Cognitive Psychology 83 (2015) 22–39

Contents lists available at ScienceDirect

Cognitive Psychology

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

# Reasoning about knowledge: Children's evaluations of generality and verifiability

Melissa A. Koenig [a,*], Caitlin A. Cole [a], Meredith Meyer [b], Katherine E. Ridge [a], Tamar Kushnir [c], Susan A. Gelman [d]

[a] University of Minnesota, United States
[b] Otterbein University, United States
[c] Cornell University, United States
[d] University of Michigan, United States

ARTICLE INFO

*Article history:*
Accepted 31 August 2015
Available online 8 October 2015

*Keywords:*
Cognitive development
Testimonial learning
Metacognition

ABSTRACT

In a series of experiments, we examined 3- to 8-year-old children's ($N = 223$) and adults' ($N = 32$) use of two properties of testimony to estimate a speaker's knowledge: generality and verifiability. Participants were presented with a "Generic speaker" who made a series of 4 general claims about "pangolins" (a novel animal kind), and a "Specific speaker" who made a series of 4 specific claims about "this pangolin" as an individual. To investigate the role of verifiability, we systematically varied whether the claim referred to a perceptually-obvious feature visible in a picture (e.g., "has a pointy nose") or a non-evident feature that was not visible (e.g., "sleeps in a hollow tree"). Three main findings emerged: (1) young children showed a pronounced reliance on verifiability that decreased with age. Three-year-old children were especially prone to credit knowledge to speakers who made verifiable claims, whereas 7- to 8-year-olds and adults credited knowledge to generic speakers regardless of whether the claims were verifiable; (2) children's attributions of knowledge to generic speakers was not detectable until age 5, and only when those claims were also verifiable; (3) children often generalized speakers' knowledge outside of the pangolin domain, indicating a belief that a person's knowledge about pangolins likely extends to new facts. Findings indicate that young children may be inclined to doubt speakers who make claims they cannot verify themselves, as well as a

* Corresponding author at: 51 East River Rd., Minneapolis, MN 55455, United States.
*E-mail address:* mkoenig@umn.edu (M.A. Koenig).

http://dx.doi.org/10.1016/j.cogpsych.2015.08.007
0010-0285/© 2015 Elsevier Inc. All rights reserved.

developmentally *increasing* appreciation for speakers who make general claims.

© 2015 Elsevier Inc. All rights reserved.

## 1. Introduction

In learning about the world, we acquire much knowledge from what others tell us (Gelman, 2009; Harris & Koenig, 2006). Our dependence on testimony is massive and far-reaching and without it, our knowledge would extend no further than what can be gained from our personal experiences. Our adult knowledge of how babies are born, how the blood circulates, or the geography of North America are not acquired by direct experience or observation (Coady, 1992; Sosa, 1994). Likewise, children's acceptance of claims concerning unobservable facts (e.g., mental states, biology of the body, round shape of the earth) and supernatural entities (e.g., omnipotence of God, efficacy of prayer) demonstrates that children's trust regularly extends to people's claims regarding hidden or non-evident properties of the world (Bering, 2006; Harris & Koenig, 2006). In this way, testimony can offer a quintessential benefit by extending the reach of our senses, giving us "vicarious" access to information from those in a better position to know (p. 50, Quine & Ullian, 1970). But, how do we determine who is in a better position to know?

This question puts into focus the epistemological problem that presents itself when we recognize that the basis or grounds for beliefs in the biology of the body or the efficacy of prayer lie only in a speaker's word. Given that testimony can be only as reliable as the beliefs that speakers report on, hearers face an epistemological vulnerability and a resulting need for reasons that substantiate or support the trust that is placed in what they're told (Faulkner, 2007; Koenig, 2012; Lackey, 2008). Importantly for our purposes, even when the basic competence of a speaker can be assumed, and her intentions are known to be good, this epistemic vulnerability remains given our reliance on other people's beliefs. Thus, in core cases, when a speaker's intention is to inform, the testimony takes the form of a single utterance, and the hearer has no privileged knowledge about the speaker, how do hearers ascribe knowledge or authority to a speaker?

In this paper, we examine children's and adults' knowledge attributions by focusing on two basic properties of testimony: generality and verifiability. Any given statement can be more or less general (as in, "Cars need gas" vs. "My Toyota needs gas"), and more or less verifiable. For example, the claim that "Mary has red hair" is easy to verify by looking at Mary; the claim that "Life continues after death" is impossible to verify given the nature of death, and statements like "Large animals have axial skeletons" lie somewhere in between in that they are in principle verifiable even if the relevant information is not immediately available. Importantly for our purposes, both generality and verifiability can be used to provide insight into the knowledge of speakers. For example, a speaker who makes a more general claim about a novel animal kind, such as "Pangolins eat insects," is claiming to know more than a speaker who presents the same content with a more specific claim, "This pangolin eats insects". Likewise, a speaker who makes a verifiable claim that can be either partly or wholly confirmed by one's own prior or current perceptual experience can be deemed knowledgeable by having made a claim that can be directly assessed as true. For example, given a picture that depicts a pangolin, a speaker who comments on a visible property makes a claim that can be verified by simply looking at the picture. This leads to our primary empirical questions: When presented with a speaker who makes general claims, do children credit her as more knowledgeable than someone who makes more specific claims? And when a speaker makes verifiable claims that can be immediately checked against direct experience, do children credit her as knowledgeable?

### 1.1. Generality

One particularly powerful means by which testimony allows us to make general statements is generic language – that is, language that refers to kinds of things (e.g., "Dogs are four-legged"). In fact,

when it comes to generic knowledge, testimony is not only an efficient way of acquiring knowledge, it is the *only* way, since generic knowledge is abstract and cannot be discovered or demonstrated non-linguistically (e.g., one cannot discover or demonstrate that dogs [as a kind] are four-legged, only that particular dogs are four-legged). Children are known to access this generic system of knowledge from an early age: generic statements are found in the speech parents direct to children as young as 20 months (Gelman, Coley, Rosengren, Hartman, & Pappas, 1998) and children produce generics themselves by 2–1/2 years of age (Gelman, Star, & Flukes, 2002; Pappas & Gelman, 1998). Children interpret generic statements with appropriate quasi-universal scope (in between 'all' and 'some', Hollander, Gelman, & Star, 2002), as defeasible (allowing for exceptions or counterexamples), and as a license for powerful inductive inferences (e.g., 'Birds lay eggs') (Shipley, 1993).

The inductive inferences that generic language affords are an important source of knowledge and indeed, even 2- to 3-year-olds recognize generics as an invitation to draw inferences about kinds and their properties. When learning about properties of novel and familiar kinds, children extend these properties more broadly when they are introduced using generic language than when they are introduced with non-generic language (Chambers, Graham, & Turner, 2008; Gelman, Ware, & Kleinberg, 2010; Gelman et al., 2002; Graham, Nayer, & Gelman, 2011; Rhodes, Leslie, & Tworek, 2012). Additionally, when properties of a novel kind are introduced using generic language ("Bants have stripes" vs. "This bant has stripes"), 4- and 5-year-olds are more likely to use that property as the basis for identifying a novel exemplar of the kind (Hollander, Gelman, & Raman, 2009). Such findings suggest that young children interpret generics as distinct from specific claims, and use generics to identify important properties of a kind and to draw inferences about unseen and newly encountered members of the kind.

Relatedly, children may recognize generics as an invitation to draw inferences about speakers. Recent research suggests that children at least recognize a potential relationship between speaker characteristics and generic statements. Children produce higher rates of generics in pedagogical contexts and when adopting a pedagogical (teacher) role than a non-pedagogical (peer) role (Gelman, Ware, Manczak, & Graham, 2013). Cimpian and Park (2014) demonstrated that when given the choice, 4- and 5-year-old children would rather learn information about a kind than an individual, but only from speakers deemed knowledgeable about the kind. Children also treat generic information (e.g., "Hedgehogs eat hexapods") as more broadly known than non-generic information (e.g., "Last night, this hedgehog ate a hexapod") (Cimpian & Scott, 2012). This finding shows that children are reasoning about how generic versus non-generic knowledge is distributed among people, and treat generic knowledge as more widespread than non-generic knowledge. The conceptual sophistication that children reveal in this literature raises the possibility that they treat a speaker's generic utterances as reflecting more knowledge than do specific claims. One question posed here is whether and when children treat generic speakers as more knowledgeable or expert about the category.

### *1.2. Verifiability*

Testimony can be checked against one's prior beliefs as well as current evidence and as we discuss here, children start to check claims in this way from a young age. By 16 months, infants check, reject, and often correct speakers who overtly violate their semantic beliefs by misnaming objects (Koenig & Echols, 2003; Pea, 1982). This ability to check claims against what they know is necessary to evaluate claims for accuracy and therefore helps children distinguish more from less reliable informants (e.g., Birch, Vauthier, & Bloom, 2008; Jaswal & Neely, 2006; Koenig & Harris, 2005). Children understand that certain claims are best verified by visual inspection of evidence and other claims are best verified either by appealing to an expert source (Clèment, Koenig, & Harris, 2004; Fitneva, Lam, & Dunfield, 2013; Koenig & Jaswal, 2011) or to others who have better access to information (Brosseau-Liard & Birch, 2011; Nurmsoo & Robinson, 2009; Robinson & Whitcombe, 2003). More naturalistic data on children's questions also show that children spot anomalies in the answers they receive from adults, and that when an answer issues an apparent conflict with other things the child believes, they often persist in their line of questioning to reconcile the conflict (Chouinard, 2007; Frazier, Gelman, & Wellman, 2009; Tizard & Hughes, 1984).

In fact, young children may be prone to doubt certain claims they can't verify themselves. In a recent review, Woolley and Ghossainy (2013) make a strong case for children's early skepticism toward claims regarding the reality status of novel entities. In one study, children were asked to evaluate the evidence regarding six rare animals, and to tell the experimenter whether each animal was real or pretend. For those children who exhibited a biased response pattern, the youngest children (age 4) were most likely to consistently deny that the animals were real (Tullos & Woolley, 2009), presumably due to the child's lack of experience with the animals. Also, in work by Bering and Parker (2006), children were asked to guess the location of a ball, and if they were wrong, some were told that they might receive help from "Princess Alice" who would send them a "sign" (e.g., flickering of the lights). Interestingly, the younger age groups (ages 3–6) were more dubious than older children; showing that the claim of an experimenter was not sufficient to boost their belief in the efficacy of random, typically non-causal signals.

Children's reliance on their prior knowledge and experience can be seen in their justifications: After watching short third-party conversations about the existence of novel animals, children justified their skepticism by saying things like "I've never heard of them before" or "I've never seen them" (Woolley, Ma, & Lopez-Mobilia, 2011). Children's doubt toward claims that they can't verify may extend to the speakers who make unverifiable claims. Thus, the second objective of the current studies is to examine whether children treat speakers who make claims that can be directly verified as more knowledgeable than those who make less verifiable claims.

## 2. Our task

In our task, we systematically varied these two properties of testimony – generality and verifiability – to examine the extent to which children make use of them when estimating a speaker's knowledge. Across studies, children (and adults in Study 2) were presented with two adult speakers, both of whom made claims about a novel animal kind – pangolins. In each study, across 4 familiarization trials, children were presented with 4 different pictures, each depicting a new and distinct pangolin, and four pairs of sentences from the two speakers. To investigate the force of generality, children were presented with a "Generic speaker" who made a series of 4 general claims about "pangolins" as a kind and a "Specific speaker" who made a series of 4 specific claims about "this pangolin" as a specific item. For example, the Generic speaker looked at the picture of the pangolin and said, "Pangolins eat insects" and the Specific speaker said, "This pangolin lives in a hollow tree". To avoid the implication that the specific claim followed from the generic claim, we always used different predicate content in each sentence (e.g., "pointy nose" from one speaker paired with "lives in Africa" from the other). To investigate the role of verifiability, we systematically varied whether the claim referred to a perceptually-obvious feature that was visible in the picture (e.g., "has a pointy nose") or a non-evident feature that was not visible (e.g., "sleeps in a hollow tree").

In Study 1A, we presented 3-, 4-, and 5-year-old children with a maximal contrast between a speaker who made generic and verifiable claims, and one who made specific, non-verifiable claims. That is, children were presented with a Generic speaker who consistently referred to perceptually-evident, visible properties (*Generic Verifiable speaker:* "Pangolins have pointy noses") and a Specific speaker who referred to non-evident properties (*Specific Non-verifiable speaker*: "This pangolin sleeps in a hollow tree"). In this case, if children appreciate that (1) generic speakers convey more knowledge than specific speakers and (2) speakers who make verifiable claims are more knowledgeable than those who make non-verifiable claims, our expectation is that children will credit the Generic Perceptible speaker as more knowledgeable given that she has two epistemic properties in her favor (generality and verifiability).

Our next step was to investigate whether children treat one property as particularly important or informative about a speaker. In Study 1B, we begin to address this question by separating the two properties of generality and verifiability, so that each of two speakers displayed only one. Thus, children were presented with a speaker who consistently referred to verifiable properties but of a specific pangolin (Specific Verifiable speaker: "This pangolin has a pointy nose") and a speaker who used generic language to refer to non-evident properties (Generic Non-Verifiable speaker: "Pangolins sleep in hollow trees"). If children appreciate the epistemic value of both properties, and see them as equally

important, then children may not systematically credit knowledge to one speaker over the other. However, if children view one of these two properties as a more potent indicator of speaker knowledge, then we may see increased attribution of knowledge to one speaker over the other. These patterns may also differ by age. For instance, children may see *both* the generality and verifiability of statements as relevant to judgments about what a speaker knows, however younger children in particular may privilege verifiability, and with age, increasingly appreciate the epistemic significance of generic statements.

In Study 2, we tested children in a larger age range (age 3–8 years, and adults), and controlled for verifiability by manipulating the visible perceptibility of claims in a between-subjects design. In Study 2, both speakers, the Generic and the Specific, jointly made claims that were both verifiable (*Verifiable condition*) or not (*Non-Verifiable condition*). If children and adults appreciate that generic statements reflect greater speaker knowledge regardless of whether they describe something perceptually obvious, then we should observe greater appeal to the Generic speaker regardless of condition. If, however, this understanding that generic statements reflect greater knowledge increases with age, or depends upon the verifiability of statements, it may be that younger children's appeal to a Generic speaker will depend upon whether the claim is perceptually verified.

Children's appeal to the more knowledgeable speaker was measured in two main ways. First, we overtly asked children who they thought was more knowledgeable in general. So, on three occasions throughout the study, after the generic speaker and the specific speaker had made their claims, children were asked to explicitly indicate, "Who do you think knows more?" Second, we were interested in children's willingness to generalize that knowledge to new instances and domains. Specifically, we were interested in whether children's preference for a more knowledgeable informant would be seen regardless of domain, or whether they would prefer the more knowledgeable informant only for matters specific to her expertise (consistent with Koenig & Jaswal, 2011; Kushnir, Vredenburgh, & Schneider, 2013; Lutz & Keil, 2002). Thus, we included 12 generalization tests to examine whether children treat the speaker as having more domain-specific knowledge (i.e., about other pangolins), or as being more knowledgeable beyond the domain in question (e.g., that extends beyond pangolins, to other animal kinds). To test the scope of children's generalizations, participants were presented with 4 new pangolin test trials, 4 novel animal test trials, and 4 novel artifact test trials. On each test trial, an item from the domain was shown (i.e., pangolin, novel animal, or novel artifact), and children were asked which speaker was likely to know a specific fact about the new item.

## 3. Study 1A

### 3.1. Method

#### 3.1.1. Participants

Forty-seven children participated in Study 1A, including 16 three-year-olds (*M*age = 3;6, range: 3;2–3;11, 11 females), 15 four-year-olds (*M*age = 4;6, range: 4;0–4;11, 6 females), and 16 five-year-olds (*M*age = 5;5, range: 5;1–5;10, 10 females). Data from an additional four participants were excluded from analysis because of failure to answer all test questions ($n = 3$) or experimental error ($n = 1$).

#### 3.1.2. Procedure

Participants watched a video in which two informants stated facts about a novel animal (pangolin) in a picture. One informant, the Generic/Verifiable (GV) speaker, consistently used the generic noun "pangolins" and commented on qualities that could be verified by looking at the picture (e.g., "Pangolins are brown"). The other informant, the Specific/Non-Verifiable (SN) speaker, consistently used the specific noun phrase "this pangolin" and commented on qualities that could not be verified by looking at the picture (e.g., "This pangolin eats insects").

*3.1.2.1. Familiarization video.* Children sat next to an experimenter in front of a computer that showed images and video of the target informants. To start, the computer displayed a still image of two infor-

mants with a photograph of a pangolin mounted in a frame between them. The experimenter placed an identical photograph of a pangolin in front of the child and introduced the task by saying, "We're going to watch a video. These people in the video are going to tell us about what's in this picture. See that? That's a pangolin! We're going to learn a bunch of things about it, so listen very carefully to what the people say and later I'll ask you who knows more. OK?"

During the video, the GV and SN speakers alternated making statements. Each speaker made a total of four statements. Whether the GV or SN speaker spoke first was counterbalanced across participants. After each pair of claims, the experimenter paused the video and repeated what each informant had said while pointing to the informant (e.g., "*She* said, 'Pangolins have pointy noses' and *she* said, 'This pangolin lives in a hollow tree.'").

*3.1.2.2. Explicit Judgment Trials.* Children were asked to explicitly evaluate the relative knowledge of the informants in three *Explicit Judgment Trials* (EJTs) that took the form: "Who do you think knows more?" The first EJT came after each speaker had made two statements. The second EJT came after each speaker had made four statements. The final EJT came at the very end of the session, after the Generalization trials.

*3.1.2.3. Generalization trials.* In 12 Generalization trials, children were asked to identify which informant would know particular facts about pangolins (four trials), other novel animals (four trials), and novel artifacts (four trials). All children received the Pangolin trials first, but the order of the Animal and Artifact trials was counterbalanced across participants. For each trial, the experimenter showed the child a photograph of a pangolin, another novel animal, or a novel artifact and asked which informant would know a particular piece of information (e.g., "I wonder how many meals it eats each day. Who do you think knows? This person or this person?").

### 3.2. Results

#### 3.2.1. Explicit Judgment Trials

We compared the proportion of the three EJTs on which children selected the GV speaker to rates expected by chance (chance = .5). Collapsed across age, a one-sample *t*-test revealed that children identified the GV speaker as more knowledgeable than the SN speaker at rates greater than expected by chance, $M = .59$, $t(46) = 1.91$, $p = .03$ (one-tailed test). Additional chance comparisons within each age group showed that only 5-year-olds selected the GV speaker at rates significantly above chance, $M = .69$, $t(15) = 2.43$, $p = .02$, while 3-year-olds, $M = .54$, $t(15) = .57$, $p = .58$, and 4-year-olds, $M = .53$, $t(14) = .37$, $p = .72$, did not differ from chance. A one-way ANOVA revealed no effect for age, ($F(2,46) = 1.16$, $p = .32$).

#### 3.2.2. Generalization trials

We next examined children's responses on the three types of Generalization trials by comparing responses on each trial type to chance. One-sample *t*-tests revealed no differences from chance on Pangolin trials, $M = .53$, $t(46) = .71$, $p = .48$, Animal trials, $M = .49$, $t(46) = -.13$, $p = .89$, or Artifact trials, $M = .51$, $t(46) = .25$, $p = .80$.

In order to assess the degree to which children's judgments of an informant's knowledge were related to their preference for that informant when evaluating information in new domains, we assessed the relationship between EJTs and Generalization trials. We grouped children according to their preference for one informant or the other on the three EJTs. Thus, children were divided into two groups: those who identified the Generic Verifiable speaker as knowing more on at least two of the EJTs ($N = 31$ or 66%, including 10 three-year-olds, 9 four-year-olds, and 12 five-year-olds), and those who identified the Specific Non-verifiable speaker as knowing more on at least two of the EJTs ($N = 16$ or 34%, including 6 three-year-olds, 6 four-year-olds, and 4 five-year-olds). A one-way ANOVA revealed significant differences between these two groups in their selection of the GV speaker as knowing new facts about pangolins, $F(1,45) = 10.15$, $p = .003$, and animals, $F(1,45) = 13.77$, $p = .001$, but not for questions about artifacts, $F(1,45) = 2.34$, $p = .13$.

Chance comparisons indicated that children who attributed 'more' knowledge to the GV speaker expected her to know facts about new pangolins ($t(30) = 2.53$; $p = .02$; $M = 60.5$, $SD = 23.1$), and about new animals ($t(30) = 2.01$, $p = .054$; $M = 58.9$, $SD = 24.6$), but not about new artifacts ($t(30) = 1.19$, $p = .24$; $M = 55.6$, $SD = 26.4$). A similar pattern emerged for those 16 children who attributed 'more' knowledge to the SN speaker; they expected the SN speaker to know facts about pangolins ($t(15) = 2.07$, $p = .056$; $M = 62.5$, $SD = 24.1$), and animals ($t(15) = 3.\ 22$; $p = .006$; $M = 68.7$, $SD = 23.2$), but not artifacts ($t(15) = 0.96$, $p = .35$; $M = 57.8$, $SD = 32.5$).

*3.2.2.1. Interim discussion.* In Study 1A, we presented children with a speaker who made 4 generic, verifiable claims, and one who made 4 specific, non-verifiable claims about a novel animal kind, pangolins. Three main findings emerged. First, consistent with our expectations, when asked, "Who knows more?," children were more likely to attribute knowledge to the speaker who made generic and verifiable claims than the speaker who made specific, non-verifiable claims. However, this ability to credit knowledge to the generic, verifiable speaker was found only in the 5-year-old group. Second, children's willingness to credit more knowledge generally to a given speaker was consistent with their attributions of domain-specific knowledge to that speaker. That is, children who consistently credited "knowing more" to the generic, verifiable speaker also treated her as more likely to know a set of facts about new pangolins and new animals but not about new artifacts. Children who credited knowledge to the specific, non-verifiable speaker, while much fewer in number, also expected her to know more specific facts about pangolins and animals, but not artifacts. Thus, thirdly, children indicated that a person's knowledge about pangolins likely extends to facts specific to the biologically-related category of new animals, but not the non-biological functioning of artifacts.

In Study 1A, children demonstrated some preference for speakers who make generic and verifiable claims, but (1) it didn't dominate children's choices, (2) it appeared only in the oldest age group, and (3) generality and verifiability were intentionally confounded within a single speaker. In Study 1B, we separated the two properties in order to investigate whether children would continue to treat a generic speaker as knowledgeable even if their claims were not immediately verifiable. Thus, children were presented with a generic speaker who used generic language to refer to non-evident properties (Generic Non-verifiable speaker: "Pangolins sleep in hollow trees"), along with a specific speaker who referred to verifiable properties (Specific Verifiable speaker: "This pangolin has a pointy nose"). If verifiability is a strong cue to knowledge, especially for young children, then we expect to see less attribution of knowledge to the generic speaker who makes unverifiable claims and instead attribution to the specific, verifiable speaker.

## 4. Study 1B

### 4.1. Method

#### 4.1.1. Participants

Forty-five children participated in Study 1B, including 15 three-year-olds (*M*age = 3;6, range: 3;1–3;11, 8 females), 15 four-year-olds (*M*age = 4;6, range: 4;1–4;11, 10 females), and 15 five-year-olds (*M*age = 5;6, range: 5;2–5;11, 8 females). Data from an additional four participants were excluded from analysis because of failure to answer all test questions ($n = 3$) or experimental error ($n = 1$).

#### 4.1.2. Procedure

The procedure for Study 1B was identical to that of Study 1A, except that in this study, one informant (Generic/Non-Verifiable, or GN) consistently used the generic noun "pangolins" and commented on non-verifiable qualities (e.g., "Pangolins eat insects") while the other informant (Specific/Verifiable, or SV) consistently used the specific noun phrase "this pangolin" and commented on verifiable qualities (e.g., "This pangolin is brown.").

### 4.2. Results

#### 4.2.1. Explicit Judgment Trials

We compared the proportion of the three EJTs on which children selected the GN speaker to rates expected by chance (chance = .5). A one-sample *t*-test showed that children had no preference for either speaker, $t(44) = -1.55$, $p = .13$ (two-tailed). Additional chance comparisons within each age group showed that while 4-year-olds, $M = .40$, $t(14) = -1.49$, $p = .16$, and 5-year-olds, $M = .53$, $t(14) = .32$, $p = .75$, did not differ from chance, 3-year-olds selected the *Specific Verifiable* speaker at rates significantly above chance, $M = .65$, $t(14) = -2.39$, $p = .03$, indicating that they rated the SV speaker as more knowledgeable. A one-way ANOVA indicated no reliable effect of age ($F(2,44) = 1.38$, $p = .26$).

#### 4.2.2. Generalization trials

We next examined children's responses on the three types of Generalization trials by comparing responses on each trial type to chance. As a group, one-sample *t*-tests revealed that children chose the Specific Verifiable speaker at rates above chance for Pangolin trials, $M = .63$, $t(44) = 2.05$, $p = .046$. Children did not choose either speaker at rates different from chance on Animal trials, $M = .47$, $t(44) = -.92$, $p = .36$, or Artifact trials, $M = .46$, $t(44) = -1.09$, $p = .28$.

As in Study 1A, in order to assess the degree to which children's judgments of an informant's knowledge were related to their preference for that informant when evaluating more specific information, we assessed the relationship between EJTs and Generalization trials. We grouped children according to their preference for one informant or the other on the three EJTs. Thus, children were divided into two groups: those who identified the Generic Non-verifiable speaker as knowing more on at least two of the EJTs ($N = 17$ or 38%, including 4 three-year-olds, 6 four-year-olds, and 7 five-year-olds), and those who identified the Specific Verifiable speaker as knowing more on at least two of the EJTs ($N = 28$ or 62%, including 11 three-year-olds, 9 four-year-olds, and 8 five-year-olds). A one-way ANOVA revealed significant differences between these two groups on their selection of the Generic Non-verifiable speaker for questions about pangolins, $F(1,43) = 5.59$, $p = .02$, and animals, $F(1,43) = 4.26$, $p = .045$, but not for questions about artifacts, $F(1,43) = 1.39$, $p = .24$.

Chance comparisons indicated that children who attributed more knowledge to the Specific Verifiable speaker expected her to know specific facts about new pangolins ($t(27) = 3.57$; $p = .001$; $M = 63.4$, $SD = 19.8$) and about new animals ($t(27) = 1.98$, $p = .057$; $M = 59.0$, $SD = 23.7$), but not about new artifacts ($t(27) = 1.61$, $p = .12$; $M = 57.2$, $SD = 23.4$). A different pattern emerged for those 17 children who attributed more knowledge to the Generic Non-verifiable speaker; they responded at chance levels for pangolins ($t(16) = 0.46$, $p = .65$; $M = 52.9$, $SD = 26.3$), animals ($t(16) = 1.07$; $p = .29$; $M = 55.8$, $SD = 22.5$), and artifacts ($t(16) = 0.25$, $p = .81$; $M = 51.5$, $SD = 24.2$).

*4.2.2.1. Discussion.* As expected, when the verifiability of the generic claims was removed, children's attributions of knowledge to the generic speaker were reduced. In fact, the youngest children systematically credited knowledge to the *Specific speaker* when her claims were immediately verifiable. So in contrast to the developmentally increasing value children placed on generality in Study 1A, verifiability was valued by the very youngest age group, even for only very specific statements made about a particular pangolin. The value of verifiability was also seen in children's generalizations of knowledge: Children expected the previously verifiable speaker to know new pangolin and animal facts whereas they had no expectations or bases for generalization for a previously non-verifiable speaker. In light of these results, this suggests that when generality and verifiability are put into conflict, children might actually place greater epistemic weight or value on verifiability. Taken together, Studies 1A and 1B lead to questions like: Does the value placed on verifiable claims change with age? Do generality and verifiability both interact with age? To shed light on these questions, we conducted cross-study comparisons on children's EJT and test trial performance below.

### 4.3. Comparison of Studies 1A and 1B

To better understand these findings across Studies 1A and 1B, we conducted follow-up analyses that (1) target children's attributions of knowledge to Generic speakers (both with and without

verifiable support) and that (2) target children's attributions of knowledge to Verifiable speakers (who make both generic and specific claims).

*4.3.1. Explicit Judgment Trials*

(1) *Holding Genericity constant*. We performed a 2 (Study: 1A vs. 1B) × 3 (Age: 3-, 4- and 5-year-olds) ANOVA on the proportion of EJTs on which the child identified the Generic speaker as knowing more. As illustrated in Fig. 1, this analysis revealed a significant main effect of Study, $F(1,92) = 5.94$, $p = .02$, $\eta\rho^2 = .07$, whereby children in Study 1A, who were *able to verify* the GV speaker's claims ($M = .59$, $SD = .32$) were more likely to identify the Generic speaker as knowledgeable than were children in Study 1B, who were unable to verify the GN speaker's claims ($M = .43$, $SD = .31$). In addition, a marginally significant main effect of Age emerged, $F(2,92) = 2.47$, $p = .09$, $\eta\rho^2 = .054$. Pairwise comparisons showed that, across both studies, 5-year-olds ($M = .61$, $SD = .36$) were more likely to identify the Generic speaker as knowledgeable than were 3-year-olds ($M = .45$, $SD = .28$, $p = .05$) or 4-year-olds ($M = .47$, $SD = .31$, $p = .08$). There was no significant interaction between Age and Study.

(2) *Holding Verifiability constant*. We also compared children's knowledge attributions to the *verifiable* speakers, who in Study 1A made generic claims, and in Study 1B, made specific claims. We performed a parallel analysis on the proportion of EJTs on which children identified the Verifiable speaker as knowing more. In this case, a 2 (Study) × 3 (Age) ANOVA on the proportion of EJTs revealed no main effect of Study ($F(1,92) = 0.07$) nor Age ($F(2,92) = 0.06$), $ps > .05$. Instead, there was a trend for an Age × Study interaction, $F(2,92) = 2.52$, $p = .08$. Pairwise comparisons conducted at each age show a shift from treating specific over generic verifiable speakers as more knowledgeable at age 3 to treating generic over specific verifiable speakers as more knowledgeable at age 5, with the most evident preference seen among 5-year-olds, (Generic Verifiable = 68.8 vs. Specific Verifiable = 46.7, $p = .09$).

*4.3.1.1. Discussion of Study 1.* Four main findings emerged from Study 1. First, generic claims are not reliably interpreted as indicating more knowledge in a speaker until age 5 or later. This was supported

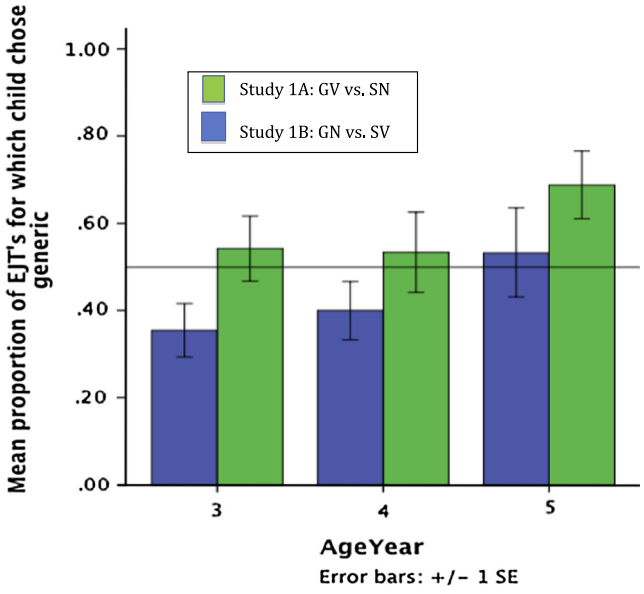


**Fig. 1.** Selection of the Generic speaker for Explicit Judgment Trials according to age in Studies 1A and 1B. Chance is 50%.

by Study 1A where 5-year-olds were the only age group to systematically credit knowledge to the Generic speaker. Also, when we collapsed across Study 1A and B, and verifiability was held constant, it wasn't until age 5 that generic speakers were seen as somewhat more knowledgeable than specific speakers. Second, as indicated by the main effect for Study, verifiability consistently supported children's knowledge attributions in the same direction, for all age groups. That is, children more readily attributed knowledge to a Generic speaker when those statements contained verifiable information about the perceptible properties of a depicted animal than when they did not. This suggests that the child's ability to independently verify statements played a role in their knowledge attributions to generic speakers. Third, it was the youngest children who were especially prone to credit knowledge to speakers who made verifiable claims, with children showing a *decreasing* reliance on verifiability with age. Fourth, children's attributions of knowledge to one informant or the other generally aligned with their appeal to that informant when seeking information about new pangolins and new classes of animals, especially when that informant made previously verifiable claims (see Study 1B).

Given that children's knowledge attributions to the Generic speaker increased with age, and related to the verifiability of her statements – especially for the youngest children – we sought further confirmation of these patterns in Study 2. In Study 2 we manipulated the verifiability of statements between subjects and randomly assigned participants to a Verifiable condition or a Non-Verifiable condition. Thus, in the Verifiable condition, the Generic and the Specific speakers made contrasting claims about perceptible properties of the depicted pangolin. In the Non-Verifiable condition, both speakers made contrasting claims about imperceptible properties of pangolins. Under these conditions, if young children, especially the 3- and 4-year-olds, make greater knowledge attributions to speakers who make statements that children can independently verify, then we expect the youngest children to credit knowledge to Generic speakers more often in the Verifiable than in the Non-Verifiable condition. Additionally, if older children and adults appreciate the knowledge expressed by the Generic speaker independent of its verifiability, we expect they will attribute greater knowledge to the Generic speaker with age across both conditions. Finally, given that differences across age groups in Study 1 were small and because not even five-year-olds demonstrated an overwhelming preference for the Generic speaker, we increased our age range in Study 2 to include children between ages three and eight as well as adults.

## 5. Study 2

### 5.1. Method

#### 5.1.1. Participants

A total of 131 children and 32 adults participated in Study 2. Participating children were recruited over the phone from a database of families in a large Midwestern city. Adult participants were undergraduates at a large Midwestern university. Participants included 40 three- and four-year-olds (*M*age = 4;0, range: 3;1–4;11, 16 females), 43 five- and six-year-olds (*M*age = 5;11, range: 5;1–6;10, 23 females), 44 seven- and eight-year-olds (*M*age = 7;8, range: 7;0–8;11, 18 females), and 32 adults (16 females). Data from an additional four children were excluded from analysis because of failure to answer all test questions.

#### 5.1.2. Procedure

The procedure for Study 2 was identical to that for Study 1 with three exceptions. First, the perceptibility of each speaker's claims differed between subjects as opposed to within subjects. Participants were assigned to one of two conditions. In the Verifiable condition, both the Generic and Specific speaker made different claims about perceptible properties of the depicted pangolin. In the Non-Verifiable condition, both speakers made different claims about imperceptible properties. The content differed across speakers to ensure that children would distinguish between the speakers. Second, at the end of the session, after the third EJT, participants completed a memory check. The experimenter repeated two of the statements made by each speaker (two generic statements, and two specific

statements), and asked participants to identify the speaker who made each statement. Third, adult participants in Study 2 were asked if they were aware of the manipulation (i.e., if they reported an explicit strategy to select either the generic or specific speaker on the basis of this difference, we considered them aware of the manipulation).

### 5.2. Results

#### 5.2.1. Explicit Judgment Trials

We performed a 2 (Condition: Verifiable vs. Non-Verifiable) × 4 (Age Group: 3-,4-year-olds; 5-, 6-year-olds; 7-, 8-year-olds; adults) ANOVA on the proportion of EJTs on which participants (adults and children) identified the Generic speaker as knowing more. As illustrated in Fig. 2, this analysis revealed a significant main effect of Age Group, $F(3,158) = 18.85$, $p < .001$, $\eta\rho^2 = .27$.

A post hoc Tukey HSD clarified the differences among age groups. Adults attributed knowledge to the Generic speaker more ($M = .89$, $SD = .22$) than 3- and 4-year-olds ($M = .43$, $SD = .27$), 5- and 6-year-olds ($M = .48$, $SD = .31$), and 7- and 8-year-olds ($M = .59$, $SD = .31$), all $ps < .001$. There was also a significant difference between the youngest children (3- and 4-year-olds) and the oldest (7- and 8-year-olds), $p = .04$.

Comparisons to chance in each age group were conducted with one-sample *t*-tests. These analyses showed that adults attributed knowledge to the Generic speaker on EJTs at rates above that expected by chance (chance = .5), $t(31) = 10.41$, $p < .001$, as did the 7- and 8-year-old children, $t(43) = 2.08$, $p = .04$. Younger children's EJT scores did not differ from chance: 3- and 4-year-olds $t(39) = -1.59$, $p = .12$; 5- and 6-year-olds $t(42) = -.44$, $p = .66$.

*5.2.1.1. Younger children.* Given the effects of age, and our age-sensitive predictions regarding the role of verifiability, we focused our next analyses on the two youngest age groups, 3- to 4-year-olds and 5- to 6-year-olds. We performed a 2 (Condition) × 2 (Age group: 3–4; 5–6) ANOVA on the proportion of EJTs on which the Generic speaker was judged as knowing more. We found a non-significant trend toward an interaction between Age and Condition, $F(1,79) = 3.\ 58$, $p = .06$. Follow-up comparisons showed that the youngest children were significantly more likely to credit knowledge to the Generic

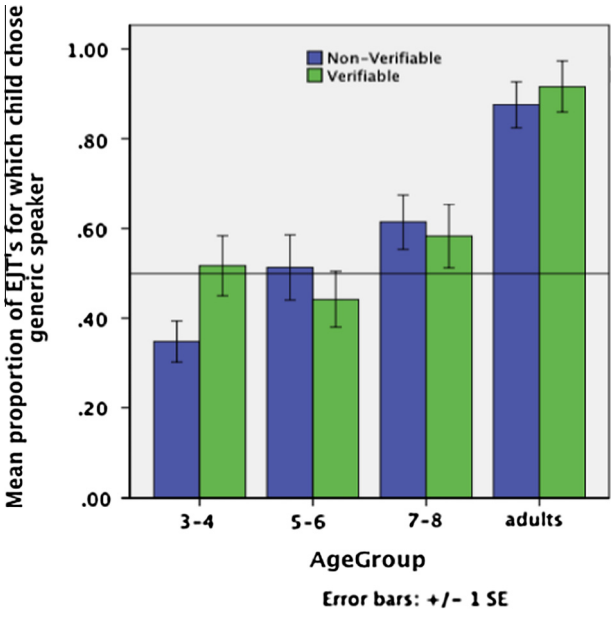


**Fig. 2.** Selection of the Generic speaker for Explicit Judgment Trials according to age in Study 2. Chance is 50%.

speaker in the Verifiable condition ($M = 51.7$; $SD = 29.7$) than in the Non-Verifiable condition ($M = 34.9$; $SD = 20.3$), $F(1,39) = 4.41$, $p = .042$. In fact, for the 3- to 4-year-old age group, without the support of verifiable statements, children systematically credited the Specific speaker with more knowledge about Pangolins in the Non-Verifiable condition, $t(19) = -3.34$, $p = .003$. Five- to six-year-old children, in contrast, showed no such condition effect, $F(1,41) = 0.55$, ns (Verifiable condition: $M = 44.3$; $SD = 28.6$; Non-Verifiable condition: $M = 51.3$; $SD = 33.8$).

*5.2.2. Generalization trials*

We next examined responses on the three types of Generalization trials by comparing responses to chance in each condition. In the Verifiable condition, one-sample *t*-tests revealed that children as a group did not choose the Generic speaker at rates above chance for Pangolin trials, $M = 47.6$; $SD = 20.1$, $t(64) = -0.92$, ns, Animal trials, $M = 48.8$, $SD = 23.1$, $t(64) = -.40$, ns, or Artifact trials, $M = 46.9$; $SD = 21.8$, $t(64) = -1.13$, ns. In contrast, adults in the Verifiable condition chose the Generic speaker at rates above chance for Pangolin trials, $M = 71.9$; $SD = 34.0$, $t(16) = 2.57$, $p = .02$, Animal trials, $M = 79.7$, $SD = 27.7$, $t(16) = 4.28$, $p = .001$, and Artifact trials, $M = 71.9$; $SD = 34.0$, $t(16) = 2.57$, $p = .02$. In the Non-Verifiable condition, a similar pattern obtained. Children did not choose the Generic speaker at above-chance rates, Pangolin trials, $M = 47.5$; $SD = 20.6$, $t(61) = -0.92$, ns; Animal trials, $M = 50.8$, $SD = 23.5$, $t(61) = 0.27$, ns; Artifact trials, $M = 52.8$, $SD = 24.4$; $t(61) = 0.91$, ns. However, adults in the Non-Verifiable condition chose the Generic speaker at rates above chance for Pangolin trials, $M = 67.2$; $SD = 25.4$, $t(16) = 2.71$, $p = .02$ and Animal trials, $M = 75.5$, $SD = 23.1$, $t(16) = 4.41$, $p = .001$, but not for Artifact trials $M = 65.6$; $SD = 35.2$, $t(16) = 1.78$, ns.

We performed a 2 (Condition) × 4 (Age Group) ANOVA on the proportion of Generalization trials on which participants said that the Generic speaker would know new facts concerning pangolins, animals, and artifacts. We found main effects of Age Group on pangolin trials, $F(3,159) = 8.62$, $p < .001$, animal trials, $F(3,159) = 12.70$, $p < .001$, and artifact trials, $F(3,159) = 4.92$, $p = .003$. Post-hoc Tukey tests revealed that all of these effects of Age Group were driven by differences between the adults, who demonstrated a clear preference for the Generic speaker on all types of trials, and all age groups of children, who did not, $ps < .03$.

We then examined the relationship between children's explicit judgments (EJTs) and their preference for that informant on Generalization trials. Because most adults chose the Generic speaker for all types of trials, only children's data were included in this analysis. We grouped children according to their preference for one informant or the other on the three EJTs as in Study 1: one group identified the Generic speaker as knowing more on at least two of the EJTs ($N = 60$ or 47%, including 16 three- and four-year-olds, 16 five- and six-year-olds, and 28 seven- and eight-year-olds), and the other group identified the Specific speaker as knowing more on at least two of the EJTs ($N = 67$ or 53%, including 24 three- and four-year-olds, 27 five- and six-year-olds, and 16 seven- and eight-year-olds). As illustrated in Fig. 3, a one-way MANOVA revealed significant differences between these two groups on their selection of the Generic speaker for questions about pangolins, $F(1,126) = 14.81$, $p < .001$, animals, $F(1,126) = 12.14$, $p = .001$, and artifacts, $F(1,126) = 10.67$, $p = .001$. Thus, the group who credited the Generic speaker as knowing more on EJTs were also more likely to expect the Generic speaker to know new facts regarding new Pangolins, Animals, and Artifacts. This effect of Explicit Judgment performance on children's generalization performance did not interact with Condition, $F(1,115) = 0.29$, $p = .59$; and held for children in the Non-Verifiable condition, $F(1,56) = 9.53$, $p = .003$, as well as the Verifiable condition, $F(1,59) = 7.79$, $p = .007$.

Comparisons to chance were conducted using one-sample *t*-tests. These analyses showed that children who preferred the Generic speaker for at least two EJTs also preferred the Generic speaker for the Generalization trials at rates marginally greater than chance for pangolin trials, $M = .55$, $t(60) = 1.89$, $p = .06$, and significantly for animal trials, $M = .57$, $t(60) = 2.66$, $p = .01$ and for artifact trials, $M = .57$, $t(60) = 2.39$, $p = .02$. In contrast, children who preferred the Specific speaker for at least two EJTs tended to choose the Specific speaker at rates greater than chance for Generalization trials, indicating that they preferred to learn from the Specific speaker. These differences from chance were significant for pangolin trials, $M = .41$, $t(66) = -3.56$, $p = .001$, animal trials, $M = .43$, $t(66) = -2.32$, $p = .02$, and artifact trials, $M = .44$, $t(66) = -2.24$, $p = .03$.

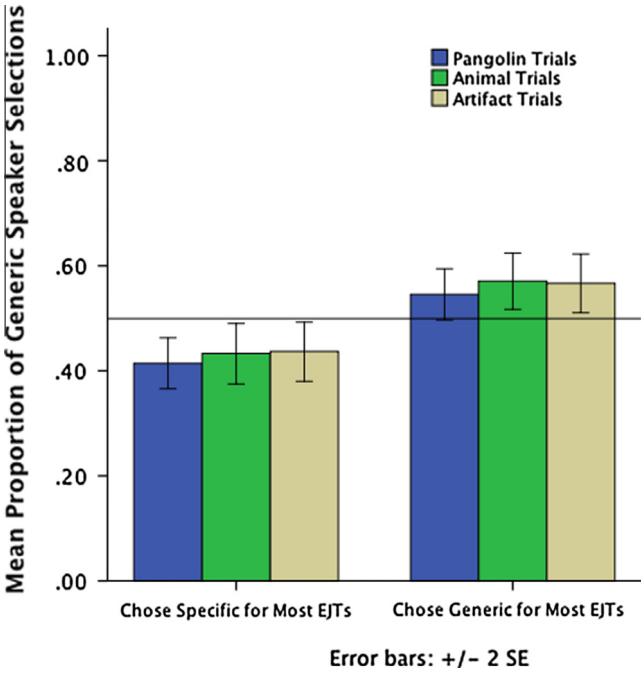


**Fig. 3.** Relationship between preference for a speaker on 2 out of 3 Explicit Judgment Trials and selection of the Generic speaker for Generalization trials in Study 2.

*5.2.3. Memory checks*

To better understand children's knowledge attributions, we examined children's memory for the speakers' original statements. Children were asked at the very end of the procedure to identify which of the two speakers had said two generic statements, and two specific statements, for a total of 4 memory checks. No differences were found between children's memory for who had produced generic (1.38 out of 2) versus specific (1.40 out of 2) statements ($t(126) = 0.35$, ns). A 2 (Condition) × 3 (Age Group) ANOVA showed that children tended to have better memory for who produced statements in the Non-Verifiable condition ($M = 3.00$, $SD = 1.13$) than in the Verifiable condition ($M = 2.58$; $SD = 1.25$), $F(1,121) = 3.73$, $p = .056$. There was no significant effect of age and no interaction. This carries the interesting suggestion that children may show better memory for statements that are more abstract, or convey new information, than for statements they can verify themselves (see also Cimpian & Erickson, 2012).

*5.2.3.1. Discussion of Study 2.* First, consistent with Studies 1A and 1B, we found that attributions of knowledge to the Generic speaker increased with age. In their explicit judgments, adults robustly credited knowledge to the Generic over the Specific speaker regardless of whether the claims were verifiable, and a similar, albeit reduced level of endorsement was seen in the 7- to 8-year-olds. Children who correctly recalled all four memory checks were also prone to credit knowledge to the generic speaker, with or without the support of verifiable statements. Second, consistent with Studies 1A and 1B, we found that the youngest children made more knowledge attributions to the generic speaker when her claims were perceptually supported than when they were not. In fact, when the speakers made comments regarding non-verifiable features (e.g., eats insects, sleep in hollow trees), the youngest children systematically preferred the Specific over the Generic speaker. Finally, children's attributions of knowledge to one informant or the other corresponded to their appeal to that informant for new facts about new pangolins, new classes of animals, and also novel artifacts. Thus, once children

have attributed knowledge to a particular informant, they seem to generalize that attribution to knowledge in other domains.

## 6. General discussion

In this paper we examined two properties of testimony that we theorized would support judgments of a speaker's reliability: generality and verifiability. We hypothesized that generic language, by virtue of its breadth of scope, would offer a quintessential benefit by giving speakers vicarious access to more observations, "more eyes to work with" (p. 50; Quine & Ullian, 1970). We also hypothesized that verifiable information – that is, information that one can check against observable features of the environment – would increase a child's confidence in the speaker. This hypothesis is consistent with the advice Hume gives in his essay "Of Miracles" (1748), where he argues "the maxim by which we commonly conduct ourselves in our [testimonial] reasonings ... is that where there is an opposition of arguments, we ought to give preference to such as are founded on the greatest number of past observations" (p. 124).

The results indicated evidence for both factors, though at different points in development. Generic speakers were judged to be more knowledgeable by adults and older children, but not by younger children. In contrast, we found that the youngest children especially treated verifiability as indicative of a speaker's knowledge. These findings suggest a developmental shift from reliance on verifiability to reliance on genericity when reasoning about speaker knowledge. An important next step would be to examine verifiability in its own right, decoupled from and unconfounded with genericity (e.g., comparing two generic statements, one of which is verifiable and one of which is not). Given that generics are themselves never directly verifiable, the privileging of generic information in older children and adults may suggest a concomitant privileging of non-verifiable information at these ages.

Thus, although we predicted that children would make use of verifiability as a cue to a speaker's greater knowledgeability, one could instead make the reverse prediction. Expertise consists of possessing knowledge or capabilities that others do not have, and thus is most evident when someone provides content that is *less* (rather than more) obvious. Little to no expertise is required to comment that a visible pangolin is brown (verifiable), whereas expertise is required to comment that a visible pangolin has a four-chambered heart (non-verifiable). There are thus competing reasons either to privilege verifiability or to privilege non-verifiability. One way to resolve this apparent contradiction is to consider the *kind of knowledge* that a learner is appraising; namely, whether a learner is assessing the basic competence of a speaker (in which making verifiable claims is reassuring of a speaker's knowledgeability) or their expertise (in which making unverifiable claims provides more pertinent evidence). Which kind of knowledge (e.g., competence, expertise, understanding) is privileged at any given moment by a learner likely rests in part on the learner's prior knowledge in a given domain and what they are being asked to estimate, judge, or infer about a speaker (Hofer & Pintrich, 1997; Wilkenfield, Plunkett, & Lombrozo, 2015). In future work it will be important to examine the effects of other kinds of attributions. For example, had we asked children not "who knows more?" but "who understands more about pangolins?", it could be that generic speakers would be favored even by younger age groups.

Verifiability may be heightened whenever a category is unfamiliar, such that even quite obvious content ("This pangolin is brown") reveals evidence of expertise, because the speaker knows something about this novel category, including its label. Thus, verifiability as manipulated here may carry decreasing weight as categories become more familiar (e.g., "This dog is brown"), including cases in which the speaker seems to know no more than the child and testimony offers no new information as a result. One implication of this would be that the epistemic value of verifiability and generality in communicating information about a speaker's knowledge interacts in important ways with what a given child or adult knows about the topic being discussed. The early fact-checking tendency found here to credit as knowledgeable those speakers who make plainly true statements may point to the prior knowledge that they do not have, and their risk-minimizing strategy of checking for information that they can partially verify themselves. If this is correct, then when young children are presented with two speakers who discuss a more familiar domain such as dogs or cats, they might show a more

adult-like pattern, crediting knowledge to generic speakers, and caring less about whether the claim is perceptibly verifiable. Similarly, when adults are presented with speakers who discuss an unfamiliar or abstract scientific domain (e.g., medicine, epigenetics, evolution), they might show a more 'child-like' pattern by crediting knowledge to speakers who make claims that confirm pieces of the adult's knowledge or seem potentially verifiable. Indeed, this idea seems to be recognized in the practice of science education. It has long been established that the best tools for explaining science, at any age, are concrete representations of abstract phenomena (models, interactive demonstrations, exhibits). An open question for research is whether these tools are effective not only because they allow students to visualize and experiment, but also because they lend support or credibility to a teacher's claims.

In light of these considerations, it is important to point out that children's verificationist tendency may not reflect a purely developmental phenomenon. Alternatively, it may more directly reflect a learner's domain knowledge at any age, and that one's assessments of knowledge to others are deeply tied to the context in which claims are made, and what the learner already knows about that domain (see also Keil, 1989). Knowing who is knowledgeable, and how far one ought to defer to another's expertise is deeply connected to our estimates of knowledge and how it is gained in the expert domain (Keil, 2008; Mills, 2013; Sobel & Kushnir, 2013). Hypothetically, it may be that whenever we are ignorant, and presented with a foreign and poorly understood domain, we verify claims as far as possible before accepting new information. And conversely, as we acquire domain knowledge, along with ideas about what it takes to be knowledgeable, we relax or open our requirements on verifiability, and show a greater capacity to defer to others. In other words, whenever we find ourselves 'at sea' with very little prior knowledge to apply, we may treat claims that confirm our observations, or what little we do know, as reassuring indicators of knowledge; and as we become increasingly knowledgeable ourselves about what distinguishes knowledgeable from less knowledgeable sources in a given domain, we show a greater willingness to defer to others.

Across studies, children were asked to generalize speaker knowledge to new pangolins, novel animals, and novel artifacts. We found that children's generalizations were generally in line with their explicit judgments about which speaker knew more about pangolins, but also found very broad generalizations in Study 2, and more circumscribed generalizations in Studies 1A and B. Given that we had no a priori predictions about the scope that children's generalizations would take, we offer three observations. First, the fact that children's generalizations tended to be narrower in Studies 1A and B (holding for pangolins and animals, but not artifacts) may relate to the fact that those speakers always represented a contrast between verifiable and non-verifiable statements. Relatedly and secondly, the only condition in which we found a complete lack of generalization was from children in Study 1B who initially credited knowledge to the Generic Non-verifiable speaker. Together, this might mean that when children's knowledge estimates are tied to their verification of a speaker's statements, these generalizations tend to be more circumscribed or cautious (because their strategy was more tied to the local truth of the utterances). Third, in Study 2, children who credited the Generic speaker as knowing more, generalized that knowledge more broadly, in both conditions, to include new facts regarding new pangolins, animals and artifacts. When verifiability no longer distinguished the speakers, the more generally knowledgeable a speaker was seen to be, the more generally their knowledge was taken to extend.

At first blush, these results may seem consistent with the suggestion that young children tend to be skeptical toward claims they cannot verify themselves (Woolley & Ghossainy, 2013). However, it's important to clarify that children were not asked to accept the claims themselves nor were they asked to judge the reality status of pangolins. Furthermore, perception was not pitted against testimony, none of the claims violated laws of nature, nor were they at all counter-intuitive (Browne & Woolley, 2004; Shtulman, 2009; Shtulman & Carey, 2007). As such, our research offers information about children's inferences from claims to speaker knowledge, not their inferences from claims to reality. Thus, we are not arguing that children shift from a skeptical to a credulous stance, nor that children require perceptible evidence to accept a claim. Instead, young children credit knowledge to speakers who make claims they can verify, and as they gain knowledge about speakers and the domain, thereby strengthening their capacities to reason about both things, they relax this verificationist strategy.

An important question for future research, one with potentially meaningful educational implications (Hofer & Pintrich, 1997), concerns whether and how children's changing criteria for knowledge attributions affect their learning and what those effects might be. One possibility is that while children's epistemic criteria for crediting knowledge may be changing, the knowledge attributions – once made – promote learning from a speaker and deliver the same general learning patterns. On this possibility, it doesn't matter how or why you credit knowledge to a speaker, knowledgeable speakers license or promote a general-purpose invitation to learn. However, given the developmental evidence (Brooker & Poulin-Dubois, 2013; Brosseau-Liard & Birch, 2011; Danovitch & Keil, 2004; Fusaro, Corriveau, & Harris, 2011; Henderson, Graham, & Schell, 2015; Koenig & Jaswal, 2011; Kushnir et al., 2013; Sobel & Corriveau, 2010; Stephens & Koenig, 2015), we think it is more likely that children's learning decisions are more selectively tied to the knowledge that the speaker displays in their behavior or language use. One goal for future research is to better specify the downstream effects of knowledge attributions for a social learner (e.g., to speakers who make general claims, verifiable claims, unusual claims, causal-explanatory claims), and why they might have these effects.

It is important to point out that features of the design provided a stringent test of children's use of genericity, and may have worked against the selection of the generic speaker. Note that the present studies were designed so that there would be no logical relation between the generic claims (e.g., "Pangolins live in hollow trees") and the specific claims (e.g., "This pangolin is brown"), because we took care to provide pairs of sentences that differed in content. Thus, the knowledge conveyed in the generic statement does not entail the knowledge conveyed in the specific statement (nor vice versa). Given that inspection of the statements themselves cannot reveal a logical reason to favor one over the other, any indication that a generic speaker should be credited as more knowledgeable stems from the understanding that claims about *kinds* likely reveal or carry more information than do claims about single individuals.

One consequence of this aspect of the methodology (with both content and scope co-varying) is that the design may have indirectly encouraged participants to focus primarily on content. We speculate that differences in content may be more immediately salient than differences in generality, given that content is explicit and central to semantic gist, whereas the generic/specific difference is implicit and often distorted in recall (Gülgöz & Gelman, 2015). If instead we had contrasted speakers who provided identical content that differed only in scope (e.g., "Pangolins are brown" vs. "This pangolin is brown"), children may have more readily discerned that the relevant distinction was the generic/specific contrast, and thus made more consistent use of this information.

In the current studies, children and adults were asked to estimate a speaker's knowledge based on the verifiability or generality of their claims. Until age 7–8, young children not only credited knowledge to speakers who verifiably discussed a novel animal kind, they generalized that speaker's knowledge to new situations, involving new information. The epistemic value given to general statements by older children and adults may relate to a developmentally increasing sensitivity to the limits of our own direct experience (Woolley & Ghossainy, 2013), changes in the nature of their epistemological understanding (Kuhn & Park, 2005), increasing exposure to conversation about entities and people children will never encounter (Harris & Koenig, 2006), and to the idea that generic statements refer to more generally-known information (Cimpian & Park, 2014). In addition to these developmental possibilities, children may be showcasing a social learning strategy that any learner would exhibit in a domain with similar knowledge and experience.

## Acknowledgments

We thank Tali Levin, Caroline Hendrickson, Megan Bausman and Sashank Cherukuri for help with data collection and scoring, and are grateful to all participating families and children. Collection of the children's data was supported by the National Science Foundation (grant 656007) to M. Koenig and collection of the adults' data by NICHD (grant HD-36043) to S. Gelman.

## References

Bering, J. M. (2006). The folk psychology of souls. *Behavioral and Brain Sciences, 29*, 453–462. http://dx.doi.org/10.1017/S0140525X06009101.

Bering, J. M., & Parker, B. D. (2006). Children's attributions of intentions to an invisible agent. *Developmental Psychology, 42*, 253–262. http://dx.doi.org/10.1037/0012-1649.42.2.253.
Birch, S. A. J., Vauthier, S. A., & Bloom, P. (2008). Three- and 4-year-olds spontaneously use others' past performance to guide their learning. *Cognition, 107*, 1018–1034. http://dx.doi.org/10.1016/j.cognition.2007.12.008.
Brooker, I., & Poulin-Dubois, D. (2013). Is a bird an apple? The effect of speaker labeling accuracy on infants' word learning, imitation, and helping behaviors. *Infancy, 18*, 46–68. http://dx.doi.org/10.1111/infa.12027.
Brosseau-Liard, P. E., & Birch, S. A. J. (2011). Epistemic states and traits: Preschoolers appreciate the differential informativeness of situation-specific and person-specific cues to knowledge. *Child Development, 82*, 1788–1796. http://dx.doi.org/10.1111/j.1467-8624.2011.01662.x.
Browne, C. A., & Woolley, J. D. (2004). Preschoolers' magical explanations for violations of physical, social, and mental laws. *Journal of Cognition and Development, 5*, 239–260. http://dx.doi.org/10.1207/s15327647jcd0502_4.
Chambers, C. G., Graham, S. A., & Turner, J. N. (2008). When hearsay trumps evidence: How generic language guides preschoolers' inferences about unfamiliar things. *Language and Cognitive Processes, 23*, 749–766. http://dx.doi.org/10.1080/01690960701786111.
Chouinard, M. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development, 72*, 1–126.
Cimpian, A., & Erickson, L. C. (2012). Remembering kinds: New evidence that categories are privileged in children's thinking. *Cognitive Psychology, 64*, 161–185. http://dx.doi.org/10.1016/j.cogpsych.2011.11.002.
Cimpian, A., & Park, J. J. (2014). Tell me about pangolins! Evidence that children are motivated to learn about kinds. *Journal of Experimental Psychology: General, 143*(1), 46–55. http://dx.doi.org/10.1037/a0031687.
Cimpian, A., & Scott, R. M. (2012). Children expect generic knowledge to be widely shared. *Cognition, 123*, 419–433. http://dx.doi.org/10.1016/j.cognition.2012.02.003.
Clèment, F., Koenig, M. A., & Harris, P. L. (2004). The ontogenesis of trust. *Mind and Language, 19*, 360–379. http://dx.doi.org/10.1111/j.0268-1064.2004.00263.x.
Coady, C. A. J. (1992). *Testimony: A philosophical study*. New York: Oxford University Press.
Danovitch, J. H., & Keil, F. C. (2004). Should you ask a fisherman or a biologist? Developmental shifts in ways of clustering knowledge. *Child Development, 75*, 918–931. http://dx.doi.org/10.1111/j.1467-8624.2004.00714.x.
Faulkner, P. (2007). What is wrong with lying? *Philosophy and Phenomenological Research, 75*, 535–557. http://dx.doi.org/10.1111/j.1933-1592.2007.00092.x.
Fitneva, S. A., Lam, N. H. L., & Dunfield, K. A. (2013). The development of children's information gathering: To look or to ask? *Developmental Psychology, 49*, 533–542. http://dx.doi.org/10.1037/a0031326.
Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2009). Preschoolers' search for explanatory information within adult–child conversation. *Child Development, 80*, 1592–1611. http://dx.doi.org/10.1111/j.1467-8624.2009.01356.x.
Fusaro, M., Corriveau, K. H., & Harris, P. L. (2011). The good, the strong, and the accurate: Preschoolers' evaluations of accurate and strong informants. *Journal of Experimental Child Psychology, 110*, 561–574. http://dx.doi.org/10.1016/j.jecp.2011.06.008.
Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology, 60*, 115–140. http://dx.doi.org/10.1146/annurev.psych.59.103006.093659.
Gelman, S. A., Coley, J. D., Rosengren, K. S., Hartman, E., & Pappas, A. (1998). Beyond labeling: The role of parental input in the acquisition of richly-structured categories. *Monographs of the Society for Research in Child Development, 63*, 1–157.
Gelman, S. A., Star, J., & Flukes, J. (2002). Children's use of generics in inductive inferences. *Journal of Cognition and Development, 3*, 179–199. http://dx.doi.org/10.1207/S15327647JCD0302_3.
Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cognitive Psychology, 61*, 273–301. http://dx.doi.org/10.1016/j.cogpsych.06.001.
Gelman, S. A., Ware, E. A., Manczak, E. M., & Graham, S. A. (2013). Children's sensitivity to the knowledge expressed in pedagogical and non-pedagogical contexts. *Developmental Psychology, 49*, 491–504. http://dx.doi.org/10.1037/a0027901.
Graham, S. A., Nayer, S. L., & Gelman, S. A. (2011). Two-year-olds use the generic/nongeneric distinction to guide their inferences about novel kinds. *Child Development, 82*, 493–507. http://dx.doi.org/10.1111/j.1467-8624.2010.01572.x.
Gülgöz, S., & Gelman, S. A. (2015). Children's recall of generic and specific labels regarding animals and people. *Cognitive Development, 33*, 84–98. http://dx.doi.org/10.1016/j.cogdev.2014.05.002.
Harris, P. L., & Koenig, M. A. (2006). Trust in testimony: How children learn about science and religion. *Child Development, 77*, 505–524. http://dx.doi.org/10.1111/j.1467-8624.2006.00886.x.
Henderson, A. M. E., Graham, S. A., & Schell, V. (2015). 24-Month-olds' selective learning is not an all-or-none phenomenon. *PLoS ONE, 10*(6), e0131215. http://dx.doi.org/10.1371/journal.pone.0131215.
Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88–140.
Hollander, M. A., Gelman, S. A., & Raman, L. (2009). Generic language and judgments about category membership: Can generics highlight properties as central? *Language and Cognitive Processes, 24*, 481–505. http://dx.doi.org/10.1080/01690960802223485.
Hollander, M. A., Gelman, S. A., & Star, J. (2002). Children's interpretation of generic noun phrases. *Developmental Psychology, 38*, 883–894. http://dx.doi.org/10.1037//0012-1649.38.6.883.
Hume, D. (1748/1957). *An enquiry concerning human understanding*. Oxford: Clarendon Press.
Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science, 17*, 757–758. http://dx.doi.org/10.1111/j.1467-9280.2006.01778.x.
Keil, F. (1989). *Concepts, kinds, and conceptual development*. Cambridge, MA: MIT Press.
Keil, F. C. (2008). Adapted minds and evolved schools. *Educational Psychologist, 43*(4), 196–202. http://dx.doi.org/10.1080/00461520802392232.
Koenig, M. A. (2012). Beyond semantic accuracy: Preschoolers evaluate a speaker's reasons. *Child Development, 83*, 1051–1063. http://dx.doi.org/10.1111/j.1467-8624.2012.01742.x.
Koenig, M. A., & Echols, C. H. (2003). Infants' understanding of false labeling events: The referential role of words and the people who use them. *Cognition, 87*, 181–210. http://dx.doi.org/10.1016/S0010-0277(03)00002-7.

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*, 1261–1277. http://dx.doi.org/10.1111/j.1467-8624.2005.00849.x.
Koenig, M. A., & Jaswal, V. K. (2011). Characterizing children's expectations about expertise and incompetence: Halo or pitchfork effects? *Child Development, 82*, 1634–1647. http://dx.doi.org/10.1111/j.1467-8624.2011.01618.x.
Kuhn, D., & Park, S. H. (2005). Epistemological understanding and the development of intellectual values. *International Journal of Educational Research, 43*(3), 111–124. http://dx.doi.org/10.1016/j.ijer.2006.05.003.
Kushnir, T., Vredenburgh, C., & Schneider, L. A. (2013). "Who can help me fix this toy?" The distinction between causal knowledge and word knowledge guides preschoolers' selective requests for information. *Developmental Psychology, 49*, 446–453. http://dx.doi.org/10.1037/a0031649.
Lackey, J. (2008). *Learning from words: Testimony as a source of knowledge*. New York: Oxford University Press.
Lutz, D. J., & Keil, F. C. (2002). Early understanding of the division of cognitive labor. *Child Development, 73*, 1073–1084. http://dx.doi.org/10.1111/1467-8624.00458.
Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology, 49*, 404–418. http://dx.doi.org/10.1037/a0029500.
Nurmsoo, E., & Robinson, E. J. (2009). Children's trust in previously inaccurate informants who were well- or poorly-informed: When past errors can be excused. *Child Development, 80*, 23–27. http://dx.doi.org/10.1111/j.1467-8624.2008.01243.x.
Pappas, A., & Gelman, S. A. (1998). Generic noun phrases in mother–child conversations. *Journal of Child Language, 25*, 19–33. http://dx.doi.org/10.1017/S0305000997003292.
Pea, R. (1982). Origins of verbal logic: Spontaneous denials by 2- and 3-year-olds. *Journal of Child Language, 9*, 597–626. http://dx.doi.org/10.1017/S0305000900004931.
Quine, W. V., & Ullian, J. S. (1970). *The web of belief*. New York: Random House.
Rhodes, M., Leslie, S. J., & Tworek, C. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences (PNAS), 109*, 13526–13531. http://dx.doi.org/10.1073/pnas.1208951109.
Robinson, E. J., & Whitcombe, E. (2003). Children's suggestibility in relation to their understanding about sources of knowledge. *Child Development, 74*, 48–62. http://dx.doi.org/10.1111/1467-8624.t01-1-00520.
Shipley, E. F. (1993). Categories, hierarchies, and induction. In D. Medin (Ed.). *The psychology of learning and motivation* (Vol. 30, pp. 265–301). New York: Academic.
Shtulman, A. (2009). The development of possibility judgment within and across domains. *Cognitive Development, 24*, 293–309. http://dx.doi.org/10.1016/j.cogdev.2008.12.006.
Shtulman, A., & Carey, S. (2007). Improbable or impossible? How children reason about the possibility of extraordinary events. *Child Development, 78*, 1015–1032. http://dx.doi.org/10.1111/j.1467-8624.2007.01047.x.
Sobel, D. M., & Corriveau, K. H. (2010). Children individuals' expertise for word learning. *Child Development, 81*, 669–679. http://dx.doi.org/10.1111/j.1467-8624.2009.01422.x.
Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review, 120*(4), 779–797. http://dx.doi.org/10.1037/a0034191.
Sosa, E. (1994). Testimony and coherence. In A. Chakrabati & B. K. Matilal (Eds.), *Knowing from words: Western and Indian philosophical analysis of understanding and testimony* (pp. 59–67). Amsterdam: Kluwer Academic Publishers.
Stephens, E. C., & Koenig, M. A. (2015). Varieties of testimony: Children's selective learning in semantic versus episodic domains. *Cognition, 137*, 182–188. http://dx.doi.org/10.1016/j.cognition.2015.01.004.
Tizard, B., & Hughes, M. (1984). *Young children learning*. London: Fontana.
Tullos, A., & Woolley, J. D. (2009). The development of children's ability to use evidence to infer reality status. *Child Development, 80*, 101–114. http://dx.doi.org/10.1111/j.1467-8624.2008.01248.x.
Wilkenfield, D. A., Plunkett, D., & Lombrozo, T. (2015). Depth and deference: When and why we attribute understanding. *Philosophical Studies*. http://dx.doi.org/10.1007/s11098-015-0497-y.
Woolley, J. D., & Ghossainy, M. (2013). Revisiting the fantasy-reality distinction: Children as naïve skeptics. *Child Development, 84*, 1491–1495. http://dx.doi.org/10.1111/cdev.12081.
Woolley, J. D., Ma, L., & Lopez-Mobilia, G. (2011). Development of the use of conversational cues to assess reality status. *Journal of Cognition and Development, 12*, 537–555. http://dx.doi.org/10.1080/15248372.2011.554929.