Frequently produced semantic features reflect principled connections

Hillary Harner[1] and Sangeet Khemlani[2]
hillary.harner@altamiracorp.com, sangeet.khemlani@nrl.navy.mil
[1]Altamira Technologies Corporation, McLean, VA 22102 USA
[2] Navy Center for Applied Research in Artificial Intelligence
U.S. Naval Research Laboratory, Washington, DC 20375 USA

Abstract

When people think about the features of common objects, like scissors, they often spontaneously recall a central feature: scissors *cut things.* They tend not to recall other features of scissors, e.g., that they *have handles.* The present paper posits a novel explanation for the behavior: the features people recall first and most often reflect semantic generalizations of kinds. A recent taxonomy of such generalizations suggests that people represent privileged links between kinds and their features known as *principled connections* (Prasada et al., 2013). Several tests diagnose principled connections: for instance, principled connections reflect norms, so one way to diagnose the presence of a principled connection is to test the acceptability of sentences of the form *all normal Xs have feature Y*, as in *all normal cars have four wheels.* We tested whether participants accept generalizations about the normality of features produced in a semantic feature production task (Experiments 1 and 2) as well as self-referential generalizations (Experiment 3). The experiments provided participants with generalizations about features listed first and most often as well as features that people list less frequently. They found that people readily accepted generalizations that diagnose the presence of principled connections. The results corroborate the view that principled connections help people recall the features of conceptual categories.

Keywords: concepts, categories, recall, principled connections, norms

Introduction

Consider all of the facts you know about birds. Before reading further, it may be worthwhile to enumerate the first three things about them that come to mind. Perhaps you remembered specific birds, such as a pet or a group of pigeons close to where you work. Or perhaps you thought of what birds are made of – they have beaks, feathers, and hollow bones. Perhaps you imagined some of their characteristic behaviors, such as the fact that they make nests, lay eggs, and fly. If you thought of one or more specific birds, you were relying on episodic memories. Instead, if you thought of birds in general, then you accessed your semantic memories of the features of prototypical birds. Features are central to theories of categorization and conceptual representation (Medin & Schaffer, 1978; Minda & Smith, 2002; Smith & Medin, 1981; Smith, Shoben, & Rips, 1974), but as Murphy (2003, p. 26) observes, features seldom serve as *defining* concepts of a kind. For example, squids possess the necessary features of being fish – having gills – but people don't consider them to be members of that kind (Hampton, 1979; Hampton, 2017).

People can nevertheless express commonsense knowledge about the features of birds using statements such as (1a-c):

1a. Birds have beaks.
b. Birds lay eggs.
c. Birds fly.

These statements are known as *bare-plural generics*: they express generalizations without the use of quantifiers such as *all* and *most* (Carlson, 2009; Chambers, Graham, & Turner, 2008; Cimpian & Markman, 2009; Gelman, 2003, 2004; Gelman & Bloom, 2007; Hampton, 2009; Leslie, 2007, 2008, 2012; Noyes & Keil, 2019; Pelletier, 2009; Prasada, 1999, 2000, 2010, 2012, 2017; Prasada & Dillingham, 2006, 2009). Generics, unlike universal quantifiers, tolerate exceptions: the statement *birds fly* is true regardless of the existence of flightless birds. As Prasada and Dillingham (2006, 2009) argue, generics serve as a window into people's commonsense conceptual representations of kinds and their features. According to Leslie (2007, 2008), generics reflect people's foundational, default mode of generalization, and they do not consistently correspond to a specific quantificational force. For instance, consider the following generics:

2a. Tigers are striped.
b. Ducks lay eggs.
c. Mosquitoes carry malaria.
d. *People are right-handed.

In some cases, generics are true when a majority of the population possesses the feature in question, as in (2a). But the other examples appear to flout any consideration of prevalence (cf. Tessler & Goodman, 2019): less than half the population of ducks lays eggs (only the fertile females; see Leslie, Khemlani, & Glucksberg, 2011), and a very small percentage of mosquitoes carry malaria. Nevertheless, (2b) and (2c) are acceptable. In contrast, even though most people are right-handed, (2d) is unacceptable. Leslie argues that these patterns reflect the fact that generics don't express quantificational information. Further evidence comes from developmental work that shows that children begin to use and understand generics such as *birds fly* earlier than when they acquire quantifiers such as *all* in *all birds are animals* (Gelman, Star, & Flukes, 2002; Hollander, Gelman, & Star, 2002). As Gelman and Bloom (2007) argue, many generics express features, e.g., *flying*, that are central, essential, and enduring to a kind, e.g., *birds*.

This paper focuses on how a recent theory of the conceptual distinctions between generics (Prasada et al., 2013) can serve as a foundation for a broader account of the organization of conceptual features in semantic memory. We first provide a brief overview of the theory, and then describe its central tenet – the idea that people maintain privileged links between certain kinds and certain features known as *principled connections*. We show how to distinguish between principled and non-principled connections, and we describe how the theory can be tested by analyzing the information people can recall about kinds and their features. We start by introducing Prasada et al.'s (2013) theory of concepts.

Semantic memory for principled features

Prasada and colleagues argue that generics can reveal the conceptual representations of kinds and their features for statements like (2a-c) above (Prasada, 2017; Prasada & Dillingham, 2006, 2009; Prasada et al., 2013). As they showed, people tend to distinguish between the following two generics:

3a. Cars have four wheels.
b. Cars have radios.

Both (3a) and (3b) are acceptable, and the percentages of cars with four wheels and cars with radios are both high. Nevertheless, (3a) seems to describe a feature of cars that is more fundamental and central to carhood – a car without four wheels would be considered abnormal or defective in some fashion, or not a car in the first place. In contrast, a car without a radio wouldn't be considered abnormal. As Prasada and Dillingham (2006; 2009) argued, (3a) reflects a *principled* connection, i.e., a privileged link between a kind and a characteristic feature of that kind, whereas (3b) reflects a *statistical* connection, i.e., an incidental feature that holds for many instances of the kind but does not contribute to its meaning or essence. Kinds can have multiple principled connections (Prasada, 2017) – some may describe the material and parts the car is made of, as in (3a), and some may describe the primary function of cars, as in *cars are for driving* (see Korman & Khemlani, 2020).

As Korman and Khemlani (2020) observe, principled connections can be distinguished from non-principled connections in four different ways:

i. *Norms.* Principled features – e.g., those features linked to a kind via a principled connection – are considered normal and normative (Prasada & Dillingham, 2006). Hence, people should consider it normal for a car to have four wheels; and they should accept the normative statement *cars are supposed to have four wheels*. In contrast, cars that don't have radios shouldn't be considered abnormal, and no expectation should exist that cars are supposed to have radios.

ii. *Self-referentiality.* Principled connections license "self-referential" generalizations of the feature, e.g., *Xs, by virtue of being Xs, are Y* (Prasada & Dillingham, 2006), e.g., people accept the statement *cars, by virtue of being cars, have four wheels* but reject the statement *cars, by virtue of being cars, have radios*.

iii. *Aspect*. Principled features should be considered an aspect of the kind (Prasada & Dillingham, 2009), and so people should agree with the statement *having four wheels is one aspect of being a car* but reject the statement *having a radio is one aspect of being a car*.

iv. *Formal explanations*. principled connections should license "formal" explanations, i.e., shallow explanations that appeal to the kind itself (Prasada, 2017). Hence, people should agree with the explanation *cars have four wheels because they're cars* but reject the explanation *cars have radios because they're cars*.

Several studies corroborate the predictions outlined above (Korman & Khemlani, 2020; Prasada et al., 2013). They show that when the four tests embodied by (i-iv) are satisfied between a kind (e.g., *cars*) and a feature (e.g., *having four wheels*), people consider the corresponding generic true (e.g., *cars have four wheels*).

If the theory of principled connections and its broader conceptual framework are true, then principled connections should serve as a representational structure fundamental to the organization of the semantic network that links kinds and their features. That is, people's conceptual representations are independent from the way they articulate those representations: hence, for a given kind, such as *cars*, its conceptual structure must be established before people can start to describe their general knowledge about cars using generic expressions. Indeed, because principled connections serve as privileged links between kinds and features, they may help direct people's recollection when individuals have to consider the commonsense knowledge they have encoded about a particular kind. Returning to our original example, when you think of everything you know about the kind *birds*, you may be more apt to recall information about principled features – e.g., that they *fly* and *lay eggs* – than incidental features of birdhood, e.g., that they *build nests* and *have eyes.* We summarize the hypothesis as follows:

Principled connection hypothesis: Those features that people produce for a given kind first and most often should be principled features, i.e., they should be linked to a kind via a principled connection.

In the remainder of the paper, we describe experiments that test the prediction. Experiments 1 and 2 tested whether the most frequent features that people generate are also those that reflect norms. Experiment 3 tested whether the most frequent responses generated permit self-referential generalizations between the kind and the property. All three studies support the theory of principled connections.

Experiment 1

Experiment 1 sought to examine whether the most frequent features listed by participants were also those features that reflected a norm, in line with the first test of principled connections described in the Introduction. When participants are presented with an experimental feature and a control feature for a particular kind, they should consider the experimental feature to have normative force more often than the control feature, because the experimental feature is hypothesized to bear a principled connection to the kind.

To test the idea, we collected data from a preliminary norming study in which participants were provided a cue, e.g., *airplane,* and were asked to list three features of that kind. The norming study collected such semantic feature production data for a total of 40 cues. Prospective principled connections (i.e., *experimental* features) concerned those features that participants generated first and most frequently; control features were those that multiple participants had listed second or third. The norming study yielded pairs of experimental and control features for 26 separate kinds. Experiment 1 used these normed experimental and control features to generate sentences designed to test whether people consider a given feature to be a normal property of the kind. Those sentences were of the form *all normal X are for Y* (see Korman & Khemlani, 2020; Prasada et al., 2013 for data validating the use of this formulation in testing for principled connections). The experiment presented participants with two statements on each trial as follows:

> *Which sentence do you think provides the best description of airplanes?*
>
> *All normal airplanes fly.* (experimental)
> *All normal airplanes are loud.* (control)

Participants responded by selecting one of the two options. On any given trial, a bias towards the experimental features would validate the prediction that the features people generate most frequently in a semantic production tasks bear principled connections to a kind.

Method

Participants. We conducted a power analysis using the `pwr` package (Champely et al., 2018) in R. The goal of Experiment 1 was to obtain .85 power to detect a large effect (d = 0.7) at .01 α error probability, and so 46 participants were required for the study. 50 participants (mean age = 36.5 years; 35 males and 15 females) volunteered through Amazon Mechanical Turk. All participants reported being native English speakers.

Design, procedure, and materials. The experiment presented participants with 26 trials based on materials generated from the norming study described above (see https://osf.io/394fc/). On each trial, participants were asked to select "the best description" of a particular kind, e.g., *airplanes.* They selected from two separate options, e.g., an experimental and a control option. Participants could not move onto the next trial without making a selection. The order of the stimuli was randomized for each participant.

Results and discussion

Figure 1 presents the distributions of responses on which participants selected the experimental option for each of the 26 items. Participants selected the experimental option 70% of the time, at a level significantly higher than chance (Wilcoxon test, $z = 5.89$, $p < .001$, Cliff's $\delta = .86$). As the figure shows, 20 of the 26 items yielded responses for which participants selected the experimental option the majority of the time (binomial test, $p = .009$).

Two factors may have explained participants' tendency to select the experimental option: one was the proportion of times people generated a particular experimental feature in the semantic feature production task, i.e., the number of times an experimental feature was produced divided by the total number of features produced. Another was an analogous measure for control features, i.e., the proportion of times people generated a particular control feature. Those data are provided in the Appendix. To test how much those factors

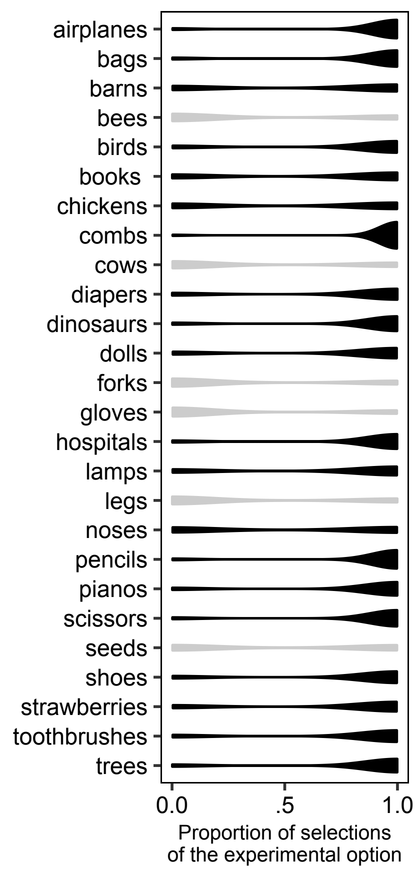

Figure 1. Violin plots of the proportions of selections of the experimental options in Experiment 1, i.e., the options that were predicted to bear principled connections to the kind, as a function of the 26 different items in the experiment. Black plots reflect items for which participants selected the experimental option more than the control option on average, and grey plots denote the opposite trend.

could explain the bias to select experimental options in Experiment 1, the data from Experiment 1 as well as the norming study were subjected to a generalized linear mixed-model (GLMM) regression that treated the proportions of experimental features and the proportion of control features as fixed effects and the items and participants as random effects. The analysis revealed that the proportion of experimental features in the norming study had little effect on participants' tendency to choose the experimental option (β = .21, t = .38, p = .70). In contrast, the proportion of control features in the norming study was a reliable predictor of their tendency to choose the experimental option (β = 2.63, t = 2.36, p = .03). In other words, the frequency with which participants described a control feature in the norming study affected which of the sentences participants selected in Experiment 1. One way to interpret this finding is that connections between kinds and their features can be along a gradient of strength for non-principled features, while the same connection is at ceiling for principled features.

A post-hoc analysis examined the difference between the artifacts and the natural kinds in the study. It found a significant difference in the tendency for participants to select the experimental option: they did so on 76% of the trials for artifacts but only 60% of the trials for natural kinds (Wilcoxon test, z = 4.82, p < .0001, Cliff's δ = .60). The prediction above did not readily predict such a difference – it may be a function of the small number of items produced in the norming study.

Experiment 1 demonstrated that participants chose an experimental option over a control option when they had to directly compare the two. One limitation of the study is that the forced-choice nature of the task make have biased participants towards the experimental option – the option hypothesized to reflect a principled connection – over the control option. Experiment 2 accordingly presented participants the two options on separate trials.

Experiment 2

Experiment 2 provided participants with the same experimental and control options as in Experiment 1. However, instead of asking which of two separate options was the best description of a kind, it presented each option in isolation, asking participants to rate its truth. For instance, participants received the following experimental statement:

All normal seeds grow. (experimental)

and they rated the extent to which they considered the statement to be true or false on a 7-point Likert scale (-3 = definitely false; 0 = I cannot be certain; 3 = definitely true).

Method

Participants. 50 native English speakers (mean age = 39.0 years; 27 male, 23 female) participated via Amazon's Mechanical Turk.

Design, procedure, and materials. Experiment 2 presented participants with 52 trials, derived from the materials created from the norming study (see Appendix). The materials were used to create sentences of the form *All normal Xs Y* where *X* describes a kind and *Y* denotes a feature. Participants rated each sentence one at a time as true or false along a 7-point Likert scale. The order of the sentences was randomized.

Results and discussion

As in Experiment 1, participants rated experimental options as more true than control items ($M_{\text{experimental}}$ = 1.29 vs. M_{control} = 0.44; Wilcoxon test, z = 6.09, p < .001, Cliff's δ = .42). The overall results corroborated the prediction above: experimental options were rated as more true than control options. Hence, participants appeared to endorse generalizations about the normality of features more often for experimental features, i.e., those features that participants produced first and most often in a semantic feature production task. The results corroborate the prediction that frequently produced features reflect norms.

A second test of the principled connection hypothesis concerns aspectival treatments of features. According to the hypothesis, if the most commonly and initially produced features of a kind reflect principled connections, they should license self-referential generalizations between the kind and the feature, as in *airplanes, by virtue of being airplanes, fly.* Experiment 3 tested that consequence, and it also addressed a major limitation of the preceding studies.

Experiment 3

One limitation of Experiments 1 and 2 is that the materials provided to participants came from a single norming study (details of the study are available at https://osf.io/394fc/) that collected feature production data for only 40 separate kinds (see Experiment 1). The results could therefore be driven by idiosyncrasies of those 40 kinds, we so sought to derive materials from a larger dataset, namely the semantic feature production database constructed by Buchanan, Valentine, and Maxwell (2019), which presents semantic feature production data for over 4000 kinds (and which incorporates data from earlier studies, e.g., McRae et al., 2005). Experiment 3 drew materials from a small subset of those 4000 kinds in a manner similar to that described in Experiment 1.

Experiment 3 otherwise mimicked Experiment 1 in its design, except for the formulation that participants evaluated, i.e., participants evaluated self-referential generalizations between kinds and features for experimental and control items. The experiment presented participants with two statements on each trial as follows:

Which sentence do you think provides the best description of trains?

Trains, by virtue of being trains, run on ***tracks***. (experimental)

Trains, by virtue of being trains, carry ***passengers***. (control)

Participants responded by selecting one of the two options. On any given trial, a bias towards the experimental option validates the prediction that the features people generate most frequently in a semantic feature production task bear principled connections to a kind.

Method

Participants. 51 participants (mean age = 38.82 years; 27 males and 24 females) volunteered through Amazon Mechanical Turk. All participants reported being native English speakers.

Design, procedure, and materials. The experiment presented participants with 22 trials. The materials were constructed on norming data provided by Buchanan et al. (2019), and the methods for their construction are available online: https://osf.io/394fc/. On each trial, participants were asked to select "the best description" of a given kind. They selected from two separate options, e.g., a sentence with a most frequent feature and one with an infrequent feature. As in the preceding studies, participants had to make a selection before moving to the next trial, and the order of the stimuli was randomized for each participant.

Results and discussion

Figure 2 presents the distribution of participants' selection of sentences with frequent features, for each of the 22 items. Overall, participants selected the experimental sentences 78% of the time, at a level significantly higher than chance (Wilcoxon test, z = 5.87, p < .001, Cliff's δ = .78). As the figure shows, participants selected the experimental sentences the majority of the time for 19 of the 22 kinds (binomial test, p < .001 assuming a prior probability of ½). The use of normed data from Buchanan et al. (2019) allowed us to examine additional factors that may have explained participants' tendency to select the experimental option, namely the relative frequencies of production for experimental and control features. To test how much those factors could explain the bias to select experimental sentences in Experiment 3, the data from Experiment 3 and the data from the corpora were subjected to a generalized linear mixed-model (GLMM) regression that treated the frequencies of experimental and control features as fixed effects and the items and participants as random effects. The analysis revealed that the relative frequency of experimental features had little effect on participants' tendency to choose the high frequency feature sentence (-.01 < β < .01, t = -.14, p = .89), and likewise for the relative frequency for control features (-.01 < β < .01, t =.51, p = .62). The intercept, however, was reliably greater than zero (β = .76, t = 5.08, p < .001), as predicted by the principled connection hypothesis.

In sum, Experiments 1-3 corroborated the predictions of the principled connection hypothesis, namely that the features people produce first and most often for kinds are those that bear a principled connection to the kind.

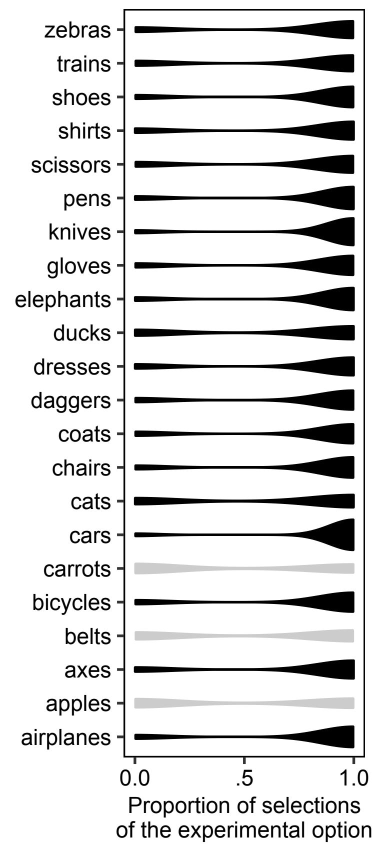

Figure 2. Violin plots of the proportions of selections of the experimental options in Experiment 3, i.e., the options that were predicted to bear principled connections to the kind, as a function of the 22 different items in the experiment. Black plots reflect items where participants selected the most frequent feature option more than the infrequent feature option on average, and grey plots denote the opposite trend.

General discussion

People's semantic knowledge of kinds such as *birds* and *airplanes* is complex. You know, for instance, that all airplanes have seats, radios, and doors. And yet, when asked to list the features of an airplane, people rarely mention such features, if ever. McRae et al. (2005) compiled a dataset of participants' natural responses in a semantic feature production task, i.e., a task designed to gather participants' spontaneous recall of the features of a kind. Their dataset shows the ten most common features of airplanes that people recall:

They *fly*.	**(25)**
They *have wings.*	(20)
They're *used for passengers.*	(15)
They're *fast*.	(11)
They *require pilots*.	(11)
They're *used for transportation.*	(10)
They're *found in airports.*	(8)
They're *large*.	(8)
They're *made of metal*.	(8)

Next to each feature, we provide the number of participants (out a possible 30 in their norming sample) that spontaneously produced the response. We highlight the top responses to suggest that there is something unique about it.

When people spontaneously recall features of various kinds, they do so in a way that is guided by the structure of their semantic network of concepts. Researchers have proposed various ways in which such a network of concepts is organized (Collins & Quillian, 1969; Collins & Loftus, 1975; Griffiths, Steyvers, & Tenenbaum, 2007; Jones, Willits, & Dennis, 2015; McClelland & Rogers, 2003; McRae & Jones, 2003; Smith, Shoben, & Rips, 1974; Siew, Wulff, Beckage, & Kenett, 2019; Szymański & Duch, 2012). But few theories explain why people produce certain features systematically more often than others.

We posit a novel solution to the problem: the features recalled first and most often reflect semantic generalizations of kinds. Hence, when people remember that *flying* is a feature of an airplane, they do so because *airplanes fly* is a true generalization about the behavior of airplanes. And it is expressed using a bare-plural generic, i.e., a generalization that makes no use of a quantifier such as *all* or *most.* A recent theory provides a conceptual taxonomy for generic assertions (Prasada, 2017; Prasada & Dillingham, 2006, 2009; Prasada et al., 2013). It posits that generics often reflect *principled* connections – privileged conceptual links -- between kinds and those features that are most central and characteristic to their meaning. For instance, the generic *cars have four wheels* reflects a principled connection, whereas the generic *cars have radios*, does not, because there's nothing strange about a car without a radio, whereas a car without four wheels may be an abnormal or broken car. Principled connections can therefore reflect norms and normative thinking – as evidenced by the fact that people agree with the statement *all normal cars have four wheels*, but they reject the statement *all normal cars have radios* (Prasada et al., 2013).

If principled connections are privileged links between kinds and their central features, then the features people recall first and most often may reflect principled connections. We find evidence for such behavior in two experiments. One experiment provided participants with statements of the form *all normal airplanes fly* – the most frequent response in a semantic feature production task – and *all normal airplanes are loud* – a less frequent response, and found that participants preferred the first to the second. A second experiment provided the statements separately, and found that overall, participants preferred the first to the second. And a third study generalized the results to self-referential descriptions. Together, these results corroborate the view that principled connections guide the recall of semantic features.

Acknowledgments

This work was supported by an NRC Research Associateship Award to the first author and support from the Naval Research Laboratory to the second author. We are indebted to Jon Baron, Tony Harrison, Laura Hiatt, Zach Horne, Laura Kelly, Joanna Korman, Greg Murphy, Sandeep Prasada, and Greg Trafton for their advice and comments. We also thank Kalyan Gupta, Danielle Paterno, Kevin Zish, and the Knexus Research Corporation for their help in data collection.

References

Abele, A., Cuddy, A., Judd, C., & Yzerbyt, V. (2008). Fundamental dimensions of social judgment. *European Journal of Social Psychology*, 38, 1063-1065.

Buchanan, E.M., Valentine, K.D., & Maxwell, N.P. (2019). English semantic feature production norms: An extended database of 4,436 concepts. *Behavior Research Methods, 51*, 1849–1863.

Carlson, G. (2009). Generics and concepts. In F. J. Pelletier (Ed.), *Kinds, things, and stuff: Mass terms and generics. New directions in cognitive science*. New York: Oxford University Press. Vol. 12.

Chambers, C.G., Graham, S.A., & Turner, J.N. (2008). When hearsay trumps evidence: How generic language guides preschoolers' inferences about unfamiliar things. *Language and Cognitive Processes*, *23,* 749–766.

Champely, S., et al. (2018). Package 'pwr'. R package version, 1(2).

Cimpian, A., & Markman, E. M. (2009). Information learned from generic language becomes central to children's biological concepts: Evidence from their open-ended explanations. *Cognition*, *113*, 14–25.

Collins, A.M., & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. Journal of Verbal *Learning & Verbal Behavior, 8*, 240–247.

Garrard, P., Lambon Ralph, M.A., Hodges, J.R., & Patterson, K. (2001). Prototypicality, distinctiveness, and intercorrelation: Analyses of the semantic attributes of living and nonliving concepts. *Cognitive Neuropsychology, 18*, 125-174.

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

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, 8, 404–409.

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

Hampton, J.A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18*, 441-461.

Hampton, J.A. (2009). Stability in concepts and evaluating the truth of generic statements. In F. J. Pelletier (Ed.), *Kinds, things, and stuff: Mass terms and generics. New directions in cognitive science*. New York: Oxford University Press. Vol. 12.

Hampton, J.A. (2017). Compositionality and concepts. In J.A. Hampton and Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 95-121). Springer International Publishing.

Hollander, M., Gelman, S., & Star, J. (2002). Children's interpretation of generic noun phrases. *Developmental Psychology*, 38(6), 883-894.

Jones, M.N., Willits, J., Dennis, S. (2015). Models of semantic memory. In J.R. Busemeyer, Z. Wang, J.T. Townsend, & A.

Eidels (Eds.), *The Oxford Handbook of Mathematical and Computational Psychology*, (pp. 232-254).

Korman, J., & Khemlani, S. (2020). Teleological generics. *Cognition, 200,* 104157.

Leslie, S.-J. (2007). Generics and the structure of the mind. *Philosophical Perspectives*, *21*, 375-403.

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

Leslie, S.-J. (2012). Generics articulate default generalizations. *Recherches Linguistiques de Vincennes*, *41*, 25-44.

Leslie, S.-J., Khemlani, S., & Glucksberg, S. (2011). Do all ducks lay eggs? The generic overgeneralization effect. *Journal of Memory and Language, 65,* 15-31.

McClelland, J.L., & Rogers, T.T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience, 4,* 310-322.

McRae, K., Cree, G., Seidenberg, M., & Mcnorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior research methods, 37,* 547-59.

McRae, K., and Jones, M. (2013). Semantic memory. In D. Reisberg (Ed.), *The Oxford Handbook of Cognitive Pyschology*.

Medin, D.L., & Schaffer, M.M. (1978). Context theory of classification learning. *Psychological Review, 85.*

Minda, J.P., & Smith, J.D. (2002). Comparing prototype-based and exemplar-based accounts of category learning and attentional allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 272-292.

Murphy, M.L. (2003). *Semantic Relations and the Lexicon: Antonymy, Synonymy and other Paradigms.* Cambridge: Cambridge University Press.

Noyes, A., & Keil, F.C. (2019). Generics designate kinds but not always essences: Implications for social categories and kinds. *Proceedings of the National Academy of Sciences, 116*, 20354-20359.

Paolacci, G., Chandler, J., & Ipeirotis, P.G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*, 411-419.

Pelletier, F.J. (2009). *Kinds, things, and stuff: Mass terms and generics. New Directions in Cognitive Science* (Vol. 12). New York: Oxford University Press.

Prasada, S. (1999). Names for things and stuff. In R. Jackendoff, P. Bloom, & K. Wynn (Eds.), *Language, logic, and concepts* (pp. 119-146). Cambridge: MIT Press.

Prasada, S. (2010). Conceptual representations and some forms of genericity. In F.J. Pelletier, (Ed.), *Kinds, Things & Stuff.* New York City: Oxford University Press.

Prasada, S. (2012). Mechanisms for characterizing kinds and classes. *Recherches Linguistiques de Vincennes*, *41*, 45-62.

Prasada, S. (2017). The scope of formal explanation. *Psychonomic Bulletin and Review*, *24*, 1478-1487.

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

Prasada, S., & Dillingham, E. (2009). Representation of principled connections: A window onto the formal aspect of common sense conception. *Cognitive Science*, *33*, 401-448.

Prasada, S., Khemlani, S., Leslie, S.-J., & Glucksberg, S. (2013). Conceptual distinctions amongst generics. *Cognition*, *126*, 405- 422.

Siew, C.S.Q., Wulff, D.U., Beckage, N.M., & Kenett, Y.N. (2019). Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity*, 1-24.

Smith, E.E., & Medin, D.L. (1981). *Concepts and categories*. Cambridge: Harvard University Press.

Smith, E.E., Shoben, E.J., & Rips, L.J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81*, 214–241.

Szymanski, J., & Duch, W. (2012). Information retrieval with semantic memory model. *Cognitive Systems Research, 14*, 84-100.

Tessler, M. H., & Goodman, N. D. (2019). The language of generalization. *Psychological Review, 126*, 395-436.

Vinson, D.P., & Vigliocco, G. (2002). A semantic analysis of grammatical class impairments: semantic representations of object nouns, action nouns and action verbs. *Journal of Neurolinguistics, 15*, 317-351.

Vinson, D.P. & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods, 40*, 183-190.