Syllogistic Reasoning with Generic Premises:
The Generic Overgeneralization Effect

Sangeet Khemlani[1], Sarah-Jane Leslie[2], Sam Glucksberg[1]
[1]Department of Psychology
[2]Department of Philosophy
Princeton University
Princeton, NJ 08540 USA

Abstract

Generics are statements that are not explicitly quantified and that express generalizations, such as 'ducks lay eggs'. Intuitively, the generic (non-quantified) form of such statements seems to be true. Furthermore, people seem to be prone to an interesting error: treating the universal form of characteristic generic assertions (e.g., 'all ducks lay eggs') as true, even though they are, upon a moment's reflection, patently false. How would people interpret generic assertions when they are used as premises in a syllogistic reasoning task? Although the normative strategy to optimize production of valid conclusions would be to treat generics as existential assertions, people displayed a strong tendency to assimilate generics to universal assertions. As in prior comprehension studies, generics behave as the prototypical default form for expressing generalizations about the world.

Keywords: generics; syllogisms; reasoning; deduction

Introduction

Suppose you are given the following syllogism:

All A are B
All B are C

What follows? There are at least two valid conclusions: 'all A are C' and 'some A are C'. This holds irrespective of content. Any deduction of this form (two universally quantified assertions as above) would be valid, even if not true in the real world. For instance,

All lions have manes
All maned animals are male
Therefore all lions are male

Given this form and content, people can readily see that the syllogism is valid, and can also see that it is empirically false, in this case because the first premise is false.

Consider, now, a case where people might not recognize that a premise is false, as in the assertion 'all ducks lay eggs'. Upon reflection, the assertion is clearly false: the majority of ducks – males, juveniles, and unfertilized females – are among the eggless. The illusion that such statements are true in universal form arises from our mode of understanding an important kind of non-quantified assertions, generics (Leslie, 2008; Prasada & Dillingham, 2006; Khemlani, Leslie, Glucksberg, & Fernandez, 2007).

Generics are generalizations that lack explicit quantifiers such as 'some' or 'all'. Leslie (2007a,b) argues that there are three kinds of generics:

i. Those that express a *characteristic* of a kind, e.g. 'ducks lay eggs'.

ii. Those that refer to a property that the *majority* of the members of a kind possess, e.g., 'cars have radios'.

iii. Those that refer to a *striking*, often dire property, e.g., 'mosquitoes carry malaria'.

As these examples illustrate, generics are claims about kinds rather than about individuals. They occur frequently in everyday speech, even in parents' speech to two-year old children who have yet to master quantifiers such as 'some' or 'all' (Hollander, Gelman, & Star, 2002; Gelman, Goetz, Sarnecka & Flukes, in press). Furthermore, they do not express universal generalizations. While it is false that all tigers are striped, the generic claim 'tigers are striped' remains true in the face of exceptions (e.g., albinos).

One might suppose that generics are interpreted as existential statements, such that the generic 'Ks are Fs' is said to be true whenever some Ks are Fs, as in 'some mosquitoes carry malaria'. However, this interpretation fails to account for statements such as 'birds are female' or 'dogs are brown', which are intuitively false even though, of course, some birds are female and many dogs are brown. As these examples suggest, generics cannot be accounted for in terms of quantifiers such as 'all' or 'some'.

These observations suggest that generics are not quantificational (see Pappas & Gelman, 1998). Quantificational statements are about how much or how many in a way that generics are not. Notice that, upon being asked 'how many tigers are striped', one might reply 'most tigers are striped', or 'some tigers are striped', but one cannot reply 'tigers are striped'. The generic is not an appropriate answer to this question (Carlson, 1977). Leslie (2007a,b) argues that the truth and falsity of generics does not depend on how many of the relevant individuals possess the predicated property. There is no sense, she claims, in which generics are dependent on such quantitative considerations.

People do, however, exhibit a bias in interpreting characteristic generic assertions such as 'ducks lay eggs'. Such assertions are patently true in existential form, but false in universal form. Therefore the assertion 'all ducks lay eggs' is false, and should be judged as such whenever any counterexamples come to mind, e.g., when it is recognized that male ducks and immature female ducks do not lay eggs. Apparently, although such counterexamples are available to people, they do not seem to be spontaneously accessible. If they were spontaneously accessible, people should not agree to universally quantified

characteristic statements such as 'all ducks lay eggs'. Yet casual observation suggests that people do agree with such statements.

To see if people do in fact judge universally quantified generic assertions to be true, Khemlani et al. (2007) asked people to judge the truth value of three types of generic assertions (characteristic, majority and striking), each in one of three forms: existential (e.g.,' some ducks lay eggs"), generic ('ducks lay eggs') and universal ('all ducks lay eggs'). There was virtually unanimous agreement that existentially quantified generic assertions were true for all three predicate types. For the generic assertions, all three types were generally viewed as true, with characteristic generics being judged true more frequently than majority or striking ones. Most interestingly, universally quantified (i.e., 'all') characteristic generics were judged to be true almost half the time (46% agreement rate), while majority and striking generics were rejected most of the time (only 7% agreement rate for either type).

People seem to view characteristic generic assertions as universal ones, not only judging universally quantified generics as true, but also paraphrasing universally quantified assertions as generics, e.g., paraphrasing 'all ducks lay eggs' as 'ducks lay eggs' (Khemlani et al., forthcoming). People thus seem prone to committing a *generic over-generalization error*: they conflate universally quantified characteristic assertions as generic ones.

Generics and Syllogistic Reasoning

How would people interpret generics in the context of syllogisms? Because generics are non-quantificational, there is no obvious and unambiguous way to interpret generic assertions when they appear as premises in a syllogism. Consider:

> *Some Xs are Ys*
> *Some Ys are Zs*
> *What follows?*

According to standard analyses of syllogistic reasoning, nothing follows. If, on the other hand, the quantifier 'some' were to be replaced by the quantifier 'all', then at least two valid conclusions follow: 'all Xs are Zs' and 'some Xs are Zs'.

How would syllogisms with generics as premises be interpreted, e.g.,

> *Xs are Ys*
> *Ys are Zs*
> *What follows?*

Because generics are non-quantificational, there is no unambiguous answer. Normatively, true generics are always true in existential form as well, and so should be viewed as existentially quantified in the context of syllogistic reasoning. Hence, in this example, nothing follows. However, if people commit an analogy of the overgeneralization effect – view the generic premises as if they were universally quantified – then they should erroneously conclude that Xs are Zs. This conclusion would not be erroneous if one could unambiguously render the

conclusion 'Xs are Zs' as 'All Xs are Zs', but this would not work. Consider the generic premises

> *Lions have manes*
> *Maned animals are male*
> *Therefore lions are male*

In this case, people would agree not only with the first premise, but also with the first premise in universal form about 50% of the time. People would agree with the second premise in both generic and universal form as well. But, despite having two premises that people judge as true, the conclusion should be perceived as false and thus the deduction should be judged invalid.

How do people treat generics as premises in the context of syllogistic reasoning problems? If the overgeneralization effect operates in such contexts, then by analogy, people should tend to view generics as universal assertions, and perhaps vice versa. In other words, conclusions drawn for generic premises should resemble conclusions drawn for universal premises. An alternative hypothesis is that people recognize that generic assertions are always true in existential form, and that by acting upon this property of generics they avoid drawing erroneous conclusions. In the context of syllogistic reasoning, unlike the context of ordinary discourse, treating a generic as an existential assertion would be normative. To determine how people use generics in the context of syllogistic reasoning, we asked people to solve syllogism problems using first and second premises in existential, universal, and generic forms. On the basis of prior work on how people comprehend generics, we expected to find generic syllogistic reasoning to resemble reasoning with universal premises rather than reasoning with existential premises, i.e., non-normatively because of the bias to commit the generic overgeneralization error.

Method

Participants. Nineteen undergraduate students at Princeton University served as participants, and none had any background in logic or computer science. They completed the experiment over the Internet online using an interface written in Ajax.

Materials. Traditionally syllogistic premises and conclusions come in four different Aristotelian "moods":

> *All A are B*   *affirmative universal*
> *Some A are B*   *affirmative existential*
> No A are B   negative universal
> Some A are not B   negative existential

The generic form permits two additional moods:

> *A are B*   *affirmative generic*
> A are not B   negative generic

The present study manipulated only affirmative moods, which are italicized above. Moreover, syllogisms must be cast in a particular "figure":

> A-B   lions – manes
> B-C   maned animals – males

There are four possible figures in which a syllogism can be cast, but the study considered only the first:

| *A-B* | B-A | A-B | B-A |
| *B-C* | C-B | C-B | B-C |

Future studies on generic syllogisms should consider different figures because of well-known "figural effects" (Johnson-Laird & Bara, 1984).

The content of a particular argument can also influence the participants' response (Oakhill & Johnson-Laird, 1985; Evans, Barston, & Pollard, 1983), and so to ensure that participants did not answer based on their prior knowledge and beliefs, the premises used were semantically empty, e.g.,

> *All comets are orthovolatile.*
> *Some orthovolatile materials are pollutants.*

The term "orthovolatile" is nonsensical and the association between a comet and a pollutant is weak, so a participant who uses these premises to make an inference cannot do so with respect to background knowledge, beliefs, or internal associations between the terms used. The stimuli are provided in Appendix A.

Design and Procedure. Participants were presented with 18 syllogisms, one at a time, and were asked to provide a conclusion for each by entering their answers on their computer keyboard. Each syllogism problem consisted of two premises, followed by the prompt, *what follows?*

Each of the two syllogistic premises appeared in universal, existential, or generic form. This generated a 3 (first premise) × 3 (second premise) repeated-measures design. Participants received each argument form twice, each time with different content. Stimuli were counterbalanced using Latin squares.

Coding. Participants' responses fell into five categories; most of the time participants produced either a universal, existential, or generic response. They would also at times provide a 'null response', e.g., "Nothing necessarily follows."

A small percentage of the responses were complex interpretations of the premises. For instance, some responses made use of modal operators, e.g., "It is *possible* that pyramidal cells are compact, but certainly not guarenteed (sic)". Other responses were statements about capability, e.g., "Septapods *might* live in warm climates," and yet others used probabilistic reasoning, e.g., "Kangaroos *probably* have gene Gamma-64." Some participants used conditionals, e.g., "*If* elephants are semiparametric, they heal quickly", and others used domain restriction, e.g., "*Septapods with ovipositors* live in warm climates". Such complex formulations comprised only 9% of the responses, and were omitted from the present analyses.

Results and Discussion

The data most relevant to the hypothesis – that people would commit an overgeneralization error – are the similarities between syllogisms with universal and generic premises. The greater the tendency to commit this error, the more similar the pattern of conclusions for generic and universal premises should be. The proportions of each type of conclusion – existential, universal, generic and null (no conclusions drawn) – are presented in Tables 1–3. There were relatively few null conclusions, almost exclusively for syllogisms in which the second premise was in existential form. Participants seemed reluctant to explicitly draw no conclusions, even when both premises were existential and a null conclusion was appropriate (see Table 1).

Table 1. Proportion of responses to an existential first premise

| 2nd premise | Existential 1st premise *(Some A-B)* | | | |
	Some A-C	*All A-C*	*A-C*	*Null*
Some B-C	76	0	6	18
All B-C	94	3	3	0
B-C	94	0	6	0

When the first premise was existential (Table 1), participants performed quite poorly, drawing incorrect existential conclusions 76% of the time when the second premise was existential; they answered correctly 94% of the time when the second premise was universal (Mann-Whitney, z = 1.99, p < .05). These are common patterns as reported in the reasoning literature (Chapman & Chapman, 1959; Johnson-Laird & Bara, 1984) and suggest that the participants were performing the reasoning task as they normally would despite the presence of generic premises.

Performance was equivalent – 94% existential conclusions – when the second premise was a generic. In this case, it is ambiguous how participants interpreted generic premises because the error would be made whether the generic was interpreted as an existential or as a universal assertion. However, because the data for universal and generic second premises are virtually identical, it is likely that generics were interpreted as if they were universals. Additional evidence for this hypothesis is provided by the data for the other two conditions, where the first premise was either universal or generic.

When the first premise was universal (see Table 2) and the second premise was existential, participants incorrectly drew existential conclusions 59% of the time, again replicating previous results (Johnson-Laird & Bara, 1984). Appropriately, 71% of conclusions were universal when the second premise was universal, with generics at 29% (Mann-Whitney, z = 2.19, p < .05). When generics appeared as the second premise, the difference between universal and generic conclusions disappeared, 46% and 41% respectively (Mann-Whitney, z = 1.05, p = .29). The pattern again suggests that generics are assimilated to universals rather than to existentials.

Table 2. Proportion of responses to a universal first premise

| 2nd premise | Universal 1st premise *(All A-B)* | | | |
	Some A-C	*All A-C*	*A-C*	*Null*
Some B-C	59	3	22	16
All B-C	0	71	29	0
B-C	5	46	41	8

The similarity between the patterns of conclusions between universals and generics is most apparent when generics appear as the first premise (Table 3). As when universals appeared as first premise, over half (58%) of the conclusions were existentially quantified. Even more striking, generic conclusions are produced 83% and 94% of the time for universals and generics as second premises, respectively. These data reflect a clear primacy effect: generics presented as the first premise dramatically increased the proportion of generic conclusions compared to those that were produced when the first premise was universal (Mann-Whitney, $z = 3.66$, $p < .001$). Further, the proportions of universal conclusions differ as a function of primacy as well, with 46% and 11% for universal-first and generic-first, respectively (Mann-Whitney, $z = 3.20$, $p < .005$; see Tables 2 and 3).

Table 3. Proportion of responses to a generic first premise

2nd premise	Generic 1st premise (A-B)			
	Some A-C	All A-C	A-C	Null
Some B-C	58	0	23	19
All B-C	3	11	83	3
B-C	3	3	94	0

These data strongly suggest the operation of a generic overgeneralization effect, where generic premises are apparently assimilated to universal assertions. Note that this is non-normative. The generic and universal conclusions in the generics-first condition are erroneous; they do not follow from the premises because the conservative normative interpretation in the context of syllogisms should be existential. If generics had been assimilated to existentials, then the generic and universal conclusions would not have been drawn.

The overall pattern of data is consistent with the hypothesis that people tend to treat generics as universals when they appear as premises in syllogistic reasoning problems. This is analogous to the generic overgeneralization effect that was found in an evaluation task (Khemlani et al., 2007), in which people erroneously agree to the universal form of a true generic such as 'duck lay eggs', even though the universal form, e.g., 'all ducks lay eggs' is false.

An alternative explanation of the present data is that syllogistic reasoning is prone to the "atmosphere effect", where the overall atmosphere constructed by the premises in syllogisms influences participants' responses (Woodworth & Sells, 1935; Wetherick & Gilhooly, 1990; Polk & Newell, 1988). That is, participants produce a generic response when given generic premises, a universal response when given universal premises, and an existential response with existential premises. The data do not support this alternative. Participants produce more generic responses when given a universal first premise and a generic second premise than when given a generic first premise and a universal second premise, while the atmosphere effect would predict roughly equivalent proportions of such responses. Moreover, participants produce reliably fewer

existential premises when given an existential first and second premise than when given any other type of second premise. The atmosphere effect has been unable to account for other phenomena in syllogistic reasoning (Johnson-Laird & Byrne, 1991), and is similarly unable to explain the results here.

The findings for generic premises that we obtained here generalize the findings from comprehension/evaluation to a new task, syllogistic reasoning. That is, the generic overgeneralization effect is not limited to comprehension, but extends to production as well. For our syllogistic reasoning task, people not only tended to interpret generic premises as if they were universal ones, they also tended to produce generic conclusions as if they were also universal.

What accounts for the generic overgeneralization effect in both comprehension and production? Several hypotheses merit consideration. The first is based on Leslie's (2007b) supposition that characteristic generic assertions require less cognitive effort to understand than their universally quantified counterparts. In order to judge that a universally quantified assertion is false, one must retrieve a counterexample from memory, e.g., to reject the assertion 'all ducks lay eggs' one must recall that male ducks do not lay eggs. On this hypothesis people should take more time to correctly reject universally quantified generics than to incorrectly accept them, because the former requires retrieval of a counterexample. This hypothesis was not supported; in three separate experiments, people took no more time to correctly reject universally quantified generics than to incorrectly accept them (Khemlani et al., forthcoming). This hypothesis also seems unlikely in the context of syllogistic reasoning tasks. If generics were computationally simpler, we would have expected more generic conclusions than universals in the case where either one is erroneous. We found no evidence for this: people were about equally likely to produce erroneous universal conclusions as erroneous generic conclusions when the first premise was a generic (Table 2). The small difference between these two conditions is most likely attributable to a primacy effect.

An alternative hypothesis is that people tend towards default options. In the case of comprehension people would treat universally quantified assertions as generic ones, and so treat them as true. And so they do (Khemlani et al., 2007). In the case of production in the context of processing syllogisms, one strategy would be to instantiate generics as quantified assertions in order to evaluate and produce conclusions. Assimilating generics to universals is the analog of treating universals as generics in comprehension tasks, and this is what our participants seemed to do. The patterns of the types of conclusions drawn are more similar between generic and universal premises than between generic and existential premises.

Conclusions

We extended the experimental study of generics from comprehension processes to production using a syllogistic

reasoning task. In comprehension tasks, people display a tendency to treat universally quantified assertions (e.g. 'all ducks lay eggs') as generics, a phenomenon that we term the generic overgeneralization effect. Do people display an analogous tendency when producing conclusions to syllogisms that employ generics as premises? We found that they do: generics behaved in ways similar to those of universals, providing a production analog to the overgeneralization effect found for comprehension. The generic overgeneralization effect, then, describes the general phenomenon of conflating a universal assertion with a generic one in either direction.

We interpret these findings as consistent with the view that generics are default interpretations, with the universal form often misinterpreted as a generic assertion, and the mirror image of this, misinterpreting the generic as a universal assertion. These errors may well reflect the role of generics as the prototypical default form for expressing important generalizations about the world.

As such, generics provide us with an extremely useful mechanism for making such generalizations. However, there is a potential disadvantage to such a powerful mechanism. As Leslie notes, generics may give voice to pernicious generalizations about race, gender, age and ethnicity, among others (Leslie, forthcoming). Upon examination, social stereotypes act very much as ordinary generic assertions about kinds. They do not employ quantifiers, e.g., 'old people tend to forget things'. They may tend to be accepted in universal form, e.g., 'all old people tend to forget things'. And perhaps most important, they resist counterexamples. Just as the existence of a three-legged dog does not affect our agreement to the assertion that 'dogs are four-legged' (Prasada & Dillingham, 2006), so does the existence of an old person with excellent memory (such as the comic George Burns who functioned at a very high level into his nineties) fail to affect our belief that old folk do forget a lot. The irony is clear: one of our most primitive and ubiquitous forms of generalization enables us to efficiently make important inferences about the world, beginning in early childhood when, for example, children are told 'flies are dirty'. At the same time, this mechanism – and its linguistic instantiation, generic assertions – makes us vulnerable to erroneous and often damaging generalizations.

Acknowledgments

References

Carlson, G. (1977). Reference to kinds in English. Ph.D. dissertation, University of Massachusetts, Amherst.

Chapman, L. J., & Chapman, J. P. (1959). Atmosphere effect reexamined. *Journal of Experimental Psychology, 72,* 193 – 204.

Evans, J., St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition, 11,* 295 – 306.

Gelman, S.A., Goetz, P.J., Sarnecka, B.W., & Flukes, J. (in press). Generic language in parent-child conversations. *Language Learning and Development*.

Hollander, M. A., Gelman, S. A., & Star, J. (2002). Children's interpretation of generic noun phrases. *Developmental Psychology, 36 (6)*.

Johnson-Laird, P. N., Bara B. (1984). Syllogistic inference. *Cognition, 16,* 1– 61.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Psychology Press.

Khemlani, S., Leslie, S.-J., Glucksberg, S., & Fernandez, P. (2007). Do ducks lay eggs? How people interpret generic assertions. In D. S. Macnamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 395 - 401). Nashville, TN: Cognitive Science Society.

Khemlani, S., Leslie, S.-J., & Glucksberg, S. (forthcoming). Do all ducks lay eggs? The generic overgeneralization effect.

Leslie, S.-J. (2007a). *Generics, cognition, and comprehension.* Ph.D. dissertation, Princeton University.

Leslie, S.-J. (2007b). Generics and the Structure of the Mind. *Philosophical Perspectives, 21 (1),* 375 – 403.

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

Leslie, S.J. (forthcoming). The original sin of cognition.

Oakhill, J., & Johnson-Laird, P. N. (1985). The effect ofbelief on the spontaneous production of syllogistic conclusions. *Quarterly Journal of Experimental Psychology, 37A,* 553 – 570.

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

Polk, T. A., & Newell, A. (1988). Modeling human syllogistic reasoning in Soar. In *Proceedings of the 10th Annual Cognitive Science Society*, 181 – 187. Hillsdale, NJ: Lawrence Erlbaum Associates.

Prasada, S., & Dillingham, E. (2006). Principled and statistical connections in common sense conception. *Cognition, 99 (1),* 73 – 11.

Wetherick, N., & Gilhooly, K. (1990). The figural effect in syllogistic reasoning. In K. Gilhooly, M. T. G. Keane, R. Logie, & G. Erdos (Eds.), *Lines of thought: Reflections on the psychology of thinking, 1*. London: Wiley.

Woodworth, R. S., & Sells, S. B. (1935). An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology, 18 (4),* 451 – 460.

Appendix A

First premise	Second premise
Alligators have enzyme I-381	Creatures with enzyme I-381 are semi-nocturnal
Dolphins are protocetaceans	Protocetaceans are viviparous
Comets are orthovolatile	Orthovolatile materials are pollutants
Hydrochlorides are paramagnetic	Paramagnetic substances cause tooth decay
Muskrats have gene BDA-23	Individuals with gene BDA-23 mark their territory with urine
Semi-automatic weapons have recoil mounts	Weapons with recoil mounts cause repetitive stress injury
Septapods have ovipositors	Individuals with ovipositors live in warm climates
Iridium tablets emit beta radiation	Substances that emit beta radiation are found in Croatia
Kangaroos are polymorphic	Polymorphic individuals have gene Gamma-64
Sensillas are trichoidal	Trichoidal subtances cause skin irritation
Elephants are semiparametric	Semiparametric individuals heal quickly
Wasps are type 6 parasites	Type 6 parasites are arthropods
Pyramidal cells are mesovoltaic	Mesovoltaic subtances are compact
Bamboo shoots are pseudospores	Pseudospores have reticulated proteins
Seasalt crystals are hypersalinic	Hypersalinic particles are gray
Snakes are ontodextrous	Ontodextrous individuals have protein G-168
Delta transmitters are quasidelphic	Quasidelphic cells can withstand superheating