing cookies. It had a sturdy handle attached at a 90° angle to a flat wooden base. The “bad” nonaffordant tool did not afford the cookie-crushing task. It had five blue fluffy cotton pom-poms secured to a metal U-shaped handle. It could, however, achieve the cookie-crushing goal a little with a great deal of effort. The physical affordances were checked with 19 college undergraduates who judged which one they would use if they wanted to crush a cookie. Every adult selected the functionally affordant tool.

#### Design and procedure

The study involved three distinct segments in a fixed sequence: an initial exploration period, a demonstration, and a test. Each segment is described below and provides the framework for all trials in these studies.

*Initial exploration period.* Prior to the main procedure, all children participated in an exploration period designed to familiarize them with the physical properties of both tools and to see whether any child



**Fig. 1.** Top panels: Cookie-crushing tools for Study 1 and the design cue trial of Studies 2 and 3 (nonaffordant tool on left). Bottom panels: Stone-pulling tools for the suboptimal tool trial of Study 2 (suboptimal tool on left).

would spontaneously use either tool to crush an Oreo cookie that was saliently located close to both the tools and the child. After introducing herself, the experimenter began by offering to show each child some “neat things” that she had brought in a box. She then handed the pair of novel objects to the child, saying, “Look at these things my friend gave me for my birthday. Aren’t they neat?” No labels or functions were provided for either of the objects. Following Bauer and Mandler (1992), the duration of this exploration period was child controlled, usually lasting between 1 and 5 min before children lost interest. No child spontaneously attempted to use the tools to crush the cookie. The cookie was never explicitly pointed out, but if children asked why it was there (as many did), the experimenter responded that they would perhaps find out later.

*Demonstration and test.* A total of 44 children (13 2-year-olds, 13 3-year-olds, and 18 4-year-olds) were randomly assigned to an experimental condition. After the exploration period, each child and the experimenter moved over to a table in a different location. At this table, the experimenter pointed out another Oreo cookie sitting on a plate and asked whether the child would mind if she crushed the cookie to put on a pie that she would be baking later. During the move to the testing table, the experimenter had surreptitiously placed the two tools on a surface that was in full view of the child but out of the experimenter’s own direct line of sight because it lay slightly behind her while she was seated. The physical layout was organized so that during the test the experimenter could convincingly reach for and unintentionally grasp the “wrong” artifact because she could not fully see the location where she was reaching.

After introducing the cookie-crushing goal and getting settled at the table, the experimenter then proceeded with the “accident” part of the demonstration. With the focus of her attention diverted elsewhere (chatting and maintaining eye contact with the child), she absentmindedly reached behind herself, grabbed the functionally affordant tool, and distractedly made brief contact with the cookie. On performing this action, she then stopped, looked down to the tool in her hand with surprise, and exclaimed, “Oops! Silly me,” briefly reestablishing eye contact with the child during this utterance. She then put down the functionally affordant tool, looked at the nonaffordant tool, pointed to it, and said “There!” to convey her intention to use that object. Picking up the nonaffordant object, she then made brief contact with the cookie while smiling and nodding. Eye contact was again briefly reestablished with the child during these actions to “check in” and retain the child’s attention.

So that there was some hint of plausibility to the model’s selection of the nonaffordant tool, initial contact with the functionally affordant tool did in fact soften the cookie such that when the non-affordant tool made contact it produced a few crumbs. However, although both tools were physically associated with the cookie, neither one completed the crushing task. Instead, children witnessed an unfulfilled intention as the experimenter interrupted herself by recollecting aloud that she had forgotten to get something important from her box. After turning away to search in the box, the experimenter turned back to the child and said, “While I do that, maybe *you* can finish crushing up the cookie for me.” The experimenter then presented the child with the test question: “Here, take a look at these. Which one do you need to crush up the cookie?” If children did not want to crush the cookie (this was true of only a small number of children), the experimenter asked them to simply select the tool that they *would* use if they were going to crush the cookie. No child refused to make a tool selection.

Despite its awkwardness when described with words, visually this whole sequence of events unfolded quite naturally. Placement of the tools (i.e., which tool was held in the experimenter’s right hand and which one was held in her left hand as she offered the tools) was counterbalanced across children. After selection, as the children attempted to crush the cookie, the experimenter made encouraging comments (e.g., “Thank you,” “You’re a good helper”) but never expressed praise for the children’s selections, success, or lack of success in performing the task.

Although prior research indicates that young children are sensitive to the kinds of structure–function relationships involved in the current work (e.g., Brown, 1990; Kemler Kemler Kemler Kemler Nelson, 1999; Kemler Nelson, Frankenfield, Morris, & Blair, 2000; Kemler Nelson, Russell, Duke, & Jones, 2000; McCarrell & Callanan, 1995), as a precaution, 25 children (7 2-year-olds, 10 3-year-olds, and 8 4-year-olds) participated in a control condition to check whether they recognized that the functionally affordant tool was a more appropriate cookie crusher than the nonaffordant tool and to rule out any

baseline preferences for the nonaffordant tool. The procedure was the same as for the experimental group except that in this version the experimenter expressed an intention to use both tools during the demonstration. Thus, rather than accidentally reaching for the first tool and then rejecting it, the experimenter reached behind herself for the first tool, saying “There!” as she made contact with the cookie. She then stopped and retrieved the second tool, again saying “There!” as she used it. The rest of the procedure was identical to the experimental condition.

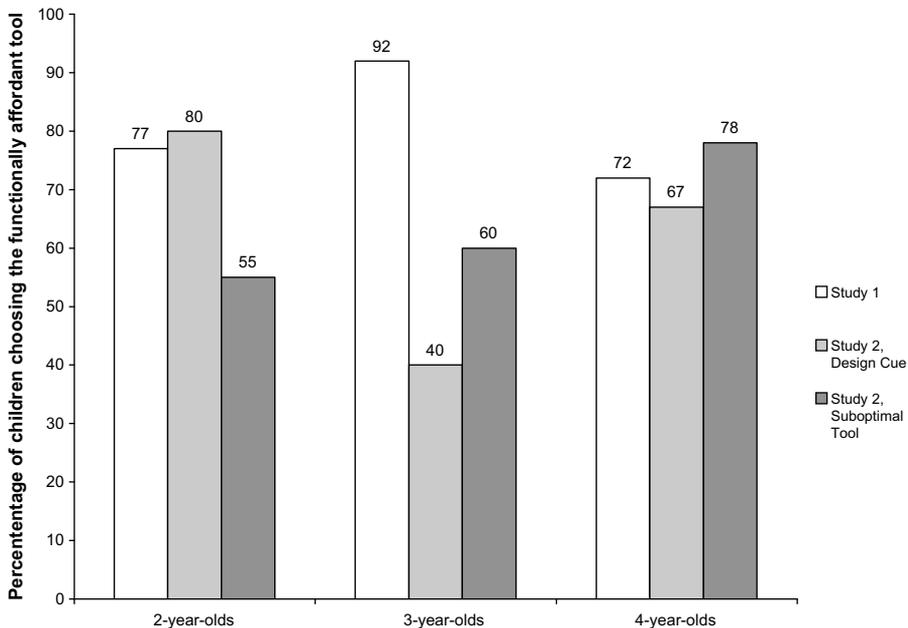
## Results

For each test, children were given a score of 1 if they chose the functionally affordant tool and a score of 0 if they chose the nonaffordant tool. Results are presented as the percentage of children selecting the functionally affordant object. Preliminary analyses indicated that children in the control condition clearly recognized the optimality of the functionally affordant tool. It was selected by 71% of 2-year-olds, 80% of 3-year-olds, and 100% of 4-year-olds. Given this high rate of selection in the control condition, remaining analyses focused on the performance of children in the experimental condition.

As Fig. 2 indicates, the results are straightforward. Despite the experimenter’s conspicuous choice of the nonaffordant tool, the majority of children at all ages disregarded her intentional choice and focused instead on their own physical intuitions. Fisher exact tests revealed no differences between age groups in the number of children selecting the functionally affordant tool. Preference for the functionally affordant tool was also above chance for each age group: 2-year-olds,  $\chi^2 = 3.77$ ,  $df = 1$ ,  $p = .05$ ; 3-year-olds,  $\chi^2 = 9.30$ ,  $df = 1$ ,  $p < .01$ ; 4-year-olds,  $\chi^2 = 3.60$ ,  $df = 1$ ,  $p = .059$ .

## Discussion

In Study 1, children were presented with an unambiguous case of a novel agent making an imprudent choice of a nonaffordant tool that threatened achievement of her stated goal. This baseline case



**Fig. 2.** Comparison of the percentages of children selecting the functionally affordant tool in Study 1 and those of children selecting the design cue and suboptimal tool trials of Study 2.

yielded clear findings. When 2- to 4-year-olds witness social cues from an agent whose actions are extremely questionable, they follow their own physical intuitions about the best course of action. In consequence, young children's sensitivity to intentional acts and desire for the interpersonal benefits of shared experience are not such that they are indiscriminate imitators of adults' artifact choices. These results are especially interesting in relation to the 2-year-old age group, whose slavish and often unnecessary fidelity in means–end imitation tasks has sometimes led them to be described as “habitual” imitators whose imitation is strongly socially motivated (Nagell et al., 1993; Nielsen, 2006). The current results indicate that, when 2-year-olds are presented with a clear goal to achieve and with a model whose instrumental acts are questionable and costly to goal achievement, interpersonal motivations such as establishing mutuality are as likely to fall by the wayside in this age group as in other age groups.

Study 1 represented an extremely strong test of children's drive to socially align with a model's intentional choice. Indeed, given the risk to goal achievement, children's lack of imitation is potentially unsurprising. However, although the instrumental context and goal threat in itself offers a clear explanation of children's tendency to ignore the model's social cues, there was also another possible explanation of the findings: We had presented children with such a salient and compelling fit between the physical structure of the functionally affordant tool and the cookie-crushing goal that perhaps children had not fully encoded the model's intentional cues at all. Some aspects of children's responses in Study 1 suggested that this was unlikely, particularly among 4-year-olds. For example, they often looked amused at the model's choice in the experimental condition. Notably, they also selected the functionally affordant tool less often in that condition compared with the control, presumably because they noted the contrasting experimental condition cues. Nevertheless, in Study 2, we decided to increase the salience of the model's intentions. To this end, we introduced a “design cue.”

Research on the “design stance”—the tendency to rationalize an artifact's structure and function by reference to its intended design—suggests that 3- and 4-year-olds, and possibly even 2-year-olds, are sensitive to the way in which intended design constrains the structure of an artifact (Asher & Kemler Nelson, 2008; Kemler Nelson, Holt, & Egan, 2004; see also Kemler Nelson, Herron, & Morris, 2002). Children with this sensitivity would have been particularly motivated to focus on the functionally affordant tool as the one designed for cookie crushing. In consequence, the design cue that we included competed with this possible design inference by having the model unambiguously inform children that her choice of the nonaffordant cookie-crushing tool was not simply personal caprice but rather because that tool was the one “made for” cookie crushing. This cue was selected because evidence suggests that children in the 3- to 5-year-old age range weigh this kind of explicit linguistic design information when judging the identity and function of an artifact (Jaswal, 2006; Kelemen, 1999; Kelemen, 2001; for a review, see Kelemen & Carey, *in press*; see also Diesendruck, Markson, & Bloom, 2003; Gelman & Bloom, 2000; but see Defeyter & German, 2003; German & Johnson, 2002; Matan & Carey, 2001). It is important to note, however, that even if children lacked a design stance (or an understanding of its connection to the “made for” phrase), the phrasing of the “made for” cue (described below) implied a normative standard that transcended the experimenter's personal preference while also strongly increasing the salience of the model's intentional decision to select a nonaffordant tool.

The Study 1 baseline results also raised another obvious follow-up question: In the context of tool use, how reasonable must a model's questionable choice be for children to accept it as worthy of imitation? As noted earlier, conventions are not always the most efficient, beneficial, or transparently motivated behavioral options because they originate for various reasons, including historical and economic ones that can become redundant over time. For example, perhaps shaking hands was once a useful precaution when issuing a verbal greeting but nowadays doing so is unlikely to reveal a concealed weapon, and religious dietary restrictions and food preparation methods persist to this day despite the elimination of many of the public health hazards that may have originally motivated them. In consequence, enculturation to a social group often requires children to tolerate or reproduce actions that are superficially relatively inexplicable or inefficient. In Study 1, the motivation for the model's choice was sufficiently opaque and inconsistent with her stated goal that it clearly brought the reliability of her actions into question. In Study 2, therefore, we wanted to explore what preschoolers would do when presented with a choice that, although relatively inefficient, could not be as readily

dismissed as entirely unreasonable because it could achieve the goal with a little extra effort. Alongside the *design cue* trial, therefore, we also included a *suboptimal tool* trial in which the experimenter again rejected a functionally affordant tool, but rather than turning to an object that was entirely non-affordant and inefficient, she instead chose one that was simply suboptimal for efficient achievement of a stated goal. In the context of real-world artifact use, this trial also had the additional benefit of ecological validity given that the nature of artifact evolution and habitual conventionalized behavior is such that during certain periods suboptimal versions of artifacts coexist and continue to be used in concert with more refined and efficient models. For example, people continued to use ink cartridge fountain pens long after roller ball pens were freely available, corks are still widely used on wine bottles, and the second author is probably not alone in persistently using a traditional “bar-and-screw” corkscrew rather than the new and improved “rabbit ears” version that lies in a drawer nearby (for a discussion of artifact evolution, see Petroski, 1992).

Finally, because the same experimenter engaged in both the extremely questionable choice of a non-affordant tool in the design cue trial and the more reasonable choice of a suboptimal tool in the suboptimal tool trial, counterbalancing of trial order allowed us to explore whether children made any dispositional attributions of reliability to the model. In particular, we were interested in seeing whether children would react differently to the model’s choice if they saw the highly costly selection in the design cue trial first rather than seeing the mildly costly selection of the suboptimal tool trial first.

## Study 2

### Method

#### Participants

The participants were 58 2-, 3-, and 4-year-olds residing in the Boston area. None of the children who had participated in Study 1 participated in Study 2. There were 20 2-year-olds (14 girls and 6 boys, mean age = 29 months, range = 22–35), 20 3-year-olds (12 girls and 8 boys, mean age = 43 months, range = 37–47), and 18 4-year-olds (11 girls and 7 boys, mean age = 54 months, range = 49–59). Children were tested in a quiet place in their home or preschool or at our laboratory. They were given a small gift for their participation.

#### Materials and procedure: Design cue trial

The design cue trial employed the functionally affordant and nonaffordant tools from Study 1 (see top panels of Fig. 1) and the same experimental procedure and design except that during the “accident” phase the experimenter changed the content of her intentional cue. After absentmindedly grabbing the functionally affordant cookie crusher and distractedly making brief contact with the cookie, she stopped, looked down at the tool in surprise, clapped her hand to her mouth, and exclaimed “Oops! Silly me” while briefly reestablishing eye contact with the child. In contrast to Study 1, however, she then explicitly turned to the child and said, “You know, [child’s name], there’s one here that’s made for crushing up cookies. I wonder where it is?” Looking around the room, she then spotted the nonaffordant fluffy tool on the table next to where the functionally affordant tool had originally been, pointed to it with a smile and a nod, picked it up, and said to the child, “Here it is! Here’s the one that’s made for crushing up cookies!” After issuing these declarative statements, she used the nonaffordant tool to make brief contact with the cookie while smiling and nodding, reestablishing eye contact with the child during these actions. The remainder of the procedure was exactly the same as in Study 1. The initial exploration period was also the same as in Study 1 except that children were simultaneously also exposed to the suboptimal tool set described below. Because the results from Study 1 clearly indicated that children of all ages recognized the differential functional affordances of the cookie-crushing tools when intentional and physical cues did not conflict, Study 2 did not include a control group.

#### Materials and procedure: Suboptimal tool trial

A functionally affordant tool and less affordant suboptimal tool were created for the task of pulling some “special stones” that were out of reach across a table. As the bottom panels of Fig. 1 show, whereas the hollow end of the functionally affordant tool could easily fit over a pile of clear marbles

and transport them toward the experimenter, the pegged end of the suboptimal tool made it a sufficient but less optimally efficient tool for the same task because many of the marbles escaped between the “teeth.” Therefore, it took more swipes to transport the stones with the suboptimal tool than with the functionally affordant tool.

Prior to Study 2, the differential physical affordances of the functionally affordant and suboptimal objects as stone pullers were checked with a group of 21 children, composed of 6 2-year-olds, 6 3-year-olds, and 9 4-year-olds, along with 18 college undergraduate adults. Children were simultaneously shown both tools and a pile of stones at a distance and asked to decide which tool would be the best one to use for the task of pulling the stones across the table. All 18 of the adults (100%) and 16 of the 21 children (76%) (5 of 6 2-year-olds, 5 of 6 3-year-olds, and 6 of 9 4-year-olds) selected the functionally affordant tool as the “better,” more efficient tool for the task. In counterbalanced order, children were also asked which one of the tools they liked better. No significant preferences were found for the functionally affordant tool. Because children clearly perceived the functional affordances of the tools in relation to the stone-pulling tasks, a control group was not included for the suboptimal tool trial.

The design and procedure were very similar to those used in Study 1. During the initial exploration period, the experimenter introduced the marbles as special stones that her roommate had brought back for her from a trip to the beach. No child spontaneously used either the stone-pulling tools or the cookie crushers for either of their respective functions during the initial exploration period despite the presence of the marbles and a cookie. After the initial exploration period was complete, the child and the experimenter relocated to a table for the demonstration and test.

To introduce the suboptimal tool task, the experimenter told the child that she needed to obtain the pile of special stones across the table from her but that because they were out of her reach, she was going to use something to help her pull them over. The experimenter then proceeded to absentmindedly reach for, begin to use, and then reject the functionally affordant tool. After pulling the stones 1 or 2 inches with the functionally affordant tool, she looked down, realized her “mistake,” and exclaimed “Oops! Silly me,” briefly reestablishing eye contact with the child during the utterance. She then turned to search for the suboptimal tool, pointed to it, and said “There!” While briefly using the suboptimal object to begin pulling some stones across the table, she smiled and nodded, also reestablishing eye contact with the child during these actions. These social cues, therefore, were equivalent to the original cues used in Study 1 rather than the enhanced ones of the design cue trial because in this more ambiguous context involving two workable tools, it seemed likely that children would be more attentive to potentially disambiguating social cues. Having produced these cues, the model then interrupted herself and stated that she had forgotten to get something else out of her box of things. The remainder of the procedure was then identical to that of Study 1. Half of the children in Study 2 participated in the design cue trial first, and the other half participated in the suboptimal tool trial first. Piloting indicated that despite its improbability from an adult perspective, young children’s task engagement was not affected by an individual who seemed positively disposed to a physically inadequate tool on more than one occasion.

## Results

In each trial, children scored 1 if they chose the functionally affordant tool and scored 0 if they selected the alternative tool. Results are expressed in terms of the percentage of children who selected the functionally affordant object despite it being rejected by the experimenter. Given that the model’s selection of an entirely nonaffordant tool in the design cue trial gave children a particularly good reason to enduringly code her as an unreliable social informant, a preliminary 2 (Trial Order)  $\times$  3 (Age) analysis of variance (ANOVA) explored whether children’s tendency to select the functionally affordant tool across both trials (0–2 range) differed as a function of seeing the design cue trial first rather than second. The analysis found no effect of trial order or age.

### Design cue trial

As Fig. 2 indicates, despite the introduction of information about the intended function of the non-affordant tool, 2-year-olds continued to perform as in Study 1, selecting the functionally affordant tool

at above chance levels,  $\chi^2 = 7.20$ ,  $df = 1$ ,  $p < .01$ . In contrast, social cues regarding the nonaffordant tool's intended function had a substantial impact on 3-year-olds' performance. Whereas 92% of 3-year-olds selected the functionally affordant tool in Study 1, this figure dropped to 40% in the design cue trial. This level of performance was statistically different from that in Study 1 and, in the only age difference found in Study 2, was also different from that of the 2-year-olds in Study 2, both Fisher exact tests,  $p < .05$ . Finally, despite prior research indicating that 4-year-olds weigh explicit information about an artifact's intended design in categorization judgments (Kelemen, 1999; Kelemen, 2001; see also Diesendruck et al., 2003; but see Defeyter & German, 2003; German & Johnson, 2002; Matan & Carey, 2001), 4-year-olds did not show as significant a strategy shift as did 3-year-olds with the introduction of design information into the intentional cue. Although the proportion of 4-year-olds selecting the functionally affordant tool was 67%—a figure that approached but was not above chance as it had been in Study 1—it was not significantly different from the number of 4-year-olds choosing the functionally affordant tool in Study 1 (72%).

#### *Suboptimal tool trial*

As Fig. 2 also indicates, 2- and 3-year-olds showed a different pattern of tool use from Study 1 after observing an adult give the same social cues but more reasonably favor a suboptimal tool rather than a nonaffordant tool. In Study 1, all age groups favored the functionally affordant tool, but in the suboptimal tool trial, 2- and 3-year-olds' preference fell to chance levels (55% of 2-year-olds and 60% of 3-year-olds chose the functionally affordant tool), with 3-year-olds' performance differing from that in Study 1 (92%), Fisher exact test,  $p = .056$ . In contrast to the performance of their younger peers, however, 4-year-olds continued to demonstrate the pattern found in all other trials. Despite the suboptimal tool's workable nature, 78% of 4-year-olds rejected the model's intentional choice and remained biased to the functionally affordant tool at a level above that expected by chance,  $\chi^2 = 7.20$ ,  $df = 1$ ,  $p < .01$ . Fisher exact tests revealed no age differences.

#### *Discussion*

Prior research indicates that, from as early as 2 years of age, children imitate the tool choice of a model who intentionally uses one functionally affordant tool rather than another to achieve a goal (Casler & Kelemen, 2005; Casler & Kelemen, 2007). The results of Studies 1 and 2 indicate, however, that preschoolers are not indiscriminate imitators of others' intentional tool use and show selectivity about who and what they copy, although their interpersonal motivation and susceptibility to social influence show subtly different patterns across the 2- to 4-year age period.

To recap, departing from characterizations of 2-year-olds as socially motivated habitual means-end imitators, most 2-year-olds in this more costly, goal-focused context ignored the model's intentional choice in both Study 1 and the more salient design cue trial. The only occasion when 2-year-olds showed any willingness to follow the model's preference was in the suboptimal tool trial, where approximately half of the children chose the suboptimal tool. In contrast, 3-year-olds showed far greater interpersonal alignment and susceptibility to social influence. Whereas 92% of 3-year-olds selected the functionally affordant tool that the model rejected in Study 1, this figure fell to 60% in the suboptimal tool trial when her questionable choice still permitted partial goal satisfaction and to 40% when the social informant engaged in the same choice but saliently invoked both design information and a normative behavioral standard in the design cue trial. Finally, 4-year-olds demonstrated a pattern that departed from that of both other age groups. Regardless of the nature and salience of the social cues or the degree of discontinuity between the agent's stated goal and her choice of tool, 4-year-olds in both Study 1 and Study 2 tended to ignore the model's information and select the functionally affordant tool—the better tool for the job. The inclusion of the design cue slightly decreased 4-year-olds' tendency to select the functionally affordant tool, but even with this decrease, the percentage of children ignoring the model's social cues in all trials across the two studies averaged 72%.

What could explain these developmental differences? With regard to the youngest children, the most straightforward interpretation is that, in a "costly choice" situation, 2-year-olds value goal achievement over social alignment. That is, 2-year-olds' drive to satisfy a goal is such that, irrespective of salience, they discount social information that is inconsistent with their physical knowledge and

that places a goal at terminal risk (as in Study 1 and the design cue trial). Therefore, they view social cues as personally relevant only when they provide information about an obviously good way of fulfilling a task (e.g., Casler & Kelemen, 2005) or a somewhat efficient, albeit suboptimal, way (e.g., Study 2, suboptimal tool trial). The finding that toddlers reject information from a social informant when it unambiguously conflicts with their own knowledge is consistent with research in the lexical domain. For example, Pea (1982) found that 18-month-olds will say “no” to an individual who mislabels a familiar artifact, exclaiming, for instance, “That’s a ball” in relation to a car (see also Koenig & Echols, 2003). Furthermore, the finding that the design cue had no extra impact on 2-year-olds’ tendencies in a context without any physical ambiguity may also be unsurprising given that, to date, there is no evidence that 2-year-olds assign meaning to explicit linguistic information about design. Whether this is simply because they are unfamiliar with linguistic cues such as “made for” or because of deeper representational limitations is unclear, although currently only one study (Kemler Nelson et al., 2004) is suggestive of any aspect of the design stance in children this young.

What shifted between 2 and 3 years of age such that the design cue led 60% of 3-year-olds to follow the model’s choice of a nonaffordant tool that they rejected in Study 1? One possibility is that children increasingly weighed the explicit “made for” cue in light of an emerging tendency to adopt the design stance (e.g., Asher & Kemler Nelson, 2008; Jaswal, 2006). Another possibility is that 3-year-olds are so increasingly concerned about normative behavior and shared experience that, even absent an understanding of design, the emphatic conventional standard implied by the “made for” phrase had an influence on them. This notion that a more general social motivation influences 3-year-olds is also consistent with the finding that 40% of 3-year-olds also followed the model’s preferred inefficient tool in the suboptimal tool trial despite the fact that no explicit design cue was associated with it and the fact that structure–function design considerations favored the functionally affordant tool.

Finally, what accounts for the fact that the 4-year-olds did not tend to follow the model’s social cues in Study 1, the design cue trial, or the suboptimal tool trial? Certainly one possibility is that they were displaying a more extreme version of 2-year-olds’ focus on goal achievement rather than intentional choices. But another possibility is that an increasingly robust understanding of artifact design and differential agent trustworthiness influenced this pattern of behavior. Around 3 to 4 years of age, children not only discriminate reliable sources of information from unreliable ones, but they increasingly also apply these evaluations when deciding what novel testimony to trust (Jaswal, 2006; Koenig et al., 2004). Furthermore, during the later preschool period, children’s construction of the design stance is such that they weigh design information and clearly recognize shape as a cue to design intent (e.g., Asher & Kemler Nelson, 2008; Kemler Nelson et al., 2002). The 4-year-olds’ lack of credulity for the model’s choices, therefore, could have occurred because children knew that the nonaffordant or suboptimal tool was unlikely to have been designed for the task at hand and, in the context of understanding this, coded the model’s action as unreliable and ignored her social cues. An additional reason to give weight to this explanation is that some of the older participants revealed their suspicions about the experimenter’s choices by laughing, looking confused, explicitly making an independent decision (e.g., “I need *that* one!” while pointing at the functionally affordant tool), or asking what the nonaffordant tool was *for* after selecting the functionally affordant tool.

To explore the replicability of the results of Studies 1 and 2 and to further understand the developmental trends, we conducted Study 3 in which a new set of preschoolers took part in new versions of the trials of Studies 1 and 2. Children’s performance was also explored in relation to assessments of sensitivity to artifact design information, agent reliability, and theory of mind.

### Study 3

#### *Method*

#### *Participants*

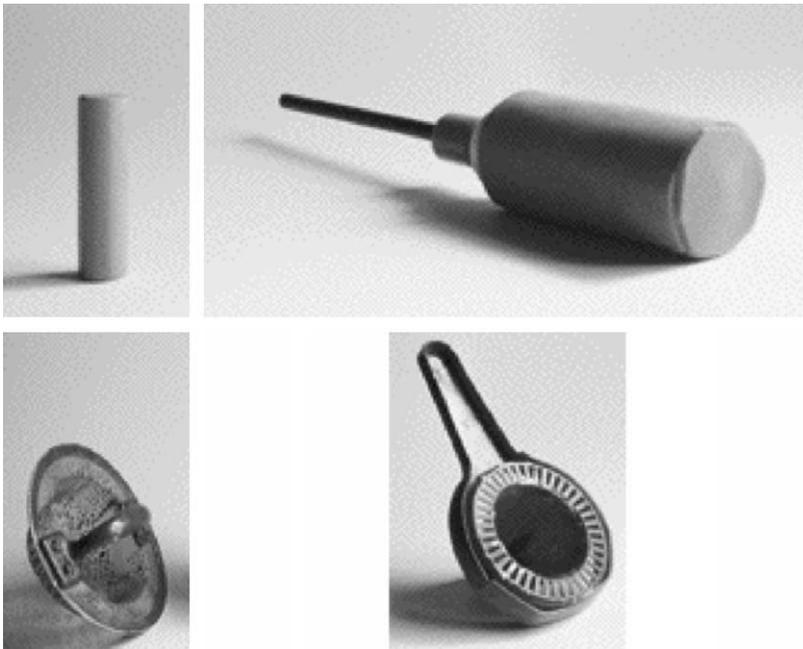
The participants were 60 preschoolers from the Boston area, none of whom had taken part in Study 1 or Study 2. The age groups in this study were more clearly differentiated than those in Studies 1 and 2. There were 30 3.5-year-olds (17 girls and 13 boys, mean age = 42 months, range = 38–46) and 30 4.5-year-olds (19 girls and 11 boys, mean age = 54 months, range = 50–58). Children were tested in

a laboratory environment or in a quiet area of their home or day care center, and they were given a small gift for their participation. It was initially our intention to also include a group of 2-year-olds in Study 3. However, piloting rapidly revealed that 2-year-olds were unable to manage the battery of tasks. In particular, many of them were unable to follow or complete the agent reliability, design stance, and theory of mind assessments. Because 2-year-olds' behavior across Studies 1 and 2 had a far more obvious interpretation than that of their older peers, a sample of 2-year-olds was not included in this study.

#### *Materials and procedures*

The study consisted of six measures. Measure 1 (Study 1) and Measure 3 (suboptimal tool) were replications of the suboptimal tool trial of Studies 1 and 2 but employed new tools and new tasks. Measure 2 (design cue) was identical to the design cue trial of Study 2. The new tools for Measures 1 and 3 are depicted in Fig. 3.

Measure 4 was an assessment of agent reliability understanding based on the design of Koenig and colleagues (2004). Measure 5, the design stance assessment, was a new procedure inspired by Markson (2001) and Diesendruck and colleagues (2003). Measure 6 was a theory of mind assessment using line-drawn versions of the deceptive containers task (e.g., Gopnik & Astington, 1988; Hogrefe, Wimmer, & Perner, 1986; Hughes & Cutting, 1999; Perner, Leekam, & Wimmer, 1987). A primary reason for including this latter measure was to explore a relevant side question concerning the relationship between children's representational theory of mind abilities and their understanding of agent reliability; construing an agent as unreliable presumably rests on recognizing that there can be disparities between what an agent believes and what is actually true (in the case of ignorance) or between what someone believes and what someone says (in the case of trickery). Therefore, we were interested in knowing more about the relationship between understanding agent reliability and understanding representational mental states.



**Fig. 3.** Top panels: Launching tools for Study 3, Measure 1 (nonaffordant tool on left). Bottom panels: Sand-moving tools for Study 3, Measure 3 (suboptimal tool on left).

### *Measure 1: Study 1 baseline*

In contrast to the cookie-crushing goal of Study 1, the Study 3 goal was to release a Styrofoam ball stuck inside an opaque tube. The ball's position within the tube's narrow diameter made it impossible for children to release the ball with their hands. As the top panel of Fig. 3 shows, whereas the long and sturdy functionally affordant tool was a highly efficient ball launcher, the nonaffordant tool was highly inefficient because it was too short to reach the ball. Prior to the study, children's ability to distinguish between the affordances of the tools was checked with a separate group of 11 3-year-olds and 8 4-year-olds. Children were shown all of the materials and asked which one of the two tools would be the best one to use for shooting the ball out of the tube. Of the 19 children, 18 (95%) selected the functionally affordant tool. In counterbalanced order with the question about affordances, children were also asked which tool they "liked best." No significant preferences were found for either tool. The procedure for Measure 1 was otherwise identical to the experimental procedure of Study 1.

### *Measure 2: Design cue*

The materials and procedure were identical to those of the design cue trial of Study 2.

### *Measure 3: Suboptimal tool*

In contrast to Study 2 stone pulling, the Measure 3 task was to transfer sand from one bowl into another bowl so as to uncover a toy frog that was hidden in the sand. As the bottom panels of Fig. 3 indicate, the functionally affordant tool was an efficient "sand mover," whereas the suboptimal tool was adequate but less efficient because small holes in its "bowl" allowed sand particles to drain. Its handle also impeded the scooping process. The procedure was otherwise nearly identical to that of the Study 2 suboptimal tool trial. Children's ability to distinguish the affordances in the tool set was checked with the same 19 children who judged the Measure 1 tools. Children were shown all of the materials and told that the goal was to scoop sand from one bowl to another so as to reveal the hidden frog. Of the 19 children, 15 (79%) selected the functionally affordant tool as the one that would be the best to use for scooping sand. In counterbalanced order, children were also asked which tool they "liked best," and no significant preferences were found for either tool.

### *Measure 4: Assessment of agent reliability understanding*

An agent reliability assessment video was created for children to watch on a computer. The procedure was based on work by Koenig and colleagues (2004) (see also Clement et al., 2004; Koenig & Harris, 2005) except that our test questions focused on the identification of novel artifact functions rather than novel object names. The video had two phases.

*Familiarization.* Children watched two agents, "Mrs. Red" and "Mrs. Blue," responding to six questions posed by an interviewer about a series of familiar objects. One agent consistently answered accurately, and the other one consistently answered inaccurately. Two different versions of the video were created, one in which Mrs. Red was the accurate informant and one in which Mrs. Blue consistently provided the correct information, and the video presentation was counterbalanced across children. In two "naming" trials, the agents were shown a familiar object (e.g., a train) and asked to name it. In two "selection" questions, the agents were presented with two familiar objects (e.g., pretzels and a jar of peanut butter) and asked to point out a requested object. In two "function" questions, the agents were asked what a familiar artifact (e.g., a toothbrush) was for. After the third and sixth questions, children were checked for their memory of the agents' answers and recognition of who was correct. Feedback was given if the children answered incorrectly at the initial check, but less than 10% of children required this. The criterion was set that if children did not pass the second postfeedback check, they would be eliminated from the final sample. All 60 children in the final sample passed the second postfeedback check.

*Test.* Children were next asked six test questions. In two "prediction" questions, they were shown a familiar object (e.g., a cup) and asked what function they believed each agent would provide for the objects. Two answers were required for each prediction question (e.g., "What would Mrs. Red say this is for?" and "What would Mrs. Blue say this is for?"), with 1 point being assigned only if both answers

were correct. For example, for the cup item, children received 1 point if they indicated that the reliable agent would say something related to the drinking function and that the unreliable agent would say something other than that. If children did not respond, seemed confused, or really did not know what kind of response the unreliable agent might give, they were prompted, “Would she say it is for drinking or for something else?” Children who responded “something else” to this question scored 0.5 rather than 1 point. In two “ask” questions, children were shown a novel object (e.g., a wooden wreath-like object with protruding prongs) and asked which agent they should approach to find out what it was really for. In two “believe” questions, posed subsequently to each “ask” question, the two agents stated equally plausible functions for a novel object (e.g., one agent claimed that the wreath-like object was for raking in the garden, and the other one said that it was for combing hair), and children judged who they believed. Two problems were found with a believe question involving a spiky soap dish item. The functions given for the item were “scratching your back” and “taking fuzzies off your clothes.” The first problem was that the item was not novel to a number of the children, who explicitly identified it as a soap dish. The second issue, possibly related to the first one, was that preliminary analyses revealed that 3.5-year-olds preferred the back-scratching function regardless of who offered the function. This bias was a concern because the point of this particular assessment was to examine children’s basic ability to track and productively use information about agent reliability rather than to test the limits of children’s tendency to take it into consideration at all. Because both of these issues raised problems for interpretation, this question was removed from analyses and the scoring range became 0 to 5.

#### *Measure 5: Design stance assessment*

Over three trials, children were presented with two physically dissimilar tools that were equally affordant for performing a task (ringing a bell in a cage, sifting sugar, or spreading glue). In each trial, the experimenter identified one of the tools (counterbalanced across children) as designed for the relevant task (e.g., “This thing was made for ringing bells”). For the other tool, she claimed ignorance but acknowledged that it might also be capable of performing the function (e.g., “I don’t know what this one is for, but maybe it can do it too”). Each tool was demonstrated after the experimenter described it so that the children could clearly see that both objects were, indeed, equally affordant for the task at hand. Children then received a reminder statement and two test questions. In a “conventionality” question, they were asked which tool a third party (e.g., a parent) needed to complete the task. In a “generalization” question, children were asked which of the tools they needed to complete another version of the task (e.g., to operate a bell in a slightly different cage). Scoring ranged from 0 to 6.

#### *Measure 6: Theory of mind assessment*

Children were shown line drawing versions of four deceptive container stories (e.g., the “Smarties” task) in which the contents of a container are not what would be expected based on the container’s appearance (e.g., Gopnik & Astington, 1988; Hogrefe et al., 1986; Hughes & Cutting, 1999; Perner et al., 1987). Children were asked three comprehension questions to monitor their ability to follow the events in each story plus a false belief test question (“What will [absent party] think is inside the box?”) and a representational change test question (“What did [actor] think was inside the box when she first picked it up?”). Children received credit for either of the test questions only if they correctly answered all three comprehension questions for that story. Children could then score 2 possible points for each story, creating a scale ranging from 0 to 8 across the four stories.

#### *Task order*

Measures were presented in two fixed orders designed to avoid boredom by minimizing consecutive presentation of trials from the same task as well as to diminish—and, in some cases, to specifically explore—the influence of earlier tasks on performance in subsequent tasks. Thus, after the initial exploration period, children received Measure 2 (design cue) and Measure 3 (suboptimal) in counterbalanced order, exactly as they did in Study 2. This allowed us to again explore whether children’s performance differed as a function of introduction to an agent who initially engaged in either a highly costly (design cue) or mildly costly (suboptimal) act. The remaining tasks occurred in the following order, which kept measure order constant to minimize experimenter error given the large number

of tasks: first trial of the design stance assessment (Measure 5); two theory of mind stories (counter-balanced order) from Measure 6, second trial of design stance assessment; Measure 1 (Study 1); two theory of mind stories (counterbalanced order); third trial of design stance assessment; and agent reliability assessment (Measure 4), which was always kept last so that we did not actively prime children to think about trustworthiness in any of the imitation tasks. Testing took approximately 15 to 20 min. One concern was that, whereas children had been unperturbed by an agent making two accidental choices in Study 2, they were now witnessing an experimenter make three accidental choices and interrupt herself three times, albeit intermittently. However, as the results indicate, our worry that this would lead them to disengage proved to be unfounded.

## Results

Scores for all measures are presented as percentages for ease of description. Preliminary analyses were conducted to explore whether children performed differently on each task depending on whether or not they saw the experimenter make the highly inefficient design cue choice first and also to check for counterbalancing effects. Interestingly, as in Study 2, there were once again no effects of trial order. There were also no other counterbalancing or item effects aside from the agent reliability item effect already noted. For Measures 1 to 3, children were given a score of 1 if they selected the functionally affordant tool. To preview, descriptions of the findings for Measures 1 to 3 are straightforward because the results from Studies 1 and 2 are replicated.

### Measures 1 to 3

Fig. 4 shows the results for Measures 1 to 3. Following the same pattern as in Studies 1 and 2, whereas 87% of 3.5-year-olds selected the functionally affordant tool in Measure 1 (Study 1), 50% selected it in Measure 2 (design cue), and 60% selected it in Measure 3 (suboptimal tool). Performance in Measure 1 was above chance,  $\chi^2 = 16.13$ ,  $df = 1$ ,  $p < .001$ . Also, replicating patterns of the prior studies, 4-year-olds consistently ignored the model's intentional cue and selected the functionally affordant

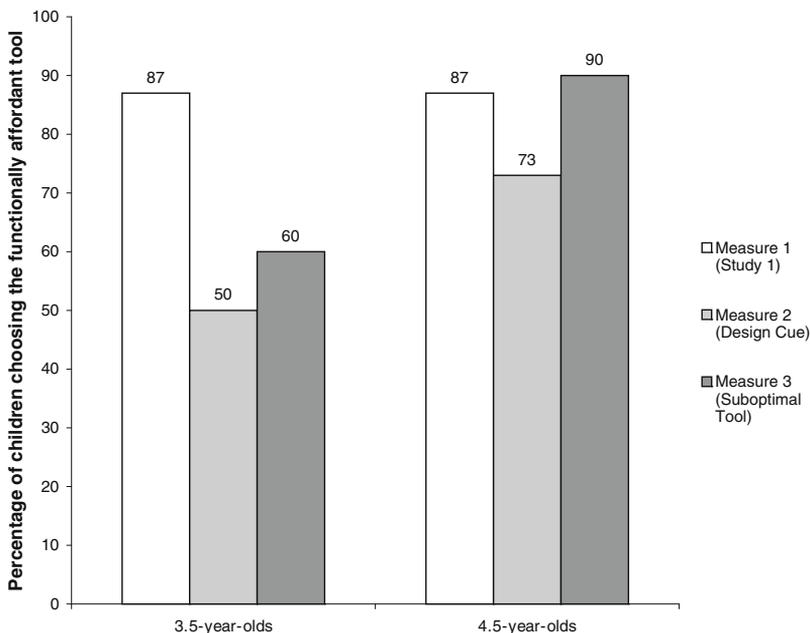


Fig. 4. Percentages of children picking the functionally affordant tool in Measure 1 (Study 1), Measure 2 (design cue), and Measure 3 (suboptimal tool) of Study 3.

tool in all trials (87, 73, and 90% in Measures 1, 2, and 3, respectively). The 4-year-olds' preference for the functionally affordant tool was above chance in all cases,  $\chi^2$  values = 16.13, 6.53, and 19.20, respectively, all  $p$ s < .05, and in Measure 3 it was significantly greater than the number of 3.5-year-olds selecting the functionally affordant tool, Fisher's exact test,  $p$  < .02.

#### *Measure 4: Assessment of agent reliability understanding*

All children demonstrated the ability to track who had been a reliable or unreliable informant about familiar objects subsequent to feedback during familiarization. During testing about unfamiliar object functions, however, 4-year-olds ( $M = 73\%$ ,  $SD = 30\%$ ) were better than 3-year-olds ( $M = 52\%$ ,  $SD = 25\%$ ) at predicting that the previously reliable informant was likely to be enduringly trustworthy,  $t(58) = 3.04$ ,  $p < .005$ . This result extends unpublished findings by Koenig and Harris that found no developmental trends in a preliminary study of children's reasoning about the reliability of agents demonstrating object functions (M. Koenig, personal communication). Also,  $t$  tests were conducted on Measures 1, 2, and 3 to see whether children who selected the functionally affordant tool scored higher on the agent reliability understanding assessment. This effect was found among 3.5-year-olds on Measure 3 (functionally affordant tool,  $M = 59\%$ ,  $SD = 26\%$ , vs. suboptimal tool,  $M = 41\%$ ,  $SD = 18\%$ ),  $t(28) = -2.07$ ,  $p < .05$ . No other significant effects were found.

#### *Measure 5: Design stance assessment*

The 3-year-olds privileged design information and chose the design tool on average 57% ( $SD = 18\%$ ) of the time. The tendency was marginally greater than chance,  $t(29) = 1.99$ ,  $p < .06$ . The 4-year-olds selected the design tool on average 66% ( $SD = 23\%$ ) of the time, which was significantly greater than chance,  $t(29) = 3.75$ ,  $p < .002$ . However, children in both age groups were more certain in their responses to conventionality questions ( $M$ s = 61% and 72% for 3- and 4-year-olds, respectively) than to generalization questions ( $M$ s = 52% and 59%, respectively). Pilot work with adults also found this difference, albeit to a lesser extent ( $M = 80\%$  vs. 72%). The difference occurred because in generalization trials, which involved selecting a tool for a new variant of the task (e.g., a second physical variant of a cage with a bell in it), child and adult participants were inclined to treat the new goal object as though it represented an entirely different task. In short, the "one tool, one function" bias demonstrated in Casler and Kelemen (2005) exerted some influence and served to slightly depress the design stance scores overall.

A  $t$  test was conducted to see whether children who selected the nonaffordant "made for cookie crushing" tool in Measure 2 (design cue) scored significantly higher on the design stance assessment than did children who selected the functionally affordant tool. They did not, and this result held whether 3.5- and 4.5-year-olds were analyzed separately or together. This finding suggests that it was the increased salience and nonidiosyncratic normative standard implied by the design cue, rather than its design content, that influenced 3.5-year-olds to increasingly select the nonaffordant tool in the design cue trial. A similar analysis did, however, find that 4.5-year-olds who selected the functionally affordant tool in Measure 1 (Study 1) scored significantly higher ( $M = 69\%$ ,  $SD = 21\%$ ) on the design stance assessment than did those who selected the nonaffordant tool ( $M = 42\%$ ,  $SD = 11\%$ ),  $t(28) = -2.45$ ,  $p < .05$ . No effects were found for Measure 3.

#### *Measure 6: Theory of mind assessment*

This assessment involved line-drawn narratives rather than acted-out versions of the deceptive container tasks due to time concerns and the desire to reduce participant fatigue by varying task formats. However, as a possible consequence of using the picture format, children's comprehension of the stories in the theory of mind measure was poor. On average, only 48% of 3.5-year-olds and 70% of 4.5-year-olds managed to follow the narrative and correctly answer all of the comprehension questions on any particular story. Children's difficulty with basic comprehension of the stories raises questions as to whether this assessment can be taken as a reliable measure of children's mental state reasoning. Indeed, relative to prior research with deceptive containers (e.g., Gopnik & Astington, 1988; Wellman, Cross, & Watson, 2001), 4.5-year-olds' correct performance on test questions was low insofar as they performed at chance ( $M = 48\%$ ,  $SD = 32\%$ ). The 3-year-olds' performance was significantly poorer ( $M = 23\%$ ,  $SD = 26\%$ ),  $t(58) = -3.25$ ,  $p < 0.005$ , and below chance,  $t(29) = -5.57$ ,  $p < .001$ . Having noted

reasons to be a little cautious about the theory of mind assessment, we also note that an age-partialled correlation revealed a significant positive correlation between children's theory of mind scores and their agent reliability understanding assessment scores,  $r = .291$ ,  $p < .03$ . Also,  $t$  tests exploring whether children who selected the functionally affordant tool in Measures 1 to 3 had higher theory of mind assessment scores than did those who selected the less efficient alternative found no differences on any task in either age group.

#### *Relationship of independent measures to children's overall tendency to imitate*

In addition to examining differences in children's performance in the individual measures, we wanted to explore the relationship between children's overall tendency to make imitative tool choices and their sensitivity to design and agent reliability information. To assess this, children received a global "model rejection" score to reflect their overall selection of the functionally affordant tool rather than the model's desired tool across Measures 1 to 3. The scores ranged from 0 to 3, with 3 representing children who rejected the model's intentional choice in all three cases. An ordinal regression was then conducted with model rejection as the dependent variable and with age, the design stance score, and the agent reliability score as the independent variables. All interactions were also considered. On backward elimination (with the significance level for model inclusion at .10), all interactions were eliminated along with the agent reliability score. The only significant variables were age and design stance score, resulting in a model with a global likelihood ratio of  $\chi^2 = 8.13$ ,  $p = .017$ . The odds ratio for age was 2.65 (95% confidence interval [0.97, 7.22]), and the odds ratio for the design stance score was 1.43 (95% confidence interval [0.95, 2.17]). In short, above and beyond the contribution of age, insights into artifact design increased children's tendency to discount the model's choice of inefficient tool and opt against imitating the model.

#### *Discussion*

Study 3 replicated the 3- and 4-year-olds' tool use imitation results of Studies 1 and 2 and yielded some important additional insights into the factors influencing preschoolers' imitative behavior. In particular, the results again demonstrated that although preschoolers are not indiscriminate imitators of others' intentional behavior (Measure 1 [Study 1]), developmental trends exist in susceptibility to social influence and desire for interpersonal alignment across the 3- to 4-year age period. Specifically, 3-year-olds showed a greater drive to imitatively align their behavior with an adult model who was making highly questionable (Measure 2 [design cue]) or moderately questionable (Measure 3 [suboptimal tool]) choices than did 4-year-olds, who instead resolutely eschewed such imitative generalization and who, on independent tests, also demonstrated greater knowledge about design aspects of the artifact domain and the differential trustworthiness of agents. Indeed, regression analysis revealed that the more knowledge children had about artifact design, the more likely they were to reject the model's choice of an inefficient artifact. This may be because knowledge about artifacts provided children with a sound basis for evaluating the model's choices as unreliable, thereby reducing the social influence of her intentional selections and children's desire to imitate them. If this interpretation is correct, it is an interesting question as to whether children's attribution of unreliability to the model was dispositional (e.g., "She is incompetent/deceptive") or situational (e.g., "She is being silly/absent-minded"). Work by [Jaswal and Neely \(2006\)](#) indicates that children have a general bias to trust adults' testimony over peers' testimony, and so it is certainly possible that children in these studies may have considered trusting our affable model outside the tool use context or on a different occasion. It will be interesting for future research to explore what happens to subsequent interactions if children see an adult engage in the current kind of questionable choice behavior and are later given a choice of people to approach to gain information on a variety of matters in addition to tool use.

#### **General discussion**

Comparative cognitive developmental research on primates' imitation of intentional acts has supported a characterization of humans as the "imitative species." Studies have consistently found that young children reproduce others' causally irrelevant, inefficient, and inexplicable intentional actions

(e.g., Call, Carpenter, & Tomasello, 2005; Horner & Whiten, 2005; Lyons et al., 2007; Nagell et al., 1993; Want & Harris, 2001). This imitative fidelity supports suggestions that humans are inherently driven to interpersonally identify and seek mutuality with each other with an adaptive consequence that, unlike any other species, they can rapidly acquire and transmit culturally shared knowledge (Tomasello et al., 2005; also Meltzoff & Moore, 1995; Užgiris, 1981). One issue that arises for this characterization, however, is that not all social informants are created equal. If children were to indiscriminately seek to satisfy their social affiliative and interpersonal drives with anybody about anything, they would run the risk of being misled by unreliable information and imitatively generalizing behavior that is idiosyncratic and potentially costly.

What the current research demonstrates is that although the characterization of children as overly zealous imitators may have some merit for certain contexts—probably those where children do not have enough knowledge to recognize that primary goal achievement has been placed at terminal risk—from early on, children are not so interpersonally motivated and socially influenced that they blindly imitate questionable intentional choices. Children at all ages in both Study 1 and Study 3 (Measure 1) consistently eschewed reproducing a model's intentional choice of a nonaffordant tool for a task.

Having noted this reason to qualify accounts emphasizing imitation as a socially and interpersonally motivated behavior, it is nevertheless the case that the current findings suggest that there may be subtle shifts in the primary goals of children's imitative acts over the 2- to 4-year age period, particularly in relation to instrumental tool use acts with well-defined goals. For example, one interpretation of 2- and 4-year-olds' bias against aligning with the model's intentional choice is that the primary motivation for imitation in these age groups is knowledge acquisition rather than social engagement, although the specific kinds of knowledge that each age group seeks to attain may differ. For 2-year-olds, it may be practical information about the "best" and most efficacious way in which to achieve a goal, as revealed by the fact that they were tempted to follow the model's intentional choice only when it involved a reasonably viable suboptimal tool that they could presume might have a hidden practical advantage that only firsthand experience could reveal. In contrast with this practical focus, it is possible that 4-year-olds may be seeking conventional knowledge about the culturally normative or "right" way in which to perform a task—an interpretation that is consistent with the fact that they were rarely tempted to reproduce the experimenter's choices, potentially because consultation of their own artifact knowledge discounted the experimenter's acts as a reliable source of generalizable cultural information. Follow-up research is currently exploring this idea that 4-year-olds become motivated to imitate suboptimal behavior when they have cues that it is conventional rather than individually capricious and when the personal cost to imitation is lower and competes with other motivations.

In comparison with both of these "model-rejecting" groups, 3-year-olds' more socially focused and affiliative pattern of imitation is quite striking. In addition to aligning with the experimenter's choice of a suboptimal tool in Studies 2 and 3 (Measure 3), this age group was the only one to be significantly induced to choose a completely nonaffordant tool when given the design cue in Studies 2 and 3 (Measure 2). Several questions arise in relation to these results. First, although this behavior is clearly more socially motivated, it would be interesting to disentangle the relative contributions of different kinds of social motivation, for example, the interpersonal desire to establish mutuality versus a concern about social authority. At an individual difference level, the fact that many 3-year-olds are transitioning from more intimate caregiving environments to larger institutional peer contexts such as preschool would certainly place pressure on them to become socially motivated on both fronts. But in addition to this, it is interesting to speculate whether 3-year-olds' level of social orientation might also reflect a more general intellectual rite of passage as young children potentially move out of being focused on the personal and practical efficacy of novel behaviors to figuring out their normativity and conventional status.

In conclusion, the current findings are consistent with a view of children's imitative acts as the selective products of an interpretive process rather than the result of an indiscriminate drive (Gergely & Csibra, 2005; Gergely & Csibra, 2006). They also contribute to a growing body of work that indicates that the fidelity of children's object-directed imitative behavior is influenced by a variety of factors, including the transparency of the model's goal (e.g., Király, 2003; Williamson & Markman, 2006), the situational constraints on the model's behavior (e.g., Gergely et al., 2002), the sociocommunicative

nature of the interaction (e.g., Gergely & Csibra, 2005; Gergely & Csibra, 2006; Nielsen, 2006), and the causal centrality of a modeled act (Schulz, Hoopell, & Jenkins, 2008). In the current instance, our results show that even very young children are not indiscriminate imitators of adults' questionable choices when goal achievement is placed at risk and that, from early on, children are well positioned to filter out potentially fallible cultural information despite powerful social motivations. Furthermore, these findings explicitly show that domain-specific knowledge directly influences children's motivation to imitate questionable behaviors. By 4 years of age, children are more willing to trust their own instincts about artifacts; far from being blind imitators, they can act like skeptics.

## Acknowledgments

This research was supported by a Clara Mayo Memorial Dissertation Fellowship to Cara DiYanni and by grants from the National Institutes of Health (NIH HD37903-01) and the National Science Foundation (NSF REC-052599) to Deborah Kelemen. Thanks are given to the many participating day care teachers, caregivers, parents, and children in the Boston area who made this study possible. Thanks also go to Krista Casler, Liz Donovan, Evie Rosset, Brenda Phillips, Alison Lyons, Richard Ely, Kim Saudino, and Helen Tager-Flusberg for their insightful comments and feedback; to Maelyn Entwistle, Ali Kozlowski, Jennifer Parks, Cassie Pestana, Nicole Shutt, and Jeneane Solz for their being wonderful undergraduate research assistants; to Nathan Demant for his help in creating and photographing the stimuli; and to Michael Pencina for his help with statistical analyses.

## References

- Asher, Y. M., & Kempler Nelson, D. G. (2008). Was it designed to do that? Children's focus on intended function in their conceptualization of artifacts. *Cognition*, *106*, 474–483.
- Bauer, P. J., & Mandler, J. (1992). Putting the horse before the cart: The use of temporal order in recall of events by one-year-old children. *Developmental Psychology*, *28*, 441–452.
- Bellagamba, F., & Tomasello, M. (1999). Re-enacting intended acts: Comparing 12- and 18-month-olds. *Infant Behavior and Development*, *22*, 277–282.
- Brown, A. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science*, *14*, 107–133.
- Call, J., Carpenter, M., & Tomasello, M. (2005). Copying results and copying actions in the process of social learning: Chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Animal Cognition*, *8*, 151–163.
- Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*, 315–330.
- Carpenter, M., Call, J., & Tomasello, M. (2002). Understanding “prior intentions” enables 2-year-olds to imitatively learn a complex task. *Child Development*, *73*, 1431–1441.
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63* (4, Serial No. 255).
- Casler, K., & Kelemen, D. (2005). Young children's rapid learning about artifacts. *Developmental Science*, *8*, 472–480.
- Casler, K., & Kelemen, D. (2007). Reasoning about artifacts at 24 months: The developing teleo-functional stance. *Cognition*, *103*, 120–130.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.
- Clement, F., Koenig, M., & Harris, P. L. (2004). The ontogenesis of trust. *Mind and Language*, *19*, 360–379.
- Csibra, G., & Gergely, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences*, *7*, 287–292.
- Defeyter, M. A., & German, T. P. (2003). Acquiring an understanding of design: Evidence from children's insight problem solving. *Cognition*, *89*, 133–155.
- Diesendruck, G., Markson, L., & Bloom, P. (2003). Children's reliance on creator's intent in extending names for artifacts. *Psychological Science*, *14*, 164–168.
- Dijksterhuis, A., & Bargh, J. A. (2001). The perception–behavior expressway: Automatic effects of social perception on social behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33, pp. 1–40). San Diego: Academic Press.
- Gelman, S. A., & Bloom, P. B. (2000). Young children are sensitive to how an object was created when deciding what to name it. *Cognition*, *76*, 91–103.
- Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*, 755.
- Gergely, G., & Csibra, G. (2005). The social construction of the cultural mind: Imitative learning as a mechanism of human pedagogy. *Interaction Studies*, *6*, 463–481.
- Gergely, G., & Csibra, G. (2006). Sylvia's recipe: Human culture, imitation, and pedagogy. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition, and human interaction* (pp. 229–255). Oxford, UK: Berg.
- German, T. P., & Johnson, S. C. (2002). Function and the origins of the design stance. *Journal of Cognition and Development*, *3*, 279–300.
- Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to understanding of false belief and the appearance–reality distinction. *Child Development*, *59*, 26–37.

- Hobson, P., & Meyer, J. (2006). Imitation, identification, and the sharing of mind: Insights from autism. In S. Rogers & J. Williams (Eds.), *Imitation and the social mind* (pp. 198–226). New York: Guilford.
- Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, *57*, 567–582.
- Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition*, *8*, 164–181.
- Hughes, C., & Cutting, A. (1999). Nature, nurture, and individual differences in early understanding of mind. *Psychological Science*, *10*, 429–432.
- Jaswal, V. K. (2006). Preschoolers favor the creator's label when reasoning about an artifact's function. *Cognition*, *99*, B83–B92.
- Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science*, *17*, 757–758.
- Johnson, S. C., Booth, A., & O'Hearn, K. (2001). Inferring the goals of a non-human agent. *Cognitive Development*, *16*, 637–656.
- Kelemen, D. (1999). The scope of teleological thinking in preschool children. *Cognition*, *70*, 241–272.
- Kelemen, D. (2001, April). Intention in children's understanding of artifact function. Paper presented at the biennial meeting of the Society for Research in Child Development, Minneapolis, MN.
- Kelemen, D., & Carey, S. (in press). The essence of artifacts: Developing the design stance. In S. Laurence & E. Margolis (Eds.), *Creations of the mind: Theories of artifacts and their representation*. Oxford, UK: Oxford University Press.
- Kemler Nelson, D. G. (1999). Attention to functional properties in toddlers' naming and problem-solving. *Cognitive Development*, *14*, 77–100.
- Kemler Nelson, D. G., Frankenfield, A., Morris, C., & Blair, E. (2000). Young children's use of functional information to categorize artifacts: Three factors that matter. *Cognition*, *77*, 133–168.
- Kemler Nelson, D. G., Herron, L., & Morris, C. (2002). How children and adults name broken objects: Inferences and reasoning about design intentions in the categorization of artifacts. *Journal of Cognition and Development*, *3*, 301–332.
- Kemler Nelson, D. G., Holt, M. B., & Egan, L. C. (2004). Two- and three-year-olds infer and reason about design intentions in order to categorize broken objects. *Developmental Science*, *7*, 543–549.
- Kemler Nelson, D. G., Russell, R., Duke, N., & Jones, K. (2000). Two-year-olds will name artifacts by their functions. *Child Development*, *71*, 1271–1288.
- Király, I. (2003, April). *Guidance by goals: The possible causes of irrelevant act omission in infants' imitative behavior*. Paper presented at the biennial meeting of the Society for Research in Child Development, Tampa, FL.
- Killen, M., & Užgiris, I. C. (1981). Imitation of actions with objects: The role of social meaning. *Journal of Genetic Psychology*, *138*, 219–229.
- Koenig, M. A., Clement, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science*, *15*, 694–698.
- Koenig, M. A., & Echols, C. H. (2003). Infants' understanding of false labeling events: The referential roles of words and the speakers who use them. *Cognition*, *87*, 179–208.
- Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, *76*, 1261–1277.
- Lyons, D., Young, A., & Keil, F. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences of the U.S.A.*, *104*, 19751–19756.
- Markson, L. M. (2001, April). *Developing understanding of artifact function*. Paper presented at the biennial meeting of the Society for Research in Child Development, Minneapolis, MN.
- Matan, A., & Carey, S. (2001). Developmental changes within the core of artifact concepts. *Cognition*, *78*, 1–26.
- McCarrell, N. S., & Callanan, M. A. (1995). Form–function correspondences in children's inference. *Child Development*, *66*, 532–546.
- Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*, 838–850.
- Meltzoff, A. N., & Moore, M. K. (1995). Infants' understanding of people and things: From body imitation to folk psychology. In A. J. Marcel, J. Bermudez, & N. Eilan (Eds.), *The body and the self* (pp. 43–69). Cambridge, MA: MIT Press.
- Nagell, K., Olguin, K., & Tomasello, M. (1993). Processes of social learning in the tool use of chimpanzees and human children. *Journal of Comparative Psychology*, *107*, 174–186.
- Nielsen, M. (2006). Copying actions and copying outcomes: Social learning through the second year. *Developmental Psychology*, *42*, 555–565.
- Pea, R. D. (1982). Origins of verbal logic: Spontaneous denials by two- and three-year-olds. *Journal of Child Language*, *9*, 597–626.
- Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*, 125–137.
- Petroski, H. (Ed.). (1992). *The evolution of useful things: How everyday artifacts—from forks and pins to paper clips and zippers—came to be as they are*. New York: Vintage.
- Phillips, B., Kelemen, D., Seston, R., & Casler, K. (2008, April). *Young children can categorize novel tools by eavesdropping on others' intentional acts*. Paper presented at the International Conference on Infancy Studies, Vancouver, Canada.
- Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*, *72*, 1054–1070.
- Schwier, C., van Maanen, C., Carpenter, M., & Tomasello, M. (2006). Rational imitation in 12-month-old infants. *Infancy*, *10*, 303–311.
- Schulz, L. E., Hoopell, C., & Jenkins, A. C. (2008). Judicious imitation: Children differentially imitate deterministically and probabilistically effective actions. *Child Development*, *79*, 395–410.
- Tomasello, M., & Carpenter, M. (2005). The emergence of social cognition in three young chimpanzees. *Monographs of the Society for Research in Child Development*, *70* (1, Serial No. 279).
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*, 675–691.
- Užgiris, I. (1981). Two functions of imitation during infancy. *International Journal of Behavioral Development*, *4*, 1–12.

- Užgiris, I. (1984). Imitation in infancy: Its interpersonal aspects. In M. Perlmutter (Ed.), *The Minnesota Symposia on Child Psychology, Parent-child interactions and parent-child relations in child development* (Vol. 17, pp. 1–32). Hillsdale, NJ: Lawrence Erlbaum.
- 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.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false beliefs. *Child Development, 72*, 655–684.
- Williamson, R. A., & Markman, E. M. (2006). Precision of imitation as a function of preschoolers' understanding of the goal of the demonstration. *Developmental Psychology, 42*, 723–731.