UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
These boots are made for walking: Teleogical generalizations from principled connections

Permalink
https://escholarship.org/uc/item/4nq120cj

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 40(0)

Authors
Korman, Joanna
Khemlani, Sangeet

Publication Date
2018

eScholarship.org
Powered by the California Digital Library
University of California

These boots are made for walking: Teleological generalizations from principled connections

Joanna Korman and Sangeet Khemlani
{joanna.korman.ctr, sangeet.khemlani}@nrl.navy.mil
Navy Center for Applied Research in Artificial Intelligence
US Naval Research Laboratory, Washington, DC 20375 USA

Abstract

Certain generalizations are teleological, e.g., *forks are for eating*. But not all properties relevant to a particular concept permit teleological generalization. For instance, forks get washed roughly as often as they're used for eating, yet the generalization, *forks are for washing*, might strike reasoners as unacceptable. What explains the discrepancy? A recent taxonomic theory of conceptual generalization (Prasada, 2017; Prasada & Dillingham, 2006; Prasada et al., 2013) argues that certain kinds of conceptual connections – known as "principled" connections – license generalizations, whereas associative, "statistical" connections license only probabilistic expectations. We apply this taxonomy to explain teleological generalization: it predicts that acceptable teleological generalizations concern concept-property pairs in which the concept bears a principled connection to a property. Under this analysis, the concept *fork* bears a principled connection to *eating* and a statistical connection to *washing*. Two experiments and a regression analysis tested and corroborated the predictions of the theory.

Keywords: teleological generalization, generics, principled connections, statistical connections

Introduction

"What are forks for?" A reasonable answer to the question may be: *forks are for eating*. The answer is a generalization, since it concerns forks in general instead of some specific instance of a fork. Moreover, the answer is teleological, because it refers to the primary purpose that forks serve. From early in their cognitive development, humans begin to understand the teleological functions of objects and actions (Atran, 1995; Carey, 1985; Csibra & Gergely, 1998; Keil, 1992). They think of both artifacts, such as *forks*, and biological parts, such as *teeth*, as existing to serve some kind of function (Keil, 1992). Toddlers appear to interpret the actions of others as purposeful and in the service of some goal (Onishi & Baillargeon, 2005; Southgate, Johnson, & Csibra, 2008). Indeed, as Kelemen discovered, children can also be "promiscuous" in their teleological conceptualizations: for instance, young children think that *mountains are for climbing* and that *lions exist for going to the zoo* (Kelemen, 1999). And even adult college students and professional scientists under time pressure find certain unwarranted teleological explanations alluring, such as: *ferns grow in forests because they provide ground shade* (Kelemen, Rottman, & Seston, 2013). The assertion is a teleological explanation because it refers to a teleological generalization (*forests exist to provide ground shade*) in order to explain a non-teleological generalization (*ferns grow in forests*). Reasoners often construct teleological explanations to interpret, not just generalizations, but specific situations as well, such as, *that particular weapon was used for that particular crime*.

The philosophy and psychology of teleological thinking

In the present paper, we focus on what makes teleological generalizations, such as *teeth are for chewing,* acceptable or unacceptable. Wright (1976) argued that teleological assertions reduce to causal relations in that they refer to the etiology of the property in question, i.e., the set of preceding causes that brought about the property. On Wright's account, *teeth are for chewing* is acceptable because chewing is a consequence of the existence of teeth, i.e., teeth exist to permit chewing. While Wright's analysis was normative – it concerned the conditions that should sanction ideal teleological assertions – Lombrozo and Carey (2006) sought to test its descriptive validity. Participants in their studies accepted teleological generalizations of specific events more often for events that resulted from a set of preceding causes than for those that resulted accidentally. Lombrozo and Carey (2006) argued that the pattern lends some credence to Wright's (1976) etiological proposal, but they further observed that participants preferred teleological explanations when the sets of preceding causes described general patterns that could be used for making predictions about the future. Their results suggest that even when reasoners consider specific situations, they often make teleological inferences and predictions by considering (perhaps ad hoc) generalizations.

An alternative view of teleological thinking comes from theorists who argue that ideal teleological generalizations, e.g., *teeth are for chewing*, should reference so-called "dispositional" properties (Bigelow & Pargetter, 1987; Brown, 1952). Dispositional properties are controversial in philosophy because they are intrinsic, unobservable, and, by definition, do not always manifest. Nevertheless, many accounts of dispositions argue that they hold in reference to a particular condition (Choi & Fara, 2016). For instance, *fragility* is considered a dispositional property because fragile items will break in the event that they're struck by some sufficiently strong force. Of course, an item need not ever break for it to be considered fragile. Bigelow and Pargetter draw a similar conclusion about teleological generalizations (1987, p. 189): the generalization *teeth are for chewing* should be true even though teeth are not engaged in the activity most of the time. Hence, *being for chewing* is a dispositional property of teeth, while *being white* is not, since it is both manifest and observable of prototypical teeth. On the dispositional account, the assertion *teeth are for being white* should be rejected.

Philosophers continue to debate the veracity and independence of both the etiological and the dispositional accounts of teleological explanation (Delancey, 2006; Kroll, 2017; Mitchell, 1993; Neander, 1991), but both accounts may have limited purchase in psychology and cognitive science, because they fail to explain what makes certain teleological generalizations unacceptable. For instance, neither account can explain why, e.g., *forks are for eating* is a reasonable teleological generalization while *forks are for washing* is not. *Being for eating* and *being for washing* are both, after all, dispositional properties. And it seems odd to think of a fork's primary purpose in terms of preceding causes. Eating is not a consequence of the existence of forks, and so Wright's (1976) analysis is incomplete at best.

Perhaps accounts of teleological thinking fail because even non-teleological generalizations present semantic challenges. For instance, people accept the generalization *cows have udders* even though only a minority of cows (mature females) develop udders (Leslie, Khemlani, & Glucksberg, 2007). A recent taxonomy proposed by Prasada and colleagues addresses representational differences that explain why people accept some generalizations and not others. In what follows, we show how the taxonomy copes with teleological generalizations, and we describe two experiments and a regression analysis that corroborate the taxonomy's predictions. We conclude by describing how teleological generalizations fit into a broad treatment of conceptual representation.

Principled and statistical generalizations

Prasada and colleagues proposed that people represent conceptual knowledge through different types of connections between kind concepts (e.g., *cow*) and properties (e.g., *udder*). These connections explain why people accept statements such as *cows have udders* (Prasada, 2017; Prasada, Khemlani, Leslie, & Glucksberg, 2013; see also Prasada & Dillingham, 2006). In particular, the researchers argue that people distinguish "principled" from "statistical" connections. Principled connections link a privileged, characteristic property to a concept. Such properties are those that are an essential element of what it means to be a member of that concept. One way to distinguish principled connections is that only principled connections license normative expectations, e.g., *cows are supposed to have udders.* In contrast, the generalization, *cows are white,* is statistical, not principled: it refers to an incidental property that is associated with the concept *cow,* not one that is an essential element of being a cow. Statistical connections yield only the probabilistic expectation that *most cows are white*; they do not yield normative expectations, and so people are less likely to believe that *cows are supposed to be white.* As Prasada et al. (2013, p. 408 et seq.) show, principled connections sanction many different types of expectations besides normative expectations: they license self-referential expectations (e.g., *cows, by virtue of being cows, have udders*), expectations about normality (e.g., *all normal cows have udders*), and aspectual expectations (e.g., *one aspect of being a cow is having udders*). Prasada and colleagues used these expectations to diagnose whether a particular generalization refers to a principled connection.

We apply Prasada et al.'s taxonomy to teleological generalizations (see also Prasada, 2017). In essence, a teleological property of a concept (e.g., *forks)* describes its primary function or purpose (e.g., *being for eating*), and a teleological generalization should be deemed acceptable when a concept (*forks*) and the property are linked by a principled connection. Principled connections should permit self-referential generalizations (e.g., *forks, by virtue of being forks, are for eating*), expectations about normality (e.g., *all normal forks are for eating*), and expectations about normativity (e.g., *forks are supposed to be for eating*). Accordingly, Prasada et al.'s theory makes the following two predictions:

> **Prediction 1:** Participants should accept teleological generalizations whenever they accept corresponding self-referential generalizations, expectations of normality, and expectations of normativity.

A corollary of Prediction 1 is that when people reject assertions diagnostic of principled connections, they should also reject the corresponding teleological generalizations. For instance, those who disagree with the diagnostic statement *all normal forks are for washing* should also disagree with *forks are for washing*. Experiments 1 and 2 sought to test prediction 1.

According to Prasada et al., generalizations depend on the conceptual link between concept and a property, not on knowledge about the probabilistic associations between the two (see also Khemlani, Leslie, & Glucksberg, 2012). In other words, people may possess relevant probabilistic beliefs about, e.g., the conditional probability of an object being for eating, given that is a fork, e.g., P(*for-eating* | *fork*), or else the likelihood of something being a fork given that it is for eating, P(*fork* | *for-eating*), but such associations should be less predictive of the acceptability of corresponding teleological generalizations, e.g., *forks are for eating*, than statements diagnostic of principled connections. Hence, the theory makes the following additional prediction:

> **Prediction 2:** Participants' estimates of relevant conditional probabilities should be less predictive of their tendency to accept teleological generalizations as compared to their endorsements of sentences diagnostic of principled connections.

A regression analysis tested prediction 2.

Experiment 1

Experiment 1 tested whether participants distinguish teleological generalizations that potentially represent principled connections from those that do not (prediction 1). Participants' task was to assess the truth of teleological generalizations such as *forks are for eating* and *forks are for* *washing*. Pairs of items were constructed such that each noun (*forks*) appeared with a potentially acceptable teleological generalization (*eating*) or a potentially unacceptable one (*washing*). Participants also assessed the same pair of items as self-referential generalizations (e.g., *forks, by virtue of being forks, are for eating* vs. *forks, by virtue of being forks, are for washing*).

Method

Participants. 40 participants (20 female) completed the task on Amazon Mechanical Turk. All but four participants had completed one or fewer courses in logic.

Materials. Each material consisted of an artifact or natural kind concept (a noun) paired with either an experimental or a control verb. Verbs described common actions performed with or by the noun, e.g., the word "forks" was paired with the verb "eating" or else "washing." Experimental and control items differed in that experimental items' verbs were constructed to yield acceptable teleological generalizations (e.g., "forks are for eating") while control items' verbs were constructed to yield unacceptable generalizations (e.g., "forks are for washing"). Half of the 22 objects were artifacts (e.g., "forks"), while the other were half were natural kinds (e.g., "stomachs"). Table 1 provides a list of 10 of the 22 objects and their corresponding verbs.

Design and procedure. Participants were instructed to evaluate the truth of statements about common objects and entities. They responded to one of two types of assertions: teleological generalizations or else self-referential generalizations. Teleological generalizations were assertions that were of the form $NP_{plural} + VP_{purposive}$, e.g., "forks" + "are for eating". Self-referential generalizations described the purpose of plural form of the given object by virtue of being that object, e.g., "forks, by virtue of being forks, are for washing." They registered their responses by moving a slider handle on a Likert scale that ranged from -3 (definitely false) to 3 (definitely true). The study implemented a design such that participants served as their own controls, i.e., 22 distinct objects x 2 types of verb (control vs. experimental) x 2 types of generalization (teleological vs. self-referential). Hence, participants assessed 88 assertions in total. Experiment 1 and all subsequent experiments were implemented in the "nodus-ponens" experimental framework (Khemlani, 2017). The study presented the items in a randomized order.

Open science. Data, code, and complete materials for Experiment 1 and subsequent experiments can be found at: https://osf.io/8v9ws.

Results and discussion

Figure 1 (left two columns) shows the results of participants' ratings for the items in Experiment 1. Their evaluations corroborated principle 1: they rated experimental items as more truthful than control items ($M_{experimental}$ = 2.42 vs. $M_{control}$ = -.71, Wilcoxon test, z = 33.93, p <.0001, Cliff's δ = 0.82). Their ratings did not differ as a function of whether the generalization was teleological or self-referential ($M_{teleological}$ = .86 vs. $M_{self\text{-}referential}$ = .85, Wilcoxon test, z = 0.09, p = .93, Cliff's δ < 0.01). And, their ratings did not yield a reliable interaction between the type of verb (experimental vs. control) and the type of generalization (teleological vs. self-referential; Wilcoxon test, z =.52, p = .60, Cliff's δ = 0.02).

Planned comparisons were conducted for experimental vs. control items for teleological generalizations and self-referential generalizations in isolation. For teleological generalizations, assertions describing experimental items were rated as more true (M = 2.43) than those describing control items (M = -0.72; Wilcoxon test, z = 23.95, p <.0001, Cliff's δ = 0.83). The result served as a manipulation check: it confirmed that participants construed experimental items as acceptable generalizations. A similar pattern held for self-referential generalizations ($M_{experimental}$ = 2.41 vs. $M_{control}$ = -.70; Wilcoxon test, z = 24.03, p <.0001, Cliff's δ = 0.82). Participants' ratings yielded a strong correlation between the teleological and self-referential generalizations, which confirmed prediction 1 (r = .65, p < .0001). Hence, participants' evaluations of the truth of self-referential generalizations strongly predicted their evaluations of teleological generalizations. Prasada et al. (2013) posit that self-referential generalizations are diagnostic of principled connections, and so the results of Experiment 1 suggest that participants represented principled connections for experimental items. Experiment 2 sought to extend the finding by exploring two additional assertion types diagnostic of principled connections.

		Verb	
	Concept	**Experimental**	**Control**
1	bag	carrying	storing
2	book	reading	packing
3	brain	thinking	sleeping
4	car	driving	painting
5	chair	sitting	dusting
6	cup	drinking	stacking
7	ear	hearing	plugging
8	eye	seeing	blinking
9	fork	eating	washing
10	hand	grasping	clapping
...	…	…	…

Table 1. A sample of the concepts and their corresponding verbs in Experiments 1 and 2. Each concept appeared as a plural noun (e.g., "bags…") and was paired with either an experimental or a control verb (e.g., "…are for carrying.") Participants received 22 concepts in both nouns, i.e., 44 items for each formulation.

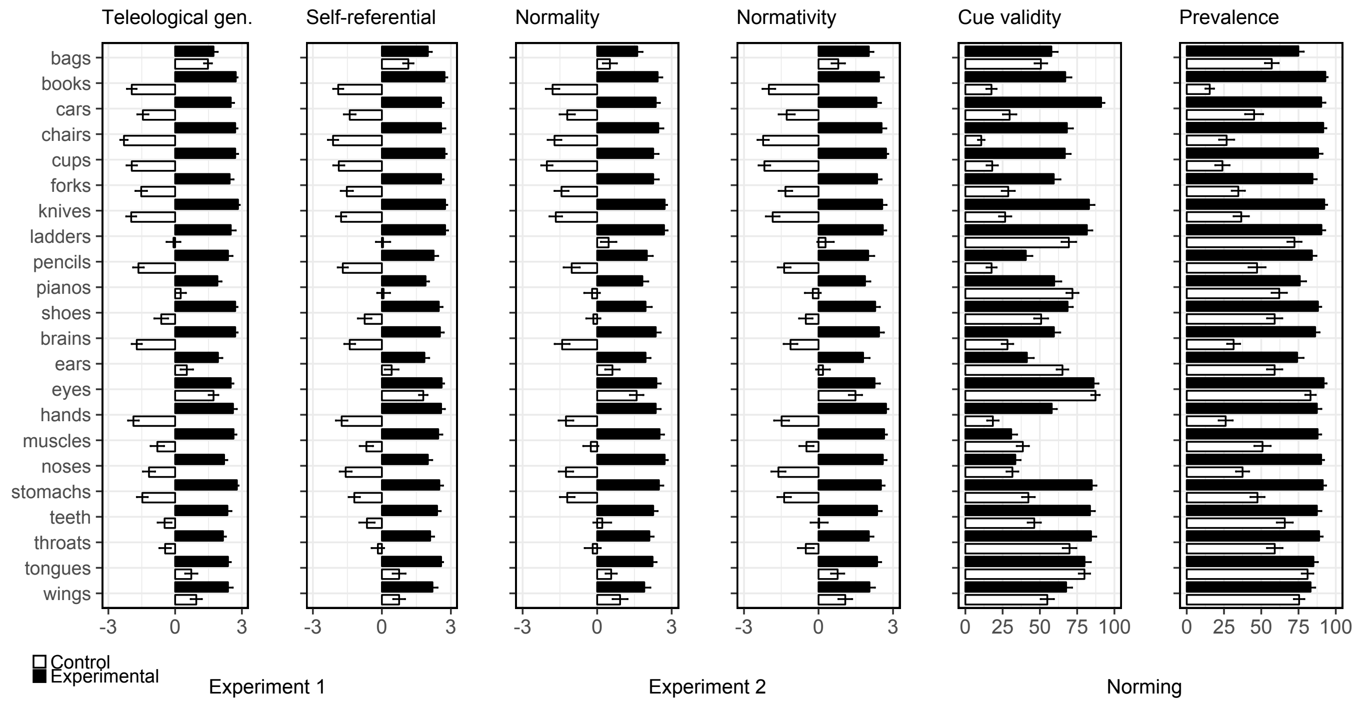

Figure 1. Participants' truth ratings on 1) teleological and self-referential generalizations (Left two panels; Experiment 1) and 2) assertions concerning normality and normativity (Middle two panels; Experiment 2). Right two panels: participants' normed conditional probability estimates (cue validity and prevalence). All panels show ratings as a function of the 22 concepts (top 11: artifacts; bottom 11: natural kinds) and the two types of teleological property (control vs. experimental).

Experiment 2

Experiment 2 was identical to Experiment 1 in all respects except that instead of teleological and self-referential generalizations, Experiment 2 provided participants with statements about normality (e.g., "all normal forks are for eating") and normativity (e.g., "forks are supposed to be for eating.") Such assertions about normality and normativity are diagnostic of principled connections (see prediction 1).

Method

Participants. 39 participants (25 female) completed the task on Amazon Mechanical Turk. All but three participants had completed one or fewer courses in logic.

Materials, design, and procedure. Materials for Experiment 2 were identical to those used in Experiment 1. The design of the experiment was also identical to Experiment 1, except that in lieu of teleological and self-referential assertions, participants received two assertions that referenced expectations of what was normal of the objects in the study (see Table 1) or else a normative property of those objects. Hence, assertions that referenced normality appeared as follows: "all normal forks are for eating." And assertions that referenced normativity appeared as follows: "forks are supposed to be for eating." As before, participants rated the extent to which each statement struck them as true on a scale that ranged from 3 to -3. They served as their own controls and rated the control and experimental items for each of the 22 objects on each of the 2 types of formulation, for a total of 88 items.

Results and discussion

Figure 1 (middle two columns) shows the results of participants' ratings for the items in Experiment 2. Participants' responses were analogous to those in Experiment 1. They rated experimental items higher than control items ($M_{\text{experimental}}$ = 2.30 vs. M_{control} = -.61, Wilcoxon test, z = 31.97, p < .0001, Cliff's δ = 0.75). Their ratings did not differ for normality vs. normativity generalizations ($M_{\text{normality}}$ = 0.86 vs. $M_{\text{normativity}}$ = .0.83, Wilcoxon test, z = 0.61, p =.54, Cliff's δ < 0.01). Their ratings yielded a small but detectable interaction between the type of verb (experimental vs. control) and the type of assertion (normality vs. normativity; Wilcoxon test, z = 2.11, p = .04, Cliff's δ = 0.06). This small effect was indicative of participants' slightly more extreme ratings for normativity assertions than for normality assertions.

Planned comparisons revealed that participants provided higher ratings for experimental items than control items, both for assertions that concerned normality ($M_{\text{experimental}}$ = 2.26 vs. M_{control} = -0.54, Wilcoxon test, z = 22.36, p <.0001, Cliff's δ = 0.73) and for those that concerned normativity ($M_{\text{experimental}}$ = 2.34 vs. M_{control} = -0.68, Wilcoxon test, z = 22.82, p <.0001, Cliff's δ = 0.77). The results further corroborated prediction 1, which states that people should distinguish principled connections from other kinds of connections based on the fact that principled connections yield expectations of both normality and normativity.

Can some mitigating factor explain why participants distinguished experimental from control items in Experiments 1 and 2? One clear alternative is that participants maintained probabilistic beliefs, e.g., they may have maintained the subjective belief that the probability is high that a thing is for eating given that it's a fork, i.e., P(*for-eating* | *fork*), or else they may have endorsed the idea that the probability that something is a fork is high given that it's for eating, i.e., P(*fork* | *for-eating*). Either of these two conditional probabilities could be correlated with participants' tendency to endorse teleological and self-referential generalizations (Experiment 1) or else assertions about normality and normativity (Experiment 2). We conducted a regression analysis to test whether probabilistic beliefs rather than principled connections best predict the acceptance of teleological generalizations.

Regression analysis

Prediction 2 above states that conditional probability estimates, e.g., P(*for-eating* | *fork*) or P(*fork* | *for-eating*) should be less predictive of participants' acceptance of teleological generalizations than statements diagnostic of principled connections. Yet, reasoners may possess relevant statistical knowledge, and indeed, as Prasada et al. (2013) observe, certain generalizations are acceptable precisely because they make statistical claims. For instance, the generalization, *cars have radios*, is true because most cars have radios. The fact may be an accident of history – there is nothing abnormal about a car without a radio – yet the generalization remains perfectly acceptable. Hence, an alternative account of the results from the previous experiments is that people endorsed assertions that are diagnostic of principled connections, not because they directly represented principled connections, but because of their underlying statistical and probabilistic knowledge. In other words, people may endorse *forks are for eating* as a true generalization not because being *for eating* bears a principled connection to *fork*, but because forks are, more often than anything else, used for eating.

To address this alternative explanation, we conducted a norming study and used it to carry out a regression analysis. If prediction 2 is true, then regression models comprised of only conditional probability estimates should fare worse at predicting the acceptability of teleological generalizations than models that include the various diagnostic assertions as predictors. If prediction 2 is false, the opposite pattern should hold.

Norming study

In a norming study on a new sample of participants through Amazon Mechanical Turk, we collected two kinds of conditional probability estimates: cue validity and prevalence (see Khemlani et al., 2012, for an analogous analysis on bare plural generalizations). Cue validity refers to the probability that an instance belongs to a kind given that it has a particular property; it can be construed as the conditional probability, P(*fork* | *for-eating*). For the relevant *fork* item, participants evaluated cue validity by answering the following question: "Suppose a particular thing is for eating. What is the probability that that thing is a fork?" In contrast, prevalence estimates refer to the probability that an object, given that it belongs to a particular kind, has a particular property. Participants generated prevalence estimates by answering the following question: "What percentage of forks are for eating?" Hence, prevalence can be construed as the conditional probability, P(*for-eating* | *fork*). Participants in the norming study received each of the 22 materials in a 2 (experimental vs. control verb) x 2 (cue validity vs. prevalence) design. For each question, they provided probability estimates on a movable slider ranging from 0 to 100 percent. As in Experiments 1 and 2, participants served as their own controls, and each received a total of 88 items in different randomized order.

Hierarchical analysis

Experiments 1, 2, and the norming study yielded numerical estimates of the truth values of teleological generalizations, self-referential generalizations, normality assertions, and normativity assertions, as well as relevant conditional probability estimates, i.e., cue validity and prevalence estimates. We conducted a regression analysis at the item level, for which we aggregated the data from the three studies and averaged them as a function of the 44 separate items (22 items x 2 verbs: experimental vs. control).

To conduct the analysis, we constructed a series of linear mixed-effects models to test both predictions 1 and 2. Prediction 1 predicts that assertions diagnostic of principled connections (i.e., self-referential generalizations, normality assertions, and normativity assertions) should significantly predict the acceptance of teleological generalizations. Prediction 2 predicts that conditional probability estimates (i.e., cue validity and prevalence) should be less predictive of acceptance of teleological generalizations.

An initial hierarchical analysis established a set of mixed-effects models as follows (where M_n is an abbreviation for "model *n*"):

M_1: `teleology ~ cue validity`
M_2: `teleology ~ cue validity` **`+ prevalence`**
M_{3A}: `teleology ~ cue validity + prevalence` **`+ self-referen.`**

Hence, M_1 describes a model in which mean ratings of teleological generalizations were regressed against mean cue validity estimates, and so on. Each model controlled for

	AIC	Deviance	R^2	χ^2	Significance
M_1	151.13	143.13	.53	–	–
M_2	83.23	73.23	.90	69.90	* (M_1)
M_{3A}	-24.10	-36.10	.99	109.33	* (M_2)
M_{3B}	34.29	22.28	.97	50.94	* (M_2)
M_{3C}	16.26	4.26	.98	68.97	* (M_2)

Table 2. Analysis of deviance for Models 1 and 2 (including only both statistical predictors) and Models 3A-3C (each of which included a single predictor diagnostic of principled connections in addition to both statistical predictors). The * denotes a significantly better fit to the data than the model in parentheses.

variation in individual materials by including it as a random effect (not shown in the formulas above). Table 2 provides an analysis of deviance for the three separate models. The table shows that the data corroborate prediction 2: models comprised of only conditional probability estimates (M_1 and M_2) were less predictive of participants' mean ratings of teleological generalizations ($R^2_{M1} = 0.53$ and $R^2_{M2} = 0.90$) and performed significantly worse than the model that included participants' mean ratings of self-referential generalizations ($R^2_{M3A} = 0.99$).

Two additional regression models were constructed as follows:

M_{3B}: `teleology ~ cue validity + prevalence` **`+ normality`**
M_{3C}: `teleology ~ cue validity + prevalence` **`+ normativity`**

They were subjected to analogous hierarchical analyses against those models that included only conditional probability estimates, M_1 and M_2, and they revealed analogous patterns (see Table 2). In general, when considered alongside conditional probability estimates, diagnostic assertions explained all of the variance in the acceptability of teleological generalizations.

General discussion

In two studies, participants evaluated teleological generalizations, e.g., *forks are for eating*, as well as assertions diagnostic of principled connections between a concept and a property, e.g., *forks, by virtue of being forks, are for eating* and *all normal forks are for eating*. They gave higher truth ratings for these statements when the corresponding teleological generalization was acceptable (*e.g., forks are for eating)* than when it was unacceptable (*e.g., forks are for washing)*. A norming study and a regression analysis revealed that assertions diagnostic of principle connections fully predicted peoples' tendency to endorse teleological generalizations. In contrast, estimates of conditional probabilities that related concepts and properties were less predictive. The data validated a taxonomy proposed by Prasada et al. (2013), and they suggest that people represent principled connections between concepts (*forks*) and teleological properties (*being for eating*). In general, the taxonomy explains why people accept certain teleological generalizations and reject others.

While recent research suggests that reasoners often invoke causal relations when interpreting teleological explanations (Lombrozo & Carey, 2006), causal relations do not provide a complete account for which teleological explanations people accept. The authors explored an additional constraint: the generality of the relevant causal relations may predict people's acceptance of teleological explanations. The present analysis parsimoniously explains why generality is important for teleological reasoning: if people represent privileged connections between concepts and their functions, they should be able to generate teleological inferences based on those connections. For instance, if you are told about a *particular* cup, you may be inclined to inductively infer its primary purpose (it's for drinking). Our ongoing work will explore whether Prasada et al.'s conceptual taxonomy accounts for the acceptability of inductive teleological inferences.

Acknowledgments

This work was supported by an NRC Research Associateship Award to JK and funding from the Office of Naval Research to SK. We thank Tony Harrison, Laura Hiatt, Zach Horne, Deb Keleman, Bertram Malle, Janani Prabhakar, and Greg Trafton. We also thank Kalyan Gupta, Kevin Zish, and Knexus Research Corporation.

References

Atran, S. (1995). Causal constraints on categories. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multi-disciplinary debate*. Oxford, UK: Clarendon Press.

Bigelow, J., & Pargetter, R. (1987). Functions. *Journal of Philosophy*, 84, 181-197.

Brown, R. (1952). Dispositional and teleological statements. *Philosophical Studies*, 3, 73-80.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

Choi, S., & Fara, M. (2016). Dispositions. In E. Zalta, *The Stanford Encyclopedia of Philosophy*.

Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science,* 1, 255-259.

Delancey, C. (2006). Ontology and teleofunctions: A defense and revision of the systematic account of teleological explanation. *Synthese*, *150*, 69-98.

Keil, F.C. (1992). The origins of an autonomous biology. In M.R. Gunnar and M. Maratsos (Eds.), *Modularity and constraints in language and cognition* (Vol. 25). Hillsdale, NJ: Erlbaum.

Kelemen, D. (1999). The scope of teleological thinking in preschool children. *Cognition,* 70, 241-272.

Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *JEP: General*, 142, 1074-1083.

Khemlani, S., Leslie, S.J. & Glucksberg, S. (2012). Inferences about members of kinds: The generics hypothesis. *Journal of Language and Cognitive Processes, 27*, 887-900.

Kroll, N. (2017). Teleological dispositions. *Oxford Studies in Metaphysics*, 10, 3-37.

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.

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition*, 99, 167-204.

Mitchell, S. (1993). Dispositions or etiologies? A comment on Bigelow and Pargetter. *The Journal of Philosophy*, 90, 249-259.

Neander, K. (1991). The teleological notion of 'function'. *Australasian Journal of Philosophy*, 69, 454-468.

Onishi, K.H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255-258.

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

Southgate, V., Johnson, M.H., & Csibra, G. (2008). Infants attribute goals to even biologically impossible actions. *Cognition*, 107.

Wright, L. (1976). *Teleological explanations*. Berkeley, CA: University of California Press.