Journal of Experimental Child Psychology 145 (2016) 64–78

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Two- and 3-year-olds integrate linguistic and pedagogical cues in guiding inductive generalization and exploration

Lucas P. Butler [a,*], Michael Tomasello [b]

[a] Department of Human Development and Quantitative Methodology, University of Maryland, College Park, MD 20742, USA
[b] Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany

ARTICLE INFO

Article history:
Received 10 April 2015
Revised 1 December 2015

Keywords:
Social learning
Pedagogy
Inductive inference
Generalization
Exploration
Cultural learning
Pragmatic reasoning

ABSTRACT

Young children can in principle make generic inferences (e.g., "doffels are magnetic") on the basis of their own individual experience. Recent evidence, however, shows that by 4 years of age children make strong generic inferences on the basis of a single pedagogical demonstration with an individual (e.g., an adult demonstrates for the child that a single "doffel" is magnetic). In the current experiments, we extended this to look at younger children, investigating how the mechanisms underlying this phenomenon are integrated with other aspects of inductive inference during early development. We found that both 2- and 3-year-olds used pedagogical cues to guide such generic inferences, but only so long as the "doffel" was linguistically labeled. In a follow-up study, 3-year-olds, but not 2-year-olds, continued to make this generic inference even if the word "doffel" was uttered incidentally and non-referentially in a context preceding the pedagogical demonstration, thereby simply marking the opportunity to learn about a culturally important category. By 3 years of age, then, young children show a remarkable ability to flexibly combine different sources of culturally relevant information (e.g., linguistic labeling, pedagogy) to make the kinds of generic inferences so central in human cultural learning.

* Corresponding author.
E-mail address: lpbutler@umd.edu (L.P. Butler).

http://dx.doi.org/10.1016/j.jecp.2015.12.001
0022-0965/

Introduction

How do children leverage the knowledge of others around them to guide their own learning about the world? Although much of the empirical work has focused on learning from what adults say, children also learn a lot from observing what others do. Others' actions, and in particular the underlying intentions that guide them, are potentially rich sources of information about the world, supporting inferences about a person's visual access, desires, preferences, and possibly even beliefs (Luo & Baillargeon, 2007; Onishi & Baillargeon, 2005; Woodward, 1998) as well as facilitating imitative learning (Williamson & Meltzoff, 2011).

Moreover, recognizing that the goal behind an action is not merely instrumental (for the actor's own benefit) but rather social (e.g., for the child's benefit) may allow for further inferences about the intended meaning of that communicative act and why an adult is choosing to communicate this information in this particular context (Csibra, 2010; Csibra & Gergely, 2009; Shafto, Goodman, & Frank, 2012; Sperber & Wilson, 1996). Specifically, if the child recognizes that a communicative act is being done for him or her, this indicates that the communicator has the specific social goal that the child attend to and learn this information for some particular reason and, thus, may license a variety of inferences about the relevance, importance, and generalizability of that information. Recent theoretical and empirical work (for reviews, see Csibra & Shamsudheen, 2015; Gergely & Jacob, 2013) has illustrated that an early sensitivity to this subtle difference, between an instrumental action and a deliberate and social demonstration, can have powerful consequences for children's learning. In particular, recognizing actions as deliberate demonstrations may help children to tackle the classic inductive problem of learning broad general knowledge from sparse evidence (Goodman, 1965). Thus, this sensitivity may have the power to shape children's conceptualizations of the world. An important next step, then, is to ask how this process is integrated with other factors that influence children's inductive inferences and how that integration may occur over the course of development. In this study, we focused specifically on one factor that has been shown to play a critical role in children's early inductive inferences—the use of linguistic kind labels—and asked specifically how children may come to integrate pedagogical intent and linguistic labeling to guide their inferences about a novel object kind. Before turning to our experiments, it is important to briefly review the evidence for how each of these factors affects inductive inference in order to provide a clear roadmap for the current research.

Sensitivity to pedagogical intent

Preverbal infants are sensitive to cues that indicate whether or not actions are being carried out communicatively, and this sensitivity affects their encoding and memory for object features (Yoon, Johnson, & Csibra, 2008), their categorization of novel objects (Futó, Téglás, Csibra, & Gergely, 2010; Hernik & Csibra, 2015), and their expectations about whether knowledge will be widely shared by other people (Egyed, Király, & Gergely, 2013). In older children, this distinction has the power to influence their active inductive inferences—both their judgments about what properties are central and defining of a category (Butler & Markman, 2014b) and their inferences about and exploration of the generalizability of novel object properties (Butler & Markman, 2012, 2014a).

Moreover, preschoolers may take an actor's choice to deliberately demonstrate an object property as indicating that this is the only function (Bonawitz et al., 2011) or the normatively correct function (Vredenburgh, Kushnir, & Casasola, 2015) of a particular object. Finally, when children already know the information being communicated, they use both an actor's demonstrations and the actor's deliberate omissions of information (when an actor fails to demonstrate what he or she could have) to make third-party judgments of whether that demonstrator is being informative and helpful (Gweon, Pelton, Konopka, & Schulz, 2014). Thus, sensitivity to whether or not actions are carried out with social pedagogical goals has the potential to have major impacts on a variety of aspects of children's inductive inference process.

Kind labels and inductive inferences

Long before this work on pedagogical cues, decades of literature have illustrated the power of linguistic kind labels to shape the way in which children make inferences from novel information. Most important, using a label signals that an individual belongs to a broader kind or category and supports inductive inferences, especially about nonobvious properties, on the basis of shared membership in that category. By 2 or 3 years of age, children take labels as referring not to loose collections of individuals but rather to kinds that share and may be defined by nonobvious properties, and they generalize information learned about an individual of a labeled kind to other individuals that share that kind membership even when kind membership conflicts with perceptual similarity (Booth & Waxman, 2002; Gelman, 2003; Gelman & Coley, 1990; Gelman & Markman, 1986; Markman, 1989). And preschoolers expect novel objects that share a label to also share invisible internal causal properties. When they learn that an individual of a labeled kind has a novel causal property, they go on to selectively explore other kind members more when they fail to have that property (Schulz, Standing, & Bonawitz, 2008). More broadly, hearing labels used generically (e.g., "birds fly") allows children to directly learn properties of entire categories (Gelman, 2004) and to infer that a given property is centrally important to the nature of that category (Cimpian & Cadena, 2010; Cimpian & Markman, 2009). Taken together, this work suggests that children use kind labels both to demark the extension of a category (even above and beyond perceptual similarity) and to guide their inferences about the importance and generalizability of novel information about an individual to the kind as a whole. Given that this process bears a good deal of similarity to the process by which pedagogical demonstration affects children's inferences about novel properties, it is important to ask how these processes might interact.

How pedagogical cues and linguistic labeling might interact

Linguistic labeling and pedagogical demonstration could theoretically interact in at least two different ways. First, labeling might facilitate inductive inference from pedagogical demonstration simply by clearly demarking the category members and thereby specifying the possible extension of any demonstrated property. In this scenario, labeling specifies the category over which inductive inferences ought to be made, thereby allowing pedagogical demonstration of a property to boost its inferred generalizability or importance to that specific category. Second, linguistic labeling may serve not only to specify what the category is but also as a cue that a category is culturally relevant, leading children to be especially attuned to social cues that might indicate whether or not a particular piece of information is in fact important.

The first of these possibilities is clear; pedagogical cues can boost the strength of children's inferences only if children know what objects that category includes—a task that labels greatly facilitate. The second of these possibilities, however, requires more explication. Even children as young as 2 years recognize that conventional labels are applied to kinds that others in a group or culture are likely to know about (Henderson & Graham, 2005), and recognizing the importance of conventionality plays a crucial role in word learning (Sabbagh & Henderson, 2007). Thus, hearing a novel linguistic label indicates the presence of a category that is not merely a collection of individuals but rather a culturally relevant category that members of the speaker's group are likely to know about. Furthermore, between 2 and 5 years of age, children come to appreciate the fact that the essential feature that defines artifact categories (as opposed to natural kinds) is the purpose or function for which it was designed (German & Johnson, 2002; Kelemen, 1999; Kelemen & Carey, 2007). Even children as young as 2 or 3 years will extend novel labels on the basis of shared function over and above shared perceptual features (Kemler Nelson, 2000; Kemler Nelson, Russell, Duke, & Jones, 2000).

Thus, we know that between 2 and 5 years of age children come to understand that linguistic labels frequently refer to culturally relevant artifact kinds that are defined by their intended functions and that other members of their group or culture likely share knowledge about that category. We also know that children face an inductive problem in reasoning about which possible functions of a novel artifact category are truly defining and which are not (Bloom, 1996; Dennett, 1989; German & Johnson, 2002; Matan & Carey, 2001). From prior work on pedagogical demonstration, we also know that before 2 years of age children understand that pedagogical cues signal information that others are

likely to share (Egyed et al., 2013) and that by 4 years of age they can use pedagogical cues to help gauge whether or not a novel function is in fact defining of an artifact category (Butler & Markman, 2014b) or is the normatively "correct" way in which to use a member of that category (Vredenburgh et al., 2015). Based on these prior findings, we hypothesized that hearing a linguistic label for a novel artifact category may lead children to be especially attuned to what object properties a knowledgeable adult explicitly demonstrates for their benefit and what object properties the adult chooses not to and that they use this distinction in order to help them solve the inductive problem of exactly what is important and defining for that category. Specifically, children may constrain their strong inferences about what is important to the category to only those artifact properties that are explicitly demonstrated for their benefit.

It is important to note that these two possibilities are not mutually exclusive. Kind labels have the potential to do both of these things—to demark the extension of the category and to signal the opportunity to learn what defines a culturally relevant category. The extent to which one or the other role of labeling comes into play is likely to depend both on the context and on children's age, as we discuss more below when we delve into the question of how this interaction between pedagogical cues and kind labels might change over the course of development.

The question of development

As we turn to the question of development, it is worth first noting that there are at least two somewhat different perspectives on the mechanism underlying the effects of pedagogical demonstration, possibly due in part to the difference in ages that have been focused on in different studies. One perspective, coming primarily from research with infants, is that the sensitivity to cues that someone is deliberately communicating information is innate and has been selected for in evolution in order to allow for the cultural transmission of important, often opaque information about the world (Csibra & Gergely, 2009). On this view, infants have a set of innate capacities and biases that support learning from others—the capacity to recognize communicative cues, the expectation that those cues are referential, and a bias to interpret information accompanied by those cues as kind relevant and generalizable as opposed to idiosyncratic or episodic. Over the course of development, this latter bias may result in pedagogical demonstrations being treated as a type of "nonverbal generic" (Csibra & Shamsudheen, 2015) that adults use to directly teach broad generalizable information about a kind as a whole by demonstrating information about a single exemplar. Absent clear cues to the contrary, children may see a pedagogical demonstration as likely conveying generic information about a kind.

Another perspective, growing more out of the research with older children, emphasizes the importance of social pragmatic reasoning in making inferences from pedagogical demonstration. This perspective suggests that children recognize communicative or pedagogical cues as signaling the social goal of conveying information—and that this recognition on the basis of simple perceptual cues may indeed be innate—but that evaluating the intended meaning of a communicative act involves a pragmatic inference based on a variety of contextual factors, including individual and shared background knowledge or common ground as well as statistical evidence (Landrum, Eaves, & Shafto, 2015; Shafto & Goodman, 2008; Shafto, Goodman, & Griffiths, 2014; Shafto et al., 2012). Indeed, there is some evidence for this. Although pedagogical cues can lead to radical reorganization of children's category knowledge, this depends on the type of property being demonstrated; if an adult pedagogically demonstrates a superficial property (e.g., a sticker on the bottom of an object), children do not take this as an important defining characteristic of this category (Butler & Markman, 2014b).

These perspectives are not mutually exclusive. Rather, their differences may largely be an issue of perspective and emphasis. The infant evidence, which by and large involves infants far too young to engage in the type of sophisticated pragmatic reasoning that even preschoolers often struggle with, is consistent with the proposal of an innate sensitivity to and default interpretation of communicative cues. The social pragmatic perspective does not conflict with this proposal but rather suggests that over the course of development this potentially innate sensitivity and interpretation bias necessarily must be integrated with other cognitive and social cognitive reasoning skills, leading to a more inferential nuanced learning mechanism later during childhood. Our goal in the current study was

not to argue for one perspective or the other but rather to take a first step in directly investigating this integration process during development.

In terms of how this integration specifically with linguistic labeling might change developmentally, recall that linguistic labeling can theoretically do at least two different things: demark the extension of a category and cue the opportunity to learn culturally relevant information. This second function depends on the first; a child first needs to know the category to which information might generalize before being able to make inductive inferences about that information. In cases where the extension of the category is not clear, the role of the label in demarking the category may be primary. But in cases where objects clearly do appear to belong to the same category, the secondary role of the kind label in signaling cultural relevance might come into play. In terms of development, it may well be that younger children are more reliant on the kind label even just to demark a kind even given perceptual similarity. In contrast, older children may more readily and easily form such kind representations, even absent explicit labeling, leaving open the possibility that they would recognize labels as also cuing the opportunity to learn important information about a culturally relevant kind.

The current research

In two experiments, we asked how linguistic labeling might affect young children's use of pedagogical cues to guide their inferences and how this might change over development. Building on prior work documenting the effect of pedagogical demonstration on inductive inference and exploration as young as 4 years of age (Butler & Markman, 2012) and possibly younger (Butler & Markman, 2014a), we presented 2- and 3-year-olds with evidence that a novel object was magnetic. This evidence was always perceptually identical but was produced pedagogically, intentionally, or accidentally. As in Butler and Markman (2012) and in earlier work (Baldwin, Markman, & Melartin, 1993; Gweon, Tenenbaum, & Schulz, 2010; Schulz et al., 2008), children were then given a set of identical but nonfunctional objects and their continued exploration of those objects was assessed as an index of how strongly they generalized that property to the other kind members. Critically, we also varied whether and how the novel objects were labeled. Children never heard the demonstration object being directly labeled. In Experiment 1, children either heard or did not hear a novel label ("doffel") uttered in an offhand non-ostensive manner prior to the demonstration, and then they either heard the additional exploration objects referred to by a shared novel label ("doffels") or simply heard them referred to as "more of those." Experiment 2 sought to further pin down the role of labeling. Children never heard any of the objects being explicitly labeled but rather either never heard a novel label at all or heard one only in passing and at the beginning of the experiment.

Based on our discussion above, we had several key predictions about the way in which labels and pedagogical cues might interact to influence children's inductive inferences in these experiments. First, we predicted that the presence of a linguistic label and whether or not information about a novel object is pedagogically demonstrated for the children would interact in driving their inferences. Specifically, we expected that young children would use the distinction between pedagogical and intentional actions in guiding their inductive generalizations about a novel object property only in the presence of a clear kind label. Second, we predicted that this effect might vary depending on the age of the children. Specifically, we expected that younger children, being reliant on the label even to make clear the extension of a novel kind, might show evidence of making strong inductive generalizations only when both pedagogical and linguistic cues are present. In contrast, older children might have higher default levels of generalization, regardless of labeling, but would still show a distinction between pedagogical and intentional actions in guiding their inferences only in the presence of a label.

Experiment 1

Method

Participants

The participants were 96 2-year-olds (M = 2.52 years, range = 2.45–2.71) and 96 3-year-olds (M = 3.54 years, range = 3.34–3.83) from a medium-sized German city. Children were randomly

assigned, 16 per condition (8 male and 8 female), equating for gender and age across conditions. Children were predominantly White and middle to upper middle class, although a variety of ethnicities and socioeconomic statuses were represented. An additional 22 children participated in the study but were excluded because of experimenter error ($n = 9$) or parental interference ($n = 11$) or because they were not attending during the demonstration ($n = 2$).

Materials

The novel objects were four cardboard tubes 3 cm in diameter $\times$ 7 cm tall. The active object was magnetic on one end, whereas the inert objects were not. All were covered with a red decorative cover with yellow electrical tape covering the magnetic or nonmagnetic end. All four objects were perceptually indistinguishable.

Procedure

The procedure was based on that used in Butler and Markman (2012, 2014b), with identical evidence produced pedagogically, intentionally, or accidentally. All children were tested in German in a private room in the lab by a trained experimenter. A parent or guardian for each child was present at all times. English glosses are provided here, but all children were tested in German.

The experimenter first warmed up with the child in a separate room and after approximately 10 min told the child that they could continue their play in another room. On entering the test room, the table was already set up with age-appropriate toys (plastic cups, rings, and a stacking pole) as well as a pile of paperclips and the active object. The experimenter asked the child to sit at the table. Then, the experimenter appeared to "notice" what was on the table and exclaimed, "Oh, what's this? Someone left the table messy! That won't do. They left the cups out, and the rings." In the *label* conditions, the experimenter then said, "and the doffel too" but did not draw attention to the object. This was the only time the child heard the novel word until after the demonstration. This cover story served two purposes. First, it allowed the experimenter (in the label conditions) to expose children to the word for the novel object in a non-pedagogical manner. Second, it established that the experimenter was simply cleaning up so that they could eventually continue their previous game, thereby removing any general expectation of being taught something by the experimenter. This allowed for a clean manipulation of pedagogical intent in that whatever general pedagogical expectations children might have about the interaction that could potentially interfere with the effect of the manipulation were stripped away (see Butler & Markman, 2014a).

While remaining standing, the experimenter then grabbed a bucket from the floor and began cleaning up. She put away each distracter object and then picked up the active object. In the *pedagogical* condition, she said "Look!" and then deliberately placed the object on the paperclips, picking it up with paperclips attached. In the *intentional* condition, the experimenter carried out the identical action but never made eye contact or otherwise interacted with the child. In the *accidental* condition, the experimenter appeared to accidentally drop the object on the paperclips as she was putting it away, exclaiming "Oops!" In all three conditions, she then looked at it and said "Cool!" and then gave the child the object to play with, saying "here."

The experimenter then brought out the three additional (inert) objects and placed them on the table. In the label conditions, the experimenter then said, "Here are some more doffels. You can play with them too. I need to quickly write something down. Go ahead and play. I'll be right back." The *no-label* conditions were the same except that the experimenter did not refer to the objects by name, saying simply, "Here are some more of those."

The child was then left alone at the table to play for 90 s. The parent was given a magazine to read and was instructed not to talk to the child during this period. The experimenter did not interact with the child during this period other than to say, "You can go ahead and play; I'll be right back" if the child got up to see what she was doing.

Coding and analysis

All sessions were videotaped and coded by a single observer. A second independent observer coded a random sample of 25% of all sessions for reliability. Both coders were blind to condition. Because our interest was not in whether children generalized but rather in the strength of children's inductive

generalizations in the face of counterevidence (see Butler & Markman, 2012), our key dependent measure was the number of attempts children made to elicit the novel property from the inert objects. The logic of this measure is that children's exploratory behavior, especially in the face of potentially surprising evidence (e.g., that an object fails to have an expected property), provides a window into their conceptualization of this object kind (see also Baldwin et al., 1993; Gweon et al., 2010; Schulz et al., 2008). This gives us an index of how resistant children's inferences were to counterevidence because it measures precisely how many individual pieces of negative evidence were required on average before children abandoned their generalization (see Butler & Markman, 2012). An attempt was coded as any intentional action done to pick up paperclips, including both placing the novel object on the paperclips and placing a paperclip on the object. Reliability for this measure was very good, with Cronbach's $\alpha = .86$.

We also coded where children were looking during the key action that produced the evidence (defined as the time from when the adult first picked up the novel object to when she gave it to the children after picking up paperclips with it). Children were coded as "attentive" if they were looking either at the object or at the experimenter during this time. All but 2 children attended during this time, and these children were not included in the final analysis. This ensured that all children saw the key action that produced the novel property.

We first used both visual inspection and Kolmogorov–Smirnov and Levene's tests to confirm that the dependent measure (mean number of attempts with the inert objects) satisfied the criteria for analysis of variance (ANOVA). Having confirmed this, we then analyzed children's exploration across the entire sample using a 3 (Condition) $\times$ 2 (Label) $\times$ 2 (Age Group) ANOVA, with mean attempts to elicit the property from the inert objects as the dependent variable. To protect against the issues involved with multiple comparison, we first compared this full model with a null model including only the intercept using a likelihood ratio test, and this comparison was highly significant ($p < .001$). This has been shown to be an effective way of keeping Type I error rates low while still allowing for later planned post hoc comparisons (Forstmeier & Schielzeth, 2011). In the full model, the key prediction was that there would be a Condition $\times$ Label interaction, with children using the distinction between pedagogical and intentional actions in guiding their inferences when there was a label but not without a label. Thus, we specifically tested for this Condition $\times$ Label interaction in the analysis. We then followed up this analysis with planned pairwise comparisons between conditions within each of the four cells of this design, based on predictions from prior work (Butler & Markman, 2012, 2014a).[1]

Results

The overall analysis revealed significant main effects of both condition, $F(2,180) = 25.5$, $p < .001$, partial $\eta^2 = .221$, and age group, $F(1,180) = 18.7$, $p < .001$, partial $\eta^2 = .094$, and a significant Condition $\times$ Label interaction, $F(2,180) = 4.20$, $p = .017$, partial $\eta^2 = .045$. These results were clear (see Fig. 1). As predicted, pedagogical demonstration and labeling significantly interacted. When the novel category was referred to by a shared novel label, children at both ages made significantly stronger inferences about the novel property when it had been pedagogically demonstrated for their benefit, even compared with seeing the identical action produce the function in an instrumental but noncommunicative manner. In contrast, children showed no such difference when the category was not labeled.

Our follow-up planned pairwise comparisons within each cell of the design confirm this picture. At both age groups, children used pedagogical cues to guide the strength of their generalizations only when the category they were learning about had been referred to using a shared kind label.

Two-year-olds with label

Consistent with previous findings with older children, we found that 2-year-olds required significantly more evidence to discard their original generalization about a labeled object category when the

[1] We also conducted these pairwise comparisons using the false discovery rate procedure to adjust *p*-values. All significant comparisons remained significant using this approach, which has been shown to be less over-conservative than standard Bonferroni corrections (Benjamini, 2010; Benjamini & Hochberg, 1995; Shaffer, 1995). We report the original *p*-values.

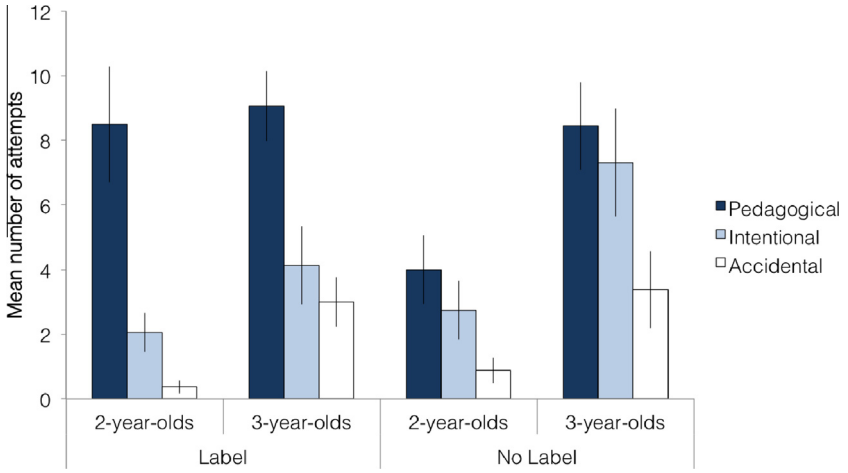

Fig. 1. Children's persistent attempts to elicit the novel property from the inert objects across condition, label group, and age group in Experiment 1. Error bars represent ± 1 standard error.

property was demonstrated pedagogically for their benefit. They made significantly more attempts to elicit the property from the inert objects in the pedagogical condition ($M = 8.50$, $SD = 7.16$) than in either the intentional condition ($M = 2.06$, $SD = 2.41$), $t(30) = 3.41$, $p = .002$, $d = 1.21$, or the accidental condition ($M = 0.38$, $SD = 0.81$), $t(30) = 4.51$, $p < .001$, $d = 1.59$. They also made significantly more attempts in the intentional condition than in the accidental condition, $t(30) = 2.66$, $p = .012$, $d = 0.934$.

Two-year-olds with no label

When the novel category was not labeled, children showed a very different pattern. Without a label, 2-year-olds made a comparable number of attempts to elicit the property from the inert objects regardless of whether the property was produced pedagogically ($M = 4.00$, $SD = 4.26$) or intentionally ($M = 2.75$, $SD = 3.64$), $t(30) = 0.89$, $p = .379$, $d = 0.315$. They made significantly more attempts in the pedagogical condition than in the accidental condition ($M = 0.88$, $SD = 1.59$), $t(30) = 2.75$, $p = .010$, $d = 0.970$, and made marginally more attempts in the intentional condition than in the accidental condition, $t(30) = 1.88$, $p = .069$, $d = 0.666$.

Three-year-olds with label

Three-year-olds also required significantly more evidence to discard their original generalization about a labeled object category when the property was demonstrated pedagogically for their benefit. They made significantly more attempts to elicit the property from the inert objects in the pedagogical condition ($M = 9.06$, $SD = 4.34$) than in either the intentional condition ($M = 4.13$, $SD = 4.84$), $t(30) = 3.04$, $p = .005$, $d = 1.07$, or the accidental condition ($M = 3.00$, $SD = 3.06$), $t(30) = 4.57$, $p < .001$, $d = 1.61$. There was no difference between the intentional and accidental conditions, $t(30) = 0.79$, $p = .438$, $d = 0.279$.

Three-year-olds with no label

This pattern did not hold with children for whom the novel object category was not labeled. Without a label, 3-year-olds made a comparable number of attempts to elicit the property from the inert objects regardless of whether the property was produced pedagogically ($M = 8.44$, $SD = 5.41$) or intentionally ($M = 7.31$, $SD = 6.69$), $t(30) = 0.52$, $p = .605$, $d = 0.186$. They made significantly more attempts in the pedagogical condition than in the accidental condition ($M = 3.38$, $SD = 4.77$), Welch's $t(30) = 2.81$, $p = .009$, $d = 0.992$, and made marginally more attempts in the intentional condition than in the accidental condition, $t(30) = 1.92$, $p = .065$, $d = 0.676$.

Discussion

As we predicted, 2- and 3-year-olds used both linguistic cues to category membership and communicative cues to pedagogical intent to guide their reasoning about whether novel information about an individual was truly important generalizable knowledge about a category—a critical task in conceptual development. Only when they heard the novel category referred to by a shared label did children use pedagogical cues to make strong inferences about the generalizability of the demonstrated function. Otherwise, children made equally strong inferences regardless of whether that function was demonstrated pedagogically or merely used in an instrumental intentional manner. Thus, this experiment provides initial evidence that the extent to which pedagogical cues influence children's inductive inferences about kinds interacts meaningfully with another key factor that we know influences those inferences—whether or not an individual is linguistically labeled as a member of a kind.

However, the way in which labels affect this process is not yet clear. As discussed previously, labels could influence this process in at least two possible ways. One is that labeling defines the scope of the category. That is, without knowing that the objects belong to the same kind, the pedagogical demonstration may have limited effect on children's inferences about generalizability because it is unclear to children what that generalization ought to apply to. On the other hand, having a label clearly demark the kind allows for pedagogical demonstration to strengthen children's inferences about that kind. Another way in which labels could influence this process is by signaling to children that there is an opportunity to learn important information about a culturally relevant kind, thereby leading them to be particularly sensitive to whether or not evidence is demonstrated for their benefit in making inferences about that evidence.

The first way in which labels could influence children's use of pedagogical cues to guide their inferences—by defining the extension of the category—seems to be particularly consistent with what the 2-year-olds did in this experiment. Here, 2-year-olds made strong generalizations, exploring more when the inert objects failed to exhibit the novel property, only when both a linguistic label and a pedagogical demonstration were present. In the no-label version, they made relatively weak inferences even in the pedagogical condition, suggesting that they needed the label to define the category before they could use the pedagogical cue to gauge the novel property's generalizability. However, this does not seem to fit with the 3-year-olds' behavior. In the no-label version, 3-year-olds made relatively strong generalizations, exploring the inert objects more, in both the intentional and pedagogical conditions.

Rather, it appears that for 3-year-olds the relatively high levels of generalization seen in the no-label intentional condition might reflect more of a baseline level of generalization. Indeed, this is consistent with research showing young children's default tendencies to generalize, and in some cases over-generalize, information about individuals to categorize (Sutherland, Cimpian, Leslie, & Gelman, 2015). In this case, it seems that the effect is one in which having a label leads children to constrain their inferences from an intentional (and non-pedagogical) action. Why might this be? As we discussed in the Introduction, given that children have heard a novel kind label, they know that there is a culturally relevant kind (in particular an artifact kind given the objects present) about which they now have the opportunity to learn. Furthermore, they are faced with the inductive problem of learning what properties are defining for that novel artifact category. Given that children can use pedagogical cues to guide this type of reasoning (Butler & Markman, 2014b), they may be particularly attuned to what the adult chooses to demonstrate and what the adult chooses not to demonstrate. This makes the adult's failure to remark on the novel object's potentially interesting property relevant. Children might reason along the lines of the following: "That seems like an interesting property, but it may not be a particularly important property of doffels or she probably would have mentioned something." Thus, when children see that the additional doffels do not share that property, they can infer that the experimenter likely did not point out that property explicitly because it is not something that is important or generalizable to that novel kind. It is important to note here that we are not speaking of "obligation" to teach; rather, given that children know that there is a potentially culturally relevant kind to be learned about, the adult's manner of interacting with an exemplar of that kind is informative—both what the adult does and what the adult chooses not to do. Indeed, there are several examples of children using such logic in the literature, including making scalar implicatures (in the classic

case inferring that "She could have said all, but she said some, so she must not mean all"), inferring possible causal functions (in Bonawitz et al., 2011, inferring that "She could have shown me more things about the toy, but she didn't, so that must be all that is important"), and the "nominal pass-over effect in language" (in Merriman, Marazita, Lorna, Evey-Burkey, & Biggins, 1995, inferring that "She could have labeled that one a blicket too, but she didn't, so it must not be one").

This explanation—that 2-year-olds required the label to define the extension of the category and, thus, open the door to the influence of pedagogical cues, whereas the label instead served to constrain 3-year-olds' inferences in the case of a deliberately non-pedagogical action—is consistent with the current results. However, because children in the label version heard the novel label both before the demonstration, when it was not directly applied to any particular objects, and after the demonstration, when it was used to refer to the three exploration objects, the results of Experiment 1 could potentially support either interpretation. It remains possible that the label primarily served to demark the category even for the 3-year-olds because it was specifically applied to the inert objects prior to the explanation period. This seems unlikely because this account ought to predict a main effect in which labeling strengthens generalizations across all conditions, with a potential additional boost from pedagogical demonstration, whereas the current experiment found no main effect of label but rather an interaction. To test these two alternatives more directly, in Experiment 2 we eliminated the direct labeling of the additional inert objects. Instead, children simply heard the novel label only at the outset of the procedure when the experimenter remarked on what was on the table, and it was never directly paired with any of the novel objects. If children are relying on the label to demark the extension of the category, they should fail to show any difference between the pedagogical and intentional conditions in this new labeling version. In contrast, if children are not relying on the label to demark the extension of the category but are instead using the label as a general cue to the opportunity to learn about a culturally relevant category, simply hearing the label at the outset of the study should be sufficient for them to now distinguish between the pedagogical and intentional conditions in this case.

Experiment 2

In Experiment 1, the presence of a label affected whether and how the manner of demonstration affected children's inductive inferences about and exploration of a novel property. But because children heard a novel label both before the demonstration and after, right before the exploration period, this leaves open the possibility that the effect of the label was not, as we claim, to signal the opportunity to learn culturally relevant information but rather simply to boost learning from pedagogical demonstration by demarking the category. Either account could be supported by the results of Experiment 1. In addition, the results pointed toward a possible developmental difference in terms of the nature of this effect. Two-year-olds in Experiment 1 made strong inferences only when both labels and pedagogical cues were present, suggesting that they may well be simply relying on labeling to demark the extension of the novel category. In contrast, 3-year-olds in the no-label versions made relatively strong inferences in both the pedagogical and intentional conditions, and they constrained their inferences from intentional but non-pedagogical action only when they had heard a label applied to the novel objects. As we have explained previously, this pattern is more consistent with our account of labeling acting as a cue to cultural relevance, thereby influencing whether or not children view the experimenter's choice of what to demonstrate and what not to demonstrate as relevant to their inference process.

In Experiment 2, children heard a novel label but only at the very beginning of the procedure, well before the demonstration, and in a non-ostensive manner that did not draw any attention to the novel object. If children still show stronger inferences in the pedagogical condition, this suggests that the pedagogical cues highlight the opportunity to learn culturally relevant information, thereby leading children to make different inferences given pedagogical demonstration or intentional action. In contrast, if children do not show stronger inference in the pedagogical condition when the label comes only at the beginning of the study, for those children labeling may simply demark the category, facilitating generalization and opening the door for stronger inferences on the basis of pedagogical

actions. Based on the pattern of inference and exploration in Experiment 1, we predicted that 3-year-olds, but not 2-year-olds, would show the pedagogical and intentional effects in this version of the labeling procedure. This would suggest that although labels influence the role of pedagogical demonstration in children's inductive inferences at both ages, the nature of this effect changes with development.

Method

Participants

The participants were an additional 32 2-year-olds (M = 2.49 years, range = 2.43–2.46) and 32 3-year-olds (M = 3.49 years, range = 3.38–3.61) from a medium-sized German city. Children were randomly assigned, 16 per condition (8 male and 8 female), equating for gender and age across conditions. Children were predominantly White and middle to upper middle class, although a variety of ethnicities and socioeconomic statuses were represented. An additional 11 children participated in the study but were excluded because of experimenter error (n = 5), parental interference (n = 5), or video malfunction (n = 1).

Materials

The materials were identical to those used in Experiment 1.

Procedure

Children were tested in either the pedagogical or intentional condition in a new label version. This procedure was identical to the label version used in Experiment 1 with one exception. When the experimenter brought out the additional objects for the child to play with, she said, "Here are some more of those," as she had done in the no-label version of Experiment 1. This was done so that children never heard the novel label directly associated with or applied to the novel objects themselves. The only time children heard the label was at the beginning of the experiment when the experimenter was saying that the table was a mess and that someone else must have left their toys out. Attention was never drawn to the novel object during this phase of the experiment.

Coding and analysis

As in Experiment 1, all sessions were videotaped and coded by a single observer. A second independent observer coded a random sample of 25% of all sessions for reliability. Both coders were blind to condition. Our key dependent measure was the number of attempts children made to elicit the novel property from the inert objects. Reliability for this measure was again very good, with Cronbach's α = .83. As in Experiment 1, we also coded where children were looking during the key action that produced the evidence. All children attended during this time, ensuring that they saw the key action that produced the novel property. As in Experiment 1, we first used both visual inspection and Kolmogorov–Smirnov and Levene's tests to confirm that the dependent measure (mean number of attempts with the inert objects) satisfied the criteria for ANOVA. Having confirmed this, we then analyzed children's exploration across the entire sample using a 2 (Condition) $\times$ 2 (Age Group) ANOVA, with mean attempts to elicit the property from the inert objects as the dependent variable. To protect against the issues involved with multiple comparison, we first compared this full model with a null model including only the intercept using a likelihood ratio test, and this comparison was again highly significant (p < .001). As discussed above, this has been shown to be an effective way of keeping Type I error rates low while still allowing for later planned post hoc comparisons (Forstmeier & Schielzeth, 2011). Given our a priori hypotheses based on Experiment 1, we then followed up this analysis with planned pairwise comparisons within each age group using independent-samples t-tests.

Results

The ANOVA revealed a significant effect of condition, $F(1,60) = 17.4$, p < .001, partial η^2 = .221, and a significant Condition $\times$ Age Group interaction, $F(1,60) = 7.20$, p = .009, partial η^2 = .107. As predicted, 3-year-olds showed the same effect of pedagogical demonstration as shown previously despite

hearing the label produced non-ostensively only at the very beginning of the study, disconnected from the demonstration events and not referentially directed toward the novel object (see Fig. 2). Three-year-olds made significantly more attempts to elicit the property from the inert objects in the pedagogical condition ($M = 8.75$, $SD = 5.80$) than in the Intentional condition ($M = 1.83$, $SD = 3.17$), $t(30) = 4.33$, $p < .001$, $d = 1.48$. In contrast, 2-year-olds did not show this effect, attempting to elicit the property from the inert objects in the pedagogical condition ($M = 4.13$, $SD = 3.81$) at levels that did not differ significantly from those in the intentional condition ($M = 2.63$, $SD = 3.16$), $t(30) = 1.21$, $p = .235$, $d = 0.428$.

Discussion

As predicted,, 2-year-olds failed to show a clear effect of pedagogical demonstration when the label was heard only at the outset of the experiment and not specifically applied to the novel objects. This suggests that for 2-year-olds the effect of labeling is primarily to demark the extension of the category, which then opens the door to pedagogical demonstration boosting their inductive inferences. Without this clear labeling, they do not make strong inductive generalizations even given clear pedagogical cues. In contrast, 3-year-olds made stronger generalizations on the basis of pedagogical demonstration than on the basis of identical but non-pedagogical intentional action even when they simply heard a label disconnected from its referent. This supports the idea that for them the label signaled the opportunity to learn something important about a culturally relevant category, thereby leading them to attend to whether or not evidence was demonstrated for their benefit in making inferences about that evidence.

General discussion

Taken together, these two experiments shed further light on the nuanced way in which pedagogical cues affect young children's inductive inferences and exploration. At both ages, children integrated pedagogical and linguistic cues in making inductive inferences about novel object properties. Only when a label was present did children use the distinction between pedagogical and intentional actions to guide their inferences about the generalizability of a novel property. This is an important contribution to our understanding of how subtle social cues affect children's inductive inferences because it begins to illustrate the way in which possibly innate sensitivity to intentional communication and demonstration is built on and integrated with other important factors that children use to guide their learning.

The current results also point toward a potentially intriguing developmental difference. In Experiment 1, children at both ages distinguished between pedagogical and instrumental actions only in guiding their inferences about a previously labeled category. But the nature of this effect was quite different for the two age groups. Consistent with previous evidence (e.g., Gelman & Coley, 1990), 2-year-olds' inferences were relatively weak in both the intentional and pedagogical conditions when the category had not been labeled. Adding a shared label greatly strengthened their inferences in the pedagogical condition but did not appear to boost the strength of their inferences in the other two conditions. In contrast, in the no-label version, 3-year-olds' inferences were relatively strong in the pedagogical and intentional conditions, whereas these children appeared to attenuate their inferences in the intentional condition of the label version, suggesting that the mechanism underlying this effect may well differ across the age groups. The results of Experiment 2 clarify this developmental difference. Two-year-olds appear to rely on linguistic labels to demark the extension of a novel category, and they are able to use pedagogical cues to make stronger inferences about a novel object property and its generalizability to the novel category only when they have heard a label applied to the members of that category. In contrast, 3-year-olds appear to be doing something different. In the absence of any label, they readily generalize novel object properties to perceptually similar objects regardless of whether or not that property was explicitly demonstrated for their benefit. But when a label is present, even if it was not specifically applied to the members of the novel category, they make inferences based on whether or not a novel property was explicitly demonstrated for them and constrain those

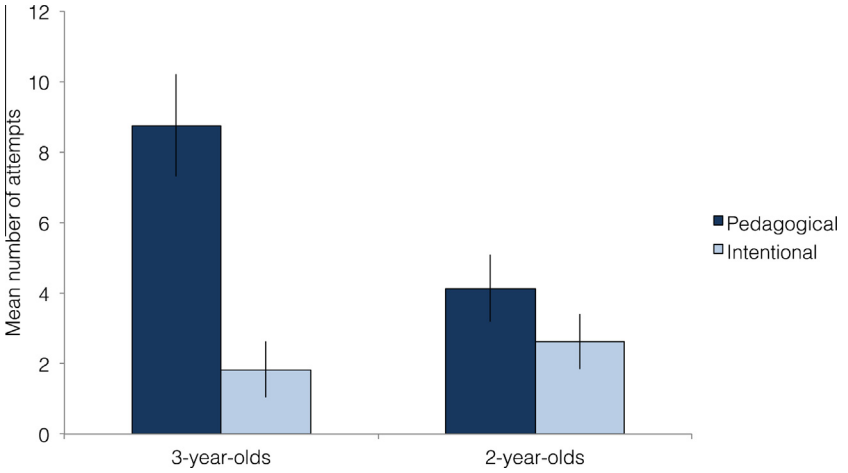

Fig. 2. Children's persistent attempts to elicit the novel property from the inert objects across condition, label group, and age group in Experiment 2. Error bars represent ± 1 standard error.

inferences when an experimenter chose not to explicitly demonstrate a potentially interesting object property.

More broadly, this pattern of results provides initial evidence for a developing flexibility in young children's use of pedagogical cues. Attending to pedagogical cues and using them to guide inductive inference may be a key learning mechanism for evaluating novel information. But such a learning mechanism can be central to children's learning only insofar as children apply it selectively and flexibly. Although parents can deliberately communicate information to help children learn general knowledge about the world, this capacity would undoubtedly lead children astray if they applied it to all communicative interactions. Rather, children must integrate their sensitivity to such cues with their broader understanding of how communication and pragmatics function in social interactions. A key theoretical question then, is whether children can use their pragmatic understanding to lead them to selectively make different inductive inferences on the basis of the same pedagogical demonstration. That is, can they flexibly and selectively learn from pedagogy? The current results suggest that they can, at least by 3 years of age, using the presence or absence of a shared kind label to modulate their use of pedagogical cues in making inductive inferences.

These results suggest that by 3 years of age, children are judiciously using pedagogical cues to guide inferences specifically about culturally relevant information. The logic here is that, faced with the problem of learning what defines a culturally relevant artifact category, children might be especially attuned to what an adult chooses to demonstrate about that category and what they choose not to demonstrate. Indeed, this is consistent with previous research illustrating the sophistication of children's understanding and inferences about teaching and teachers (Bonawitz et al., 2011; Gweon et al., 2014; Landrum et al., 2015; Shafto & Goodman, 2008) as well as more broadly with evidence on children's pragmatic reasoning in other domains (Merriman et al, 1995). In the current study, 3-year-olds may take the experimenter's deliberate choice not to teach potentially relevant information (as in the intentional condition) as evidence of that information's lack of importance in terms of assessing the essential defining properties of the novel artifact category. The results of the current experiments support this account and suggest the way in which labeling and pedagogical cues interact to guide children's inductive inference changes between 2 and 3 years of age. However, further follow-up research is necessary in order to explore in more depth the exact nature of this effect with 3-year-olds as well as to explore whether and how it affects older children's inductive inferences.

These results shed important new light on children's social learning. They illustrate that children's social learning is driven by an active nuanced integration of multiple sources of information in order to capitalize on what others know about the world. Thus, children have a powerful tool at their disposal

for guiding their learning about the world, and they employ that tool flexibly in the service of smart selective social learning. There are several important future directions that this work opens up. First, how is this specific capacity—to use pedagogical cues to guide inductive inference—more broadly integrated with children's developing social understanding and sociopragmatic competence? Here, we have investigated just one contextual manipulation—a label that cues potential cultural relevance. How might this effect interact with other aspects of social cognitive development such as more sophisticated understanding of communicative alternatives (e.g., what a communicator could have said or done but did not; Barner, Brooks, & Bale, 2011), understanding of the importance of who is doing the teaching and what that person's prior knowledge is (Landrum et al., 2015; Shafto & Goodman, 2008), and developing understanding of how one needs to select evidence in order to accomplish various pedagogical goals (Rhodes, Bonawitz, Shafto, Chen, & Caglar, 2015)? Second, how does this social learning process play out in children's actual development across various social contexts? Children's environments vary tremendously on numerous dimensions, many of which are likely to affect and interact with this social learning process. Investigating the functioning of this learning mechanism and seeing when and how it impacts children's conceptual development across childhood and in different areas of learning present exciting avenues for research in many directions.

Acknowledgments

This research was supported by a postdoctoral fellowship from the Alexander von Humboldt Foundation to the first author. We thank Tuende Malomvoelgyi, Elena Rossi, and Katharina Walter for assistance with coordination, data collection, and coding.

References

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.
Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition, 118*, 84–93.
Benjamini, Y. (2010). Discovering the false discovery rate. *Journal of the Royal Statistic Society B: Statistical Methodology, 72*, 405–416.
Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistic Society B: Methodological, 57*, 289–300.
Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition, 60*, 1–29.
Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition, 120*, 322–330.
Booth, A. E., & Waxman, S. (2002). Object names and object functions serve as cues to categories for infants. *Developmental Psychology, 38*, 948–957.
Butler, L. P., & Markman, E. M. (2012). Preschoolers use intentional and pedagogical cues to guide inductive inferences and exploration. *Child Development, 83*, 1416–1428.
Butler, L. P., & Markman, E. M. (2014a). Learning to learn from others: Developmental differences in children's ability to navigate pedagogical interactions. Manuscript under review.
Butler, L. P., & Markman, E. M. (2014b). Preschoolers use pedagogical cues to guide radical reorganization of category knowledge. *Cognition, 130*, 116–127.
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. (2009). Information learned from generic language becomes central to children's biological concepts: Evidence from their open-ended explanations. *Cognition, 113*, 14–25.
Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language, 25*, 141–168.
Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*, 148–153.
Csibra, G., & Shamsudheen, R. (2015). Nonverbal generics: Human infants interpret objects as symbols of object kinds. *Annual Review of Psychology, 66*, 689–710.
Dennett, D. C. (1989). *The intentional stance*. Cambridge, MA: MIT Press.
Egyed, K., Király, I., & Gergely, G. (2013). Communicating shared knowledge in infancy. *Psychological Science, 24*, 1348–1353.
Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology, 65*, 47–55.
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*. Oxford, UK: Oxford University Press.
Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8*, 404–409.
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.

Gergely, G., & Jacob, P. (2013). Reasoning about instrumental and communicative agency in human infancy. In J. B. Benson (Ser. Ed.), F. Xu, & T. Kushnir (Vol. Eds.), *Rational constructivism in cognitive development* (pp. 59–94). San Diego: Elsevier.
German, T., & Johnson, S. (2002). Function and the origins of the design stance. *Journal of Cognition and Development, 3*, 279–300.
Goodman, N. (1965). *Fact, fiction, and forecast*. Indianapolis, IN: Bobbs-Merril Company.
Gweon, H., Pelton, H., Konopka, J. A., & Schulz, L. E. (2014). Sins of omission: Children selectively explore when teachers are under-informative. *Cognition, 132*, 335–341.
Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 9066–9071.
Henderson, A. M. E., & Graham, S. A. (2005). Two-year-olds' appreciation of the shared nature of novel object labels. *Journal of Cognition and Development, 6*, 381–402.
Hernik, M., & Csibra, G. (2015). Infants learn enduring functions of novel tools from action demonstrations. *Journal of Experimental Child Psychology, 130*, 176–192.
Kelemen, D. (1999). The scope of teleological thinking in preschool children. *Cognition, 70*, 241–272.
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. (2000). Young children's use of functional information to categorize artifacts: Three factors that matter. *Cognition, 77*, 133–168.
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.
Landrum, A. R., Eaves, B. S., & Shafto, P. (2015). Learning to trust and trusting to learn: A theoretical framework. *Trends in Cognitive Sciences, 19*, 109–111.
Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition, 105*, 489–512.
Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.
Matan, A., & Carey, S. (2001). Developmental changes within the core of artifact concepts. *Cognition, 78*, 1–26.
Merriman, W. E., Marazita, J. M., Lorna, H., Evey-Burkey, J. A., & Biggins, M. (1995). What can be learned from something's not being named. *Child Development, 66*, 1890–1908.
Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255–258.
Rhodes, M., Bonawitz, E., Shafto, P., Chen, A., & Caglar, L. (2015). Controlling the message: Preschoolers' use of information to teach and deceive others. *Frontiers in Psychology, 6*. http://dx.doi.org/10.3389/fpsyg.2015.00867.
Sabbagh, M. A., & Henderson, A. M. (2007). How an appreciation of conventionality shapes early word learning. *New Directions in Child and Adolescent Development, 115*, 25–37.
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.
Shaffer, J. P. (1995). Multiple hypothesis testing. *Annual Review of Psychology, 46*, 561–584.
Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science, 7*, 341–351.
Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology, 71*, 55–89.
Shafto, P., & Goodman, N. D. (2008). Teaching games: Statistical sampling assumptions for learning in pedagogical situations. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual meeting of the Cognitive Science Society* (pp. 1632–1637). Austin, TX: Cognitive Science Society.
Sperber, D., & Wilson, D. (1996). *Relevance: Communication and cognition* (2nd ed.). Oxford, UK: Wiley-Blackwell.
Sutherland, S. L., Cimpian, A., Leslie, S. J., & Gelman, S. A. (2015). Memory errors reveal a bias to spontaneously generalize to categories. *Cognitive Science, 39*, 1021–1046.
Vredenburgh, C., Kushnir, T., & Casasola, M. (2015). Pedagogical cues encourage toddlers' transmission of recently demonstrated functions to unfamiliar adults. *Developmental Science, 18*, 645–654.
Williamson, R. A., & Meltzoff, A. N. (2011). Own and others' prior experiences influence children's imitation of causal acts. *Cognitive Development, 26*, 260–268.
Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1–34.
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 United States of America, 105*, 13690–13695.