Goals Affect the Perceived Quality of Explanations

Nadya Vasilyeva (Vasilyeva@Berkeley.Edu)
Daniel Wilkenfeld (Daniel.Wilkenfeld@Berkeley.Edu)
Tania Lombrozo (Lombrozo@Berkeley.Edu)
Department of Psychology, 3210 Tolman Hall
Berkeley, CA 94720 USA

Abstract

Do people evaluate the quality of explanations differently depending on their goals? In particular, are explanations of different kinds (formal, mechanistic, teleological) judged differently depending on the future judgments the evaluator anticipates making? We report two studies demonstrating that the perceived "goodness" of explanations depends on the evaluator's current goals, with explanations receiving a relative boost when they are based on relationships that support anticipated judgments. These findings shed light on the functions of explanation and support pragmatic and pluralist approaches to explanation.

Keywords: explanation, inference, goals, context, pragmatic factors

Do people evaluate the quality of explanations differently depending on their goals? Suppose, for instance, that Ana and Bob are both interested in marsupials. Ana is studying marsupials because she hopes to diagnose their ailments; Bob is interested because he hopes to understand their biological adaptations. When it comes to explaining why kangaroos have tails, will Ana find mechanistic explanations (for instance, in terms of development or genes) more compelling than Bob? Will Bob find teleological explanations (for instance, that appeal to balance) more compelling than Ana? What if their goals are more transient and context-specific?

On the one hand, accounts of explanation from psychology suggest that judgments of explanation quality should track a person's goals. Lombrozo and Carey (2006), for instance, suggest that one function of explanation is to support future reasoning and behavior – including novel inferences – by highlighting generalizable or "exportable" relationships (see also Craik, 1943; Heider, 1958). Given that different kinds of inferences are differentially useful in the context of different goals, one might expect judgments about the quality of explanations (in a given context) to be similarly sensitive to goals (see also Leake, 1995, for a relevant discussion). More broadly, there's increasing support for the idea that (many) mental representations are sensitive to context and goals (e.g., Barsalou, 1983; Markman & Ross, 2003), raising the possibility that constraints on explanations may be similarly flexible.

On the other hand, mainstream accounts of explanation from philosophy have often set pragmatic and contextual considerations to the side, focusing instead on a specification of formal relationships or features that are constitutive of explanations, such as deductive arguments of a particular form (Hempel & Oppenheim, 1948) or causal processes that generate an effect (Salmon, 1984). On these views, pragmatic factors have a limited influence, perhaps in what one chooses to explain or in the level at which an explanation is pitched.

Importantly, however, another family of accounts of explanation within philosophy, known as pragmatic accounts, allow for context effects not only in what is explained, but also in what counts as a (good) explanation. For example, van Fraassen (1980) proposes that context fixes the contrast class – that is, the implicit set of possible alternatives to the target observation that the explanation needs to account for – and narrows down the range of relevance relationships that count as explanatory in that context. Such proposals raise the possibility that different contexts call for different *kinds* of explanations to account for one and the same observation. More concretely: Ana might be right, given her context and goals, to favor a mechanistic explanation for the kangaroo's tail, and Bob might be right, given his context and goals, to favor a teleological explanation.

Here we investigate whether a person's goals have an impact on evaluations of explanation quality, and in particular, whether people evaluate different kinds of explanations differently depending on the kinds of judgments that they anticipate making. We report two experiments in which we experimentally manipulate participants' goals (i.e., the kinds of judgments that they anticipate making) and have them evaluate explanations of different kinds: formal (which appeal to category membership; see Prasada & Dillingham, 2009), mechanistic (which appeal to proximate causes), and teleological (which appeal to goals or functions). In particular, the goals we specify call for generalizations based on the relations that underwrite each type of explanation – that is, between a property and category membership (formal), its proximate causes (mechanistic), or its function (teleological). If judgments of explanation quality are sensitive to contextually-defined goals, then explanations should receive higher ratings under congruent than incongruent goals.

Experiment 1

Participants evaluated formal, mechanistic, and teleological explanations in the context of one of three goals: category-, cause- or function-based generalization. As additional reference points, we included uninformative circular explanations and included a condition in which participants evaluated explanations in the absence of any explicitly-specified goal. We predicted that ratings of explanation

Table 1. Sample instructions, explanations, and goal-reinforces used in Experiment 1. Text referring to artifacts appears in square brackets. Goal was manipulated between subjects; explanation type and domain were manipulated within subjects.

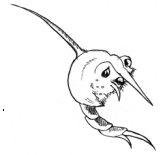
1. Generalization instructions (Goal manipulation)

(After evaluating each explanation) you will make a *prediction* about a new object or organism:

Category-based goal: The new organism [object] will *either be of the same species [kind]* as each original that you read about or *of a different species [kind]*.	**Cause-based goal:** The new organism [object] will *either have the same internal characteristics [parts] and processes* as each original that you read about or *have different internal characteristics [parts] and processes.*	**Function-based goal:** The new organism [object] will *either have the same needs [purpose]* as each original that you read about or *have different needs [purpose]*.

Based on this information you'll have to guess whether the organism has the same properties as the original or different properties.

2. Description



Glenta are microorganisms in the ocean. Their motion is controlled by a set of light-seeking photoreceptors, which makes them rise towards the ocean's surface during the day. Spending some time at the ocean's surface helps them replenish their oxygen reserves.

3. Explanation evaluation (each participant rates one explanation per item)

Below is a picture of one particular specimen, ID-Zd89u0002, from a research facility.

Why does this specimen *rise to the ocean's surface during the day?*

Formal explanation: Because it's a glenta, and glentas rise to the ocean's surface during the day.	**Mechanistic explanation:** Because its motion is controlled by a set of light-seeking photoreceptors, which makes it rise to the ocean surface during the day.	**Teleological explanation:** Because rising to the ocean surface during the day helps it replenish oxygen reserves.	**Circular explanation:** Because some things can rise to the ocean's surface.

Very bad explanation (1) – Very good explanation (9)

4. Goal reinforcer

Behind this box there is a microorganism. Click HERE to find out…

Category-based goal: ...if it's a glenta. *(Text appears on click)* Yes, it's a glenta *or* No, it's NOT a glenta.	**Cause-based goal:** ...if its motion is controlled by a set of light-seeking photoreceptors. *(Text appears on click*) Yes, its motion is controlled by a set of light-seeking photoreceptors *or* No, it its motion is NOT controlled by a set of light-seeking photoreceptors.	**Function-based goal:** ...if it needs to replenish oxygen reserves. *(Text appears on click*) Yes, it needs to replenish oxygen reserves *or* No, it does NOT need to replenish oxygen reserves.

Do you think it rises to the ocean during the day? Definitely no (1) - Definitely yes (9)

"goodness" would be affected by goals, with a boost for goal-congruent explanations.

Method

Participants Four-hundred-and-twelve participants were recruited on Amazon MTurk in exchange for $1.65; an additional 95 participants were excluded for failing a memory check. In both experiments, participation was restricted to workers with an IP address within the United States and with a HIT approval rating of 95% or higher from at least 50 previous HITs.

Materials, Design and Procedure Participants were presented with descriptions of 16 fictional living things and artifacts[1], each described with a label and three features organized into a causal chain (see Table 1). For each entity, participants evaluated one of four possible explanations for the middle feature in the causal chain (formal, mechanistic, teleological, or circular) using a 9-point scale anchored at "very bad explanation" (1) and "very good explanation" (9). Each participant evaluated four explanations of each type, with item-explanation pairings counterbalanced across participants.

Crucially, participants rated explanations under one of four goal conditions: category-based, cause-based, function-based, or no goal. In each goal-based condition, participants were informed that after evaluating explanations (as illustrated with two training trials), they would be making predictions about new objects and organisms, where the predictions involved categories, causes, or functions and served as the manipulation of goals.

Participants completed 16 alternating trials: explanation-evaluation, in which they rated the quality of an explanation, and (for participants in one of the three goal-based conditions) goal reinforcers, in which they were given information about an entity behind a black box and had to rate how likely it was that the target feature of the original item generalized to the obscured item. The information provided varied across goals: participants were told whether the obscured entity belonged to the same category as the original (category-based), whether it shared the same cause feature (cause-based), or whether it shared the same function (function-based).[2] Importantly, as shown in Table 1, for a given item the explanation evaluation always preceded the goal reinforcer.

[1] Domain was not a variable of central theoretical interest, and it did not interact with the effect of goal in either experiment; due to space limitations, we omit analyses of domain.

[2] The main purpose of this task was to maintain the goal focus. The data are not reported due to space limitations.

Table 2. Mean explanation goodness ratings as a function of explanation type and goal in Exp. 1 and 2 (*SD* in brackets).

	Goal	Formal	Mechanistic	Teleological	Circular	All
Expt. 1	Category-based	4.03 (*1.95*)	7.46 (*1.41*)	7.27 (*1.52*)	2.04 (*1.31*)	5.19 (*0.96*)
	Cause-based	3.64 (*1.74*)	7.44 (*1.41*)	6.57 (*1.73*)	1.85 (*1.05*)	4.88 (*0.80*)
	Function-based	3.52 (*1.91*)	6.64 (*1.75*)	7.49 (*1.40*)	1.97 (*1.27*)	4.91 (*0.94*)
	No goal	3.04 (*1.47*)	6.76 (*1.55*)	7.22 (*1.43*)	1.63 (*0.98*)	4.66 (*0.73*)
	All goal conditions	3.56 (*2.43*)	7.07 (*1.58*)	7.15 (*1.55*)	1.88 (*1.17*)	
Expt. 2	Categorize	3.88 (*2.43*)	7.38 (*1.69*)	7.23 (*1.70*)	2.08 (*1.62*)	5.14 (*1.07*)
	Identify causal origin	3.8 (*2.39*)	7.29 (*1.88*)	6.90 (*2.07*)	2.13 (*1.51*)	5.03 (*1.13*)
	Identify function	3.43 (*2.04*)	6.19 (*2.33*)	7.52 (*1.78*)	1.99 (*1.44*)	4.78 (*1.13*)
	No goal	3.42 (*2.05*)	6.95 (*2.04*)	7.03 (*1.99*)	2.06 (*1.27*)	4.87 (*1.03*)
	All goal conditions	3.63 (*2.23*)	6.95 (*2.08*)	7.17 (*1.90*)	2.07 (*1.46*)	

Results and Discussion

Explanation ratings were analyzed in an ANOVA with explanation type as a within-subjects factor and goal as a between-subject factor. This revealed significant main effects of both explanation type, $F(3,1224)=1365.60$, $p<.001$, $\eta_p^2=.770$, and goal, $F(3,408)=6.81$, $p<.001$, $\eta_p^2=.048$. Overall, participants preferred mechanistic and teleological explanations over formal explanations, all of which were preferred over circular explanations, all p's$<.001$ (see Table 2). Causal and teleological ratings did not differ, $t(411)=.63$, $p=.531$. Ratings were also higher under the categorical goal than the causal goal (Tukey HSD $p=.039$) and no goal ($p<.001$) conditions.

Most importantly, we found a significant interaction between explanation type and goal, $F(9,1224)=5.73$, $p<.001$, $\eta_p^2=.040$. We analyzed the interaction with a series of planned contrasts motivated by our prediction that explanation ratings would be higher in the context of a congruent goal. Three separate contrasts compared ratings of formal, mechanistic, and teleological explanations in the context of the congruent goal versus the average of ratings for that explanation type in the other three goal conditions. As predicted, each of the explanation types was rated significantly better under the congruent goal compared to the rest of the goal conditions: formal $F(1,408)=9.85$, $p=.002$, $\eta_p^2=.024$; mechanistic $F(1,408)=7.36$, $p=.007$, $\eta_p^2=.018$; teleological $F(1,408)=7.23$, $p=.006$, $\eta_p^2=.019$ (see Figure 1a). Circular explanations were not significantly influenced by goals, as revealed by a one-way ANOVA, $F(3,408)=2.48$, $p=.061$.

As a further test of the relationship between explanatory preferences and goals, we classified participants based on the explanation type for which they gave the highest average ratings. Twenty ties (18 between causal and teleological explanations) were excluded. As shown in Figure 2a, the distribution of explanation preferences varied significantly across goals, $\chi^2(9,N=392)=31.87$, $p<.001$. Based on examination of standardized residuals, the effect was driven by participants being more likely to favor mechanistic and, marginally, teleological explanations under the corresponding congruent goals (standardized residuals 2.5, 1.9), and less likely to favor these explanations under incongruent goals (standardized residuals -2.3, -2.4). This suggestive pattern of competition between cause- and function-based reasoning was additionally supported by a negative correlation between ratings of mechanistic and teleological explanations, $r(410)=-.19$, $p<.001$. No other pair of explanation ratings was significantly negatively correlated (ps $> .05$).

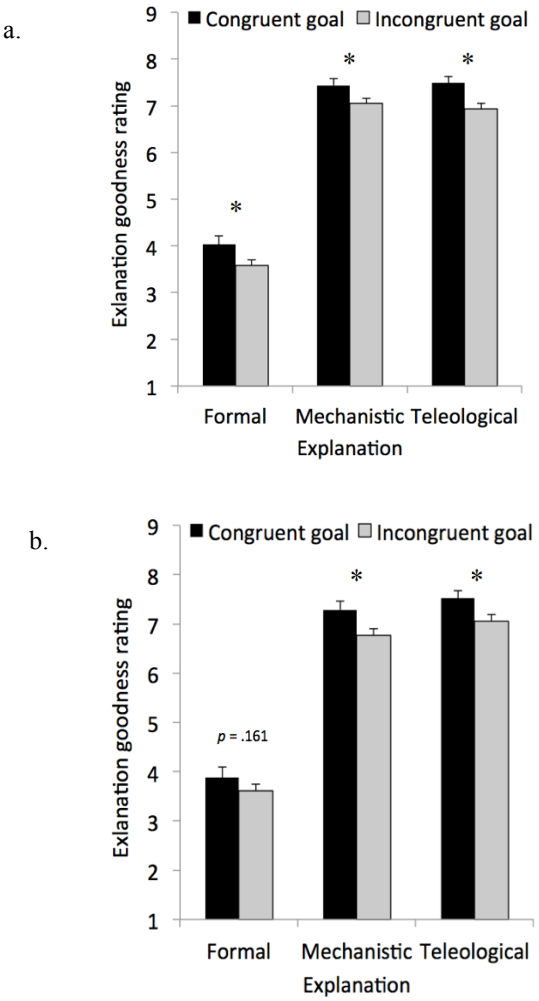


Figure 1: Explanation goodness ratings as a function of explanation type and goal in Experiments 1 (a) and 2 (b); error bars represent 1 SEM; stars indicate contrasts significant at < .05.

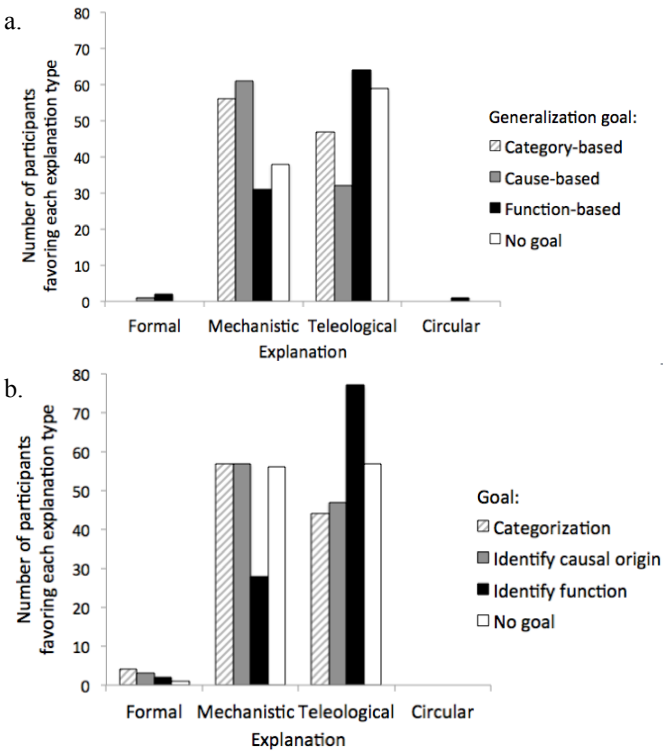


Figure 2: Explanation preferences as a function of goals in Experiments 1 (a) and 2 (b).

In sum, we find that the kinds of inferences that one anticipates making influence the perceived quality of different kinds of explanations. Statements that explained an observation in terms of category membership, in terms of proximal causal mechanisms, or in terms of goals and purposes were perceived as better explanations in the context of goals that called for the information provided by these explanations.

Experiment 2

Experiment 2 had two main objectives: to replicate the interaction between goals and explanation types from Experiment 1 with a different set of goals, and to rule out the possibility that the effect of goals was due to changes in the implied contrast class of the questions. To address the first objective, we introduced a different goal manipulation: participants were given a task as an assistant to the director of a museum, where the task involved classification (grouping items), proximate causes (identifying how something came about), or functions (identifying functions). To address the second objective, we added a clarification to the explanation requests specifying the contrast class.

Method

Participants Four-hundred-and-ninety-six participants were recruited on Amazon MTurk in exchange for $1.65. An additional 317 participants were excluded for failing a memory check.

Materials, Design and Procedure The materials, design and procedure were the same as in Experiment 1 with the following exceptions. First, we introduced a cover story that the participant was a museum assistant, and participants were told that they would need to figure out one of three things: how new objects or organisms should be grouped in the museum (categorization goal), how it is that objects or organisms come to possess certain properties (causal origin goal), or what functions the properties of objects or organisms serve (functional goal). The goal reinforcers were adapted accordingly (Table 3 illustrates all changes). As in Experiment 1, the no goal condition served as a baseline. Second, to rule out the possibility that effects of goals on explanation judgments were produced (only) by a shift in the implied contrast class of the questions, we added a clarification to the explanation probes specifying the contrast class, for instance: "Why does this specimen *rise to the ocean's surface during the day*? (as opposed to not rising to the ocean's surface during the day)." Finally, domain was manipulated between subjects.[3]

Results and Discussion

Explanation ratings were analyzed in an ANOVA with explanation type as a within-subjects factor and goal as a between-subjects factor. The main effect of explanation type was replicated, $F(3,1476)=946.06$, $p<.001$, $\eta_p^2=.658$: as shown in Table 2, participants preferred mechanistic and teleological explanations over formal explanations, which were all preferred over circular explanations, $ps<.001$. Mechanistic and teleological explanation ratings did not differ, $t(495)=1.64$, $p=.102$. The goal manipulation also produced a significant main effect, $F(3,492)=2.71$, $p=.004$, $\eta_p^2=.016$, driven by higher ratings under the categorical goal than the teleological goal (Tukey HSD $p=.046$, all remaining p's$\geq.190$).

Most importantly, there was a significant interaction, $F(9,1476)=4.00$, $p<.001$, $\eta_p^2=.024$ (see Figure 1b). The planned contrasts showed that mechanistic and teleological explanations were rated significantly higher under the congruent goal compared to the rest of the goal conditions: mechanistic $F(1,492)=4.49$, $p=.035$, $\eta_p^2=.009$; teleological $F(1,492)=5.59$, $p=.018$, $\eta_p^2=.011$; however, the contrast did not reach significance for formal explanations, $F(1,492)=1.97$, $p=.161$. Circular explanations were not influenced by goals, as revealed by a one-way ANOVA $F(3,492)=.20$, $p=.898$.

[3] Experiment 2 ended with an additional exploratory task that examined whether the effect of goals extends to judgments of an explanation's *probability* in addition to its quality, as might be anticipated if an explanation's "loveliness" is used as a cue to its "likeliness" (Lipton, 2004). At the end of the study participants were shown 16 additional living things and artifacts, each described by one feature, and asked the to evaluate the probability of a formal, mechanistic, teleological, or circular explanation for that feature. We found no evidence of a goal effect on evaluations of explanation probability, $F(9,1440)=1.04$, $p=.407$, $\eta_p^2=.006$. However, given that this task occurred at the end of the experiment, it is possible that the effects of the goal manipulation were too weak; we therefore hesitate to draw conclusions from this null result.

Table 3. Sample instructions and goal-reinforces used in Experiment 2. Text referring to artifacts appears in square brackets.

Instructions (Goal manipulation): In this experiment, you will be the assistant to a museum director. The museum will present the public with little-known organisms and objects. Your job will be to figure out…

Goal: Categorization:…how these organisms [objects] should be *grouped* with others in the museum. For example, zoos often group animals of the same kind together [stores often put objects of the same kind next to each other]. Your job will be to figure out how the organisms [objects] should be organized in the museum.

Goal: Identify causal origin:…*how it is* that organisms [objects] come to have certain traits. For example, botanists and zoologists often try to figure out [when engineers encounter novel objects, they may need to "reverse-engineer" them to figure out] what produces some characteristic that they observed in a plant or animal [that object]. Your job at the museum will be to figure out what produces features of living organisms [objects]: how they do certain things, or come to have certain characteristics.

Goal: Identify function: … what the biological traits (such as parts or behaviors) of each organism *are for* [what each object (or some feature of an object) *is for*]. For example, biologists often identify the functions of the biological traits of animals or plants that they are studying [archaeologists often identify the functions of the objects that they find]. Your job at the museum will be to identify the functions of exhibited organisms [objects].

Goal reinforcer: Now you receive two completely new animal specimens [objects]: specimen [item] A and specimen [item] B. Each one may or may not be a glenta. Both of them rise to the ocean's surface during the day.

Specimen	A	B
Rises to the ocean's surface during the day?	Yes	Yes

Goal: Categorization: Do you think specimens [items] A and B both belong in the *same part of the museum*?

Goal: Identify causal origin: Do you think the *same factor* produces this characteristic in both specimen [item] A and specimen [item] B?

Goal: Identify function: Do you think this characteristic serves the *same function* for specimen [in item] A and specimen [item] B?

Definitely no (1) - Definitely yes (9)

As in Experiment 1, we also found that the distribution of explanation preferences varied significantly as a function of goal, $\chi^2(6, N=433)=26.19$, $p<.001$. (This analysis excluded 60 ties, 50 of which were between causal and teleological and evenly spread across conditions.) As shown in Figure 2b, the effect was driven by the functional goal condition, where fewer participants preferred mechanistic explanations and more participants preferred teleological explanations (standardized residuals -3.0 and 2.9). Under the causal goal, the differences were in the predicted direction but did not reach significance (standardized residuals 1.2, – 1.2). Once more, ratings of mechanistic and teleological explanations were negatively correlated, $r(494)= -.19$, $p<.001$. No other pair of explanation ratings was significantly negatively correlated ($ps > .05$).

Taken together, these results indicate a pattern of interaction between goals and explanation type similar to that observed in Experiment 1.

General Discussion

Across two studies involving different manipulations of goals, we found that the judgments a person anticipates making influence the perceived quality of explanations of different kinds. In particular, we found evidence that people preferred explanations congruent with their goals: this was the case for formal, mechanistic, and teleological explanations in Experiment 1, and for mechanistic and teleological explanations in Experiment 2.

Importantly, we found that goal context does not simply shift the explanandum (which was specified in Experiment 2), but instead affects the relative ratings for different *kinds* of explanations, with goal-congruent explanations receiving a relative boost. This supports the idea that explanations are tailored to context by supplying information with high anticipated utility (Lombrozo & Carey, 2006). These findings also have implications for philosophical accounts of explanation. One of the main critiques of pragmatic accounts is the lack of constraint on the relation between candidate explanans and the explanandum (Kitcher & Salmon, 1987). Our work demonstrates that the goals of the explainer can systematically constrain that relation, which raises the possibility of a pragmatic approach that is sufficiently constrained and descriptively adequate as an account of human judgments.

That said, our findings do not rule out more traditional accounts of explanation. For instance, accounts that allow for incomplete (Hempel & Oppenheim, 1948) or partial explanations (Railton, 1978; Kitcher, 1989) could accommodate our results if our manipulation impacted which parts of the underlying or ideal explanation were selected (but see Woodward, 2003). Alternatively, our results could be accommodated by allowing for pluralism in the patterns, covering laws, or other structures governing explanations. One possibility is that people represent both teleological and mechanistic explanatory patterns as subsuming a given phenomenon, and switch between the two depending on their goals.

Our findings also provide potential evidence for competition between mechanistic and function-based reasoning (see also Heussen, 2010; Lombrozo & Gwynn, 2014). In Experiment 1, teleological explanations were rated significantly lower under the cause-based goal compared to other conditions, and in Experiment 2, mechanistic explanations were rated significantly lower under the functional goal relative to other conditions, suggesting that in addition to boosting goal-congruent explanations, goals can also penalize goal-incongruent explanations. Notably, this pattern of competition was restricted to mechanistic versus function-based reasoning: only causal and functional goals produced suppression effects, and only ratings of mechanistic and teleological explanations were significantly negatively correlated.

Relationship to Prior Work Our findings are consistent with prior work suggesting a close relationship between explanation and inference. For example, Lombrozo and Gwynne (2014) and Vasilyeva and Coley (2013) found that different types of explanations predicted different patterns of property generalization (for similar effects in categorization, see Lombrozo, 2009; Ahn, 1998). These studies, however, did not investigate a relationship in the reverse direction, with (anticipated) inferences affecting explanation judgments.

Prior work also suggests that the production of teleological and mechanistic explanations can depend on context. Hale and Barsalou (1995) had participants complete a task with an initial system-learning phase followed by a trouble-shooting phase. They found that the types of explanations generated varied across phases. However, their goal manipulation (system-learning vs. trouble-shooting) was confounded with several factors, including task order, changes in background knowledge, and task instructions (think aloud vs. explanation). Chin-Parker and Bradner (2010) also found that the frequency with which participants generated mechanistic and teleological explanations was influenced by changing background conditions, but they did not vary participants' goals. To our knowledge, our studies provide the first demonstration that goals affect the perceived quality of explanations.

Future Directions Our findings demonstrate that anticipated inferences can affect the perceived quality of different kinds of explanations, but further work is needed to specify the basis and limits of this effect. For example, do goals induce different stances (Dennett, 1987)? Are people responsive to the goals of others in generating explanations? Is the effect of goals restricted to anticipated inferences, or does it extend to other markers of utility, such as past inferences or even the salience of particular information in a given context? And finally, what are the psychological processes responsible for such changes? Although much work remains to be done, our studies take an important step towards developing a psychological account of explanation that recognizes the context-sensitive and flexible nature of human explanatory judgments.

Acknowledgments

This work was supported by the *Varieties of Understanding* Project, funded by the John Templeton Foundation. We are grateful to Daria Serrano Cargol and Marc Collado-Ramírez for assistance in preparation of materials and data collection.

References

Ahn, W. (1998). Why are different features central for natural kinds and artifacts? *Cognition, 69,* 135–178.

Barsalou, L.W. (1983). Ad hoc categories. *Memory & Cognition, 11,* 211-227.

Chin-Parker, S., & Bradner, A. (2010). Background shifts affect explanatory style: how a pragmatic theory of explanation accounts for background effects in the generation of explanations. *Cognitive Processing*, *11*(3), 227–49.

Craik, K. J. W. (1943). *The Nature of Explanation.* Cambridge: Cambridge University Press.

Dennett, D.C. (1996). *The Intentional Stance*. Cambridge, Massachusetts: The MIT Press.

Hale, C. R., & Barsalou, L. W. (1995). Explanation Content and Construction During System Learning and Trouble-shooting. *Journal of the Learning Sciences*, *4*(4), 385–436.

Harman, G. H. (1965). The inference to the best explanation. *Philosophical Review*, *74*(1), 88-95.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: John Wiley & Sons.

Hempel, C., & Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science*, *15*(2), 135–175.

Heussen, D. (2010). When functions and causes compete. *Thinking & Reasoning*, *16*(3), 233–250.

Kitcher, P. (1989). Explanatory Unification and the Causal Structure of the World. In P. Kitcher & W. Salmon (Eds.), *Scientific explanation*. Minneapolis: University of Minnesota Press.

Kitcher, P., & Salmon, W. (1987). Van Fraassen on Explanation. *Journal of Philosophy,* 84, 315-330.

Leake, D. (1995). Abduction, experience, and goals: A model of everyday abductive explanation. *Journal of Experimental & Theoretical Artificial Intelligence*, *7*(4), 407–428.

Lipton, P. (2004). *Inference to the best explanation*. Psychology Press.

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition*, *99*(2), 167–204.

Lombrozo, T., & Gwynne, N. Z. (2014). Explanation and inference: mechanistic and functional explanations guide property generalization. *Frontiers in Human Neuroscience*, *8*, 700.

Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin, 129*(4), 592.

Prasada, S., & Dillingham, E. M. (2009). Representation of principled connections a window onto the formal aspect of common sense conception *Cognitive Science 33*(3), 401-48.

Railton, P. (1978). A deductive-nomological model of probabilistic explanation. *Philosophy of Science, 45 (2)*:206-226.

Rehder, B., & Kim, S. (2006). How causal knowledge affects classification: A generative theory of categorization. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *32*(4), 659–83.

Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton University Press.

van Fraassen, B. (1980). *The Scientific Image.* Oxford: Clarendon Press.

Vasilyeva, N. & Coley, J.C. (2013). Evaluating two mechanisms of flexible induction: Selective memory retrieval and evidence explanation. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society*). Austin, TX: Cognitive Science Society.

Woodward, James (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.