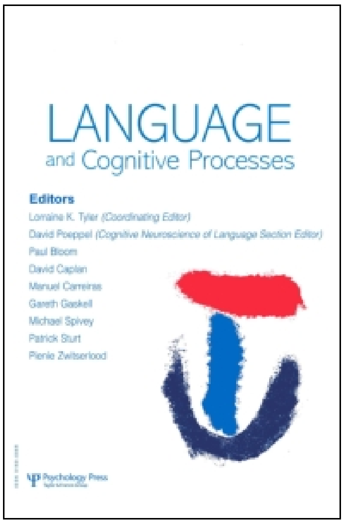



**Language and Cognitive Processes**
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683153

# Theory-based considerations influence the interpretation of generic sentences

Andrei Cimpian [a]; Susan A. Gelman [b]; Amanda C. Brandone [b]
[a] University of Illinois at Urbana-Champaign, Champaign, IL, USA [b] University of Michigan, Ann Arbor, MI, USA

First published on: 03 August 2009

**To cite this Article** Cimpian, Andrei, Gelman, Susan A. and Brandone, Amanda C.(2010) 'Theory-based considerations influence the interpretation of generic sentences', Language and Cognitive Processes, 25: 2, 261 — 276, First published on: 03 August 2009 (iFirst)

**To link to this Article: DOI:** 10.1080/01690960903025227

**URL:** http://dx.doi.org/10.1080/01690960903025227

LANGUAGE AND COGNITIVE PROCESSES
2010, 25 (2), 261–276

Psychology Press
Taylor & Francis Group

# Theory-based considerations influence the interpretation of generic sentences

Andrei Cimpian
*University of Illinois at Urbana-Champaign, Champaign, IL, USA*

Susan A. Gelman and Amanda C. Brandone
*University of Michigan, Ann Arbor, MI, USA*



Under what circumstances do people agree that a kind-referring generic sentence (e.g., 'Swans are beautiful') is true? We hypothesised that theory-based considerations are sufficient, independently of prevalence/frequency information, to lead to acceptance of a generic statement. To provide evidence for this general point, we focused on demonstrating the impact of a specific theory-based, essentialist expectation – that the physical features characteristic of a biological kind emerge as a natural product of development – on participants' reasoning about generics. Across three studies, adult participants ($N = 99$) confirmed our hypothesis, preferring to map generic sentences (e.g., 'Dontrets have long tails') onto novel categories for which the key feature (e.g., long tails) was absent in all the young but present in all the adults rather than onto novel categories for which the key feature was at least as prevalent but present in some of the young and in some of the adults. Control conditions using 'some'- and 'most'-quantified sentences demonstrated that this mapping is specific to generic meaning. These results suggest that generic meaning does not reduce to quantification and is sensitive to theory-based expectations.

***Keywords:*** Generic language; Language comprehension; Psychological essentialism; Quantifiers.

---

Correspondence should be addressed to Andrei Cimpian, Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel St., Champaign, IL 61820, USA. E-mail: acimpian@psych.illinois.edu

This research was supported in part by an Elizabeth Munsterberg Koppitz Child Psychology Graduate Fellowship to Cimpian, NICHD Grant HD-36043 to Gelman, and an NSF Graduate Research Fellowship to Brandone. We would like to thank Denise Larabell for her assistance in collecting the data. Joe Robinson,Vikram Jaswal, and two anonymous reviewers provided very useful comments on the manuscript.


http://www.psypress.com/lcp DOI: 10.1080/01690960903025227

Generic sentences (e.g., 'Swans are beautiful') express broad generalisations about categories in the world (Carlson, 1977; Carlson & Pelletier, 1995; Gelman, 2004; Leslie, 2008). Intuitively, such sentences may be taken to mean that most, if not all, members of a category share the property in question. Although generics imply that the properties they refer to are prevalent (Gelman, Star, & Flukes, 2002), we argue that they do not mark quantification per se (see also Carlson, 1977; Leslie, 2008; Prasada, 2000). Generics are a linguistic means of expressing knowledge about categories, and as such their interpretation is unlikely to be based solely on frequency information. There is considerable evidence suggesting that category representations consist of more than just feature tabulations – that they are embedded in 'naive theories' (Murphy & Medin, 1985) that incorporate a wealth of causal-explanatory knowledge about the concept's features (e.g., about their origins, functions, centrality) and the links between them (e.g., Ahn, Kim, Lassaline, & Dennis, 2000; Barrett, Abdi, Murphy, & Gallagher, 1993; Chapman & Chapman, 1969; Gelman, 2003; Keil, 1992; Rehder & Hastie, 2001). We hypothesise that this 'theory'-based knowledge also factors into how people reason about generic sentences – the linguistic structures that instantiate their conceptual representations. More precisely, we claim that theory-based considerations are *sufficient*, even in the absence of corroborating prevalence information, to generate agreement with a generic sentence. They may not be *necessary*, however: The existence of true 'statistical' generics such as 'Barns are red' or 'Taxis are yellow' (Prasada & Dillingham, 2006) suggests that feature prevalence by itself, in the absence of a theory-based connection to the category (since there is no deep reason why, e.g., barns are red), can lead to a generic's acceptance as well.

In this paper, we illustrate the general argument about the sufficiency of theoretical knowledge by showing how a specific theory-based expectation drives adults' interpretation of generic sentences. We use as a test case a causal assumption that is part of essentialist thinking (Gelman, 2003; Gelman & Wellman, 1991; Hickling & Gelman, 1995) – that the physical features characteristic of a biological kind often emerge as an *inherent outcome of development*, as a consequence of the gradual unfolding of the kind's essence over time. Consider a familiar example: As the children's story makes clear, swans are not born beautiful. Yet, perhaps because their beauty emerges without fail during the course of normal development, the generic 'Swans are beautiful' is uncontroversial. Now imagine a different species (say, 'snaws') in which some of the babies are born beautiful and remain so through adulthood, while the others are born less-than-beautiful and remain so through adulthood. Let's also assume that, when considering individuals of all ages, exactly the same proportion of swans and snaws are beautiful. Despite the equal prevalence of beauty in these two categories, people may be more willing to accept 'Swans are beautiful' than 'Snaws are beautiful' because swans' beauty is a normal function of growth (and is thus compatible with the assumption above), whereas for snaws beauty is more of an individual characteristic with no principled basis underlying its distribution in the population.

Gelman and Bloom (2007) recently suggested that adults are sensitive to property *origins* when judging the truth of a generic sentence. In their experiments, participants were shown a sample of four animals from a novel kind (e.g., 'dobles') and were told that these animals were either born with a feature or acquired it (e.g., born with claws vs. put on claws). When asked whether this key feature applies to the category as a whole (e.g., 'Do dobles have claws?'), adult participants who heard that dobles were born with claws were much more likely to say 'yes' than those who heard that dobles put on claws (86% vs. 0%). What's more, the magnitude of this difference held up even when the sample of dobles *lost* their claws. That is, even when none of the dobles they saw in front of them had claws, adults overwhelmingly agreed that 'dobles have claws' if they had been told that dobles are born with claws. Thus, adults find generics about *intrinsic* properties (i.e., properties that originate within the animals themselves) much more acceptable than generics about *extrinsic* properties, even when frequency is controlled for.

The distinction we investigate here is orthogonal to the intrinsic vs. extrinsic dimension, in that it concerns only intrinsic properties. Some intrinsic features are distributed randomly within a population, with some individuals possessing them and others lacking them from birth (e.g., beauty for snaws, stripes for cats). In contrast, other intrinsic properties are distributed along biologically meaningful lines and in accord with essentialist thinking about kinds. For example, some are absent in the young of a species and present in the adults because they emerge as a function of development (e.g., beauty for swans, sweetness for apples, legs for frogs). We predict that, when deciding on the truth of a generic sentence, people will be sensitive to the *distribution* of the feature with respect to age. Although in both cases only some of the category members possess the relevant feature, the acceptability of a generic should be higher when the distribution of the relevant feature suggests it emerges with development and lower when its distribution crosscuts developmental stages. Such a result would illustrate the powerful effect of theory-based expectations on the interpretation of generic sentences. Experiment 1 provided a first test of this idea.

Note that we will test our claim in the context of novel categories, which afford much more control over both the statistical information and the prior knowledge participants bring to bear on their judgements. Although there are certainly many familiar generics that seem to be in agreement with our argument (e.g., 'Swans are beautiful'), their familiarity makes it difficult to determine precisely the relative contribution of prevalence vs. theoretical knowledge to their acceptance. For example, it may be that people agree with 'Swans are beautiful' just because it is something they have been told – and not necessarily because the distribution of the relevant property matches their essentialist expectations.

# EXPERIMENT 1

## Method

*Participants.* Twenty-four undergraduate students (16 females, 8 males) from introductory psychology courses at the University of Michigan participated in this study for course credit.

*Materials, design, and procedure.* On each of the 8 trials in the session, the experimenter showed participants two samples taken from two different novel categories of animals 'that live on a planet that is far away' (see Figure 1). Exactly half of the animals in each sample displayed a key property, which was a distinctive part or marking (e.g., long tail, stripes). Similarly, half of the animals in each sample were small (described as 'the babies') and half were large (described as 'the grown-ups'). The crucial difference between the two samples on each trial was in the distribution of the key property relative to the babies and the grown-ups. In the *age-varying* sample, all the grown-ups possessed the distinctive property, but none of the babies did. In the *random-varying* sample, half of the babies and half of the grown-ups possessed the property. Which of the two novel categories on a trial was age-varying and which was random-varying was randomised across participants.

The participants were then told, for example, 'OK, now one kind is called dontrets. I don't know which ones are dontrets, but I have a clue here.' On half of the trials, participants were given a generic 'clue' (e.g., 'Dontrets have long tails'). On the other half (the control trials), the clue was a 'some'-quantified sentence (e.g., 'Some dontrets have long tails'). Each item set (e.g., dontrets) appeared in either the generic or the 'some' wording for a given participant, but not both. Next, participants were asked, 'Can you find the dontrets?' The generic and 'some' trials were blocked, and the order of the blocks was counterbalanced across participants. The order of the item sets was randomised for each participant.

## Results and discussion

We predicted that participants should be more likely to choose the age-varying groups when they hear the generic sentences, despite the fact that the frequency of the key property is equated across the two samples. Moreover, if participants' preference for the age-varying sets on the generic trials is driven by their assumptions about generic meaning, they should not choose these sets as often on the 'some' trials. In fact, given that the meaning of 'some' seems

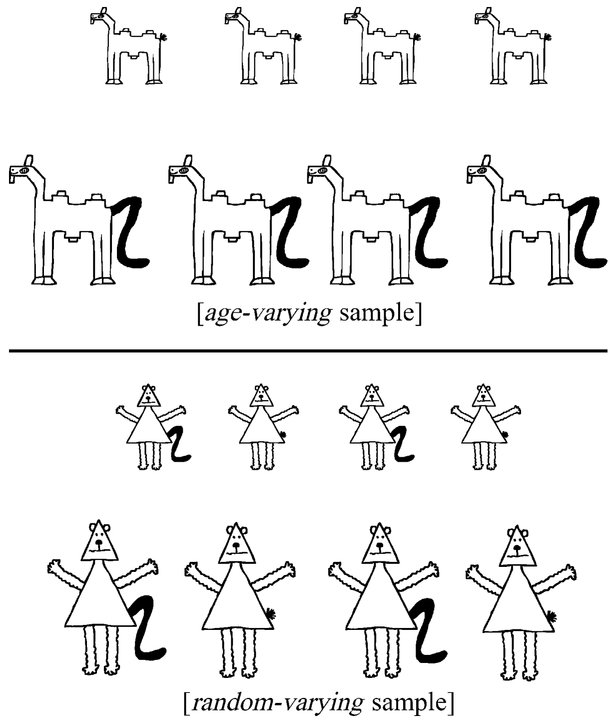



**Figure 1.** Sample stimuli for a trial in Experiment 1. On this trial, the subjects were told that 'Dontrets have long tails' (generic condition) or 'Some dontrets have long tails' ('some' condition) and asked to 'find the dontrets'.

quite compatible with the structure of the random-varying sets (where some of the individuals seem to possess the key feature across developmental stages), it may even be that the random-varying sets will be preferred.

On each trial, participants' responses were scored as a 1 if they chose the age-varying sample and as a 0 if they chose the random-varying sample. Scores within the generic and 'some' blocks were summed and submitted to a 2 (wording: generic vs. 'some'; within subject) $\times$ 2 (block order: generic-first vs. 'some'-first; between subjects) analysis of variance (ANOVA). The ANOVA uncovered only a main effect of wording: As predicted, participants were more likely to choose the age-varying samples in the generic condition ($M = 74\%$) than in the 'some' condition ($M = 14\%$), $F(1, 22) = 33.44$, $p < .001$ (see Figure 2). Participants chose the age-varying groups more often than expected by chance on the generic trials (74% vs. 50%), $t(23) = 3.29$, $p = .003$. On the 'some' trials, though, the age-varying groups were chosen at below-chance levels (14% vs. 50%), $t(23) = 6.26$, $p < .001$ – that is, participants had a reliable preference for the random-varying samples.

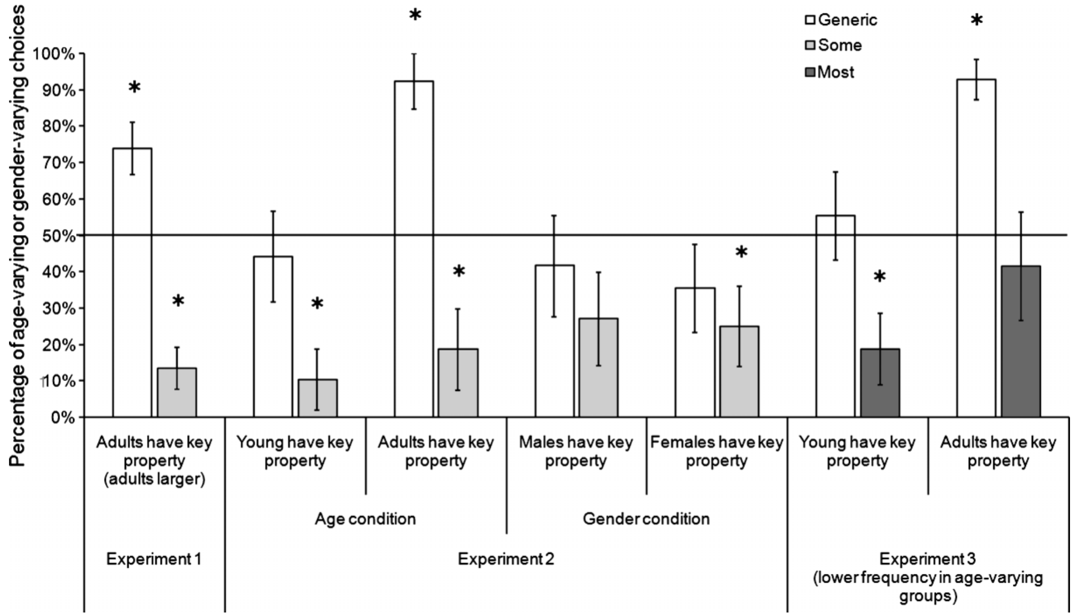



**Figure 2.** The percentage of age-varying (Experiments 1–3) and gender-varying (Experiment 2) choices. The error bars represent the standard error of the mean. * $p < .05$ different from chance (50%).

The results of this first experiment suggest that people prefer to map generic meaning onto inherent properties that emerge with development rather than onto properties that, although equally prevalent, occur randomly at both developmental stages. The 'some' control trials confirmed that this preference is not an artifact of our stimulus design: When told that, for example, '*some* dontrets have long tails', participants chose the random-varying samples instead.

Experiment 2 was designed to address three additional questions. First, would participants' preference for the age-varying sets on generic trials hold if we removed the correlation between age and size by making all animals on a page the same size? Although in reality age and size are often confounded, it could be that participants selected the age-varying samples because the animals displaying the key property were larger and thus provided more evidence of the property (e.g., more 'long-tailedness'). Second, would participants' preference for the age-varying sets on generic trials hold if the *babies* displayed the property instead of the grown-ups? If people assume that the properties that are characteristic of kinds – and are thus likely candidates for generic meaning – are the ones that *develop* rather than *disappear* with age, we predict they would restrict their preference to just the age-varying groups in which adults display the key feature. Third, are there other meaningful, theory-laden biological distinctions (besides growth/age) that people take into account when reasoning about generic meaning? For instance, would people be sensitive to the distribution of a property with respect to *gender*? That is, would they prefer gender-varying samples over random-varying samples when asked to find the referents of a generic sentence?

## EXPERIMENT 2

### Method

*Participants.* Forty-nine undergraduate students (25 females, 24 males) from introductory psychology courses at Stanford University participated in this study for course credit.

*Materials, design, and procedure.* The materials were identical to Experiment 1, with the following exceptions: First, all 8 animals in a picture were the same size. To prevent confusion, we changed how we referred to the younger group from 'babies' (who are always smaller than adults) to 'young ones.' Second, instead of size, the top and bottom four animals in each picture were differentiated by a few other features (e.g., hair shape) in order to make it more plausible that they differ in age or gender (see Figures 3 and 4). The

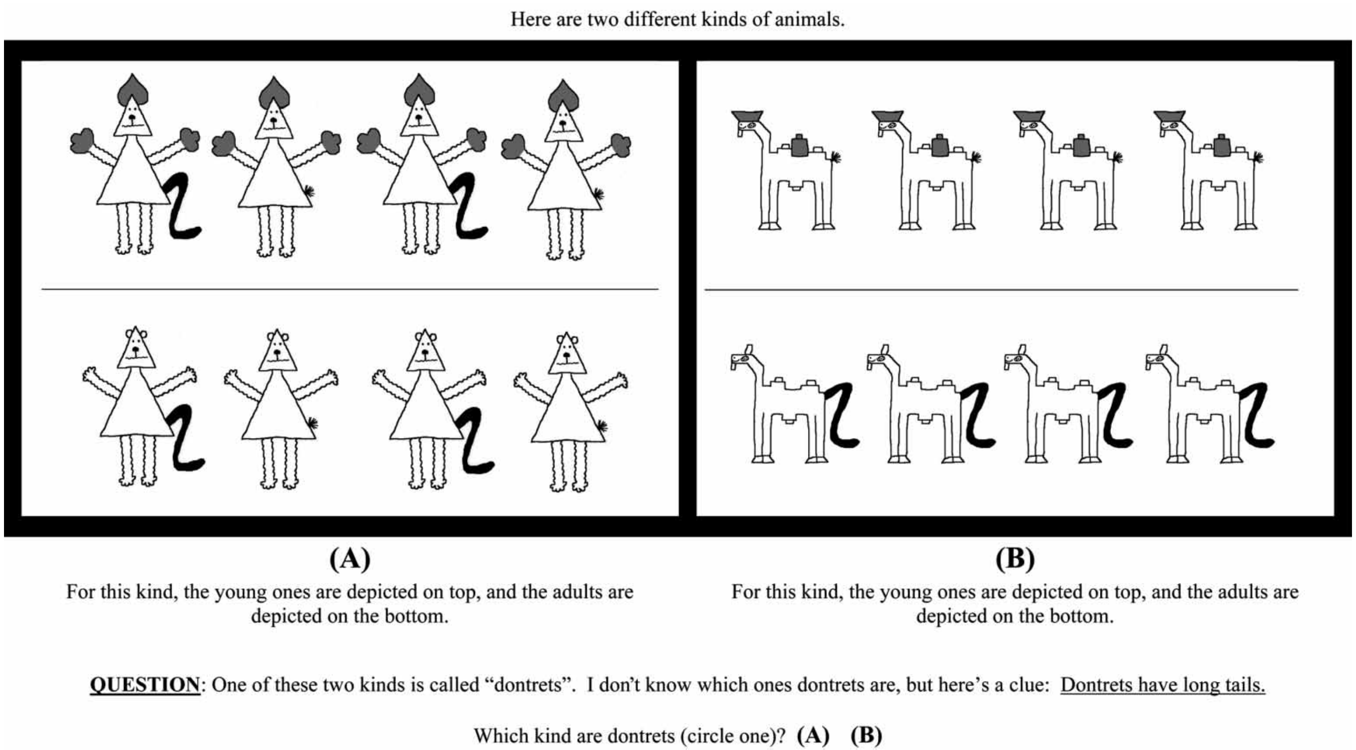
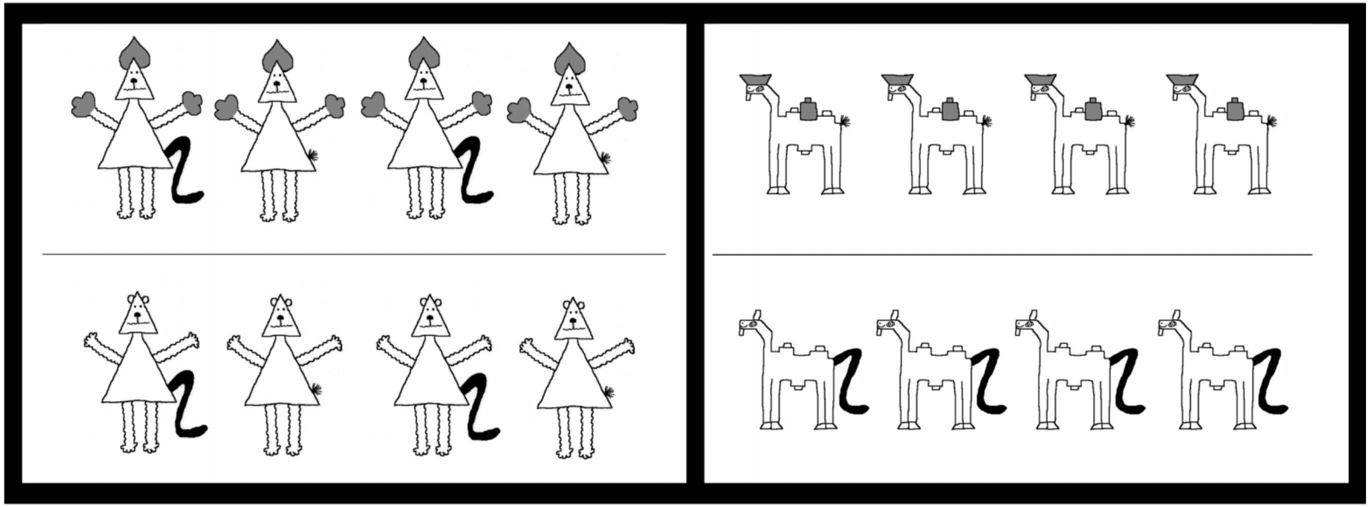



**Figure 3.** Sample booklet pages from the age condition in Experiment 2.

Here are two different kinds of animals.



**(A)**

For this kind, the females are depicted on top, and the males are depicted on the bottom.

**(B)**

For this kind, the females are depicted on top, and the males are depicted on the bottom.

**QUESTION**: One of these two kinds is called "dontrets". I don't know which ones dontrets are, but here's a clue: Dontrets have long tails.

Which kind are dontrets (circle one)? **(A)** **(B)**

**Figure 4.** Sample booklet pages from the gender condition in Experiment 2.

assignment of the top and bottom animals in each picture to the young/adult or male/female categories was counterbalanced across participants.

Participants were randomly assigned to one of four conditions determined by the wording of the 'clues' (generic vs. 'some') and the biological distinction tested (age vs. gender). In the age condition, half of the trials had age-varying samples with the key feature present in the young subgroup, and the other half had age-varying samples with the key feature present in the adults. Similarly, in the gender condition, the key property was present in the males of the gender-varying sample on four trials and in the females on the other four. The stimuli (pictures, clues, etc.) were printed in colour and presented to the subjects in booklet form (see Figures 3 and 4 for two sample pages).

## Results and discussion

On each trial, participants' responses were scored as a 1 if they chose the age- or the gender-varying sample and as a 0 if they chose the random-varying sample. Recall that we manipulated whether it is the young or the adults (age condition) or the males or the females (gender condition) that display the key feature on each trial. Given that the two levels of this manipulation are different for the age and gender conditions (i.e., young/adult vs. male/female), we analysed these conditions separately.

A 2 (wording: generic vs. 'some'; between subjects) $\times$ 2 (age group displaying the key property: young vs. adult; within subject) ANOVA on the number of age-varying choices in the *age* condition revealed significant main effects of wording ($M_{generic} = 69\%$ vs. $M_{\text{'some'}} = 15\%$), $F(1, 23) = 18.79$, $p < .001$, age group displaying the key property ($M_{adult} = 57\%$ vs. $M_{young} = 28\%$), $F(1, 23) = 15.35$, $p = .001$, and an interaction between the two, $F(1, 23) = 7.62$, $p = .011$. As seen in Figure 2, participants who heard clues containing 'some' chose the age-varying groups equally infrequently regardless of which age subgroup displayed the key property ($M_{adult} = 19\%$ vs. $M_{young} = 10\%$), paired-$t(11) = 1.30$, $p = .220$. The frequency of age-varying responses on both of these types of trials was below chance, $t$s$(11) > 2.80$, $p$s $< .017$, indicating a preference for the random-varying samples. As expected, however, participants in the generic condition clearly differentiated between the trials in which the young group vs. the adults displayed the key property ($M_{adult} = 92\%$ vs. $M_{young} = 44\%$), paired-$t(12) = 3.85$, $p = .002$. As in Experiment 1, the age-varying groups were preferred at above-chance levels when *adults* had the key property (92% vs. 50%), $t(12) = 5.50$, $p < .001$; in contrast, when the young displayed the key property, selection of the age-varying samples was at chance (44% vs. 50%), $t(12) = 0.47$, $p = .650$.

A similarly structured ANOVA performed on the number of gender-varying responses in the *gender* condition revealed no significant main effects or interactions. In the generic condition, participants selected the gender-varying samples at chance levels both when males displayed the key property ($M=$ 42%), $t(11)=0.60$, $p=.560$, and when females did ($M=35\%$), $t(11)=1.21$, $p=.253$ (see Figure 2). In the 'some' condition, participants' gender-varying choices were below chance when females displayed the key property ($M=$ 25%), $t(11)=2.25$, $p=.046$, and showed a non-significant trend to occur below chance when males displayed it ($M=27\%$), $t(11)=1.78$, $p=.102$.

The three questions that motivated this study found clear answers in our results. First, the mapping of generic meaning onto inherent properties that emerge with development was unaffected by removing the correlated size information: Participants chose the age-varying sets in which adults had the key property on 92% of trials. Second, people's preference for the age-varying sets is in fact limited to the sets in which *adults* display the property: On trials where only the young subgroup of the age-varying sample displayed the property, participants' age-varying choices were at chance (44%). Third, it does seem that the age/growth dimension has a special status in people's reasoning about generics: When presented with a generic clue, participants did not prefer the *gender*-varying samples over the random-varying samples. This is an intriguing finding, since there are many generics about familiar categories in which the relevant property is true of only one gender (e.g., 'Birds lay eggs'). To speculate, it may be that participants were unable to connect the gender dimorphisms in these unfamiliar kinds to the theory-based knowledge that might make familiar gendered generics acceptable. For example, people know that laying eggs is (a) an important characteristic feature of birds (only certain kinds of animals lay eggs) and (b) a feature that males cannot display because they are not equipped with the right biological parts (Leslie, 2008). It may only be in light of this knowledge that people agree with generics such as 'Birds lay eggs.' Our novel stimuli allowed no inferences about, for instance, why only one gender possessed the key feature (e.g., why only male dontrets had long tails) or how common the feature was on this 'planet that is far away.' In contrast, it seems that the age dimorphisms in our stimuli were in fact sufficient to allow essentialist inferences about the growth-related emergence of category-typical features.

Experiment 3 was designed to provide an even stronger test of our hypothesis. Two crucial elements distinguish it from the first two studies. First, the key property was *less frequent* in the age-varying samples than in the random-varying samples (see Figure 5). Would people actually override the frequency information and maintain a preference for the samples where the distribution of the key property is aligned with age/growth? Second, we used sentences quantified with 'most' (e.g., 'Most dontrets have long tails') as our controls. 'Most'-quantified sentences are arguably closer in meaning to generics and thus provide a stronger control.

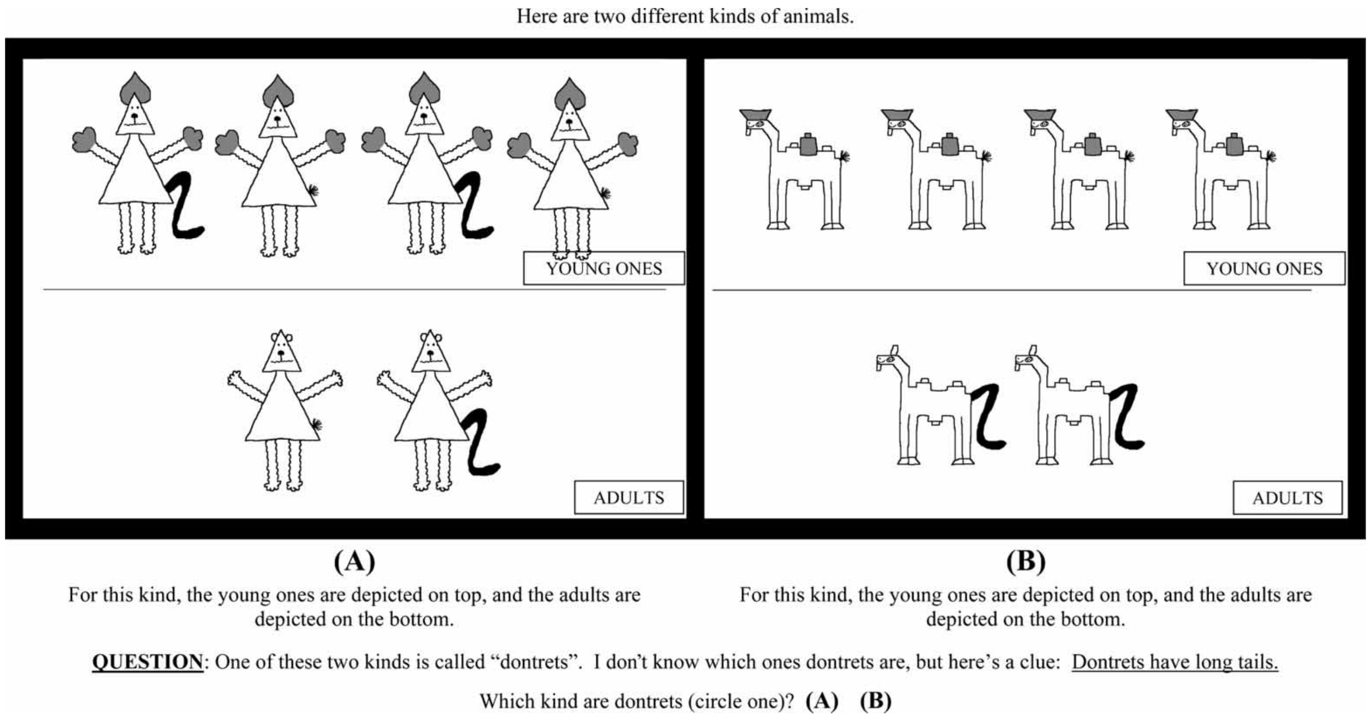



**Figure 5.** Sample booklet page from Experiment 3.

# EXPERIMENT 3

## Method

*Participants.* Twenty-six undergraduate students (20 females, 6 males) from the University of Michigan and Stanford University participated in this study for course credit or a \$5 gift certificate.

*Materials, design, and procedure.* The materials, design, and procedure were identical to the age condition in Experiment 2, with the following exceptions: Each picture contained only 6 animals, 4 in one row and 2 in the other. The top/bottom position of these rows was counterbalanced across participants. Three of the animals in the random-varying samples (50%) and only 2 of the animals in the age-varying samples (33%) displayed the key property. To ease the processing load, we added a label ('adults' or 'young ones') next to each of the age groups (see Figure 5). Participants were randomly assigned to the generic condition ($n = 14$) or the 'most' condition ($n = 12$).



## Results and discussion

A 2 (wording: generic vs. 'most'; between subjects) × 2 (age group displaying the key property: young vs. adult; within subject) ANOVA on the number of age-varying choices revealed significant main effects of wording ($M_{generic} =$ 74% vs. $M_{\text{'}most\text{'}} = 30\%$), $F(1, 24) = 14.08$, $p = .001$, and age group displaying the key property ($M_{adult} = 69\%$ vs. $M_{young} = 38\%$), $F(1, 24) = 9.02$, $p = .006$, but no significant interaction between the two, $F(1, 24) = 0.53$, $p = .476$.

As in the previous experiment, participants in the generic condition were more likely to choose the age-varying samples when the adults displayed the key property ($M = 93\%$) than when the young group did ($M = 55\%$), paired-$t(13) = 3.07$, $p = .009$ (see Figure 2). The preference for age-varying groups was above chance in the former case (93% vs. 50%), $t(13) = 7.77$, $p < .001$, but not in the latter (55% vs. 50%), $t(13) = 0.44$, $p = .664$.

In the 'most' condition, the frequency of age-varying choices did not differ between the trials where the adult group had the property ($M = 42\%$) and the trials where the young group had it ($M = 19\%$), paired-$t(11) = 1.40$, $p = .190$. In contrast to the generic condition, participants who heard that, e.g., '*most* dontrets have long tails' were at chance in choosing the age-varying groups in which the adults displayed the key property (42% vs. 50%), $t(11) = 0.56$, $p =$ .586. The frequency of age-varying choices dropped below chance for the trials where the young subgroup had the key property (19% vs. 50%), $t(11) = 3.19$, $p = .009$.

The results of Experiment 3 are clear. First, people prefer to map generic meaning onto samples where the key property emerges as a function of growth, *even when the key property is actually more prevalent in the other,* *random-varying samples*. Second, this sensitivity to the distribution of a feature within a category seems *specific to generic sentences*. Participants who heard 'most'-quantified sentences did not map them onto the age-varying choices, even though – at least in terms of implied property frequency – these sentences are rather similar to generics.

## GENERAL DISCUSSION

The three studies reported support the argument that theory-based knowledge and expectations play an important role in adults' interpretation of generic sentences. Specifically, we showed that participants' choice of referent for a generic sentence was influenced by the theory-based, essentialist assumption that the physical properties of a biological kind emerge as a natural outcome of development. Participants consistently preferred to map generic meaning onto the samples that conformed to this expectation (i.e., samples in which the adult animals displayed the property but the young did not), despite the fact that the prevalence of the key property in these samples was either the *same* as (Experiments 1 and 2), or *lower* than (Experiment 3) in the random-varying samples. These results suggest that theory-based knowledge is *sufficient*, independently of prevalence, to lead to acceptance of a generic statement.

Relatedly, our studies lend further support to the claim that generic meaning cannot be reduced to probabilities or quantification (see also Carlson, 1977; Carlson & Pelletier, 1995; Gelman & Bloom, 2007; Leslie, 2008; Prasada, 2000): Not only did our participants have a consistent theory-based preference for the age-varying sets regardless of prevalence, but they also clearly distinguished between generics and 'most'-quantified sentences in their responses (Experiment 3). Even though a 'most'-quantified sentence (e.g., 'Most swans are beautiful') seems like an intuitive, reasonable quantificational translation of a generic sentence (e.g., 'Swans are beautiful'), our participants' interpretation of these two sentence types was quite divergent. For generics, their referent choices were driven by the compatibility of the samples with their essentialist expectations; for 'most'-quantified sentences, on the other hand, theory-based considerations appeared to be much less influential, as participants no longer preferred the age-varying sets where the adults displayed the property over the random-varying sets.

Although these studies focused on a particular theory-based expectation relevant to a particular domain (biological natural kinds), our conclusions are likely to generalise more broadly. First, we expect other types of theoretical knowledge to contribute to the acceptance of generic statements about biological natural kinds. Gendered generics such as 'Birds lay eggs' may be evidence for this point, since they too are true despite a lack of converging prevalence information (only healthy adult female birds lay eggs). Second, we expect theoretical information to influence the interpretation of generic sentences in other ontological domains (e.g., inanimate natural kinds, social kinds, artifacts). For example, people's causal knowledge about artifacts most likely includes an expectation of breakdown or degradation after sustained use. Thus, if asked to map a generic such as 'Blickets are bright red' either onto (a) a 'usage-varying' sample in which the *new* objects were bright red and the *used* ones were a faded pink or onto (b) a random-varying sample with red and pink objects present in both the new and the used subsets, adults might well prefer the usage-varying sample.[1] However, in light of arguments that our causal-explanatory theories of artifacts are rather shallow relative to our theories of natural and social kinds (e.g., Brandone & Gelman, 2009; Gelman, 2003), it is also possible that the influence of theoretical knowledge on artifact-related generics is weaker. Exploring these alternatives may provide fertile ground for future research.

To return to the point we opened with, the fact that there are so many types of information that are relevant to the interpretation of a generic sentence (e.g., property distribution, property origins, frequency) may reflect the complexity of the category knowledge expressed by these utterances. We know many types of facts about the categories that make up our world. Some of these facts may just be statistical summaries of the experienced exemplars (e.g., barns tend to be red). Others may go 'deeper' and connect with our lay theories about why the world is structured as it is (e.g., swans grow to be beautiful; see Gelman, 2003; Gelman & Bloom, 2007). It stands to reason that our ability to interpret *sentences* about categories would draw on the full spectrum of information (statistical, theory-laden, etc.) that comprises our *knowledge* about categories.

Manuscript received November 2008
Revised manuscript received May 2009
First published online August 2009



## REFERENCES

Ahn, W., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology*, *41*, 361–416.

Barrett, S. E., Abdi, H., Murphy, G. L., & Gallagher, J. M. (1993). Theory-based correlations and their role in children's concepts. *Child Development*, *64*(6), 1595–1616.

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*, 1–22.

Carlson, G. N. (1977). *Reference to kinds in English*. Doctoral dissertation, University of Massachussetts, Amherst.

[1] We are grateful to an anonymous reviewer for this example.

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

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, *74*(3), 271–280.

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

Gelman, S. A. (2004). Learning words for kinds: Generic noun phrases in acquisition. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 445–484). Cambridge, MA: MIT Press.

Gelman, S. A., & Bloom, P. (2007). Developmental changes in the understanding of generics. *Cognition*, *105*(1), 166–183.

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

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, *38*, 213–244.

Hickling, A. K., & Gelman, S. A. (1995). How does your garden grow? Early conceptualization of seeds and their place in the plant growth cycle. *Child Development*, *66*, 856–876.

Keil, F. C. (1992). The origins of an autonomous biology. In M. R. Gunnar & M. Maratsos (Eds.), *Minnesota symposia on child psychology: Modularity and constraints in language and cognition* (pp. 103–137). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Leslie, S. J. (2008). Generics: Cognition and acquisition. *Philosophical Review*, *117*, 1–47.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.

Prasada, S. (2000). Acquiring generic knowledge. *Trends in Cognitive Sciences*, *4*, 66–72.

Prasada, S., & Dillingham, E. M. (2006). Principled and statistical connections in common sense conception. *Cognition*, *99*, 73–112.

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General*, *130*, 323–360.