CHILD DEVELOPMENT

Child Development, July/August 2012, Volume 83, Number 4, Pages 1416–1428

Preschoolers Use Intentional and Pedagogical Cues to Guide Inductive Inferences and Exploration

Lucas P. Butler and Ellen M. Markman
Stanford University

Children are judicious social learners. They may be particularly sensitive to communicative actions done pedagogically for their benefit, as such actions may mark important, generalizable information. Three experiments (N = 224) found striking differences in preschoolers' inductive generalization and exploration of a novel functional property, depending on whether identical evidence for the property was produced accidentally, intentionally, or pedagogically and communicatively. Results also revealed that although 4-year-olds reserved strong generalizations for a property that is pedagogically demonstrated, 3-year-olds made such inferences when it was produced either intentionally or pedagogically. These findings suggest that by age 4 children assess whether evidence is produced for their benefit in gauging generalizability, giving them a powerful tool for acquiring important kind-relevant, generic knowledge.

A fundamental aspect of human cognition is our ability to learn from and teach others. Our ability to read others' intentions and engage in collaborative learning may provide the foundation for human culture, from law and government to industry and education (Gergely & Csibra, 2005; Tomasello, 1999). Children's understanding of intentions is inherent to many domains, including word learning (Baldwin, 1991, 1993) and imitation (Carpenter, Akhtar, & Tomasello, 1998; Meltzoff, 1995). Recent work has elaborated on the importance of *pedagogy*, which fundamentally relies on children's ability to read the social intentions underlying action (Csibra, 2010). Indeed, an indispensible component of our social cognition is the capacity not only to reason *about* others in terms of their mental states but also to learn *from* them. And although our closest primate relatives have been shown to have a surprisingly wide array of abilities to reason about others, learning *from* others via various forms of communication seems to be beyond them (see Tomasello, 2008).

Building on the importance of learning from others, Csibra and Gergely (2006, 2009) have presented a theory of *natural pedagogy* to explain why sensitivity to pedagogical, communicative acts might be evolutionarily adaptive, and what implications it might have for learning and development. On this view, infants take a ''pedagogical stance'' toward acts of intentional, ostensive communication and infer not only that information being communicated is *relevant* in Sperber and Wilson's (1986) terms, but more specifically that information communicated ostensively is *kind*-relevant and likely generalizable. Following this, having something intentionally and ostensively communicated for your benefit by a knowledgeable adult acts as a tacit guarantee not only of its relevance to you in the given situation but of its generalizability to the world at large.

From the first months of life, children are sensitive to communicative cues, such as eye gaze and establishing joint attention, which signal that others are deliberately sharing information with them (Farroni, Csibra, Simion, & Johnson, 2002; Grossmann & Johnson, 2010; Grossmann et al., 2008). Moreover, infants treat information differently

Portions of this work were presented at the 2010 CEU Cognitive Development Center Opening Conference, the 2010 biennial meeting of the International Conference on Infant Studies, the 32nd annual meeting of the Cognitive Science Society, and the 2011 biennial meeting of the Society for Research in Child Development. This research was supported by an NSF Graduate Research Fellowship to the first author. We are grateful to the teachers, staff, parents, and children at Bing Nursery School and the Arboretum Child Care Center for participating in this research and to Hannah Jaycox, Cole Murphy-Hockett, and Sam Saenz for their assistance with data collection and coding. We thank Andrei Cimpian for helpful comments on a previous draft.

Correspondence concerning this article should be addressed to Lucas P. Butler, Department of Psychology, Stanford University, 450 Serra Mall, Jordan Hall, Building 420, Stanford, CA 94305. Electronic mail may be sent to lpbutler@stanford.edu.

 0009-3920/2012/8304-0022
DOI: 10.1111/j.1467-8624.2012.01775.x

when they encounter it in communicative contexts. For example, infants follow an adult's eye gaze only when it occurs in a communicative context (Csibra & Volein, 2008; Senju & Csibra, 2008), and infants expect communicative cues to be directed to specific objects (Csibra & Volein, 2008; Gliga & Csibra, 2009; Senju, Csibra, & Johnson, 2008). Beyond this, infants seem to expect others to communicate information that is relevant and generalizable beyond the current object or situation (Egyed, Király, Krekó, Kupán, & Gergely, 2007; Futó, Téglás, Csibra, & Gergely, 2010; Gergely, Egyed, & Király, 2007; Yoon, Johnson, & Csibra, 2008). Put another way, infants seem to expect that when others communicate, they intend to share general or generic knowledge about the world.

Infants are clearly sensitive to communicative cues such as eye gaze and joint attention, but do children use this sensitivity to guide inductive inferences and aid in the acquisition of generic knowledge about kinds and categories? There is a long tradition of research with older, preschool-aged children on the process of kind-based inductive inference and the role this plays in conceptual development (see Gelman, 2003). It is as yet unclear what role this sensitivity to pedagogical cues that is present in infancy might play in the rich, category-based inductive inferences that older children are capable of. By 3 or 4 years of age, and even earlier, children take labels as referring to kinds that share nonobvious properties, and generalize novel properties learned about objects on the basis of shared labels, rather than on the objects' perceptual similarity (e.g., Booth & Waxman, 2002; Gelman & Coley, 1990; Gelman & Markman, 1986, 1987; Graham, Baker, & Poulin-Dubois, 1998). Moreover, recent work has demonstrated that preschoolers expect novel objects that share a label to share a novel causal property and selectively explore those objects more when that property fails to extend to additional kind members (Schulz, Standing, & Bonawitz, 2008). Beyond mere labeling, children also understand that novel information conveyed in a generic statement (e.g., ''snakes have holes in their teeth'') has greater inductive potential than information conveyed nongenerically (e.g., ''this snake has holes in his teeth''; Cimpian & Markman, 2008; Gelman, Star, & Flukes, 2002; Hollander, Gelman, & Raman, 2009), and information conveyed generically becomes more central to their kind representations (Cimpian & Cadena, 2010; Cimpian & Markman, 2009).

This process of assessing whether novel information should be generalized to a kind, and whether it is an essential property of that kind, is critical in category and concept formation, where children rely on others to impart information about the world that might otherwise be difficult or impossible to learn (Gelman, 2009; Harris, 2002; Harris & Koenig, 2006; Leslie, 2007). Transmission of knowledge that supports inductive inferences about generalizability is often carried out linguistically, using language that refers to kinds and categories, and children make a variety of inductive inferences on the basis of kind-referring language. Might analogous nonlinguistic cues play a similar role in category and concept formation as language clearly does? Csibra and Gergely (2009) suggest that at the core of generic knowledge transmission is the intention to communicate new and important information. On a rational, Gricean analysis (Grice, 1957, 1969; see also Clark, 1996; Sperber & Wilson, 1986), children may infer that when adults intentionally communicate information *for them*, it is likely because it is important and relevant, and thus they may use ostensive, communicative cues to gauge generalizability. Given this, we hypothesize that preschoolers may take information demonstrated pedagogically as both more generalizable and conceptually central, as they do for information conveyed generically (Cimpian & Cadena, 2010; Cimpian & Markman, 2009; Gelman et al., 2002; Hollander et al., 2009). Furthermore, although linguistic cues such as kind labels are powerful in driving generalization, they are not *always* used pedagogically. Thus, even when objects share a label, children might reserve stronger inductive inferences about generalizability for properties conveyed pedagogically, treating them as more conceptually central and more likely to be shared by other kind members.

Prior research has established that children's exploratory play is a window into their implicit inductive processes. Specifically, having learned that an exemplar of a kind has a particular causal property, young children, even infants, explore more upon encountering exemplars lacking that property (Baldwin, Markman, & Melartin, 1993; Schulz et al., 2008). In the current research, we tapped preschoolers' spontaneous exploration of identical but *inert* kind-members to investigate whether they would form different expectations about generalizability depending on whether a novel property was demonstrated pedagogically or produced in a nonpedagogical manner. Furthermore, we investigated whether such a pedagogical versus nonpedagogical difference might exist even given shared kind labels.

Our prediction is that if cues that evidence is being pedagogically demonstrated lead children to treat a property as generic information about the kind, then they ought to make a stronger inference about its generalizability and importance to the kind, and will explore *more* when they encounter kind members that fail to have that property. It is not that children who witness evidence for a property produced nonpedagogically will fail to draw any inferences at all. They may well pick up on the property, want to carry out for themselves the action that produced it, and be interested in testing whether it generalizes to other kind members. After all, there is no real cost to making a *tentative* generalization. Even when a property is produced in a nonpedagogical manner, children may well at least try another kind member to see whether it generalizes, but this inference likely will be quickly abandoned in the face of negative evidence.

Experiment 1

In Experiment 1, we taught children a name for a novel object, and either pedagogically demonstrated or accidentally produced a novel causal property (magnetically picking up paperclips). We then presented children with an identical set of exemplars, all of which shared the same label, but lacked the property (they were not magnetic), and let them play.

Method

Participants. Thirty-two 3-year-olds (16 girls; M = 42 months, range = 36–46 months) and thirty-two 4-year-olds (16 girls; M = 54 months, range = 48–61 months) from a university preschool participated. These ages were chosen for consistency with prior research on kind-based inductive inference in preschoolers (e.g., Cimpian & Cadena, 2010; Cimpian & Markman, 2009; Gelman & Markman, 1986, 1987; Gelman et al., 2002; Hollander et al., 2009; Schulz et al., 2008). Children came from predominantly middle- and upper-middle-class families, representing a variety of ethnic groups. Children were randomly assigned to condition, equating for gender and age.

Materials. The novel objects were 11 small wooden blocks, 2.5 cm × 2.5 cm × 5 cm. The *active* block had thick magnetic tape on one end; the 10 *inert* blocks had similar but nonmagnetic tape on the same end. All were covered with black electrical tape, with green electrical tape covering the magnetic or nonmagnetic end. All 11 blocks were perceptually indistinguishable.

Procedure. All children were tested in a private room in their preschool by a trained experimenter. Children first learned a novel label (*blicket*) for the active block. When asked for the blicket, all children successfully selected it from an array of four distractors on two successive trials, without error.

After learning the word, children engaged in a short distracter task. The experimenter said, ‘‘I have something else fun to show you. I'm going to show you how to make a house by folding colored paper.’’ He then introduced several colored pencils, a pile of metal paperclips, and two pieces of colored paper. He showed children how to fold the paper to make a house, and then allowed them to construct their own house. This task served two goals. First, it distanced the word-learning phase, which was clearly pedagogical, from the demonstration phase. Second, it provided a plausible cover story for placing a pile of paperclips on the table.

The experimenter then started to clean up the toys. He put away the distractors, then picked up the *active* blicket. In the pedagogical condition, he said, ‘‘Look, watch this,’’ and then deliberately placed the blicket on the paperclips, picking it up with paperclips attached. In the accidental condition, he appeared to accidentally drop the blicket on the paperclips as he was putting it away, exclaiming ‘‘Oops!’’ In *both* conditions, he then picked the blicket up and looked at it, saying ‘‘Hmmm’’ in a neutral tone, and then placed it next to the paperclips. Finally, he placed 10 *inert* blocks on the table, saying, ‘‘and here are some blickets, and told the child to ‘‘go ahead and play’’ while he left the table and sat facing away from the child for 60 s.

Coding and data analysis. Two independent judges coded each child's exploration. Importantly, each coder viewed only the portion of the video immediately following the demonstration, and so was blind to condition.

Given our argument, we were particularly interested in children's exploration in the face of negative evidence, and thus focused on how children explored the inert blickets. Specifically, we analyzed three aspects of this exploration. First, we measured the *time* children spent exploring the inert blickets. This was coded as the cumulative number of seconds that children were actively exploring; each period of exploration began with when children picked up a blicket that they then tried to use to pick up paperclips, and ended when and if they put the blicket down or began to do something else with it (e.g., stacking). Second, we measured the *number of attempts* children made to elicit the property from the inert blickets. An attempt was coded as any intentional action clearly done to pick up paperclips, including both placing the blicket on the paperclips and placing a paperclip on the blicket. Third, we measured the *number of inert blickets* children explored. This was coded as how many individual objects children deliberately used to try to pick up paperclips.

We also examined the possibility that children might simply be more interested in the property or see the activity as more relevant when it is demonstrated for their benefit, or might even pick up on and learn the property better. To test this, we measured how many children explored at least one inert blicket, how quick they were to reproduce the action that had elicited the property, and how they explored the original, active magnetic blicket (both in terms of time and number of uses). The logic of these measures is as follows. First, trying to elicit the property from at least one inert blicket is evidence that children noticed and remembered the property, were interested in reproducing the action that produced evidence for that property, and have made at least a tentative generalization. Second, quickly using a blicket to reproduce the action (or trying equally quickly across conditions) is further evidence that children noticed the property and were interested in carrying out the action that produced it, as is exploring the active blicket and reproducing the activity of magnetically picking up paperclips.

Agreement on all measures was high (for categorical measures, Cohen's κs > 0.82, for continuous measures, Pearson's *r*s between .91 and .97, all $ps < .001$), and any disagreements were resolved by discussion. Because children's exploration tended to be bimodal, and violated assumptions of normality and homogeneity of variance, we analyzed all data using nonparametric chi-square and Mann–Whitney *U* tests, which provide a conservative test of the differences across conditions. All analyses reported in this article were also carried out using standard *t* tests, yielding the same results in every case.

Results

None of the 3-year-olds explored the blickets in the accidental condition, leading to zero variance in that cell of the design and precluding analyses of continuous measures of their exploration. Moreover, due to such stark differences between overall patterns of exploration in the two age groups, we analyzed 3- and 4-year-olds' responses separately.

Four-year-olds. These children inferred that the property was kind-relevant and should generalize to other kind-members when it was demonstrated pedagogically, as evidenced by their continued exploration in the face of negative evidence (see Figure 1). When the property was demonstrated pedagogically, rather than produced accidentally, 4-year-olds spent significantly more time trying to pick up paperclips with the inert blickets ($M_{pedagogical} = 38.56$ s, $SD = 19.37$ vs. $M_{accidental} = 18.69$ s, $SD = 21.83$), Mann–Whitney $Z = 2.25$, $p = .024$, Cohen's $d = 0.96$, and made significantly more attempts to pick up paperclips with the inert blickets ($M_{pedagogical} = 9.25$, $SD = 7.62$ vs. $M_{accidental} = 2.94$, $SD = 3.45$) condition, Mann–Whitney $Z = 2.54$, $p = .011$, Cohen's $d = 1.07$. They also explored marginally more of the inert blickets ($M_{pedagogical} = 2.13$, $SD = 2.06$ vs. $M_{accidental} = 1.31$, $SD = 2.06$), Mann–Whitney $Z = 1.78$, $p = .08$, Cohen's $d = 0.40$.

Importantly, it is not simply that children did not notice or were not interested in the property or in reproducing the activity in the accidental condition, nor is it that children failed to make at least a tentative generalization. Just as many children in both

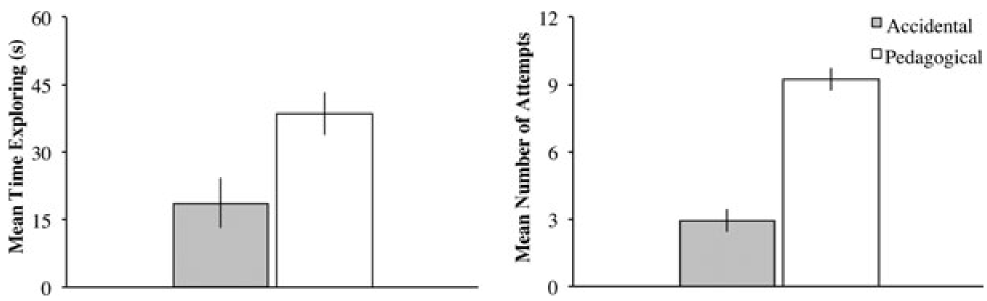

Figure 1. The mean amount of time 4-year-olds spent exploring (left) and the mean number of times they attempted to elicit the property from the inert blickets (right) in Experiment 1.
Note. Error bars represent standard errors of the mean.

conditions tried at least one inert blicket (pedagogical: 14 children, 87.5% vs. accidental: 12 children, 75%), $\chi^2(1, n = 32)$, = 0.82, $p = .37$, suggesting that they had made a tentative generalization and were motivated to test it, but for children in the accidental condition, this inference was quickly abandoned in the face of negative evidence. Additionally, although there was high variance on these measures, there were no significant differences in children's latency to produce the target action ($M_{pedagogical} = 2.15$ s, $SD = 7.77$ vs. $M_{accidental} = 0.45$ s, $SD = 1.51$, Mann–Whitney $Z = 0.60$, $p = .95$) or their exploration of the active blickets (time exploring: $M_{pedagogical} = 9.25$ s, $SD = 12.75$ vs. $M_{accidental} = 5.13$ s, $SD = 6.63$, Mann–Whitney $Z = 0.34$, $p = .74$; number of uses: $M_{pedagogical} = 1.94$, $SD = 3.75$ vs. $M_{accidental} = 0.94$, $SD = 1.18$, Mann–Whitney $Z = 0.20$, $p = .84$). Additionally, although many children did engage in other actions with the blickets (e.g., stacking them), all children who explored did so *before* performing any other actions with the blickets.

Three-year-olds. These children showed an analogous effect of condition, but one that was starker than for the 4-year-olds. Not one of the 16 children in the accidental condition explored at all, compared to 8 of 16 in the intentional condition, $\chi^2(1, n = 32) = 10.67$, $p = .001$, nor did any of them explore the active blicket either. This suggests that 3-year-olds in the accidental condition may have failed to notice the property, or might have been hesitant to engage in exploratory play.

Discussion

These results demonstrate that by age 4, children use pedagogical cues to guide their inductive inferences and exploration. When shown a causal property of a novel kind in a pedagogical manner, 4-year-olds explored more upon discovering that it did not obtain for additional kind members, relative to when the property was produced accidentally. They appeared to have made a generic inference about the kind (e.g., ''Blickets are magnetic'') on the basis of pedagogical demonstration, and showed continued exploration in the face of conflicting evidence, indicating that they had formed a relatively strong expectation that the property would generalize. Although children in the accidental condition made a similar inference—quickly picking up the objects and exploring to see if the additional exemplars shared the property—this inference was weaker and easily disconfirmed.

Interestingly, although children in the pedagogical condition did explore marginally more of the inert objects, children in all conditions tended to try one or two inert blickets, rather than trying many. But children in the pedagogical condition *persisted* in trying to elicit the property from those few inert objects. Although this contrasts with children's exploratory behavior on previous tasks (Schulz et al., 2008), in which children tended to explore significantly more *objects* when they had made a strong inductive generalization, the fact that children showed continued exploration of those few inert objects may be even stronger evidence that children took the property as particularly generalizable and important information about the kind. It seems that, in this study, children who witnessed the property demonstrated for their benefit not only inferred that the property *would* generalize to other kind members, they may even have inferred that even the few kind members they tried *should* share the property.

Two additional factors beyond the pedagogical cues may have influenced children's inferences and exploration, especially for the younger children who engaged in very little exploration overall. First, to convey that it was accidental, the experimenter said, ''Oops!'' after producing the property in the accidental condition. But this may have also marked the property as negative or as something children should not reproduce. This could potentially have inhibited exploration, especially in the younger children. Additionally, the conditions may have produced slightly different evidence—videotape analysis suggested that more paperclips stuck to the blicket during demonstration in the pedagogical condition, making the property potentially more salient. Experiment 2 explored the possible effect of these factors on children's inferences and exploration, in particular whether this could explain the especially low level of exploration in the younger children. To address these issues, we added an enthusiastic exclamation (''Wow!'') in both conditions to mitigate any influence of negative effect and to actively encourage children's exploration, and also equated the number of paperclips picked up across conditions.

Experiment 2

Method

Participants. An additional thirty-two 3-year-olds (16 girls; $M = 41$ months, range = 39–46 months) and thirty-two 4-year-olds (16 girls; $M = 52$ months, range = 48–57 months) participated, with comparable backgrounds to children in Experiment 1.

Procedure. The procedure was identical to Experiment 1 with several modifications. First, while maintaining the manipulation of saying either ''Look, watch this'' or ''Oops!'' the experimenter said, ''Wow!'' (instead of ''Hmm''), after producing the property in *both* conditions. This served to mitigate any inhibitory effect that exclaiming ''Oops!'' in the accidental condition might have had on children's exploration. Second, we controlled for the number of paperclips picked up across conditions. The experimenter always picked up two paperclips in the pedagogical condition, whereas in the accidental condition the mean was 2.41 paperclips.

Coding and data analysis. As in Experiment 1, two independent judges coded each child's exploration. Agreement on all measures was high (for categorical measures, Cohen's κs = 1.00, for continuous measures, Pearson's *r*s between .90 and .96, all *p*s < .001), and any disagreements were resolved by discussion.

Results

Unlike Experiment 1, in which not one 3-year-old in the accidental condition explored, some 3-year-olds in both conditions of Experiment 2 did try at least one inert blicket. However, as in Experiment 1, violations of assumptions of normality and homogeneity of variance precluded valid comparisons across age groups. Instead, we again report results from each age group separately.

Four-year-olds. As in Experiment 1, 4-year-olds made stronger inferences about the generalizability of the novel property when it was pedagogically demonstrated, as measured by their exploration of the inert blickets (see Figure 2). When the property was demonstrated pedagogically, 4-year-olds spent significantly more time exploring the inert blickets ($M_{pedagogical}$ = 35.47 s, SD = 18.55 vs. $M_{accidental}$ = 15.38 s, SD = 16.56), Mann–Whitney Z = 2.76, p = .006, Cohen's d = 1.14, and made significantly more attempts to pick up paperclips with the inert blickets ($M_{pedagogical}$ = 7.63, SD = 4.53 vs. $M_{accidental}$ = 2.88, SD = 2.40) condition, Mann–Whitney Z = 3.14, p = .002, Cohen's d = 1.17. They also again explored marginally more of the inert blickets ($M_{pedagogical}$ = 1.88, SD = 2.36 vs. $M_{accidental}$ = 0.81, SD = 0.66), Mann–Whitney Z = 1.86, p = .06, Cohen's d = 0.62.

As in Experiment 1, this was not due to differences across conditions in whether 4-year-olds learned the property, were interested in reproducing it, or inferred that it might generalize to other kind members. Equal numbers of children in both conditions explored at least one inert blicket (pedagogical: 14 children, 87.5% vs. accidental: 12 children, 75%), $\chi^2(1, n = 32) = 0.82$, p = .37, suggesting that even children in the accidental condition made at least a tentative generalization. And again, there were no significant differences in children's latency to perform the target action ($M_{pedagogical}$ = 2.50 s, SD = 9.35 vs. $M_{accidental}$ = 3.75 s, SD = 11.51, Mann–Whitney Z = 0.74, p = .46), or their exploration of the active blickets (time exploring: $M_{pedagogical}$ = 6.60 s, SD = 7.59 vs. $M_{accidental}$ = 10.44 s, SD = 16.35, Mann–Whitney Z = 0.22, p = .83; number of uses: $M_{pedagogical}$ = 1.44, SD = 1.75 vs. $M_{accidental}$ = 1.25, SD = 2.21, Mann–Whitney Z = 0.80, p = .42).

Three-year-olds. The younger children also appeared to make stronger inferences about generalizability in the pedagogical condition (see Figure 2). When the property was demonstrated pedagogically, they spent more time exploring the inert blickets ($M_{pedagogical}$ = 32.38 s, SD = 23.48 vs. $M_{accidental}$ = 11.56 s, SD = 21.09), Mann–Whitney Z = 2.40, p = .017, Cohen's d = 0.93, and made more attempts to elicit the property from the inert blickets ($M_{pedagogical}$ = 5.63, SD = 6.29 vs. $M_{accidental}$ = 1.00, SD = 2.76), Mann–Whitney Z = 2.80,

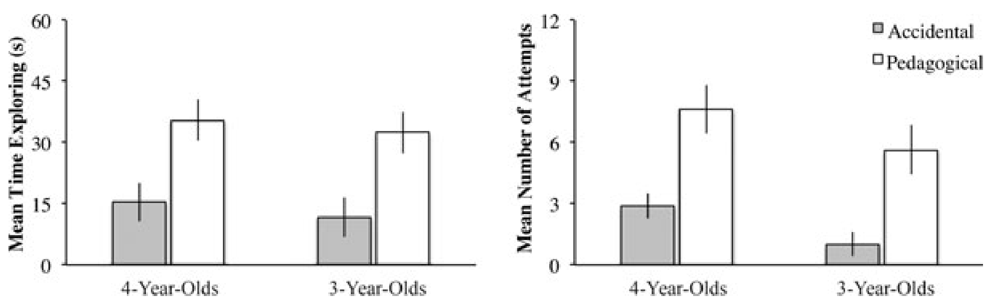

Figure 2. The mean amount of time 3- and 4-year-olds spent exploring (left) and the mean number of times they attempted to elicit the property from the inert blickets (right) in Experiment 2.
Note. Error bars represent standard errors of the mean.

$p = .005$, Cohen's $d = 0.95$. They also explored significantly more of the inert blickets in the pedagogical condition than the accidental condition ($M_{pedagogical} = 0.75$, $SD = 0.58$ vs. $M_{accidental} = 0.25$, $SD = 0.45$), Mann–Whitney $Z = 2.50$, $p = .013$, Cohen's $d = 0.96$.

However, as in Experiment 1, the effect of condition was starker for the younger children, as significantly more 3-year-olds explored at least one inert blicket in the pedagogical condition (12 children; 75%) than in the accidental condition (5 children; 31%), $\chi^2(1, n = 32) = 6.15$, $p = .013$. This suggests that unlike the older children, 3-year-olds may have been less likely to infer that the property might generalize to other kind members when it was produced accidentally. Nevertheless, our results do not appear to be because children failed to learn the property or were less motivated to reproduce it, as there were no differences in children's exploration of the original, active blicket. Three-year-olds were equally likely to explore the active blicket in the pedagogical (8 children, 50%) and accidental (7 children, 44%) conditions, and there were no significant differences in their exploration of the active blickets (time exploring: $M_{pedagogical} = 12.38$ s, $SD = 19.40$ vs. $M_{accidental} = 4.25$ s, $SD = 9.29$, Mann–Whitney $Z = 1.49$, $p = .14$; number of uses: $M_{pedagogical} = 1.94$, $SD = 3.15$ vs. $M_{accidental} = 0.38$, $SD = 0.45$, Mann–Whitney $Z = 1.68$, $p = .10$). The differences in exploration across the two conditions was thus driven entirely by children's exploration of the inert blickets, suggesting that as with the older children, 3-year-olds did use pedagogical cues to assess the generalizability of new information and guide their exploration.

Discussion

When controlling for the emotional valence of the event and the salience of the property, children again showed different patterns of inductive inference and exploration on the basis of whether a property was demonstrated pedagogically. Moreover, having controlled for these issues, 3-year-olds in the accidental condition explored more than in Experiment 1 and showed patterns of exploration similar to the older children. However, the pattern of exploration suggests that older children's inductive reasoning on the basis of pedagogical cues may be more nuanced—although they appeared to *make* the inductive inference equally in all conditions, exploring at least one inert blicket, this inference was *stronger* when the evidence had been deliberately demonstrated for their benefit. The younger children were also sensitive to whether or not evidence was produced pedagogically, but in their case this distinction appeared to influence whether or not they made an inductive generalization in the first place.

These two experiments suggest that children as young as 3 are sensitive to the distinction between actions done accidentally and those done pedagogically for their benefit. However, there are two possible explanations for this pattern of results. On one hand, children's sensitivity to pedagogical intent—their understanding that someone is carrying out an action with the intention of communicating relevant information—may be guiding their inferences and exploration. But it is important to note that although the pedagogical condition in Experiments 1 and 2 is both intentional and pedagogical, the accidental condition is neither. Thus, it could be that children are simply responding to the fact that the demonstration is itself an intentional act and that this leads them to more strongly infer that the property should be generalized. Indeed, there is ample evidence that children are in fact sensitive to the intentionality of actions in deciding what to imitate, and will selectively imitate actions done intentionally over those done accidentally (Carpenter et al., 1998). As we argued earlier, it seems unlikely that the effect seen here is due to the effect of intentionality on children's imitative learning—children were just as likely to learn the property and reproduce it, and by age 4 they were just as likely to make at least a tentative generalization. Nevertheless, it remains a possibility that children will take information produced in an intentional manner to be kind relevant and generalizable, regardless of whether this was done with the intention of explicitly teaching them. To test this, in Experiment 3, we added a third condition in which the property evidence was produced intentionally, but in the absence of any communicative or pedagogical cues.

Experiment 3

Method

Participants. An additional forty-eight 3-year-olds (16 girls; $M = 43$ months, range = 40–46 months) and forty-eight 4-year-olds (16 girls; $M = 54$ months, range = 48–59 months) participated, with comparable backgrounds to children in Experiments 1 and 2.

Procedure. The procedure was identical to Experiment 2, which controlled for both emotional valence and saliency of the property evidence, with one modification. In addition to the accidental and pedagogical conditions, we added a third, intentional condition. In this condition, after putting away the distractor items, the experimenter picked up the active blicket and placed it deliberately on the pile of paperclips. He then picked it up with paperclips attached, looked at it, and said, ''Wow!'' before placing it on the table with the magnetic side facing the child. Unlike in the pedagogical condition, the experimenter did not make eye contact with the child or establish joint attention during this phase of the study. The perceptual evidence produced by this action was identical to that in the pedagogical condition—the only difference between these two conditions is the presence of clear communicative and pedagogical cues.

Coding and data analysis. As in the previous experiments, two independent judges coded each child's exploration. Agreement on all measures was high (for categorical measures, Cohen's κs > 0.82, for continuous measures, Pearson's *r*s between .90 and .98, all *p*s < .001), and any disagreements were resolved by discussion.

Results

Four-year-olds. These children made a stronger inference about the kind-relevance and generalizability of the property when it was demonstrated pedagogically for their benefit, as measured by their exploration of the inert blickets, more so even compared to when they saw the property produced in an intentional, but nonpedagogical manner (see Figure 3). They spent longer exploring the inert blickets in the pedagogical condition ($M_{\text{pedagogical}}$ = 33.69 s, *SD* = 20.55) than in either the intentional condition ($M_{\text{intentional}}$ = 16.44 s, *SD* = 18.04), Mann–Whitney *Z* = 2.31, *p* = .021, Cohen's *d* = 0.89, or the accidental ($M_{\text{accidental}}$ = 12.44 s, *SD* = 15.51) condition, Mann–Whitney *Z* = 2.83, *p* = .005, Cohen's *d* = 1.16. The difference in time exploring between the intentional and accidental conditions was not significant, Mann–Whitney *Z* = 0.689, *p* = .515, Cohen's *d* = 0.23. Additionally, 4-year-olds also made more attempts to elicit the property from the inert blickets in the pedagogical condition ($M_{\text{pedagogical}}$ = 9.25, *SD* = 7.77) than in either the intentional condition ($M_{\text{intentional}}$ = 3.50, *SD* = 3.46), Mann–Whitney *Z* = 2.18, *p* = .030, Cohen's *d* = 0.95, or the accidental condition ($M_{\text{accidental}}$ = 3.63, *SD* = 5.10), Mann–Whitney *Z* = 2.25, *p* = .024, Cohen's *d* = 0.85. The difference in number of attempts between the intentional and accidental conditions was not significant, Mann–Whitney *Z* = 0.374, *p* = .724, Cohen's *d* = 0.02. Finally, they also explored significantly more of the inert blickets in the pedagogical condition ($M_{\text{pedagogical}}$ = 2.44, *SD* = 2.45) than in either the intentional condition ($M_{\text{intentional}}$ = 1.06, *SD* = 1.18), Mann–Whitney *Z* = 2.00, *p* = .046, Cohen's *d* = 0.72, or the accidental condition ($M_{\text{accidental}}$ = 0.88, *SD* = 1.50), Mann–Whitney *Z* = 2.60, *p* = .009, Cohen's *d* = 0.77. The difference in number of blickets explored between the intentional and accidental conditions was not significant, Mann–Whitney *Z* = 0.85, *p* = .394, Cohen's *d* = 0.13. Thus, 4-year-olds made stronger inferences about generalizability on the basis of whether or not the property was demonstrated pedagogically for their benefit, relative to when it was produced in an accidental, or even an intentional, but nonpedagogical manner.

Once again, this was not due to difference across conditions in whether 4-year-olds learned the property, viewed it as relevant, were motivated to reproduce it, or inferred that it might generalize to other kind members. Critically, equal numbers of 4-year-olds in all conditions explored at least one inert blicket (13 children [81%] in all conditions), suggesting they had made at least a tentative generalization. And again, there were no significant

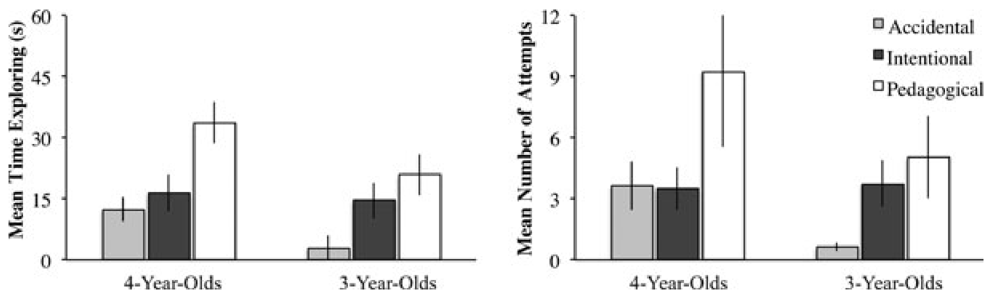

Figure 3. The mean amount of time 3- and 4-year-olds spent exploring (left) and the mean number of times they attempted to elicit the property from the inert blickets (right) in Experiment 3.
Note. Error bars represent standard errors of the mean.

differences in children's latency to perform the target action ($M_{pedagogical}$ = 1.09 s, SD = 3.62; $M_{intentional}$ = 1.85 s, SD = 5.12 $M_{accidental}$ = 2.17 s, SD = 5.29; pedagogical vs. intentional: Mann–Whitney Z = 0.54, p = .65; pedagogical vs. accidental: Mann–Whitney Z = 0.93, p = .35; intentional vs. accidental: Mann–Whitney Z = 0.51, p = .61). There were also no significant differences in time spent exploring the active blickets ($M_{pedagogical}$ = 14.50 s, SD = 14.82; $M_{intentional}$ = 13.81 s, SD = 17.55 $M_{accidental}$ = 21.63 s, SD = 20.94; pedagogical vs. intentional: Mann–Whitney Z = 0.16, p = .88; pedagogical vs. accidental: Mann–Whitney Z = 0.91, p = .36; intentional vs. accidental: Mann–Whitney Z = 0.92, p = .36). Finally, there were no significant differences in the number of uses of the active blickets ($M_{pedagogical}$ = 3.08, SD = 4.92; $M_{intentional}$ = 1.50, SD = 1.71 $M_{accidental}$ = 2.31, SD = 2.55; pedagogical vs. intentional: Mann–Whitney Z = 0.48, p = .63; pedagogical vs. accidental: Mann–Whitney Z = 0.09, p = .93; intentional vs. accidental: Mann–Whitney Z = 0.82, p = .41).

Three-year-olds. Unlike the older children, 3-year-olds made stronger inferences about generalizability, and showed continued exploration of the inert blickets, when the property was produced via an intentional action, regardless of whether or not that action was done pedagogically for their benefit (see Figure 3). They spent longer exploring the inert blickets in both the pedagogical condition ($M_{pedagogical}$ = 20.87 s, SD = 20.55), Mann–Whitney Z = 2.93, p = .007, Cohen's d = 1.14, and the intentional condition ($M_{intentional}$ = 14.56 s, SD = 18.48), Mann–Whitney Z = 2.25, p = .024, Cohen's d = 0.80, compared to the accidental condition ($M_{accidental}$ = 3.00 s, SD = 8.25). In contrast to the 4-year-olds, the difference in time exploring between the intentional and pedagogical conditions was not significant, Mann–Whitney Z = 0.957, p = .358, Cohen's d = 0.32. Finally, 3-year-olds also made more attempts to elicit the property from the inert blickets in both the pedagogical condition ($M_{pedagogical}$ = 5.06, SD = 6.38), Mann–Whitney Z = 2.63, p = .009, Cohen's d = 0.95, and the intentional condition ($M_{intentional}$ = 3.75, SD = 5.01), Mann–Whitney Z = 2.23, p = .026, Cohen's d = 0.84, compared to the accidental condition ($M_{accidental}$ = 0.63, SD = 1.54). Again, the difference in number of attempts between the intentional and pedagogical conditions was not significant, Mann–Whitney Z = 0.586, p = .590, Cohen's d = 0.22. They explored significantly more inert blickets in both the pedagogical condition ($M_{pedagogical}$ = 0.94, SD = 1.12), Mann–Whitney Z = 2.37, p = .018, Cohen's d = 0.89, and the intentional condition ($M_{intentional}$ = 0.63, SD = 0.40), Mann–Whitney Z = 2.21, p = .027, Cohen's d = 1.10, than in the accidental condition ($M_{accidental}$ = 0.19, SD = 0.40). The difference in number of blickets explored between the pedagogical and intentional conditions was not significant, Mann–Whitney Z = 0.57, p = .57.

Although 3-year-olds were equally likely to explore the inert blickets in both the pedagogical and intentional conditions (12 children, 75%), they were again less likely to explore in the accidental condition (6 children, 38%), $\chi^2(1, n = 32) = 4.57$, p = .033. This suggests that, at least in the pedagogical and intentional conditions, 3-year-olds were just as likely to make at least a tentative inference that the property might generalize, although they may have been less so in the accidental condition. Nevertheless, there were no significant differences across conditions in whether children viewed the property as relevant and were motivated to produce it themselves. In the pedagogical and intentional conditions, in which equal numbers of 3-year-olds did explore at least one blicket, there were no significant differences in their latency to perform the target action ($M_{pedagogical}$ = 1.83 s, SD = 6.35; $M_{intentional}$ = 7.60 s, SD = 17.82, Mann–Whitney Z = 1.23, p = .22). There were also no significant differences in their exploration of active blickets. Roughly equal numbers of children in pedagogical (8 children, 50%), intentional (8 children, 50%), and accidental (7 children, 44%) conditions explored the active blicket. Moreover, children spent a comparable amount of time exploring the active blicket across the conditions ($M_{pedagogical}$ = 11.69 s, SD = 15.17; $M_{intentional}$ = 13.19 s, SD = 16.34 $M_{accidental}$ = 6.88 s, SD = 16.06; pedagogical vs. intentional: Mann–Whitney Z = 0.14, p = .88; pedagogical vs. accidental: Mann–Whitney Z = 1.47, p = .14; intentional vs. accidental: Mann–Whitney Z = 1.60, p = .11). Finally, there were no significant differences in the number of uses of the active blickets ($M_{pedagogical}$ = 1.31, SD = 1.99; $M_{intentional}$ = 1.06, SD = 1.53 $M_{accidental}$ = 0.50, SD = 1.32; pedagogical vs. intentional: Mann–Whitney Z = 0.13, p = .90; pedagogical vs. accidental: Mann–Whitney Z = 1.52, p = .13; intentional vs. accidental: Mann–Whitney Z = 1.47, p = .14).

As in Experiment 2, the differences in 3-year-olds' exploration were driven entirely by their exploration of the inert blickets, suggesting that they were using the experimenter's social cues to guide their inductive inferences. For 3-year-olds, then, it seems to be that the intentionality of the action influenced their inferences about generalizability and guided their exploratory behavior, rather than whether or not the action was done pedagogically for their benefit.

Discussion

From Experiment 3, we can conclude that by age 4 children modulate the strength of their inferences about the kind-relevance and generalizability of a novel property on the basis of whether it was demonstrated pedagogically for their benefit. Even when the property was produced in an intentional manner, with the experimenter deliberately using the object for that novel function, but in the absence of communicative, pedagogical cues, 4-year-olds made weaker inferences about its generalizability. This confirms that, at least for the older children, the stronger inferences and increased exploration of the inert kind members, now seen here for a third time, are truly the result of children's reasoning about others' communicative and pedagogical intentions.

These results also reveal an interesting developmental difference. Whereas 4-year-olds appear to be conservative in how strongly they will infer that a novel property should extend to other kind members, and do so only when that property is demonstrated pedagogically, 3-year-olds seem to make a strong inference about generalizability simply on the basis of seeing a novel object is used intentionally for a particular function. In essence, it seems that 3-year-olds base such inferences on whether the object is *used* intentionally, regardless of whether that action is done for their benefit, where 4-year-olds make a more fine-grained distinction between such intentional *actions* and acts of intentional, pedagogical *communication*.

General Discussion

The experiments reported here provide initial purchase on the question of how preschool children use pedagogical and intentional cues to guide their acquisition of kind-relevant, generic knowledge about the world. This work synthesizes findings from previous research that document an early sensitivity to pedagogy (Csibra & Volein, 2008; Egyed et al., 2007; Gergely et al., 2007; Yoon et al., 2008), with research on the foundational process of theory-based categorization and concept formation in early childhood (e.g., Booth & Waxman, 2002; Cimpian & Cadena, 2010; Cimpian & Markman, 2009; Gelman & Coley, 1990; Gelman & Markman, 1986, 1987; Gelman et al., 2002; Hollander et al., 2009; Schulz et al., 2008). By age 4, children take pedagogically demonstrated evidence to be generalizable, and make stronger inferences about the generalizability of that evidence, even than perceptually identical evidence produced in a nonpedagogical manner.

Although it is clear that children use intentional and pedagogical cues to guide their inferences about the importance of new information, what exactly are children inferring from seeing a property demonstrated for their benefit? One interpretation of children's behavior in these experiments, and the one that we have argued for in this article, is that they infer that the property being demonstrated is a kind-relevant property that is likely to generalize to other kind members. This inference then guides their continued exploration of the objects in the face of negative evidence. A second interpretation of the data is that children make an inference not simply about the *kind* (e.g., that blickets magnetically pick up paperclips), but rather something about what we as members of a group or society *do* with this kind of object (e.g., that one uses blickets to magnetically pick up paperclips). Children may infer that this is how they *should* interact with these objects, because it is how one is supposed to use them. They thus may continue to try to use the objects in this way even when they find that the additional kind members fail to work. Indeed, these interpretations are not mutually exclusive, and both are consistent with the view that pedagogical demonstration serves to facilitate the transmission of culturally important knowledge (Csibra & Gergely, 2009). It is entirely plausible that children may be making both types of inferences and that together they drive children's continued exploration of the objects.

Whichever interpretation best characterizes children's inferences, these results support the hypothesis that, much as generic language conveys information about the generalizability and conceptual importance of new information (Cimpian & Markman, 2009; Gelman et al., 2002; Hollander et al., 2009; Leslie, 2007), so too may pedagogical demonstration. When presented with identical evidence that a novel object possessed a particular causal property, children made stronger inferences about generalizability when the property was explicitly demonstrated for their benefit, and those inferences drove their exploration in the face of counterevidence. Even when seeing identical evidence produced in an intentional, but noncommunicative and nonpedagogical, manner, by age 4 they made weaker inferences about generalizability that were more easily abandoned in the face of counterevidence. Furthermore, this was true even though the objects in every condition shared a label. Kind labels are known to license category-based inductive inferences, (e.g., Gelman & Markman, 1986), and having shared rather than distinct kind labels does influence exploration in the face of counterevidence (Schulz et al., 2008). Our research suggests that pedagogical demonstration may play an important role in children's developing conceptions of object kinds, above and beyond that of kind labels.

More broadly, what is the role of pedagogy and direct instruction in young children's acquisition of generic knowledge about kinds and categories? The current research shows that children do not merely learn more when they observe something being pedagogically demonstrated—indeed, they seem to *learn* the same thing regardless of how evidence is produced: that this object has a particular function and that other members of the same kind might share it. But children go beyond this, treating information produced pedagogically for their benefit as better evidence for making a kind-based inference, and at least by age 4 they make that inference more strongly than when they simply *observe* someone produce that information even in an intentional manner.

Of course, there are also downsides to such explicit instruction, and it may in some cases be a ''double-edged sword,'' potentially dampening children's natural curiosity and constraining learning to only what is being taught (Bonawitz et al., 2011). The key distinction here is between learning about and exploration of an *individual*, and learning about and exploration of a *kind*. Although pedagogical demonstration may limit children's open-ended exploration, as Bonawitz et al. (2011) suggest, it has the advantage of marking information as kind-relevant, facilitating rapid learning about an object kind that might otherwise take repeated exposure to many members of the kind. As researchers focused on conceptual development have pointed out (Cimpian & Cadena, 2010; Gelman, 2003, 2009; Leslie, 2007), generic knowledge about kinds and categories is never directly observable—one can only observe individuals or collection of individuals, and must make inductive inferences in order to draw conclusions about the kind more generally. In addition to kind-referring language, pedagogical demonstration may provide children with an important source of information about the kind-relevance of new information. And although pedagogical instruction might lead children to miss out on some opportunities for learning more about a *particular* object, it may also help guarantee that the information that does make into children's developing conception of an object *kind* is more likely to support accurate inductive inferences and facilitate a coherent causal understanding of the world. This is manifested in the domain of artifacts, where children's task is not to learn simply everything one *can* do with an object, but rather what such artifacts are *for* (Kelemen, 1999; Kelemen & Carey, 2007).

Additionally, these results yield several intriguing developmental differences. First, older children appear more nuanced in their inductive reasoning. Four-year-olds made the inductive generalization equally regardless of how evidence was produced, but modulated the strength of these inferences on the basis of whether or not the evidence was demonstrated pedagogically. The younger children, in contrast, appeared to use intentional and pedagogical cues to guide *whether* or not they should generalize the property to additional kind members. Second, the results of Experiment 3 suggest that 3-year-olds base their inferences about generalizability solely on whether the object was intentionally used in a manner that produced novel property evidence. In contrast, 4-year-olds were more discriminating, only making a strong inference about generalizability, and showing increased exploration in the face of counterevidence, when the property had been deliberately and pedagogically demonstrated for their benefit.

What might account for this developmental difference? One possible explanation for the developmental difference seen here is a development in the nature of children's artifact representations. It has been shown that during the later preschool years, children start to develop what has been termed a ''design stance'' toward artifacts (see Kelemen & Carey, 2007). Over this period, children begin to understand that the essence of artifacts is what they were *designed* for, and begin to recognize that simply seeing someone use an artifact in a particular way does not necessarily entail that this is its originally intended function (Defeyter & German, 2003; German & Defeyter, 2000; Kelemen, 1999, 2001; Matan & Carey, 2001). Several studies have shown that children begin to grasp this distinction around 4 or 5 years of age (Kelemen, 1999, 2001); 3-year-olds, on the other hand, may not (Kemler Nelson, Herron, & Morris, 2002). In our study, then, it may be that the older children inferred that pedagogically demonstrated information is better evidence for what an artifact kind is for, recognizing that intentional use of an artifact for a particular function does not guarantee that it is for that purpose, whereas the younger children may have failed to make this distinction.

A second possible explanation is a development in children's communicative inference abilities. Although it seems unlikely that children as old as 3 simply fail to recognize when an adult is communicating something for their benefit, given that such recognition has been shown even in very young infants (see Csibra, 2010; Csibra & Gergely, 2009, for a review), it is conceivable that they might be less nuanced in their understanding of communicative interactions. It is worth pointing out that in all conditions, children were pedagogically taught the novel word for the object at the beginning of the study. This was followed by a short distractor task, which served as an endpoint to the stretch of time in which the experimenter was clearly and deliberately teaching the child about the novel object. In a sense, the experimenter attempted to ''deactivate'' the pedagogical interpretation of the situation by making a clear break between the tasks. In the 4-year-olds' case, then, the pedagogical demonstration may have served to ''reactivate'' the pedagogical interpretation of the situation, leading children to interpret the information that followed as important and generalizable. Following this line of reasoning, 3-year-olds may have failed to recognize that the pedagogical situation had been interrupted. This seems plausible given that the distractor task was conducted in a pedagogical manner (showing the child how to make a house by folding colored paper), and thus 3-year-olds may have failed to differentiate between the pedagogical contexts. They thus may have viewed any information coming next, including the intentional use of the object, as teaching done for their benefit, unless and until it was marked otherwise by the experimenter exclaiming, ''Oops!'' in the accidental condition. On this account, the developmental difference might reflect not a change in the inferences that children make on the basis of pedagogical cues, but rather a change in how children demarcate pedagogical situations.

When facing the inductive problem of assessing kind-relevance and generalizability, children have many sources of information available to them, both nonsocial (e.g., observation, exploration, and prior knowledge) and social (e.g., labels, generic language, and intentional and pedagogical cues). Children's ability to integrate these sources of information—especially when they conflict—is an important skill. The current research suggests that this ability is developing during the preschool years, and that by as young as 4 children are particularly sensitive to intentionally communicated information as they form and test hypotheses about the world.

References

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*, 875–890.

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, *20*, 395–418.

Baldwin, D. A., Markman, E. M., & Melartin, R. L. (1993). Infants' ability to draw inferences about nonobvious object properties: Evidence from exploratory play. *Child Development*, *64*, 711–728.

Bonawitz, E. B., Shafto, P., Gweon, H., Goodman, N., Spelke, E., & Schulz, L. E. (2011). The double-edged sword of pedagogy: Teaching limits children's spontaneous exploration and discovery. *Cognition*, *120*, 322–330.

Booth, A. E., & Waxman, S. R. (2002). Word learning is ''smart'': Evidence that conceptual information affects preschoolers' extension of novel words. *Cognition*, *84*, B11–B22.

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.

Cimpian, A., & Cadena, C. (2010). Why are dunkels sticky? Preschoolers infer functionality and intentional creation for artifact properties learned from generic language. *Cognition*, *117*, 62–68.

Cimpian, A., & Markman, E. M. (2008). Preschool children's use of cues to generic meaning. *Cognition*, *107*, 19–53.

Cimpian, A., & Markman, E. M. (2009). Information learned from generic language becomes central to children's biological concepts: Evidence from their open-ended explanations. *Cognition*, *113*, 14–25.

Clark, H. H. (1996). *Using language*. New York: Cambridge University Press.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, *25*, 141–168.

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In Y. Munakata & M. H. Johnson (Eds.), *Processes of change in brain and cognitive development* (Vol. 21, pp. 249–274). New York: Oxford University Press.

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

Csibra, G., & Volein, Á. (2008). Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology*, *26*, 1–11.

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

Egyed, K., Király, I., Krekó, K., Kupán, K., & Gergely, G. (2007, March). *Understanding object-referential attitude expressions in 18-month-olds: The ''Human Pedagogy'' as interpretation switch*. Poster presented at the biennial meeting of the Society for Research in Child Development, Boston.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the USA*, *99*, 9602–9605.

Futó, J., Téglás, E., Csibra, G., & Gergely, G. (2010). Communicative function demonstration induces kind-based artifact representation in preverbal infants. *Cognition*, *117*, 1–8.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.

Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology*, *60*, 115–140.

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology*, *26*, 796–804.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*, 183–209.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, *58*, 1532–1541.

Gelman, S., Star, J., & Flukes, J. (2002). Children's use of generics in inductive inferences. *Journal of Cognition and Development*, *3*, 179–199.

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., Egyed, K., & Király, I. (2007). On pedagogy. *Developmental Science*, *10*, 139–146.

German, T. P., & Defeyter, M. A. (2000). Immunity to functional fixedness in young children. *Psychonomic Bulletin & Review*, *7*, 707–712.

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*, *20*, 347–353.

Graham, S. A., Baker, R. K., & Poulin-Dubois, D. (1998). Infants' expectations about object label reference. *Canadian Journal of Experimental Psychology*, *52*, 103–113.

Grice, H. P. (1957). Meaning. *Philosophical Review*, *66*, 377–388.

Grice, H. P. (1969). Utterer's meaning and intentions. *Philosophical Review*, *78*, 147–177.

Grossmann, T., & Johnson, M. H. (2010). Selective prefrontal cortex responses to joint attention in early infancy. *Biology Letters*, *6*, 540–543.

Grossmann, T., Johnson, M. H., Lloyd-Fox, S., Blasi, A., Deligianni, F., Elwell, C., et al. (2008). Early cortical specialization for face-to-face communication in human infants. *Proceedings of the Royal Society of London, B: Biological Sciecnes*, *275*, 2803–2811.

Harris, P. L. (2002). What do children learn from testimony? In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 316–334). New York: Cambridge University Press.

Harris, P. L., & Koenig, M. A. (2006). Trust in testimony: How children learn about science and religion. *Child Development*, *77*, 505–524.

Hollander, M. A., Gelman, S. A., & Raman, L. (2009). Generic language and judgements about category membership: Can generics highlight properties as central? *Language and Cognitive Processes*, *24*, 481–505.

Kelemen, D. (1999). Why are rocks pointy? Children's preference for teleological explanations of the natural world. *Developmental Psychology*, *35*, 1440–1453.

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. (2007). The essence of artifacts: Developing the design stance. In E. Margolis & S. Laurence (Eds.), *Creations of the mind: Theories of artifacts and their representation* (pp. 212–230). New York: Oxford University Press.

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.

Leslie, S. J. (2007). Generics: Cognition and acquisition. *Philosophical Review*, *117*, 1–47.

Matan, A., & Carey, S. (2001). Developmental changes within the core of artifact concepts. *Cognition*, *78*, 1–26.

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

Schulz, L. E., Standing, H. R., & Bonawitz, E. B. (2008). Word, thought, and deed: The role of object categories in children's inductive inferences and exploratory play. *Developmental Psychology*, *44*, 1266–1276.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*, 668–671.

Senju, A., Csibra, G., & Johnson, M. H. (2008). Understanding the referential nature of looking: Infants' preference for object-directed gaze. *Cognition*, *108*, 303–319.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.

Yoon, J. M. D., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences of the USA*, *105*, 13690–13695.