Language Learning and Development

ISSN: 1547-5441 (Print) 1547-3341 (Online) Journal homepage: https://www.tandfonline.com/loi/hlld20

Do Children Recall Numbers as Generic? A Strong Test of the Generics-As-Default Hypothesis

Susan A. Gelman, Sarah-Jane Leslie, Rochel Gelman & Alan Leslie

To cite this article: Susan A. Gelman, Sarah-Jane Leslie, Rochel Gelman & Alan Leslie (2019) Do Children Recall Numbers as Generic? A Strong Test of the Generics-As-Default Hypothesis, Language Learning and Development, 15:3, 217-231, DOI: 10.1080/15475441.2019.1571418

To link to this article: https://doi.org/10.1080/15475441.2019.1571418

Published online: 09 Feb 2019.

LANGUAGE LEARNING AND DEVELOPMENT
2019, VOL. 15, NO. 3, 217–231
https://doi.org/10.1080/15475441.2019.1571418

Do Children Recall Numbers as Generic? A Strong Test of the Generics-As-Default Hypothesis

Susan A. Gelman [a], Sarah-Jane Leslie [b], Rochel Gelman [c], and Alan Leslie [c]

[a]Department of Psychology, University of Michigan, Ann Arbor, MI, USA; [b]Department of Philosophy, Princeton University, Princeton, NJ, USA; [c]Department of Psychology, Rutgers University, Piscataway, NJ, USA

ABSTRACT

A striking characteristic of human thought is that we form representations about abstract kinds (Giraffes have purple tongues), despite experiencing only particular individuals (This giraffe has a purple tongue). These generic generalizations have been hypothesized to be a cognitive default, that is, more basic and automatic than other forms of generalization, including those involving quantifiers such as "all" or "most." In support of this hypothesis, children often recall quantified statements (e.g., "All/many/most bears climb trees") as generic ("Bears climb trees"), and do so more frequently than the reverse error of recalling generics as quantified. The present study provides a strong test of the generics-as-default position by testing whether even numerically quantified statements (e.g., "Five giraffes have purple tongues") are recalled as generic. Two groups of three-year-old children (N = 74) were tested: those who held a correct numerical interpretation of five ("5-knowers"), and those who did not ("non-5-knowers"). Results indicate that non-5-knowers often defaulted to the generic in recall after hearing a number (e.g., recalling "Five giraffes." as "Giraffes."), whereas 5-knowers did not. Thus, consistent with the generic-as-default hypothesis, before children have a specific numerosity assigned to "five," they display a tendency to recall numerically quantified phrases as generic.

A commonplace yet remarkable aspect of human cognition is that we readily form abstract representations on the basis of small samples of particular instances. Experience with one or two cats may be sufficient for a child to reliably identify cats in the future, and a single mouthful of lima beans may be sufficient for a child to conclude that lima beans are disgusting. As these examples illustrate, the kinds of generalizations that we most commonly express in language are generic (Carlson & Pelletier, 1995; Gelman, Hollander, Star, & Heyman, 2000; Prasada, 2000)—that is, they are generalizations about kinds (cats, lima beans). Generics can be distinguished from generalizations based on logical quantifiers, such as all, most, or some.

Generics have captured the interest of linguists, philosophers, and psychologists, because their semantics belie any simple characterization based strictly on statistical frequency, and instead require domain-specific knowledge and causal theories (Brandone, Cimpian, Leslie, & Gelman, 2012; Brandone & Gelman, 2009, 2013; Cimpian, Brandone, & Gelman, 2010; Leslie, 2007, 2008; Tasimi, Gelman, Cimpian, & Knobe, 2017). For example, although most sharks don't attack humans, it is permissible to assert that sharks attack humans. Generics privilege features of a category that emerge with development, again even when contrasted with frequency (Cimpian, Gelman, & Brandone, 2010). Similarly, adults accept generics that involve properties for which an animal has an innate propensity, even if no instances currently have that feature (e.g., "Dobles have claws" if they were

CONTACT Susan A. Gelman gelman@umich.edu Department of Psychology, University of Michigan, 530 Church St., Ann Arbor, MI 48109-1043

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hlld.

born with claws that they then lost due to ingesting a bad chemical) (Gelman & Bloom, 2007). Moreover, generic language shapes children's expectations about novel categories beyond frequency considerations: when information about a new category is provided in generic form, children are more likely to treat the category as a stable natural kind with an internal essence (Gelman, Ware, & Kleinberg, 2010; Rhodes, Leslie, & Tworek, 2012).

Given the link between generics and kind concepts, given the importance of kind representations for human reasoning, and given the intriguing puzzle of generic acquisition, it is important to understand when and how children acquire generic generalizations. One theory that has received support is that generics are in fact a default mode of generalization. Leslie (2007, 2008) and Leslie & Gelman (2012) suggest that people most readily and easily generalize on the basis of generics rather than quantifiers. The evidence for this proposal comes from converging lines of research with young children. Generics are acquired early in development, by about 2.5 years of age (Gelman, Goetz, Sarnecka, & Flukes, 2008; Gelman & Raman, 2003; Graham, Gelman, & Clarke, 2016; Graham, Nayer, & Gelman, 2011) and well before quantifiers (English: Hollander, Gelman, & Star, 2002; Mandarin: Tardif, Gelman, Fu, & Zhu, 2012; Quechua: Mannheim, Gelman, Escalante, Huayhua, & Puma, 2011). Children's interpretation of generics adheres to some of the same distinctive patterns found in adults, such as reflecting conceptual centrality rather than frequency per se (e.g., "Lions have manes" is true but "Lions are male" is not; Brandone et al., 2012; Cimpian & Markman, 2009, 2011), and displaying an asymmetry between evidence required for assertion vs. prevalence implied by assertion (Brandone, Gelman, & Hedglen, 2015). Additionally, 3-year-old children interpret "all"-quantified sentences as if they were generic, but not vice versa (Hollander et al., 2002). And in a memory paradigm, participants (adults as well as children) show an asymmetric bias to recall quantifiers as if they were generic (not vice versa). For example, participants recall "All bears climb trees" as "Bears climb trees," and not the reverse (Gelman, Leslie, Was, & Koch, 2015; Leslie & Gelman, 2012).

This asymmetry in recall does not simply reflect an information-processing shortcut in recall (e.g., dropping the initial word). First, participants do not show the asymmetry for all quantifiers: adults and 4-year-olds do so only for *all* and *most* (not *some*) (Leslie & Gelman, 2012). Second, the asymmetry is restricted to generalizable properties, such as eating fruits and vegetables, and not idiosyncratic properties, such as getting mud in one's hair (Sutherland, Cimpian, Leslie, & Gelman, 2015). Third, Spanish-speaking participants (children as well as adults) show the asymmetry even when converting a quantified statement to a generic requires the addition of a linguistic marker (e.g., in Spanish, recalling *Many bears climb trees* [*Muchos osos trepan arboles*] as *Bears climb trees* [*Los osos trepan arboles*] requires dropping the quantifier *muchos* and replacing it with the article *los*; Gelman, Tapia, & Leslie, 2016).

One prediction of the generics-as-default hypothesis is that of developmental change. Specifically, we would expect to see an initially broad tendency to convert any sort of non-generic generalization to generic, but that this broad expectation would increasingly narrow with age, as additional semantic and conceptual understandings come into play that would block a generic interpretation. For example, a young child may recall any logical quantifier as generic, whereas an older child may restrict this to quantifiers that apply broadly to the kind (e.g., *many, most, all*), excluding *some*. And, indeed, the available evidence does reveal a systematic developmental narrowing of generics as a default: from a broad range of quantifiers in young children (*all, most, some* at age 3), to increasingly more restricted with age (*all, most* at age 4; and then *most* [and *all*, to a lesser extent] for adults). This pattern accords with the idea that the youngest children are most likely to default to generics—although, notably, even adults display this pattern with certain quantifiers.

The data thus far, then, accord with the idea that generics are a default interpretation, one that is broadly applied in early childhood, and increasingly narrows with age. What is unknown, however, are the limits to this default process, as well as the mechanisms that lead older children and adults to restrict their treatment of generics as default. Examining whether children default to generics when hearing sentences with numerical expressions (e.g., "Five bears climb trees") provides a novel and critical test of these issues.

On the one hand, children may treat numbers differently from quantifiers. There is a fundamental conceptual distinction between numbers, which in these constructions pick out specific, concrete individuals ("Five giraffes have purple tongues" makes a claim about five particular giraffes), and quantifiers, which lend themselves to more open-ended interpretations ("All giraffes have purple tongues" makes a broad and unbounded claim about the entire set of giraffes). This can be seen in the structure of counting sequences, where individual items are tagged with unique number-words, a principle which even very young children honor (Gelman & Gallistel, 1978). Given that generics, like quantifiers, express claims about kinds, not specific individuals ("Giraffes have purple tongues" also applies broadly to the entire kind), this distinction may be sufficient to block even young children from recalling numerical expressions (e.g., "Five giraffes have purple tongues") as generic. The number word in particular may be a clear and overt signal that would block recall using a generic.

On the other hand, if children have a powerful tendency to treat generics as default, they may do so despite these differences. Numbers and quantifiers are outwardly similar in a variety of ways, particularly in English (Barner, Libenson, Cheung, & Takasakic, 2009). They can appear in many of the same syntactic frames (e.g., prior to a plural noun; *Five/some bears climb trees*). Indeed, Barwise and Cooper (1981) originally argued that numbers, just like "some," "all," and "none," function as quantifying determiners within the framework of generalized quantifiers. Further, of the four linguistic properties distinguishing number words from adjectives listed by Bloom and Wynn (1997), three are shared with quantifiers: they cannot appear with modifiers (**the very five/some dogs*); they precede adjectives within the NP and cannot follow them (**the big five/some dogs*); and they can occur in the partitive construction (*five/some of the dogs*). The fourth property of numbers is shared with a subset of quantifiers (they cannot be used with mass nouns; **five/many milk*). Numbers and quantifiers are not fully interchangeable (e.g., numbers can't precede mass nouns, but some quantifiers can; numbers but not quantifiers appear in counting sequences), but from a strictly formal analysis, noun phrases headed by numerals are similar to those headed by quantifiers in many respects. Semantically, too, numerals and quantifiers are similar, in that they are internally ordered from smallest to largest (e.g., 1, 2, 3…; *some, most, all*) (Hurewitz, Papafragou, Gleitman, & Gelman, 2006).

We are unaware of any research to date examining whether children recall number sentences as generics. However, studies examining children's recall of statements regarding individuals is informative. Three- and four-year-old children appropriately distinguish between generic and specific noun phrases in recall (e.g., "Bears climb trees" vs. "This bear climbs trees") (Gelman & Raman, 2007). Also, three- and five-year-olds are less likely to default to generics when hearing quantifiers that refer to sets of individuals (e.g., "All *of these* bears…") than quantifiers that refer to abstract sets (e.g., "All bears…") (Gelman et al., 2015). On the other hand, an asymmetric tendency to misremember non-generic information as generic (more than the reverse error) is found even when children hear facts about individuals (Gelman & Raman, 2007; Gülgöz & Gelman, 2015). It is thus unclear how children will treat numerals, which not only are specific in meaning, but are used when counting, which involves a highly specific procedure of pointing to items in a specific context (Gelman, 2015; Gelman & Gallistel, 1978).

To this point we have been discussing findings as a function of child age, but in the case of numerical reasoning, there is variation among children within an age group that also must be considered. Specifically, preschoolers vary in their performance when asked to select a certain number of items out of a larger array (e.g., "Give five apples to the frog"; also known as the "give-a-number" task) (Sarnecka, 2015; Sarnecka & Carey, 2008; Wynn, 1990, 1992). Success is defined as giving the correct amount when hearing the corresponding number, and *not* giving that amount when hearing a different number. Children who fail to provide the correct number of items on this task do still understand that unmapped numbers refer to quantities; for example, they judge that the number of a set changes only when items are added or removed, and not when items are shaken or rotated (Gelman, 2006; Sarnecka & Gelman, 2004) or moved (see also Gelman, 1972). Moreover, those who fail the give-X task nonetheless solve addition and subtraction problems with small numbers (e.g., Gelman, 1972; Starkey & Gelman, 1982; Wynn, 1992). They even can predict the result of adding and subtracting one item to a display (Zur & Gelman, 2004). However, it is not until children pass the number-knower levels that they understand that adding one item to a given number inside a container is the next number in the count list, or that identical sets must have identical numbers (Sarnecka, 2015). Thus, there is considerable variation within an age group in children's number-knower status, as well as across different tasks that assess understanding of counting and arithmetic principles (see a review in Danan & Gelman, 2018). But it does seem as if success on the Give-x task indexes a robust understanding of cardinality.

Our primary question is whether children default to generics upon hearing numerically quantified statements, as well as the possible developmental function of this effect. We therefore wished to assess children's number knowledge using a measure that would allow us to assess developmental variation in number concepts, and to do so at an age when children are producing generics in recall. The number-knower task is ideal for these purposes, given that it is relatively more demanding than other measures. An additional benefit is the ubiquity and reliability of results obtained with the number-knower task.

Therefore, to test whether children default to generics upon hearing numerically quantified statements, we chose to assess their number-knower status. We would predict that knowing the precise numerosity of the number provided would block a generic interpretation. The more interesting question is whether children who lack that rich number knowledge default to a generic interpretation. Being able to compare children's recall performance as a function of their number-knower status would provide important insight into the developmental mechanism underlying the asymmetry in generic recall that was observed in prior studies.

The present experiment

In this study, we pose the following questions. (a) Do children default to generics in recall, after hearing a numeral (e.g., recalling "Five horses…" as "Horses…")? (b) Are those who do not yet have a specific numerosity assigned to a number, as assessed by the Give-X task, more likely to default to generics in their recall? (c) Are children more likely to default to generics in their recall if they do not yet apply a scalar implicature to "some"? Those who apply a scalar implicature to "some" (i.e., treating it as contrasting with "all") may view it as fairly limited in scope and restricted to a relatively small number of particular members, whereas those who do not yet apply a scalar implicature may interpret "some" as quite broad in scope, and thus conceptually more related to generics.

We chose to focus on 3-year-olds, as in prior research they showed a strong tendency to recall *some*-quantified sentences as generic (Leslie & Gelman, 2012), and because they are in transition, in terms of their number-knower status, thus permitting a comparison of children who do vs. do not succeed on the give-a-number task (Sarnecka & Carey, 2008; Sarnecka & Gelman, 2004). Adults were included as well, as a comparison case. We also focused on the numeral 5, since knowledge of 5 corresponds to having knowledge of the cardinal principle of numbers (i.e., that the last word in a count sequence corresponds to the numerosity of the set; see Sarnecka & Carey, 2008).

Method

Participants

Seventy-four three-year-olds (3.06–3.99, mean age 3.66), 41 girls and 33 boys, participated. Mean ages in the two conditions were 3.65 for children in the Generic/Some condition and 3.68 for children in the Generic/Five condition (described below). Parents identified their children as White or European-American (n = 37), Black or African-American (n = 2), Asian or Asian-American (n = 2), Hispanic (n = 1), multiracial (n = 6), or not reported (n = 26). Participants were recruited in a midwestern U.S. university town. An additional 27 children were tested but not included, for the following reasons: did not complete the task (4), limited English proficiency (1), experimenter error (1), parental interference (1), and insufficient data (20; see Coding section in the Results, below).

Thirty-two college students (undergraduates at a large Midwestern University who received course credit in Introductory Psychology for their participation) were included as a comparison group. They included 18 men and 14 women, ranging in age from 18–22 (mean age 20.6). They self-identified as White or European-American (n = 18), Black or African-American (n = 3), Asian or Asian-American (n = 8), and Biracial or Multiracial or Other (n = 3). Four additional college students were tested but not included, due to language fluency issues (2), experimenter error (1), and having already participated in the experiment previously (1).

Items

Items consisted of eight sentences (see Table 1), each with a corresponding set of five exemplar photographs (e.g., the sentence "Giraffes have purple tongues" was accompanied by photographs of five different giraffes; see Figure 1). The properties in the sentences were not visible in the pictures (e.g., the giraffes' tongues were not visible). We included five exemplars of each animal so that there was a pragmatically sensible context for making a claim using the word *five*. Additionally, the inclusion of multiple instances provides a conservative test of the generics-as-default hypothesis, since prior research indicates that parents generate generic language more often in the context of a single exemplar than in the context of multiple instances (Pappas & Gelman, 1998).

Procedure

Child participants were tested individually, either at an on-campus child development laboratory or local preschools. The procedure had five tasks, in this order: a warm-up, a memory task, two tasks to assess children's number knowledge (Counting; Give-a-Number), and a scalar implicature task. The fixed order was designed to minimize carryover effects from one task to the other. We reasoned that presenting the number knowledge tasks first could interfere with the memory task, whereas the memory task was unlikely to affect performance on the number knowledge tasks.[1] A second consideration is that presenting the number knowledge task would mean that any carryover would differentially impact memory across the two conditions (i.e., carryover would likely be higher in the Generic/Five condition than the Generic/Some condition), and we wished not to introduce differential overlap across conditions between the number task and the memory task.

Table 1. Test sentences, in generic, five, and some wording.

Generic	Five	Some
Snails lay pink eggs.	Five snails lay pink eggs.	Some snails lay pink eggs.
Elephants walk on tiptoe.	Five elephants walk on tiptoe.	Some elephants walk on tiptoe.
Hummingbirds can fly backward.	Five hummingbirds can fly backward.	Some hummingbirds can fly backward.
Snakes swallow food whole.	Five snakes swallow food whole.	Some snakes swallow food whole.
Owls eat at night.	Five owls eat at night.	Some owls eat at night.
Giraffes have purple tongues.	Five giraffes have purple tongues.	Some giraffes have purple tongues.
Zebras sleep standing up.	Five zebras sleep standing up.	Some zebras sleep standing up.
Ants have strong legs.	Five ants have strong legs.	Some ants have strong legs.

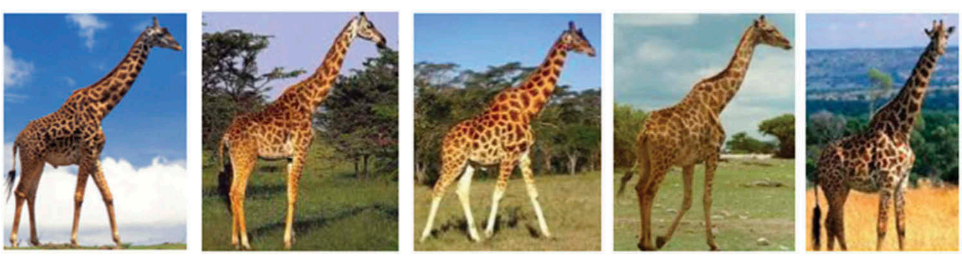

Figure 1. Pictures used in giraffe item.

Warm-up

The warm-up was designed to get children comfortable with talking to the puppet ("Mr. Squirrel"), and with repeating a sentence provided by the researcher. On the first warm-up trial, the child was asked what his/her favorite color was, then the researcher said what her favorite color was, and then the child was asked to tell both favorite colors to Mr. Squirrel, who had been out of earshot and didn't hear what they had said. On the second warm-up trial, the researcher stated a fact about a picture ("This ice cream is yummy"), and then asked the child to tell Mr. Squirrel what she had said. If the child did not provide the full sentence, the information was repeated and the child was prompted again. At the end of the warm-up, children were told, "Mr. Squirrel is going to go away again. I'm going to tell you things about pictures in this book. When Mr. Squirrel comes back later, you can tell him what I said about the pictures, okay?" This served as a segue to the memory task.

Memory task

Children were randomly assigned to one of two conditions: generic/*five* or generic/*some* (corresponding to the language provided in the test sentences; see Table 1). Participants in the generic/*five* condition heard four sentences with generic wording and four sentences with *five* wording; participants in the generic/*some* condition heard four sentences with generic wording and four sentences with *some* wording. The assignment of sentence to either generic or non-generic wording was counterbalanced across participants (e.g., within each condition, half the participants heard the giraffe sentence as generic, and half heard the giraffe sentence as non-generic). The items were presented in one of two random orders.

The method was modeled after a task used in prior research (e.g., Gelman et al., 2015, 2016; Leslie & Gelman, 2012), with some differences to reflect the novel research question. During the teaching phase, children saw four item sets, one at a time. For each item set, the researcher showed photographs of five animals of a given type (e.g., five giraffes, similar in appearance but distinct individuals) and provided a novel fact, in one of three formats: generic (e.g., "Giraffes have purple tongues"), *some* (e.g., "Some giraffes have purple tongues"), or *five* (e.g., "Five giraffes have purple tongues"). Each fact was stated twice before moving on to the next item. The teaching phase was followed by a 4-min distractor task in which the child played with a set of Legos, and then the recall phase, in which the child saw the same pictures as before, one set at a time, and were asked to tell Mr. Squirrel what the researcher had said about each set. If a child only provided a label, they were prompted to tell Mr. Squirrel everything the researcher had said; if they provided a pronoun (e.g., "they"), they were asked to elaborate (e.g., "Can you tell me who 'they' is?"). Responses were audio-recorded and transcribed.

Counting

The child was asked to count up to ten and their response was recorded. If they were unable to do so, the researcher prompted them to count together. This served both to remind children that the focus was numbers, and to assess whether they spontaneously produced the number "five."

Give-a-number

This task is based on Wynn (1990, 1992). The child was told, "Mr. Squirrel is now very hungry! In this game, you will give him acorns." Then a bucket of toy acorns was placed on the table, and for each of a series of trials, the child was asked to give a certain number of acorns to Mr. Squirrel, by placing them on a tray. For each numerosity, the child was first asked to place the amount (e.g., "Can you give Mr. Squirrel three acorns?"), and then to confirm it (e.g., "Is that three?"). If the child did not confirm, then the trial restarted. Testing started with the numbers 1 and 3, and then based on the child's performance moved up by increments of 1 (if correct), or down by increments of 1 (if incorrect). Testing continued until the child had at least two successes at a given number N and at least 2 failures at N + 1, up to N = 6 (indicating "N-knower" status).[2] Children were free to count the acorns if they wished.

Scalar implicature

Children were asked, "Are some of the crayons in the box?" three times (order counterbalanced across participants): when 2 out of 6 were in the box, when 4 out of 6 were in the box, and when 6 out of 6 were in the box. If children interpret "some" according to a scalar implicature, then they should say "yes" for 2 and 4, but "no" for 6.

Coding

Based on Leslie and Gelman (2012) and expanded to include the new generic/*five* condition, each response on the Memory task was coded as generic (e.g., "Giraffes have purple tongues"), number (e.g., "Five giraffes have purple tongues"), quantifier (e.g., "Some giraffes have purple tongues"; "Elephants sometimes tiptoe on their tiptoes"), or other (including ambiguous responses, such as "They have purple tongues"). We did not require participants to recall the predicates accurately or even fully (e.g., "Snakes hide in holes," "Zebras can walk at the night," and "Snakes um…snakes um…have…" were all coded as generic). However, responses that simply provided a label (e.g., "Giraffes"; "Strong legs and ants") or did not include a relevant noun phrase (e.g., "I don't know"; "I forgot"; "I can't find any guesses") were considered uncodable. Children who did not provide at least one codable response per sentence type (generic and non-generic) were dropped from the study. Inter-rater agreement exceeded 97% on each code, with kappas per coding category ranging from .95–.99.

Coded responses were then classified as correct (generic after hearing a generic, quantifier after hearing *some*, and number after hearing *five*) or opposite (generic after hearing *some* or *five*; a quantifier after hearing a generic in the generic/*some* condition; a number after hearing a generic in the generic/*five* condition). We did not require that participants provide the same quantifier or number as they heard (for example, a child who said "Three ants …" rather than "Five ants …" was coded as "correct"; a child who said "Many hummingbirds…" instead of "Some birds…" was coded as "correct").[3]

Adult participants

The college students who participated in the task received only the memory task and the scalar implicature task, sans puppet. The distractor task for college students was not playing with Legos, but rather involved listing as many words as they could within different (non-animal) categories, such as furniture, clothing, college majors, tools, units of measurement, sports, holidays, etc. Participants were tested individually in the same on-campus laboratory in which the children were tested.

Results

Number knowledge

Based on their performance on the Give-a-number task, children were classified as either 5-knowers (n = 38) or non-5-knowers (n = 36). Knower-level (N-knower) was defined as at least two successes at a given number N (i.e., providing the exact number of acorns for the number requested, for example three acorns when asked, "Can you give Mr. Squirrel 3 acorns?") and at least 2 failures at N + 1, up to N = 6. Children were classified as 5-knowers if they succeeded on the Give-a-number task for all numbers up to and including 5 (some of whom also succeeded on 6, which was the highest number tested). Non-5-knowers reached an upper limit of success on the Give-a-number task that ranged from 1–4. Most children in both groups were able to count to 5 without the researcher's help (95%, 74%) and spontaneously mentioned 5 (95%, 77%). Thus, the non-5-knowers were at least familiar with the word "five." These groups did not significantly differ in age, Ms = 3.69 and 3.64, respectively, $t(72) = 1.09$, $p = .279$.

Correct responses (producing a generic after hearing a generic, a number after hearing five, or a quantifier after hearing some)

We conducted a 2 (condition: generic/*five*, generic/*some*) x 3 (group: child 5-knower, child non-5-knower, and adult) x 2 (sentence type: generic, non-generic) repeated measures ANOVA. Condition and group were between-subjects factors, and sentence type was a within-subject factor. Results indicated a main effect of group, $F(2, 100) = 37.40$, $p < .001$, $\eta_p^2 = .43$. Setting our alpha level to .017 to account for multiple comparisons (.05/3 tests), post-hoc analyses revealed that adults were more accurate than both groups of children, $ps < .001$, who did not differ from one another, $p = .45$. There was also a main effect of sentence type, $F(1, 100) = 14.82$, $p < .001$, $\eta_p^2 = .13$, and a sentence type x group interaction (see Figure 2), $F(2, 100) = 4.95$, $p = .009$, $\eta_p^2 = .09$. No other effects were significant. As predicted, and consistent with the generics-as-default hypothesis, participants were more often correct on the generic sentences than the non-generic sentences. However, setting our alpha level to .017 to account for multiple comparisons (.05/3 tests), post-hoc analyses of the interaction revealed that this difference in recall between generic and non-generic sentences was exclusively due to child non-5-knowers, $p < .001$; child 5-knowers and adults displayed no significant difference in correct recall of generic versus non-generic sentences, $ps = .61$ and .155, respectively. Post-hoc comparisons revealed that for non-5-knowers, the effect of sentence type held up for both conditions ($Ms_{\text{generic/five}} = 2.68$ [generic], 0.53 [five], $p < .001$; $Ms_{\text{generic/some}} = 2.06$ [generic], 0.71 [some], $p = .014$).

Opposite responses (producing a generic after hearing some or five, a number after hearing a generic in the generic/five condition, or a quantifier after hearing a generic in the generic/some condition)

We conducted a 2 (condition: generic/*five*, generic/*some*) x 3 (group: child 5-knower, child non-5-knower, and adult) x 2 (sentence type: generic, non-generic) repeated measures ANOVA. Condition and group were between-subjects factors, and sentence type was a within-subject factor. As in the correct responses above, there was a main effect of sentence type, $F(1, 100) = 22.35$, $p < .001$, $\eta_p^2 = .18$, and a sentence type x group interaction (see Figure 3), $F(2, 100) = 6.16$, $p = .003$, $\eta_p^2 = .11$. No other effects were significant. As predicted, and consistent with the generics-as-default hypothesis, participants more

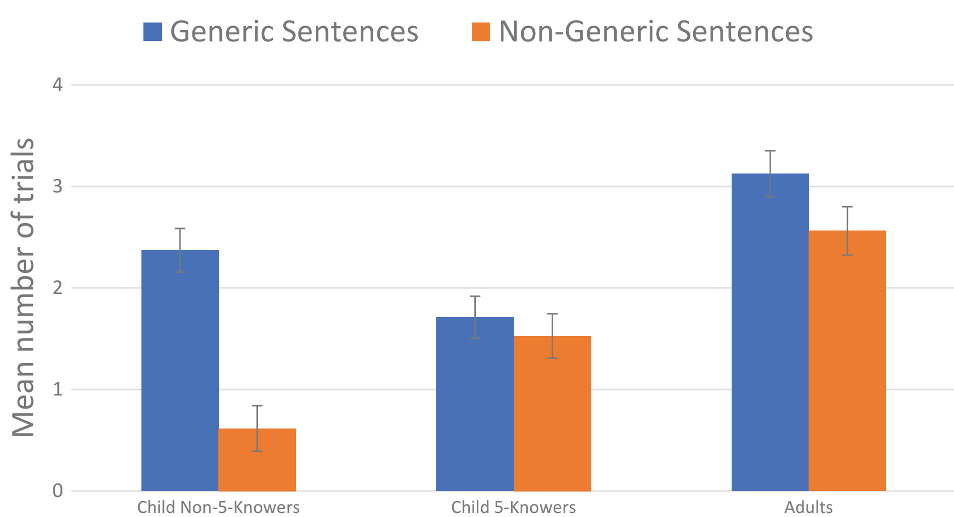

Figure 2. Correct responses: Mean number of trials (out of four) on which participants correctly recalled generic and non-generic sentences, as a function of group (child non-five-knowers, child five-knowers, adults). **Correct** responses were defined as producing a generic after hearing a generic, a number after hearing *five*, or a quantifier after hearing *some*.

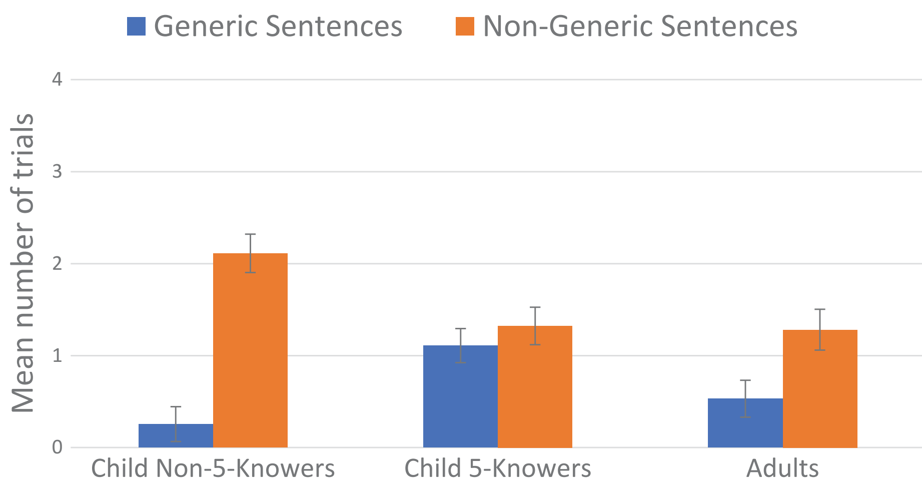

Figure 3. Opposite responses: Mean number of trials (out of 4) on which participants produced opposite responses in recall of generic and non-generic sentences, as a function of group (child non-5-knowers, child 5-knowers, adults). **Opposite** responses were defined as producing a generic after hearing *some* or *five*, a number after hearing a generic in the generic/*five* condition, or a quantifier after hearing a generic in the generic/*some* condition.

often produced opposite responses on the non-generic sentences than the generic sentences. However, setting our alpha level to .017 to account for multiple comparisons (.05/3 tests), post-hoc analyses revealed that this difference was exclusively due to child non-5-knowers, $p < .001$; child 5-knowers and adults displayed no significant difference in opposite recall of non-generic vs. generic sentences, *ps* = .52 and .04, respectively. Post-hoc comparisons revealed that for non-5-knowers, the effect of sentence type held up for both conditions ($Ms_{\text{generic/five}} = 0.16$ [generic], 2.11 [five], $p < .001$; $Ms_{\text{generic/some}} = 0.35$ [generic], 2.12 [some]; $p = .001$).

Five vs. some

In order to test whether 5-non-knowers' performance could be explained by a general lack of attentiveness in the study or whether it truly reflects the extent to which they default to generics relative to 5-knowers, we examined whether responses differed as a function of condition (generic/*five* vs. generic/*some*), for each group separately. As described in the Coding section of the ms., responses included any number or quantifier (i.e., they did not have to be the same number or quantifier that had been presented during the teaching phase). If children were being attentive, they were expected to be more likely to use number responses in the generic/*five* condition than in the generic/*some* condition, and more likely to use quantifier responses in the generic/*some* condition than in the generic/*five* condition. As can be seen in Table 2, all three groups distinguished between the conditions in their frequency of quantifiers, which were used almost exclusively in the generic/*some* condition. Furthermore, the child 5-knowers and adults (but not child non-5-knowers) distinguished between the conditions in their frequency of numbers.

Scalar implicatures

Responses to the scalar implicature questions were analyzed with non-parametric tests, given that the data consisted of three binary responses per participant ("yes" or "no" to each of the three questions). First, each participant was classified as either displaying the predicted implicature pattern (responding "yes" to the question "Are some of the crayons in the box?" when there were 2 or 4 crayons, and "no" when there were 6 crayons) or not (all other response patterns). Using the binomial test, we compared the rates at which participants showed the implicature pattern (8% of

Table 2. *T*-tests comparing across generic/*five* vs. generic/*some* conditions, as a function of group. Number responses and quantifier responses refer to the use of numbers/quantifiers in recall statements. These are composite scores out of eight trials.

Group:	Generic/*Five* Condition (*M, SE*)	Generic/*Some* Condition (*M, SE*)	*t*-test
Non-5-Knowers (age 3)			
Number responses	0.68 (.33)	0.12 (.08)	$t(34) = 1.61, p = .12$
Quantifier responses	0.11 (.07)	1.06 (.39)	$t(34) = -2.55, p = .016$
5-Knowers (age 3)			
Number responses	2.55 (.69)	0.00 (.00)	$t(36) = 3.52, p = .001$
Quantifier responses	0.00 (.00)	2.72 (.75)	$t(36) = -3.82, p = .001$
Adults			
Number responses	3.63 (.45)	0.00 (.00)	$t(30) = 8.13, p < .001$
Quantifier responses	0.00 (.33)	2.56 (.46)	$t(30) = -5.62, p < .001$

child non-5-knowers, 18% of child 5-knowers, and 16% of adults) to chance (12.5%), and found that none of the three groups exceeded chance, $ps \geq .19$. A series of chi-square tests further revealed no differences as a function of group, of condition, or of condition within each group, all $ps > .19$.

Discussion

Given prior research that children default to generics in their recall of quantifiers (*all, most, some*), we asked whether an analogous pattern obtains even in the strong case when the quantified information is provided not with a quantifier but with a number (*five*). Numbers are conceptually distinct from quantifiers in referring to precise numerosities, they are formally distinct from quantifiers (and other words) in their distributional patterns in speech (Bloom & Wynn, 1997; Gelman & Gallistel, 1978; Leslie, Gelman, & Gallistel, 2008), and they differ in their processing and pragmatic implications (Barner, Chow, and Yang 2009; Gallistel, 2018; Gelman, 2015; Hurewitz, Papafragou, Gleitman, & Gelman, 2006). Thus, one possibility was that children would not default to generics when hearing numerical statements. In contrast, a distinct possibility was that the tendency to represent generalizations as generic may be sufficiently powerful to exert an effect on the memory of numbers as well as quantifiers, at least for a subset of the children. Second, and relatedly, we were interested in comparing children who did versus did not succeed on the Give-a-number task for the target number (*five*). This provides an opportunity to gain insights into the mechanisms that lead children to restrict their treatment of generics as default.

To test these questions, we provided three-year-old children and adults with a verbal memory task (modeled on Leslie & Gelman, 2012), in which participants heard a series of statements, both generic and non-generic (either with the quantifier *some*, or the numeral *five*), and after a 4-min delay, were asked to recall the sentences, given a set of five pictures as a prompt. The children were also tested on their number knowledge, given a counting task and a Give-a-Number task.

Below, we provide a recap of the results and their interpretation, organized in terms of the three key questions listed in the Introduction.

Did children default to generic in recall (e.g., "Horses…"), after hearing a number (e.g., "Five horses…")?

Yes; this finding was very robust. Children who have not learned the precise numerosity of "five" were more correct on generic sentences than on quantified sentences, and this held when looking separately at children in the condition involving *five*. Furthermore, children who have not learned the precise numerosity of "five" were more likely to produce opposite responses on quantified sentences than on generic sentences; again, this held when looking separately at the condition involving *five*. That children defaulted to generic after hearing noun phrases that included *five* was striking, given that there were five visible instances of the category present during both the teaching phase and the recall phase, thus serving as a potential reminder of the number *five*. We suggest that the mechanism underlying this result is knowledge of the meaning of "five," which allows children to focus on particular instances and thus override their existing tendency to default to generics.

Did non-5-knowers default to generic more than 5-knowers?

Yes. In fact, *only* those children who did not know the precise numerosity of "five" defaulted to generic, and they did so both for sentences starting with *some* and for sentences starting with *five*. That non-5-knowers defaulted to generic after hearing 5 was not due to unfamiliarity with the number: 74% of non-knowers could count to at least 5 without help, and 77% spontaneously provided the number 5 on the counting task. The greater rate of defaulting to generics among non-knowers than knowers is reminiscent of the developmental findings reported in Leslie and Gelman (2012). In that earlier work, younger children were more likely than adults to recall quantified statements as generic, and 3-year-olds were more likely than 4-year-olds to recall quantified statements as generic. Although the present study did not compare different ages, it did find that the less firm a child's knowledge about the meaning of a numeral, the more likely they were to default to generics in their recall. (Note that the prior work did not include a test of children's number knowledge.)

One puzzle is why 3-year-old 5-knowers did not default to generic in this study. In prior research, 3-year-olds defaulted to generic in their recall of sentences quantified with *some* (Leslie & Gelman, 2012). However, that study didn't assess knower-status. In the current dataset, 3-year-olds overall defaulted to generic for *some*, thus replicating the finding from Leslie and Gelman (2012) when examining 3-year-olds as a whole – but this collapsed over two distinct patterns: a strong tendency to default on *some* for non-knowers, and no tendency to default on *some* for knowers. We suspect this same effect may have been operating in Leslie and Gelman (2012). Furthermore, in that article, 4-year-olds (who can be assumed to know the numerosity of 5) defaulted to generic for *all* and *most* but not for *some*—in other words, showing the same pattern with *some* as found in the present experiment with 5-knowers (at age 3).

A related puzzle is why 5-knowers differed from non-5-knowers in *both* conditions (generic/some as well as generic/five). We speculate that two factors may be at play. First, prior research has found that quantifier knowledge and number knowledge develop in tandem within this age group. Specifically, (Barner, Chow, and Yang, 2009 ; Barner, Libenson, Cheung & Takasaki, 2009) found a positive, age-independent correlation between quantifier knowledge (including "some") and number-knower levels in children in this age range (mean age 45 months). They suggest that this link is due to syntactic bootstrapping based on overlap between the constructions in which these words can appear in the input language (e.g., "*five/some* of the toys"). Given their findings, it is likely that the 5-knowers in the present study differed from non-5-knowers not only in their understanding of "five", but also in their understanding of "some". In both cases, knowing the meanings of these words allows children to override their tendency to default to generics (which is still present when hearing quantifiers such as "most" or "all").

Additionally, the context of the present memory task may have encouraged children with more advanced number concepts to focus on individual instances even in the absence of number labels. Recall that children saw five animal instances on each trial (in both the teaching and the testing phases). Five-knowers may have been more attuned to the numerosity of these sets, and thus more likely to interpret "some Xs" (e.g., "some giraffes") as referring to these five instances (e.g., these five giraffes) rather than the more general notion (e.g., some unspecified subset of all giraffes).

Did differences between 5-knowers and non-5-knowers reflect different rates of applying scalar implicatures?

We found no evidence that differences in performance between 5-knowers and non-5-knowers were due to different levels of using scalar implicatures. 5-knowers did not show evidence of using a scalar implicature when interpreting *some* (i.e., they reported that "some of the crayons are in the box" when all of the crayons were in the box—6 out of 6). Therefore, differences in performance on the memory task were not due to changes in their understanding of "some," at least on this measure. This is particularly surprising, given that knowers and non-knowers differed in how often they defaulted to generic on *some* as well as *five*. We speculate that perhaps the scalar implicature task that we employed did not sufficiently signal that the relevant set was just the six crayons in the context. For example, some participants may have assumed that "the crayons" referred to the crayons in their original packaging (which typically would include a minimum set size of 8—one of each primary color). If so, this could also account for why adults also failed to show a scalar implicature effect. Had we emphasized the relevant set (e.g., "Look at these crayons. Are some of these crayons in the box?"), this may have provided a less ambiguous test. Alternatively, it may be that adults (who were college students tested in a psychology lab) applied a more strictly logical interpretation in this context than they would normally do in everyday conversation. In summary, we cannot entirely rule out differences in scalar implicatures as an explanation for the 5-knower vs. non-5-knower difference, and would welcome further research to test this possibility more convincingly.

Conclusions

These results suggest that the tendency to represent generalizations as generic emerges early in development and is broadly encompassing in scope. Extending beyond prior work finding that 3-year-old children tend to recall sentences with a range of verbal quantifiers (e.g., *all, most, some*) as generic, we found that 3-year-old children who fail the Give-a-number task for the number "five" also display a tendency to recall numerically quantified phrases (e.g., "Five giraffes…") as generic. This is in contrast to 3-year-olds who understood the precise numerosity of *five*, suggesting that learning the precise meaning of a number term can block a generic interpretation in recall. These data thus add to our understanding of the mechanisms that permit or restrict unwarranted generic generalizations. Specifically, we suggest that knowing "five" allows children to focus on particular instances, and thus override the default generic interpretation. To be clear, the claim is *not* that learning precise numerosities for count words such as "five" eliminates the generics-as-default tendency. Rather, we suggest that there is a fundamental propensity to default to generics in human reasoning, throughout the lifespan, but that its expression lessens with age. The youngest children are most susceptible to showing this tendency, but become increasingly capable of suppressing it, particularly as they acquire linguistic expressions that directly compete with generic meaning (such as "five" or "some"). This proposal is consistent with converging evidence for adults treating generics as the default on tasks ranging from sentence recall (Gelman et al., 2016; Leslie & Gelman, 2012; Sutherland et al., 2015), to semantic interpretations (Gelman et al., 2015; Hollander et al., 2002; Mannheim et al., 2011; Tardif, Gelman, Fu, & Zhu, 2012), to judgments about individuals (Leslie et al., 2011) to linguistic analyses (Leslie, 2008).

Although Gelman et al. (2015) also provided evidence that specific quantified statements function in the same way, the innovation in this article is two-fold: (a) children nonetheless default to generic even in the case of numbers, which is a substantially stronger test than any prior research (all of which included quantifiers, and thus could instead reflect something about the semantics of quantifiers per se); and (b) we have documented a direct link between children's number concepts and their defaulting to generics.

Given the implications of generic generalizations for reasoning about significant real-world categories (Gelman et al., 2010; Rhodes et al., 2012), we hope that future research will examine other mechanisms that modulate generic recall, throughout development.

Notes

If that were the case, then we would expect better performance on the Give-A-Number task for those in the Generic/Five condition than those in the Generic/Some condition. However, no such differences were obtained (including counting consecutively to at least 5 without the researcher's help, spontaneously mentioning 5, and highest consecutive number reached without the researcher's help), all $ps > .22$.

For one child, test trials were administered in the incorrect order; this child nonetheless was classified as a 1-knower, due to two successes on "1" and two failures on "2."

When reviewing the transcribed responses, we noticed that some children referred to the pictures during the recall phase (e.g., pointing to each animal in turn while saying, "This one and this one and this one and this one have long mouths"). Referring to the set of instances in the picture is plausibly an alternate way of expressing the 5-statements provided during the teaching phase. We therefore conducted a supplementary set of analyses in which the specific plural responses (also including, e.g., "These guys have strong tongues"; "The snakes swallow food in holes") were coded as Number responses, and thus correct when in response to 5-sentences, and opposite when in response to generic-sentences in the generic/*five* condition. We found that precisely the same patterns obtained, whether or not the specific-plural responses were included, and thus these responses are not considered further.

Acknowledgment

The research was supported by NICHD grant HD-36043 to S. Gelman. We are grateful to the children, parents, and teachers of Annie's Preschool and Gretchen's House for participating in the research. We also thank Natalie Davidson, Megan Martinez, Merranda McLaughlin, and members of the Conceptual Development Lab at the University of Michigan for their research assistance.

Disclosure

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development [HD-36043].

ORCID

Susan A. Gelman http://orcid.org/0000-0003-1005-2691

References

Barner, D., Chow, K., & Yang, S. J. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology*, *58*(2), 195–219. doi:10.1016/j.cogpsych.2008.07.001

Barner, D., Libenson, A., Cheung, P., & Takasaki, M. (2009). Cross-linguistic relations between quantifiers and numerals in language acquisition: Evidence from Japanese. *Journal of Experimental Child Psychology*, *103*(4), 421–440. doi:10.1016/j.jecp.2008.12.001

Barwise, J, & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, *4*(2), 159-219.

Bloom, P., & Wynn, K. (1997). Linguistic cues in the acquisition of number words. *Journal of Child Language*, *24*(3), 511–533.

Brandone, A. C., Cimpian, A., Leslie, S.-J., & Gelman, S. A. (2012). Do lions have manes? For children, generics are about kinds rather than quantities. *Child Development*, *83*(2), 423–433. doi:10.1111/j.1467-8624.2011.01708.x

Brandone, A. C., & Gelman, S. A. (2009). Differences in preschoolers' and adults' use of generics about novel animals and artifacts: A window onto a conceptual divide. *Cognition*, *110*(1), 1–22. doi:10.1016/j.cognition.2008.08.005

Brandone, A. C., & Gelman, S. A. (2013). Generic language use reveals domain differences in children's expectations about animal and artifact categories. *Cognitive Development*, *28*(1), 63–75. doi:10.1016/j.cogdev.2012.09.002

Brandone, A. C., Gelman, S. A., & Hedglen, J. (2015). Children's developing intuitions about the truth conditions and implications of novel generics versus quantified statements. *Cognitive Science*, *39*(4), 711–738. doi:10.1111/cogs.12176

Carlson, G. N., & Pelletier, F. J. (Eds.). (1995). *The generic book*. Chicago, IL: University of Chicago Press.

Cimpian, A., Brandone, A. C., & Gelman, S. A. (2010). Generic statements require little evidence for acceptance but have powerful implications. *Cognitive Science*, *34*(8), 1452–1482. doi:10.1111/j.1551-6709.2010.01126.x

Cimpian, A., Gelman, S. A., & Brandone, A. C. (2010). Theory-based considerations influence the interpretation of generic sentences. *Language and Cognitive Processes*, *25*(2), 261–276. doi:10.1080/01690960903025227

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*(1), 14–25. doi:10.1016/j.cognition.2009.07.004

Cimpian, A., & Markman, E. M. (2011). The generic/nongeneric distinction influences how children interpret new information about social others. *Child Development*, *82*(2), 471–492. doi:10.1111/j.1467-8624.2010.01525.x

Danan, J. A. J., & Gelman, R. (2018). The problem with percentages. *Phil. Trans. R. Soc. B*, *373*(1740), 20160519. doi:10.1098/rstb.2016.0519

Gallistel, C. R. (2018). Finding numbers in the brain. *Philosophical Transactions of the Royal Society B*, *373*(1740), 20170119. doi:10.1098/rstb.2017.0119

Gelman, R. (1972). Logical capacity of very young children: Number invariance rules. *Child Development*, 75–90. doi:10.2307/1127873

Gelman, R. (2006). Young natural-number arithmeticians. *Current Directions in Psychological Science*, *15*(4), 193–197. doi:10.1111/j.1467-8721.2006.00434.x

Gelman, R. (2015). Learning in core and non-core number domains. *Developmental Review*, *38*, 185–200. doi:10.1016/j.dr.2015.07.010

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gelman, S. A., & Bloom, P. (2007). Developmental changes in the understanding of generics. *Cognition*, *105*(1), 166–183. doi:10.1016/j.cognition.2006.09.009

Gelman, S. A., Goetz, P. J., Sarnecka, B. S., & Flukes, J. (2008). Generic language in parent-child conversations. *Language Learning and Development*, *4*(1), 1–31. doi:10.1080/15475440701542625

Gelman, S. A., Hollander, M., Star, J., & Heyman, G. D. (2000). The role of language in the construction of kinds. In D. Medin (Ed.), *Psychology of learning and motivation* (Vol. 39, pp. 201–263). New York, NY: Academic Press.

Gelman, S. A., Leslie, S. J., Was, A. M., & Koch, C. M. (2015). Children's interpretations of general quantifiers, specific quantifiers and generics. *Language, Cognition, and Neuroscience*, *30*(4), 448–461. doi:10.1080/23273798.2014.931591

Gelman, S. A., & Raman, L. (2003). Preschool children use linguistic form class and pragmatic cues to interpret generics. *Child Development*, *74*(1), 308–325. doi:10.1111/1467-8624.00537

Gelman, S. A., & Raman, L. (2007). This cat has nine lives? Children's memory for genericity in language. *Developmental Psychology*, *43*(5), 1256–1268. doi:10.1037/0012-1649.43.5.1256

Gelman, S. A., Tapia, I. S., & Leslie, S. J. (2016). Memory for generic and quantified sentences in Spanish-speaking children and adults. *Journal of Child Language*, *43*(6), 1231–1244. doi:10.1017/S0305000915000483

Gelman, S. A., Ware, E., & Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cognitive Psychology*, *61*(3), 273–301. doi:10.1016/j.cogpsych.2010.06.001

Graham, S. A., Gelman, S. A., & Clarke, J. (2016). Generics license 30-month-olds' inferences about the atypical properties of novel kinds. *Developmental Psychology*, *52*, 1353–1362. doi:10.1037/dev0000183

Graham, S. A., Nayer, S. L., & Gelman, S. A. (2011). Two-year-olds use the generic/non-generic distinction to guide their inferences about novel kinds. *Child Development*, *82*(2), 493–507. doi: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. doi:10.1016/j.cogdev.2014.05.002

Hollander, M. A., Gelman, S. A., & Star, J. (2002). Children's interpretation of generic noun phrases. *Developmental Psychology*, *38*(6), 883–894. doi:10.1037/0012-1649.38.6.883

Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development*, *2*(2), 77–96. doi:10.1207/s15473341lld0202_1

Leslie, A. M., Gelman, R., & Gallistel, C. R. (2008). The generative basis of natural number concepts. *Trends in Cognitive Sciences*, *12*, 213–218. doi:10.1016/j.tics.2008.03.004

Leslie, S., & Gelman, S. A. (2012). Quantified statements are recalled as generics: Evidence from preschool children and adults. *Cognitive Psychology*, *64*(3), 186–214. doi:10.1016/j.cogpsych.2011.12.001

Leslie, S. J. (2007). Generics and the structure of the mind. *Philosophical Perspectives*, *21*(1), 375–403. doi:10.1111/j.1520-8583.2007.00138.x

Leslie, S. J. (2008). Generics: Cognition and acquisition. *Philosophical Review*, *117*(1), 1–47. doi:10.1215/00318108-2007-023

Leslie, S. J, Khemlani, S, & Glucksberg, S. (2011). Do all ducks lay eggs? the generic over generalization effect. *Journal Of Memory and Language*, *65*(1), 15-31.

Mannheim, B., Gelman, S. A., Escalante, C., Huayhua, M., & Puma, R. (2011). A developmental analysis of generic nouns in Southern Peruvian Quechua. *Language Learning and Development*, *7*(1), 1–23. doi:10.1080/15475441003635620

Pappas, A., & Gelman, S. A. (1998). Generic noun phrases in mother-child conversations. *Journal of Child Language*, *25*(1), 19–33. doi:10.1017/S0305000997003292

Prasada, S. (2000). Acquiring generic knowledge. *Trends in Cognitive Sciences*, *4*(2), 66–72. doi:10.1016/S1364-6613(99)01429-1

Rhodes, M., Leslie, S. J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences*, *109*(34), 13526–13531. doi:10.1073/pnas.1208951109

Sarnecka, B. W. (2015). Learning to represent exact numbers. *Synthese*, 1–18. doi:10.1007/s11229-015-0854-6

Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, *108*(3), 662–674. doi:10.1016/j.cognition.2008.05.007
Sarnecka, B. W., & Gelman, S. A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition*, *92*(3), 329–352. doi:10.1016/j.cognition.2003.10.001
Starkey, P., & Gelman, R. (1982). The development of addition and subtraction abilities prior to formal schooling in arithmetic. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 99–116). Hillsdale, NJ: Erlbaum.
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. doi:10.1111/cogs.12189
Tardif, T., Gelman, S. A., Fu, X., & Zhu, L. (2012). Acquisition of generic noun phrases in Chinese: Learning about lions without an '-s'. *Journal of Child Language*, *30*, 1–32. doi:10.1017/S0305000910000735
Tasimi, A., Gelman, S. A., Cimpian, A., & Knobe, J. (2017). Differences in the evaluation of generic statements about human and non-human categories. *Cognitive Science*, *41*, 1934–1957. doi:10.1111/cogs.12440
Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*(2), 155–193. doi:10.1016/0010-0277(90)90003-3
Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220–251. doi:10.1016/0010-0285(92)90008-P
Zur, O., & Gelman, R. (2004). Young children can add and subtract by predicting and checking. *Early Childhood Research Quarterly*, *19*(1), 121–137. doi:10.1016/j.ecresq.2004.01.003