Available online at www.sciencedirect.com
SCIENCE DIRECT
COGNITION
Cognition 99 (2006) 167–204
www.elsevier.com/locate/COGNIT

Functional explanation and the function of explanation

Tania Lombrozo*, Susan Carey

Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA

Received 3 August 2004; accepted 23 December 2004

Abstract

Teleological explanations (TEs) account for the existence or properties of an entity in terms of a function: we have hearts because they pump blood, and telephones for communication. While many teleological explanations seem appropriate, others are clearly not warranted—for example, that rain exists for plants to grow. Five experiments explore the theoretical commitments that underlie teleological explanations. With the analysis of [Wright, L. (1976). *Teleological Explanations*. Berkeley, CA: University of California Press] from philosophy as a point of departure, we examine in Experiment 1 whether teleological explanations are interpreted causally, and confirm that TEs are only accepted when the function invoked in the explanation played a causal role in bringing about what is being explained. However, we also find that playing a causal role is not sufficient for all participants to accept TEs. Experiment 2 shows that this is not because participants fail to appreciate the causal structure of the scenarios used as stimuli. In Experiments 3–5 we show that the additional requirement for TE acceptance is that the process by which the function played a causal role must be general in the sense of conforming to a predictable pattern. These findings motivate a proposal, Explanation for Export, which suggests that a psychological function of explanation is to highlight information likely to subserve future prediction and intervention. We relate our proposal to normative accounts of explanation from philosophy of science, as well as to claims from psychology and artificial intelligence.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Explanation; Teleology; Function; Causal reasoning; Philosophy of science

* Corresponding author.
E-mail address: lombrozo@wjh.harvard.edu (T. Lombrozo).

0022-2860/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2004.12.009

> “Knowledge is the object of our inquiry, and men do not think they know a thing till they have grasped the ‘why’ of it (which is to grasp its primary cause).”
>
> ‘Why is he walking about?’ we say. ‘To be healthy’, and, having said that, we think we have assigned the cause.”

- Aristotle (*Physics*, II.3.194B17; II.3.194B31)

1. Introduction

Aristotle believed that why-questions could be addressed by appeal to four kinds of causes, each providing insight into some aspect of the question. Among them was the final cause, which explains ‘that for the sake of which’ something is the case by supplying a function or goal, also called a *teleological* or *functional* explanation. Aristotle explained a variety of physical and biological phenomena by appeal to final causes, suggesting that rocks fall in order to assume their natural place and that plants have leaves for the sake of shading their fruit (*Physics* II,8). Today we continue to invoke final causes in a variety of contexts. We typically explain the existence and properties of artifacts by appeal to their functions, and people’s behavior by appeal to their goals. But what Aristotle saw as a universal form of explanation, relevant to addressing an aspect of any why-question, today enjoys a more restricted scope. To most Western adults, it sounds ludicrous to explain the properties of a mountain by appeal to its function, or the fall of a rock by appeal to its goal.

Not only have intuitions about the acceptability of teleological explanations changed historically from Aristotle to the present, so too do they change in the course of development (Kelemen, 1999a; Kelemen, 1999b) and across cultures (Casler & Kelemen, 2003). In marked contrast to their elders, young American and British children will happily accept and provide teleological explanations for just about everything, claiming that lions exist for going in the zoo and clouds for raining (Kelemen, 1999a; Kelemen, 2003). This variability in teleological intuitions across populations suggests that the acceptability of teleological explanations tracks differences in beliefs about the world. Aristotle presumably found teleological explanations of falling objects appealing as a result of particular assumptions about the nature of physics—assumptions not shared by most contemporary physicists. Similarly, the developmental change and cultural variability of teleological intuitions may result from underlying differences in folk theories about the structure of the world.

In this paper we explore the commitments underlying teleological explanations, and in particular how the acceptability of teleological explanations relates to conceptual domains, causal beliefs, and general constraints on explanation. Providing a mapping between beliefs and the form of preferred explanations can shed light on the sources of historical, cultural, and developmental differences in teleological intuitions. But more generally, understanding the conditions under which certain explanations are warranted can elucidate the role of explanation in the application and acquisition of folk-theoretic knowledge. Several researchers have suggested that causal-explanatory principles are at the heart of folk theories (Carey, 1985; Gopnik & Meltzoff, 1997; Keil, 1994; Wellman & Gelman, 1992), but the precise relationship between explanation and theoretical commitments has yet to be elaborated. Furthermore, explanatory knowledge has been posited as the "glue" that gives categories conceptual coherence (Murphy & Medin, 1985), making an account of explanation central to a theory of concepts and categorization.

We begin by reviewing four accounts of teleological explanation, three intended as descriptive accounts of human intuitions and the fourth as a normative account from philosophy. These accounts raise a number of questions about the relationship between the acceptability of teleological explanations, the domain of what is being explained, and the causal history of what is being explained. We address these questions in Experiment 1. In the remaining four experiments we consider other factors that might modulate the acceptability of teleological explanations, such as the generality of the explanatory pattern.

Overall, we find that people accept teleological explanations when two conditions obtain: the function invoked in the explanation played a causal role in bringing about what is being explained, and the process by which it did so can be subsumed under some kind of pattern or causal schema. As our results were not anticipated by existing accounts, we develop a hypothesis—Explanation for Export—that explains our findings and suggests a function for explanation in learning and restructuring theoretical knowledge. We conclude by relating the Explanation for Export hypothesis to claims from philosophy, psychology, and artificial intelligence.

1.1. The psychology of teleological explanation

Within psychology, researchers have focused on the extent to which teleological explanation is tied to a particular domain of reasoning, whether the relationship between teleology and particular domains is innate, and at what age the selective application of teleological explanation emerges For example, Atran (1995) suggests that teleology is specific to an innate "living thing" module for biological reasoning. Keil (1992, 1994 1995) instead argues that teleological reasoning is an autonomous mode of construal, but like Atran, believes that it becomes a privileged form of explanation within the domain of intuitive biology very early in development. Kelemen (1999d) presents yet a third view, promiscuous teleology, according to which teleology is intimately related to goal-directed action. Her data suggest that children initially extend teleological explanations to everything they can imagine serving some function, and only come to restrict the scope of teleology with age and education. These views emphasize conceptual domains and causal properties to varying degrees.

According to Atran (1995), teleological and essentialist assumptions form the core of an innate "living thing" module, an evolutionary adaptation that extends initially to "phenomenally apparent organisms" (p. 219) and underlies biological classification and inference. By implementing domain-specific, teleo-essentialist causal reasoning, the module leads children to believe that morphological properties of organisms are caused by underlying essences, and that these properties play a functional role for the organism. With sufficient cultural prodding teleology can also extend to artifacts, but Atran regards living things as the initial and proper domain of teleological explanation.

On Atran's account, the acceptability of a teleological explanation depends on the domain of the object whose existence or properties are being explained: teleology will be warranted for the properties of living things, and other entities to which cultural practice extends teleology, like artifacts. As support for the "living thing" module, Atran (1994) has shown impressive convergence in the structure of taxonomic classifications among American populations and the Itza-Maya of Guatemala, as well as reasonable agreement with scientific taxonomies for both groups. Combined with a wealth of evidence for essentialized biological kinds (e.g. Gelman et al., 1994; Gelman, 2003; Medin & Atran, in press), the data support the claim that essentialism plays a role in the representation of "living things", but do not speak to the role of teleology.

Keil provides direct data in support of his (and Atran's) contention that young children recruit teleological explanations exclusively in reasoning about biological or artifactual properties (see also Slaughter, Jaakkola, & Carey, 1999). Keil suggests that the teleological mode of construal is particularly useful for understanding the functional significance of properties within complex systems, and therefore becomes associated with biology and artifacts early in development. In one experiment, for example, kindergarten and second grade children were required to choose between two explanations for a property of a familiar object (Keil, 1992). The provided explanations, one teleological and the other mechanistic, were presented as a forced choice. For instance, some children were shown a plant, and told the following: "Two people are talking about why plants are green. This person says that plants are green because it is better for the plants to be green and it helps there be more plants. This person says it is because there are little tiny parts in plants that when mixed together give them a green color. Which reason is a better one?" Other children heard the same description, but for an emerald (a non-biological natural kind) rather than a plant (a biological natural kind). Keil found that second graders preferred teleological explanations for biological kinds and mechanistic explanations for non-biological kinds.

In contrast to Keil's and Atran's views that teleological explanation are applied selectively in early childhood, Kelemen argues that children initially overextend teleological explanations to all domains. Kelemen (1999c) questions Keil's evidence for the selective application of teleology on the grounds that the teleological explanations provided in Keil's experiments involve the phrases "better for X" or "helps there be more X"—locutions that are highly associated with and limited to living things. Thus children may have restricted their inferences based on this linguistic cue and the fact that it makes more sense to "help" something that is alive and has goals, and not as the result of sensitivity to the teleological form of the explanations.

Kelemen's own studies suggest that without linguistic cues like "helps" children are not selective in their application or acceptance of teleological explanations. In a variety of tasks, she finds that children provide and prefer teleological to mechanistic explanations for non-biological natural kinds as well as biological kinds and artifacts. For example, Kelemen (1999a) found that preschoolers willingly produced functions for items from any domain, even when explicitly given the option of claiming that the object is not "for" anything. Adults restricted functions to artifacts, artifact parts like handles, and biological parts like noses. Kelemen (1999b) obtained the same general finding with a task that resembled Keil's explanation choice experiments. Children were presented with two explanations for a property of some object, and asked to choose between them. For example, they might be shown a pointy rock and asked to decide if it is pointy because little bits of stuff piled up on it, or because being pointy keeps animals from sitting on it. Consistent with Kelemen (1999a), children chose the teleological explanation for all kinds of objects, including non-biological natural kinds, like rocks, and whole biological kinds, like tigers. In contrast, adults only accepted the teleological explanation for artifacts and biological parts.

These results have led Kelemen to propose a theory of "promiscuous teleology" (Kelemen, 1999c; Kelemen, 1999d), which shares with several analyses in philosophy (e.g. Hempel & Oppenheim, 1948; Wright, 1976) the idea that teleological explanations are intuitive because of their close association with intentional explanations of purposeful behavior. On this view, the initial proper domain of teleological reasoning is human goal-directed action. The theory has two main tenets: first, that "the tendency to view objects as designed for a purpose develops as part of our ability to view intentional agents as having purposes", and second, that "because of the way our minds are designed, intention-based teleological explanations come easily to us" (Kelemen, 1999d, p. 287). Kelemen thus predicts the overextension of teleological explanations in childhood, and explains this finding by appeal to a basic mode of intentional explanation. Only late in development and only in some cultural contexts are teleological explanations restricted to behavior, artifacts, and biological parts (Kelemen, 1999a).

In sum, Atran, Keil, and Kelemen disagree about the origin and selectivity of teleological explanation, but agree in their characterization of the Western, adult state: teleological explanations are licensed as a function of the domain of objects or properties being explained. When explaining a biological part like eyes, teleological explanations are appropriate; when explaining the fall of a rock, they may not be. This formulation raises important questions about what constitutes a domain, and in virtue of which properties certain domains license teleological explanation. These theorists take domains to be sets of objects and properties in the world, united by falling under some systematic conceptual structure like an intuitive theory. Atran sees the connection between teleological reasoning and intuitive biology as resulting from innate conceptual constraints, while Keil and Kelemen suggest that domains are individuated by distinctive causal/explanatory principles that in turn license teleological explanation. For example, while Keil is not explicit about the properties of biological kinds and artifacts that license teleology, he does relate explanatory modes to underlying causal similarities, noting that beliefs about causal laws cluster in ways that suggest "domains of explanation" (1995, p. 259). So while preferred explanation types may differ as a function of the "kind of thing involved" (1994, p. 247), kinds are in turn individuated by causal properties. Similarly, Kelemen writes that differences across domains and between children and adults could be related to causal properties—specifically, to "beliefs about the causal history of different entities" (Kelemen, 1999a).

A view like Keil's or Kelemen's, in which domains are individuated by causal properties, still requires an account of why some causal processes and not others license teleological explanations. In recent work, Kelemen hints at what such an account might look like (Kelemen, 2004; Kelemen & DiYanni, in press). Children are asked whether they believe someone or something is responsible for the object or phenomenon being explained, or whether it "just happened". She finds a significant correlation between the items for which children accept a teleological explanation and those for which they invoke a creator, typically another human or God. This suggests that the causal commitment underlying a teleological explanation may be related directly to intentional agency (Kelemen, 2004). Specifically, being intentionally created or designed towards some end is a causal process that will warrant teleological explanation, and that will also render the object in question an artifact—albeit in a liberal sense that includes God-created animals. While this possibility is suggested by Kelemen's work, neither Keil nor Kelemen provide an explicit account of the causal conditions under which teleological explanations are warranted. Fortunately, causal theories of teleological explanation in philosophy, discussed in Section 1.2, provide precisely such an account.

1.2. The philosophy of teleological explanation

Early discussion of teleology within the explanation literature, such as Hempel and Oppenheim (1948), tended to emphasize the prima facie acausal nature of teleological explanation. Specifically, in explaining the existence or properties of an object by appeal to a *future* goal or a function that is only *later* realized, teleological explanations seem to get the temporal order wrong: they explain the present by appeal to the future. Hempel and Oppenheim point out that in explanations of intentional behavior, a goal can be shorthand for an intention that precedes the behavior being explained. For example, we might explain that Aristotle crossed the road to get to the other side, with "the other side" being the goal of the behavior. However, if we regard the goal as shorthand for an intention to get to the other side of the road, then the temporal order problem is solved: the road-crossing is explained by a preceding intention, not a future goal.

In his 1976 book, *Teleological Explanations*, Larry Wright defended the legitimacy of teleological explanations (TEs) by arguing that they are in fact causal, and can accommodate a wider range of cases then those involving goal-directed behavior. Specifically, he suggested that a function F can appear in a teleological explanation of X if and only if F is a *consequence* of X and X is there *because* it results in F. If Aristotle crossed the road to get to the other side, his getting to the other side of the road is a *consequence* of his road crossing, and the road crossing occurs *because* it results in getting to the other side of the road. Thus the road crossing satisfies Wright's criteria, and we can explain Aristotle's behavior by appealing to a goal. In general, goal-directed behavior will automatically satisfy Wright's conditions. If we ask why a goal-directed behavior is occurring or takes the form it does, we can explain it (indirectly) by appealing to the goal it will ultimately fulfill and (directly) by an intention on the part of the agent to satisfy that goal.

Wright's analysis is more general than the solution offered by Hempel and Oppenheim in that some non-intentional processes, like natural selection and operant conditioning, will also satisfy his conditions. A crude example is that seeing is a *consequence* of having eyes, and we have eyes *because* they result in seeing. The fact that eyes perform the function of seeing explains why they spread throughout the population, were maintained, and involve some of the features they do. More generally, processes will yield the kinds of functions that appear in teleological explanations if they have what Wright calls a "consequence etiology": a causal process sensitive to the consequences of changes it produces. This condition is important in extending teleological explanation beyond the paradigmatically human cases, as it clarifies how non-intentional teleological explanations can be causal, even though they are not preceded by an intention to perform a function or achieve a goal. Specifically, if a process has a consequence etiology, then the fact that a feature or object X has a particular consequence (=function) will be the result of a historical process partially caused by X's having that consequence (=function). In this way the function invoked in a legitimate teleological explanation will have played a causal role in bringing about what is being explained.

While Wright's causal account of teleological explanation has been defended and refined, it is not without critics. In particular, some accounts of teleology reject the idea that teleological explanation can be causal in the way Wright requires, because they interpret the functions invoked in explanations ahistorically: as claims about potential causal contributions towards some goal rather than backwards-looking pointers to a specific causal history (e.g. Cummins, 1975). Advocates of this position would suggest, for example, that the function of the heart is to pump blood not because blood-pumping causally contributed to our having hearts (a historical claim), but rather because hearts play a causal role in blood-pumping (an ahistorical claim). If functions are interpreted ahistorically, then it once more becomes problematic to explain the existence of an object or the presence of some of its properties by appeal to a function determined by potential *future* contributions. Nonetheless, such ahistorical accounts can make sense of straightforward function ascriptions ("hearts are for pumping blood") and functional explanations of the *behavior* of a system ("hearts beat a particular way because they are for pumping blood"). Defending the causal account against ahistorical critics would go well beyond the scope of this paper, but it is worth noting that these philosophical accounts are intended as normative rather than descriptive characterizations of explanatory practice, whereas the question in this paper is the descriptive one of what people take to license teleological explanation. Given that people *do* in fact invoke teleological explanations for the existence and properties of objects, accounts that deny the legitimacy of such practices cannot be descriptively accurate.[1] Thus we focus on a descriptive re-interpretation of Wright's theory as a possible account of actual teleological intuitions.

1.3. Integrating domains and causal histories

Given Wright's machinery for describing the causal processes that license TEs, we can revisit the discussion from Section 1.1 and reconsider claims about the relationship between domains, causal processes, and teleological explanation. Atran, Keil, and Kelemen discuss TE acceptance in terms of domains, but Keil and Kelemen emphasize that domains are individuated with respect to causal processes. So how different are these accounts from Wright's? If domain-individuation is such that domains licensing TEs involve a consequence etiology while those that do not license TEs do not, then the views correspond nicely. However, there is good reason to doubt a simple correspondence. Domains and causal processes have a one-to-many mapping, and nothing about the way

[1] At least one ahistorical account of functions, Bigelow and Pargetter (1987), does attempt to accommodate teleological explanations of existence and properties, but see Mitchell (1993) for an excellent reply.

domains are individuated guarantees that the many corresponding processes will share a particular etiology. For example, some causal processes within intuitive biology (such as natural selection or a divine creator) license teleological explanation, whereas others (such as genetic drift) do not. Even with artifacts, for which a single process (human intervention) is sufficient for domain membership, teleological explanations will not always hold. A teapot's handle can be explained by appeal to its function, but an accidental scratch on the spout cannot. So while domains may be individuated with respect to causal processes, they need not respect the causal distinctions relevant to TE acceptability. The idea that a TE is licensed for some property in virtue of its *domain's* causal history is thus quite different from believing a TE is licensed for some property in virtue of that *property's* causal history.

Even appreciating this distinction, Keil and Kelemen may have tended to discuss teleological explanation in terms of domain because of the high correlation between domain membership and the presence of a consequence etiology. Artifact properties are typically regarded as the result of intentional design, and biological properties as the result of natural selection or divine creation. This makes domain a useful if defeasible heuristic for inferring causal history, and it could well play an independent psychological role in the development of teleological intuitions.

Kelemen's recent work suggests that she endorses another possibility—that properties of an entity's history (not its domain) are responsible for licensing TEs, but that Wright is wrong about the requisite etiology. Specifically, it could be that people have a more restricted understanding of the conditions that license TEs, corresponding to Hempel and Oppenheim's earlier discussion: they may believe that teleological explanations are legitimate only when the property or entity being explained was preceded by an intention to produce some function or achieve some goal. This is the possibility tested in Kelemen and Di Yanni (in press), where TE acceptance was correlated with claims about an intentional creator. We call the subset of consequence etiologies that satisfy the requirement of being preceded by an intention to have a function or achieve a goal *design* etiologies. Kelemen and DiYanni did not attempt to distinguish whether design etiologies or consequence etiologies more generally license teleological explanation.

In Experiment 1, we test the psychological reality of Wright's analysis, and additionally explore the possibility that design etiologies or entities from particular domains are privileged in determining TE acceptability. This experiment is the first to attempt to disentangle domain membership and causal history in determining the acceptability of teleological explanations. In the explanation choice experiments done to date, participants are provided with the domain of the entity for which explanations are given, but not the causal history. Thus it could be that domain itself is responsible for the applicability of teleological explanations, or that participants make a reasonable inference to the kind of process responsible for the object or property being explained, and accept or reject the teleological explanation on the basis of inferred causal history. To test whether causal history determines the acceptability of teleological explanations, and if so whether the relevant history is a design etiology or any consequence etiology, it is necessary to provide participants with information about the causal history of the property being explained. This is what we do in Experiment 1.

2. Experiment 1: domain versus causal history

To examine the role of domains and causal histories in determining TE acceptability, we constructed scenarios in which biological kinds, non-biological natural kinds, and artifacts are intentionally created or modified. In some cases the resulting functional features satisfy Wright's consequence etiology, and in other cases they do not. When they do, half of the time they also satisfy a design etiology: the creator had the intention to create the object with the resulting function. In a final third of cases, a functional property comes about as the result of an accident, and hence fails to satisfy either a design etiology or a consequence etiology.

If causal history determines TE acceptance, then we might expect judgments to align with the presence of a design etiology or any consequence etiology. Specifically, if intentional design is the relevant requirement, then only those scenarios with a design etiology should license teleological explanation. And if a consequence etiology is the relevant causal history, then the two types of scenarios with this structure should yield acceptable teleological explanations. Alternately, if domain membership *independently of the causal history of the feature in question* is a relevant factor in the acceptance of teleological explanations, TEs should be accepted for biological objects and artifacts, but not non-biological natural kinds. If belonging to a particular domain is sufficient for TEs to be licensed, then another possibility is that the presence of human intervention may lead participants to construe all of these entities as artifacts, and hence to accept TEs in all scenarios.

2.1. Methods

2.1.1. Participants

Participants were 36 Harvard undergraduates (21 male; mean age = 20, SD = 1) who completed the study in exchange for candy. There were an additional eight participants who were replaced as a result of failing to follow directions.

2.1.2. Materials

The experimental stimuli consisted of short causal stories followed by why-questions and candidate answers to the why-questions (see Appendix for full set of stimuli). The causal stories involved an object from one of three domains (artifact, biological part, non-biological natural kind[2]) undergoing one of three causal processes (Design, non-intentional consequence etiology or NICE, and Accident), for a total of nine stimulus types. The first two causal histories involved a function that played a causal role in bringing about what was being explained—what Wright calls a consequence etiology. Only the first causal history, however, involved an intention to produce something with a particular function, and hence satisfied a design etiology as well. Below are sample stimuli

[2] Our non-biological natural kinds were a cave, a rock, and an unspecified element, so the term "kind" should be interpreted loosely.

from the domain of non-biological natural kinds in each of the three causal history types, followed by the corresponding why-question and answers labeled in brackets:

[History: Design, Domain: Non-bio natural]
Sally owns a canyon park that attracts tourists who come to have picnics in the many caves. The best caves are those that are large enough to produce an audible echo, so she enlarges all the caves so that they produce louder echoes.
[History: NICE, Domain: Non-bio natural]
Sally owns a canyon park that attracts tourists who come to have picnics in the many caves. The best caves are those that are large enough to produce an audible echo, so tourists tend to prefer these. Sally does not realize that these are popular because of the echo, but she does notice that larger caves are everyone's favorites. As a result she decides to enlarge all the caves.
[History: Accident, Domain: Non-bio natural]
Sally owns a canyon park that attracts tourists who come to have picnics in the many caves. When reinforcing the caves to make them safer, Sally accidentally enlarges them, yielding only large caves. However, large caves are more popular because they produce audible echoes, so the tourists love the modified caves.

Why are Sally's caves large?
[teleological explanation]
(A) Because larger caves produce audible echoes.
[intention-based explanation]
(B) Because that's the way Sally wanted them.
[mechanistic explanation]
(C) Because Sally enlarged them.
[true-irrelevant filler]
(D) Because tourists go site seeing.
[false-irrelevant filler]
(E) Because Sally dislikes sandwiches.

The main question of interest is whether participants will accept the teleological explanation: "Because larger caves produce audible echoes". Note that in both the Design and NICE stories, the fact that larger caves produced audible echoes led to all of the caves becoming large. Had the larger caves not produced audible echoes, Sally would not have enlarged any of the caves. In the Design case this is because Sally realizes the relationship between size and echoes, and intervenes with the intention of producing more caves with audible echoes. In the NICE case, audible echoes cause popularity, which in turn causes Sally to enlarge the caves. While she does enlarge them intentionally, she does not do so with the intention of having them produce audible echoes, as she is unaware of the relationship between cave size and echo volume. Finally, in the Accident cases the fact that larger caves produce audible echoes plays no role in bringing about more large caves, but the caves nonetheless exhibit the functional property of producing audible echoes.

We were also concerned with the degree to which participants' judgments are systematic and sensitive to the causal structure of the scenario. For each story, four candidate answers were presented in addition to the *teleological* explanation. The *mechanistic* explanation, acceptable for all scenarios, mentioned the most immediate cause of the feature, while the *intention-based* explanation, acceptable for the Design and NICE scenarios, simply stated that the agent intended the intervention that brought about the property in question. Two filler explanations were also included: one true but largely irrelevant, the other potentially false and also irrelevant. For the accident scenarios, the mechanistic explanation specified that the intervention was accidental. For example, for the story above, participants in the accident condition saw "Because Sally accidentally enlarged them" for the mechanistic explanation.

In order for stories that varied in domain and causal history to remain as similar as possible in all other respects, three versions of each of the nine stimulus types were created. For each version, the corresponding stimuli involved a similar functional feature. For example, the three stories above are from a version involving a change in size that affects sound. The corresponding stories for artifacts involved a satellite dish that was enlarged and became more sensitive, while the biological part stories involved a cat's ears that increased in size and became more effective for hearing mice. A second version involved nine stimuli in which a change in color rendered the item more visible at night, and the third version involved nine stimuli in which a change in form made an object better for destroying weed roots. The version manipulation was necessary so that participants did not see stories from different domains and causal histories that involved similar properties with similar functional consequences. Thus, a given participant would see one story involving a change in size, one involving a change in color, and a third involving a change in form.

2.1.3. Design and procedure

Participants received a one-page questionnaire with the following instructions: "Below you will read three short scenarios each followed by a question. After each question are five possible answers. For each answer, please indicate whether you think it is *true of the scenario* and *addresses the question*. If so circle '*y*', otherwise circle '*n*'. If you circle '*y*', indicating that the answer is true and addresses the question, then please rate how *satisfying* you find the answer by circling a number from 1 to 7, where 1 corresponds to 'not at all satisfying', 7 corresponds to 'very satisfying', and so on for intermediate values. Please note that more than one answer per question may be indicated '*y*', and that you may repeat satisfaction rating values for different answers". Participants were then presented with three causal stories with their corresponding why-questions and explanations, and indicated both which explanations were acceptable and how satisfying they found them.

Story types were counterbalanced such that the three stories on each questionnaire involved entities from different domains, different causal histories, and different versions. For example, one participant might have a story about an artifact undergoing a NICE process with a change in form, a biological part undergoing a Design process with a change in color, and a non-biological natural kind undergoing an Accident process with a change in size. The order of the stories was randomized for each participant, as was the order of the explanations following each why-question.

Table 1
Explanation acceptance and satisfaction for Experiment 1 as a function of causal history

Explanation type	Acceptance			Satisfaction		
	Design	NICE	Accident	Design	NICE	Accident
Teleological	**86**	**50**	**17**	**5.77** (0.29)	**4.17** (0.33)	**3.83** (0.48)
Intention-based	**89**	**92**	**17**	4.72 (0.34)	4.21 (0.32)	3.50 (0.85)
Mechanistic	100	92	97	5.28 (0.29)	5.27 (0.33)	6.14 (0.22)
True-irrelevant	11	6	8	3.00 (0.53)	3.00 (1)	2.67 (0.67)
Potentially false	0	0	0	–	–	–

The percent of participants accepting each explanation type, and the average satisfaction rating for accepted answers, are shown as a function of causal history. The standard error of the mean follows each satisfaction rating average in parentheses. Significant main effects are indicated in bold.

2.2. Results

Table 1 shows the acceptance rates for the five explanation types for each of the three types of causal histories, along with their average satisfaction ratings. Although our primary interest is what makes teleological explanations acceptable and satisfying, the other four explanation types provide a manipulation check. As can be seen in Table 1, participants very rarely accepted the irrelevant or false explanations, and almost always accepted the mechanistic explanations. They also appropriately accepted the intention-based explanations for the Design and NICE scenarios, but not the Accident scenarios. Thus for these four explanation types, participants accepted an explanation when it was part of the causal story leading to the property to be explained. Participants found the task natural and provided very systematic judgments.

Acceptance rates for the explanation type of primary interest, teleological explanation, appear in Table 1 and Fig. 1. A 3 (domain) X 3 (causal history) ANOVA[3] with acceptance of the teleological explanation as a dependent variable revealed a main effect of causal history ($F(2, 105)=24.5$, $P<0.01$), but no effect of domain ($F(2,105)=1.5$, $P=0.24$) nor an interaction between causal history and domain. The main effect of causal history was driven by the fact that TEs were accepted by most participants in the intended function scenarios (86%), half of participants in the NICE scenarios (50%), and a minority of participants in the accident scenarios (17%). All three acceptance rates were significantly different from each other. The 50% acceptance rate for TEs in the NICE scenarios is not different from chance. However, in light of that fact that every other acceptance rate in Table 1 differed from chance, we do not think participants were merely guessing, but rather were genuinely ambivalent or had individual differences in opinion about the acceptability of TEs in the NICE scenarios. In Experiments 2–5 we explore some of

[3] Because this analysis involves a dichotomous outcome variable, one of the assumptions of ANOVA, normality, is violated. For this reason all such ANOVAs for Experiment 1 were repeated as log-linear analyses to ensure that a deviation from normality was not distorting the results. In all cases the log-linear results revealed identical patterns of significance.

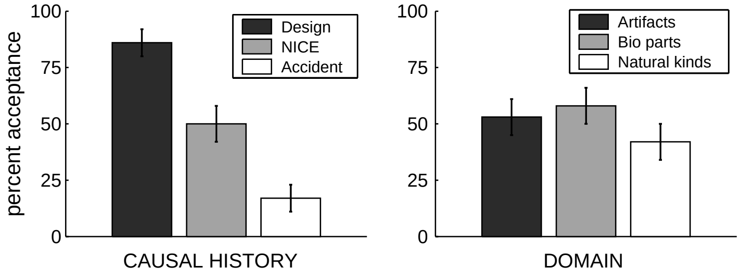

Fig. 1. Explanation acceptance for teleological explanations in Experiment 1 as a function of causal history and domain.

the factors that influence the acceptance rates of TEs in NICE scenarios and confirm this interpretation.

Similar ANOVAs were carried out for the other four explanation types in Table 1. There were no effects of domain. The only effect of causal history emerged for intention-based explanations: acceptance rates for intention-based explanations in the Design scenarios (89%) and NICE scenarios (92%) were significantly higher than in the Accident scenarios (17%).

In addition to acceptance ratings, satisfaction ratings for *accepted* TEs were analyzed. A 3 (domain) X 3 (causal history) ANOVA with satisfaction ratings for accepted TEs as the dependent measure revealed a significant main effect of causal history ($F(2, 52)=9.4$, $P<0.01$), no effect of domain ($F(2, 52)=0.8$, $P=0.46$), and no interaction between causal history and domain. TEs were found to be more satisfying in the Design scenarios ($\mu=5.77$, SD = 1.61) than either the NICE scenarios ($\mu=4.17$, SD = 1.38) or Accident scenarios ($\mu=3.83$, SD = 1.17). Because very few participants ($N=6$) accepted the teleological explanation for the Accident scenarios, the average satisfaction rating is not particularly meaningful in this case. However, the difference between the Design and NICE scenarios is significant and involved the ratings of most participants. Design and NICE scenarios were thus not only distinguished in overall acceptance rates, but also in overall satisfaction among participants already accepting the teleological explanation for these cases. When considered for each domain, teleological explanations were found most satisfying for artifact scenarios ($\mu=5.32$, SD = 1.73) followed by biological part ($\mu=4.90$, SD = 1.73) and non-biological natural kind ($\mu=4.87$, SD = 1.69) stories, though these differences did not approach significance.

Similar ANOVAs for satisfaction of accepted intention-based and mechanistic explanations revealed a marginally significant ($P<0.06$) main effect of causal history for the mechanistic explanations. Specifically, there was a trend to find the mechanistic explanation more satisfying for the accident stories, presumably because no other explanation was reasonable, leaving this explanation with no competition. There were no significant order effects for either explanation acceptance or explanation satisfaction.

2.3. Discussion

Consistent with Keil's and Kelemen's results, our adults accepted teleological explanations selectively. Experiment 1 explored two factors that might influence teleological intuitions: the domain of the entity whose properties are being explained, and the causal history of the property being explained. We found that causal history had a significant effect on the acceptability of TEs, with high acceptance for intentional function-driven processes, moderate acceptance for non-intentional processes with a consequence etiology, and low acceptance for accidental processes. We found no effect of domain, either in the acceptability or satisfaction ratings of TEs.

Our data rule out a particular version of the view that domain influences teleological intuitions: namely the idea that domain membership per se, independently of causal history, determines TE acceptability. For adults at least, teleological explanations are not more natural for functionally relevant properties of animals or artifacts than for properties of caves or rocks. One might object to this interpretation of the data on the grounds that all of our causal stories involve a human manipulating an entity, and that in some sense this renders all the stimuli artifacts. This may be, but our data nonetheless demonstrate that domain membership cannot be sufficient for a teleological explanation to be acceptable. Participants almost never accepted TEs for the accident stories involving prototypical artifacts like hats and tools, and only accepted TEs for such artifacts in NICE stories half the time, even though the manipulation that created the property was fully intentional. These exceptions can only be understood by appeal to the causal history of the property in question.

The question remains: what exactly about causal history licenses TE acceptance? While our data support Wright's prediction that the function invoked in a TE must have played a causal role in bringing about what is being explained, this condition does not appear to be sufficient for a majority of participants to accept TEs. Consistent with the view we attribute to Kelemen, significantly more participants accepted the TE for the Design stories than the NICE stories, no matter that both involved a process with a consequence etiology. There are three reasons why this difference may have been observed. First, it could be that for some participants, only a design etiology licenses teleological explanation, while other participants have a more abstract understanding corresponding to a consequence etiology. Second, it could be that teleological explanations for Design and NICE causal histories vary with respect to some other virtue of explanation, like those considered within philosophy of science. For example, it could be that TEs for Design stories are more general or have greater predictive value than TEs for NICE stories. Finally, it could be that some participants failed to appreciate that NICE stories in fact involve a consequence etiology. Because NICE stories are more complex and unfamiliar than Design stories, we would expect misunderstandings to be more frequent. Experiment 2 examines this last possibility before we consider the first two in Experiments 3–5.

3. Experiment 2: counterfactuals

In Experiment 1 we found significantly lower acceptance of TEs in NICE scenarios (50%) compared with Design scenarios (86%). Because both scenarios involve a function that played a causal role in the objects' coming to have the feature in question, this difference is prima facie inconsistent with Wright's causal theory. However, this difference between Design and NICE scenarios could be reconciled with Wright's view if the participants who did not accept TEs for NICE scenarios did so because they failed to realize the causal role of the function invoked in the TE. This hypothesis—call it the causal misrepresentation hypothesis—makes two predictions: (a) that an independent assessment of participants' causal representation of the scenario should reveal a relationship between TE acceptance and noting the causal role of the function, and (b) that if the causal role of the function is made more salient, more participants will accept the TE.

To test these predictions, participants read two NICE stimuli from Experiment 1 and evaluated the acceptability of corresponding answers to a why-question. Participants were also asked to answer true or false counterfactuals following each scenario. The counterfactuals were intended to index participants' causal representation, and highlight the causal role played by the function if it had not been noted. The causal misrepresentation hypothesis would predict that (a) there should be a relationship between acceptance of TEs and answers on counterfactuals that involve the causal role of the function, and (b) that some of the participants who do not accept the TE on the first question should accept it on the second as a result of having answered counterfactuals for question one that emphasized the causal role of functions.

3.1. Methods

3.1.1. Participants

Participants were 32 students (24 Harvard Undergraduates and 8 Harvard summer school students; 24 female; mean age = 20, SD = 2) who completed the study in exchange for candy or course credit.

3.1.2. Materials

The experimental stimuli consisted of two NICE stories from Experiment 1 (NICE/biological part/form and NICE/artifact/color, see Appendix) in addition to counterfactual questions. The counterfactuals were constructed to correspond to the teleological, mechanistic, and potentially false answers such that if playing a causal role is governing explanation acceptance, acceptance of the explanation should be perfectly correlated with the corresponding counterfactual response. The mechanistic and false counterfactuals were included to insure that participants were willing to respond both 'true' and 'false' to counterfactuals and understood their logical structure. They further provided a baseline for a reasonable level of correspondence between explanation acceptance and response on the counterfactuals. Below is a sample scenario followed by the three counterfactuals:

> Fred is a genetic engineer who creates plants and animals that facilitate agriculture, specializing in weed-eating gophers. His clients tend to purchase gophers with pointy claws, because they help destroy weeds not only by nibbling on them, but also by damaging the roots as they dig. Fred does not realize that

pointy claws allow gophers to damage weed roots, but he does notice that the pointy claws are the most popular. As a result he decides to create all of his gophers with pointy claws.

Why do Fred's gophers have pointy claws?

	Circle choices:
[intention-based explanation]	
(A) Because that's the way Fred wanted them.	y/n
[true-irrelevant filler]	
(B) Because gophers aren't marsupials.	y/n
[mechanistic explanation]	
(C) Because Fred genetically modified the gophers to be like that.	y/n
[false-irrelevant filler]	
(D) Because Fred doesn't like vegetables.	y/n
[teleological explanation]	
(E) Because the pointy claws damage weed roots.	y/n

Consider the following sentences and decide whether they are true or false:

[Teleological counterfactual]

"If gophers with pointy claws didn't damage weed roots, Fred probably wouldn't have decided to create all of his gophers with pointy claws."

Circle one: *TRUE FALSE*

[Mechanistic counterfactual]

"If Fred hadn't genetically modified his gophers to have pointy claws, they probably wouldn't all have pointy claws."

Circle one: *TRUE FALSE*

[False counterfactual]

"If Fred liked vegetables, he probably wouldn't have made gophers with pointy claws."

Circle one: *TRUE FALSE*

3.1.3. Design and procedure

Participants received a two-page questionnaire with instructions similar to those from Experiment 1. Each page of the questionnaire contained one of the two NICE stories with the corresponding why-question, candidate answers, and two counterfactuals. Half of the participants evaluated the teleological and mechanistic counterfactuals; the remaining half evaluated the teleological and false counterfactuals. The order of the NICE stories and counterfactuals was counterbalanced. The answers to the why-questions were presented in one of several random orders.

3.2. Results

Table 2 displays acceptance rates for each explanation type. The results replicate those from the NICE scenarios of Experiment 1. Acceptability judgments were highly systematic: participants rejected the irrelevant and false explanations and accepted the mechanistic and intention based explanations. The acceptance rates for TEs also replicated

Table 2
Explanation and counterfactual acceptance for Experiment 2

Explanation type	Explanation acceptance			Counterfactual judgment		
	First	Second	Overall	First	Second	Overall
Teleological	56	50	53	78	84	81
Intention-based	75	84	80	–	–	–
Mechanistic	94	97	95	100	100	100
True-irrelevant	3	0	2	–	–	–
Potentially false	0	3	2	0	0	0

The percent of participants accepting an explanation and corresponding counterfactual are shown for each explanation type. The data are divided into the first and second question seen by each participant. There were no significant differences between responses for the questions seen first and second.

those of Experiment 1's NICE scenarios, with TEs accepted 53% of the time in Experiment 2, compared to 50% of the time in Experiment 1.

To examine the prediction that acceptance of the TE would correlate with acceptance of the teleological counterfactual, we first looked for relationships between acceptance of the mechanistic and false explanations and counterfactuals to establish a baseline. All participants who saw the mechanistic counterfactual accepted both the mechanistic explanation and judged the mechanistic counterfactual true. Similarly, all participants who saw the false counterfactual rejected both the false explanation and the false counterfactual. The identical patterns of response on the explanations and counterfactuals for both the mechanistic and false cases suggest that the explanation and counterfactual tasks tap similar aspects of the how the story is represented. More importantly, the consistent and correct responding on the counterfactuals implies that participants understood their logical structure in this simple task.

With this baseline for performance on the counterfactual task, we can examine the relationship of primary interest: acceptance of the teleological explanation and the teleological counterfactual. The first result worth noting is that participants judged the counterfactual true (81% overall) significantly more often than they judged the TE acceptable (53% overall) as revealed by a chi-square test of independence ($\chi^2(1)=13.74$, $P<0.01$). As a group, participants thus appreciated the causal role of the function in bringing about the property being explained. Nonetheless, there could still be a relationship between explanation and counterfactual responses. There was a high correlation between how individual participants responded on the first and second questions, so we conducted two independent analyses. For the question seen first, 56% of participants accepted the teleological explanation and 78% accepted the teleological counterfactual. A 2 (explanation accept versus reject) $\times$ 2 (counterfactual true versus false) chi-square test of independence revealed a marginally significant relationship between responses on the two tasks ($\chi^2(1)=2.79$, $P<0.1$). For the question seen second, 50% of participants accepted the teleological explanation and 84% accepted the teleological counterfactual, again resulting in a suggestive but not significant relationship ($\chi^2(1)=2.13$, $P<0.15$). In both cases the relationship was such that more participants who accepted the TE judged the counterfactual true compared with participants who rejected the TE. Specifically, 90% of participants who accepted the TE also accepted the counterfactual, compared with 70% of participants who accepted the counterfactual despite rejecting the TE. While suggestive, these effects are quite small. The counterfactual responses accounted for less than 10% of the variance in acceptance of the teleological explanation, suggesting that the causal misrepresentation hypothesis is unlikely to provide the full explanation for the difference between TE acceptance in Design and NICE scenarios observed in Experiment 1.

Finally, we compared acceptance of the teleological explanation for the two scenarios to see if more participants accepted the teleological explanation for the scenario seen second. There was no significant difference ($t(62)=0.494$, $P>0.6$ two-tailed). Indeed, 28 of the 32 participants responded consistently in their acceptance of the teleological explanation for both questions.

3.3. Discussion

Experiment 2 was designed to test the causal misrepresentation hypothesis, which would explain the lukewarm acceptance of TEs in NICE scenarios while being consistent with Wright's causal account of teleological explanation. However, the results go against both predictions of the hypothesis. Acceptance of the teleological counterfactual was greater than chance, suggesting that overall the participants did understand the causal structure of the story. There was only a small and non-significant relationship between acceptance of the TE and response on the teleological counterfactual. Furthermore, the nearly identical acceptance rates for the TE in the first and second scenarios suggests that making the causal role of the function salient had no or little effect on TE acceptance.

Two positive claims can be made on the basis of results from Experiment 2. First, that the vast majority of participants who accepted the TE also judged the teleological counterfactual true suggests that a function's appropriate causal role may be a necessary if not sufficient condition for TE acceptance—only 3 of 32 participants accepted the teleological explanation but judged the teleological counterfactual 'false'. And second, the consistent pattern of responding for individual participants suggests that people have stable intuitions about TE acceptability. The replicated pattern of about 50% acceptance is not due to uncertainty and guessing.

If the causal misrepresentation hypothesis is false, then we lack an explanation for the differential TE acceptance in Design and NICE causal histories from Experiment 1. As suggested in Section 2.3, it could be that some participants have an understanding of teleological explanation that is restricted to design etiologies, or that some virtue of explanation, like generality, is better exemplified by the Design stories. The possibility that the generality of a causal process influences TE acceptability can be motivated in terms of theories of explanation from philosophy. Several theories suggest that an explanation must demonstrate why what you are explaining was to be *expected*, or how it can be *subsumed* under a law-like pattern (e.g. Salmon, 1989; Hempel & Oppenheim, 1948; see Strevens, in preparation for nice discussion of different explanatory relations). Because Design causal histories conform to a known pattern, providing a function can both suggest why something was to be expected and be subsumed under an intended-function causal schema. Take the sample Design story from Experiment 1, where Sally modified the caves in her park to increase audible echoes. Because participants would in general be responsive to the *reason* large caves are popular, knowing that large caves produce audible echoes provides some basis on which to expect that Sally will enlarge the caves. In contrast, the teleological explanation does not have this predictive thrust for the NICE story unless one knows the park's visitors are selective in choosing caves and that Sally is responsive to popularity when she does not understand its basis—expectations which are hardly the default.

In sum, the greater TE acceptance for Design compared with NICE scenarios could mean that for some participants, a design etiology is a requirement for TEs to apply. On the other hand, the greater generality of design processes may be the critical factor. These views generate different predictions about TE acceptability for artificial and natural selection. If a design etiology is important, participants should find TEs more acceptable for the products of artificial selection, an intentional process, than those of natural selection, a non-intentional process. On the other hand, if generality mediates the observed difference between Design and NICE scenarios, participants should accept TEs equally and at high rates for both natural and artificial selection, since both processes conform to a general causal pattern and satisfy a consequence etiology. Experiment 3 tests these alternative hypotheses.

4. Experiment 3: artificial and natural selection

In Experiment 3, participants were presented with scenarios and why-questions like those of previous experiments, but involving two familiar situations: artificial and natural selection. These causal processes were selected because they match Design and NICE scenarios in the causal role of intentions. Specifically, both artificial selection and Experiment 1's Design scenarios involve an intention to change an object's feature such that it have a particular function, and both natural selection and Experiment 1's NICE scenarios lack a Design etiology while involving a function-driven process. Unlike the NICE scenarios from Experiments 1 and 2, however, both artificial and natural selection are general processes in the sense that they can each be subsumed under familiar causal patterns.

4.1. Methods

4.1.1. Participants

Participants were 36 student (16 Harvard undergraduates, 20 students recruited over the summer from other selective universities) who completed the study in exchange for candy or course credit. Two additional participants completed the questionnaire but were replaced as a result of failing to follow directions.

4.1.2. Materials

The experimental stimuli consisted of short causal stories describing natural or artificial selection, followed by a why-question and candidate answers. There were three matched pairs of artificial and natural selection stories, like the sample below, with the same corresponding why-question and answers:

> ***[artificial selection]***
> The purple-frocked lemur is a small, nocturnal primate that eats bugs. Biological anthropologists visiting Madagascar at the turn of the century were worried that the species would go extinct, since the lemurs were not very good at catching bugs in dim light. As a result the anthropologists decided to selectively breed the lemurs with the largest eyes, since lemurs with bigger and hence more sensitive eyes are better at catching bugs. As a result of this process of artificial selection over several generations, modern-day purple-frocked lemurs have very large eyes.
> ***[natural selection]***
> The purple-frocked lemur is a small, nocturnal primate that eats bugs. Early biological anthropologists who visited Madagascar investigated the lemur's current behavior and evolutionary history. They found that because bugs are difficult to spot in dim light, those lemurs with bigger and hence more sensitive eyes were historically better at catching bugs, thus living longer and producing more offspring. As a result of this process of natural selection, modern-day purple-frocked lemurs have very large eyes.

Why do today's purple-frocked lemurs have large eyes?		***Circle choices:***
[teleological explanation]		
(A) Because large eyes are better for catching bugs.	y/n	− 1 2 3 4 5 6 7 +
[human intention-based explanation]		
(B) Because that's the way the anthropologists wanted them.	y/n	− 1 2 3 4 5 6 7 +
[animal intention-based explanation]		
(C) Because that's the way the purple-frocked lemur wanted them.	y/n	− 1 2 3 4 5 6 7 +
[mechanistic explanation]		
(D) Because over several generations purple-frocked lemurs with bigger eyes produced more offspring.	y/n	− 1 2 3 4 5 6 7 +
[true-irrelevant filler]		
(E) Because lemurs are from Madagascar.	y/n	− 1 2 3 4 5 6 7 +

The other two matched pairs involved a fish that is dark blue for camouflage, and a bird with a sharp beak for pecking at worms.

4.1.3. Design and procedure

Participants received a one-page questionnaire with instructions similar to those from Experiment 1. Each questionnaire contained a single artificial or natural selection stimulus. The scenarios involved different creatures and features, which were matched for the artificial and natural selection conditions. Following the stimulus participants read the corresponding why-question and indicated the acceptability of candidate answers, along with the satisfaction ratings for accepted answers. The answers were presented in one of several random orders.

4.2. Results

TE acceptance was well above 50% for both the artificial and natural selection conditions (see Table 3), Participants accepted the TE 78% of the time in artificial selection cases and 94% of the time for natural selection cases. These acceptance rates were not significantly different from each other or from the 86% TE acceptance for Design scenarios in Experiment 1. Further, the intention-based explanation (e.g. "Because the anthropologists wanted them that way") was always and only accepted for the artificial selection cases, suggesting that participants appropriately distinguished between the two selection types. Interestingly, the mechanistic explanation was accepted more often for natural than artificial selection: 94% compared with 67%. This difference was significant ($\chi^2(1)=4.4$, $P<0.05$). Differential reproduction is mentioned more explicitly in the natural selection scenarios than in the artificial selection scenarios, so perhaps this result is not surprising. It could also be that participants found the mechanistic explanation for the artificial selection scenario pragmatically awkward because it did not mention the role of the anthropologists, or that participants do not understand the mechanism of artificial selection (for some evidence to this effect see Shtulman, under review).

Finally, we looked at satisfaction ratings for those answers that were accepted. There were no significant satisfaction differences for the teleological explanation in the natural ($\mu=5.18$, SD=1.69) and artificial ($\mu=4.64$, SD=1.13) selection cases ($t(29)=1.05$, $P>0.3$). Mirroring the acceptance rates, there was a significant difference between satisfaction ratings for accepted mechanistic explanations in the two scenarios. Specifically, the mechanistic explanation was rated as more satisfying for natural selection ($\mu=5.06$, SD=1.48) than for artificial selection ($\mu=3.58$, SD=2.02; $t(27)=2.275$, $P<0.05$).

Table 3
Explanation acceptance and satisfaction for Experiment 3

Explanation type	Acceptance		Satisfaction	
	Artificial	Natural	Artificial	Natural
Teleological	78	94	4.64 (0.44)	5.18 (0.26)
Intention-based (human)	**100**	**0**	5.17 (0.46)	–
Intention-based (animal)	0	0	–	–
Mechanistic	**67**	**94**	**3.58** (0.54)	**5.06** (0.35)
True-irrelevant	0	0	–	–

The percent of participants accepting each explanation type, and the average satisfaction rating for accepted answers, are shown for the artificial and natural selection scenarios. For satisfaction ratings, the standard error of the mean follows each average in parentheses. Significant differences between the artificial and natural selection conditions are indicated in bold.

4.3. Discussion

Participants accepted teleological explanations at high levels for both artificial and natural selection. The acceptance rates for both kinds of selection were well over 50%, and not significantly different from acceptance rates for Experiment 1's Design scenarios. Although artificial selection is both a general process and involves an intention to produce something with a particular function, TEs were not accepted more often for artificial than natural selection. In fact, more participants accepted the TE for natural than artificial selection (94 versus 78%). These findings suggest that the generality of the causal process, and not the presence of a design etiology, is the critical variable modulating TE acceptance in Experiments 1 and 2. However, another possibility is that participants misconstrue natural selection as an intentional process (e.g. Evans, 2000), or that they have simply heard teleological explanations applied to the products of natural selection in the past. Experiment 4 more rigorously examines whether the generality of the causal process per se is a critical factor.

5. Experiment 4: general processes

To test whether generality per se is driving the increase in TE acceptance for natural selection over NICE scenarios, we made NICE scenarios "general" by providing evidence that the process conforms to a predictable pattern. We did this by giving participants examples of a similar scenario before having them evaluate the critical scenario, which was the same as a NICE scenario seen alone by participants in Experiments 1 and 2. If the generality of a function-driven process increases TE acceptance, then significantly more participants should accept the TE after seeing the similar examples compared to evaluating the critical scenario in isolation.

5.1. Methods

5.1.1. Participants

Participants were 72 members of the Harvard summer community (38 male; mean age = 22, SD = 6) who completed the study in exchange for candy or course credit. Several additional participants were replaced as a result of failing to follow directions: they provided satisfaction ratings for rejected explanations.

5.1.2. Materials

Experimental stimuli varied as a function of condition. Thirty-six participants were in the *baseline* condition, in which they saw a single NICE scenario and were asked to evaluate the same explanations for the same why-questions as in Experiment 1. The baseline condition is essentially a partial replication of Experiment 1, as the scenarios were two NICE stories from that experiment (NICE/biological part/form and NICE/biological part/color, see Appendix). Thirty-six participants were in the *generality* condition, in which they saw two NICE stories before evaluating a third, which was the same as what participants in the *baseline* condition evaluated alone. Below is a sample generality stimulus set, which appears as participants would have seen it with the exception of the explanation labels in brackets:

Fred is a genetic engineer who creates plants and animals that facilitate agriculture, specializing in weed-eating gophers. Because many of his clients have large gardens, they prefer gophers who can eat many weeds before becoming satiated. Fred does not realize that his clients prefer gophers with large appetites because they eat more weeds, but he does notice that those gopher lineages with the highest metabolism are the best sellers. As a result Fred decides to exclusively engineer gophers with high metabolism, all of which are a success.

Fred has also noted that his clients prefer gophers with large noses. They prefer large noses because gophers with large noses are better at distinguishing the scent of weeds, which they should eat, from other plants that they should not eat. Fred does not realize that large noses are popular because gophers with large noses have a better sense of smell, but he does realize that large-nosed gophers sell best. As a result he decides to make all of his gophers large-nosed—a move that increases his sales.

Fred's clients also tend to purchase gophers with pointy claws, because they help destroy weeds not only by nibbling on them, but also by damaging the roots as they dig. Fred doesn't realize that pointy claws allow gophers to damage weed roots, but he does notice that the pointy claws are the most popular. As a result he decides to create all of his gophers with pointy claws.

Why do Fred's gophers have pointy claws?		***Circle choices:***
[teleological explanation]		
(A) Because the pointy claws damage weed roots.	y/n	− 1 2 3 4 5 6 7 +
[intention-based explanation]		
(B) Because that's the way Fred wanted them.	y/n	− 1 2 3 4 5 6 7 +
[mechanistic explanation]		
(C) Because Fred genetically modified the gophers to be like that.	y/n	− 1 2 3 4 5 6 7 +
[true-irrelevant filler]		
(D) Because gophers aren't marsupials.	y/n	− 1 2 3 4 5 6 7 +
[potentially false-irrelevant filler]		
(E) Because Fred doesn't like vegetables.	y/n	− 1 2 3 4 5 6 7 +

5.1.3. Design and procedures

Participants received a one-page questionnaire with instructions similar to those from Experiment 1. Each questionnaire contained a single *baseline scenario* or a single *generality* set of scenarios. After the last scenario, participants read the corresponding why-question and indicated the acceptability of candidate answers, along with the satisfaction ratings for accepted answers. Participants were randomly assigned to condition, and the answers were presented in one of several random orders.

5.2. Results

In the *baseline* condition, 58% of participants accepted the teleological explanation (see Table 4). This was comparable to the results from Experiments 1 and 2. For reasons that are unclear, acceptance rates for the intention based (69%) and mechanistic (75%)

Table 4
Explanation acceptance and satisfaction for Experiment 4

Explanation type	Acceptance		Satisfaction	
	Baseline	Generality	Baseline	Generality
Teleological	**58**	**81**	4.81 (0.41)	5.28 (0.29)
Intention-based	69	83	4.64 (0.34)	4.37 (0.32)
Mechanistic	**75**	**97**	5.11 (0.36)	5.97 (0.22)
True-irrelevant	11	3	1.25 (0.25)	2 (–)
Potentially false	0	0	–	–

The percent of participants accepting each explanation type, and the average satisfaction rating for accepted answers, are shown for the *baseline* condition as well as the *generality* condition. For satisfaction ratings, the standard error of the mean follows each average in parentheses. Significant differences are indicated in bold.

explanations were lower than in Experiment 1, where acceptance was 92% for both explanation types in the NICE scenarios. Satisfaction ratings for all explanation types in the *baseline* condition were comparable to ratings for the NICE scenarios from Experiment 1.

Of most interest, we examined whether TE acceptance was elevated in the *generality* condition compared with the *baseline* condition (see Table 4). A chi-square test of independence revealed a significant difference in the predicted direction ($\chi^2(1)=4.19$, $P<0.05$). This suggests that when participants had more evidence that a function-driven process corresponded to a predictable pattern, they were more likely to accept the teleological explanation. However, one concern is that participants may have increased TE acceptance for some other reason. For example, it could be that seeing three stories that each mentioned a function simply made the function invoked in the teleological explanation more salient. To test this alternative explanation, we showed an additional 18 participants a "*generality*" condition but for the accident scenarios rather than the NICE scenarios. These participants saw two similar accident scenarios before evaluating a third that was just like those seen by participants in Experiment 1. As in the NICE *generality* condition, a function was mentioned in each story, so any salience or pragmatic effect due simply to multiple functions was mimicked in this condition. But because the function invoked in the TE did not play a causal role in bringing about what was being explained, we would not expect an increase in TE acceptance in an accident *generality* condition compared to an accident *baseline* condition. Indeed, no increase in TE acceptance was observed. Three of the 18 participants accepted the TE in the accident *generality* condition, compared with 3 out of 12 for the same stimuli in Experiment 1. This difference was not significant.

We also found differences between conditions for the other explanation types. Specifically, participants accepted the mechanistic explanation significantly more often in the *generality* condition (97%) than in the *baseline* condition (75%). Of course, the repeated scenarios emphasized the generality of all relevant causal factors. There were no significant differences between conditions in satisfaction ratings for accepted answers.

5.3. Discussion

Experiment 4 examined whether providing evidence that a process is general—in the sense of belonging to a predictable pattern—increases TE acceptance. When participants saw three NICE scenarios that were highly similar they accepted TEs significantly more often than when seeing only one. This is compelling evidence that the generality of a process increases TE acceptance. The control experiment with repeated Accident scenarios supports this interpretation of the findings, as it demonstrates that the increase was not due to mere repetition of functions.

These data suggest that the greater acceptance of TEs for the Design scenarios of Experiment 1 than for the NICE scenarios is not due to a restriction, for some participants, to causal histories involving intentional design. When NICE scenarios are made to seem more general, TE acceptance rose to the rates of the Design scenarios of Experiment 1. In the General discussion we suggest an explanation for this finding: that by providing evidence that a NICE process is typically function-driven, the function becomes not only causally relevant, but also predictively useful.

6. Experiment 5: individual differences in teleological intuitions

Given the findings from Experiments 1–4, it appears that two conditions must obtain for a majority of participants to accept a teleological explanation. First, the function invoked in the explanation must have played a causal role in bringing about what is being explained, as per Wright's analysis. And second, the process by which the function played a causal role must belong to a process that is general in the sense of conforming to a predictable pattern. This suggests that individual differences in TE acceptance may be attributable to differing perceptions of the generality of a process. In Experiments 1, 2, and the *baseline* condition of Experiment 4, about half the participants tested accepted the TE for NICE scenarios. It could be that those participants who accepted the TE took the relevant process to be more general than those who did not. If this is the case, then participants who accepted the TE should also find a similar NICE story more plausible, as they take it to follow a predictable pattern. Experiment 5 investigates this prediction.

6.1. Methods

6.1.1. Participants

Participants were 36 Harvard students (12 male; mean age = 21, SD = 4) who completed the study in exchange for course credit or candy. An additional 6 students participated but were replaced as a result of failing to follow the directions: they provided satisfaction ratings for rejected explanations.

6.1.2. Materials

Three NICE stories from Experiment 1, with corresponding why-questions and candidate answers, were employed. Two were the biological scenarios from Experiment 4, with the third involving an artifact. In addition to these NICE stories with why-questions, participants rated the plausibility of a second NICE story, which was constructed to match the first as follows: it involved a NICE process undergone by the same agent in the same context, but with a different object and property. Below is an example of the stimuli a participant with the artifact scenario would have seen:

> ***[page one]***
> Carl manufactures and sells hats at a sporting club. Most of his clients are joggers who worry that car drivers don't see them when they jog at night, so they buy the brightest colored hats Carl sells. Carl doesn't realize that they buy these hats because they're attention grabbing, but he does notice that bright hats sell well. As a result he decides to make a series of glow in the dark hats by dying them with glow-in-the-dark dye. The new hats are an instant hit.

Why are Carl's hats glow-in-the-dark?

	Circle choices:
(A) Because hats are accessories.	y/n −1 2 3 4 5 6 7+
(B) Because they're attention grabbing at night.	y/n −1 2 3 4 5 6 7+
(C) Because Carl is worried that joggers might develop knee-problems.	y/n −1 2 3 4 5 6 7+
(D) Because Carl dyed them like that.	y/n −1 2 3 4 5 6 7+
(E) Because that's the way Carl wanted them.	y/n −1 2 3 4 5 6 7+

> ***[page two]***
> Another popular item in Carl's store is introductory yoga books. The joggers often buy such books to get ideas about how to stretch before and after running. Carl has no idea that the joggers, and not novice yoga enthusiasts, are behind the boost in introductory yoga book sales, but he notices that these books sell quickly. As a result he orders many such books to sell, all of which are a success.

How plausible is this scenario? (*circle one*):								
"not at all"	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**"very"**

6.1.3. Design and procedures

Participants received a two-page questionnaire, with instructions like those from Experiment 1. The first page contained a single NICE story with the corresponding why-question and candidate answers. The answers were presented in one of several random orders. Participants were explicitly told not to go on to the second page until the first page was completed. On the second page they saw the following instructions: "Given what you know from page one, please rate how plausible you think it is that the following scenario would occur. Use a scale from 1 to 7, where 1 corresponds to "not at all plausible", 7 corresponds to "very plausible", and so on for intermediate values".

6.2. Results

We first looked at acceptance of the various explanation types. The teleological explanation was accepted 72% of the time, which is higher than in Experiments 1, 2, and the baseline condition of 4, but not significantly so. The intention-based explanation was accepted 81% of the time and the mechanistic explanation 97%, comparable to the average results from the previous experiments.

A *t*-test comparing plausibility ratings as a function of TE acceptance revealed a significant difference in the predicted direction ($t(34)=1.88$, $P<0.05$, one-tailed). Specifically, participants who accepted the TE for the first scenario judged the second scenario significantly more plausible than those who did not accept the TE on the first scenario ($\mu=6.23$, SD = 1.03 versus $\mu=5.50$, SD = 1.08), consistent with the hypothesis that participants who accept the TE also find the causal process to be general. However, it is possible that some participants simply have a lower threshold for accepting answers than others. Such a tendency could lead to a higher probability of accepting the TE and of judging the second scenario plausible, whether or not there is of a real relationship between explanation and generality. To rule out this possibility, we also checked for positive relationships between acceptance of other explanation types and the plausibility rating of the second scenario. There were no effects for either the intention-based ($t(34)=0.698$, $P=0.094$, two-tailed) or the mechanistic ($t(34)=0.909$, $P=0.376$, two-tailed) explanation. In fact, there was a trend in the opposite direction for the intention-based explanation.

Even among participants who accepted the TE, the explanation may have seemed more or less satisfying as a function of how general they judged the process. Indeed, TE satisfaction ratings were significantly and positively correlated with the plausibility judgments ($N=26$, $r=0.400$, $P<0.05$, two-tailed). To rule out the possibility that individual participants simply tended to use the rating scales in consistent ways, we also looked for correlations between plausibility and the satisfaction ratings for other explanation. These correlations were neither sizable nor significant for either the intention-based ($N=29$, $r=-0.022$, $P=0.909$, two-tailed) or the mechanistic ($N=35$, $r=0.169$, $P=0.332$, two-tailed) explanations.

6.3. Discussion

Experiment 5 suggests that individual differences in TE acceptance are partially attributable to differing assumptions about the generality of a process. Participants who accepted the TE rated a similar scenario more plausible than those who did not, suggesting that they thought the causal process in the first story belonged to a more predictable pattern than participants who rejected the TE. The correlation between TE satisfaction and plausibility further suggests that even among those participants who accepted the TE, satisfaction reflected the extent to which the process was deemed general. These findings provide additional support for the hypothesis that generality contributes to the acceptability of teleological explanations.

7. General discussion

We began this paper by considering the commitments that underlie teleological explanations, and in particular how the conceptual domain of what is being explained, the causal history of what is being explained, and general constraints on explanation relate to the acceptability of teleological explanations. We found that adults accept teleological explanations when two conditions obtain: (a) the function invoked in the explanation played a causal role in bringing about what is being explained and (b) the process by which the function played a causal role seems general, in the sense that it conforms to a predictable pattern. Experiment 1 established the important role of (a), a prediction of Wright's theory, as judgments of TE acceptability and satisfaction were found to depend strongly on properties of the causal history responsible for what was being explained. The importance of (b), an unanticipated factor, was established in Experiments 3 and 4: processes known to be general were shown to license TEs, and providing evidence that a particular process is general increased TE acceptance. Experiment 5 confirmed the role of generality and further suggested that individual differences in teleological intuitions were due to differences in the extent to which participants took the relevant processes to be general.

These findings provide some support for a descriptive re-interpretation of Wright's account of teleological explanation, according to which processes with a consequence etiology are both necessary and sufficient to yield functions that are acceptable as teleological explanations. For our participants this requirement was indeed necessary, but not sufficient: the causal process also had to be general. We failed to find support for other proposals requiring that the object being explained belong to a particular domain, or result from a design etiology. However, the generality of the causal process interacts in interesting ways with these proposals. Teleological explanations are typically used in domains that have a dominant process with a consequence etiology, such as intentional design for artifacts or natural selection for biology. Because the generality requirement will always be met in such canonical cases, teleological explanations will be common. But for other domains, in which consequence etiologies are rare, the generality requirement will rarely be met. Hence teleological explanations will be uncommon, even when a consequence etiology is present. Generality may thereby exaggerate differences in the acceptability of teleological explanations across domains.

7.1. The development of teleological intuitions

Our results suggest several potential sources of developmental, historical, and cultural difference in teleological intuitions. One possibility is that such variation is due to differences in what constitutes an explanation—that is, that causal history and generality are not always the relevant criteria. This possibility is most plausible for the case of developmental change, where evidence from children's understanding of artifacts suggests an initial, ahistorical notion of function, which only becomes historically nuanced around age 5 (e.g. German & Johnson, 2002; Kelemen & Carey, in press; Matan & Carey, 2001). If this is the case, then children may not understand the importance of causal history in teleological explanations, and this understanding is what changes with age.

Another possibility is that the explanatory relation remains fixed, but beliefs about causal history change. If this is the correct developmental story, then we would expect children's promiscuous acceptance of teleological explanations to result from anomalous beliefs about causal origins: beliefs that any property or object with a plausible function resulted from a process with a consequence etiology, consistent with the findings reported in Kelemen and DiYanni (in press). Similarly, Aristotle's endorsement of teleological explanations for particular physical movements and biological properties would commit him to certain beliefs about the ontogeny of what he sought to explain. This possibility makes the strong prediction that cross-cultural differences in teleological intuitions necessarily entail differing assumptions about the causal structure of the world. In particular, only those processes believed to operate in a function-driven manner should warrant teleological explanations.

A final possibility is that either the domain of what is being explained or a design etiology is privileged for some populations. For example, children may initially constrain their application of teleological explanations on the basis of domain, and only later come to understand the causal commitments underlying teleological explanations. Another possibility is that early understanding of teleological explanation is historical, but conforms to a design etiology. This possibility could also plausibly account for historical and cultural variation. Before Darwin, few non-intentional processes with a consequence etiology were well characterized. Indeed, the difficulty in understanding natural selection for contemporary adults (e.g. Evans, 2000; Shtulman, in preparation) suggests that consequence etiologies that are not also design etiologies present a special challenge. For most people, education and exposure may be the only way to generalize an understanding of design etiologies to encompass non-intentional cases, and in turn free teleological explanations from intentional entailments.

7.2. Explanation for export: a framework for explanation

Why might teleological intuitions depend both on the causal role of the function and on the generality of the causal process? While these factors nicely account for our data, they are somewhat unparsimonious and post hoc. However, they motivate a more cohesive proposal about the nature of explanation, which extends beyond our data on teleological explanation. This proposal, Explanation for Export, is a hypothesis about the psychological function of explanation in general, and predicts the effects of both causal history and generality.

Explanation for Export (EFE) claims that the function of explanation is to provide the kind of information likely to subserve future intervention and prediction—that is, to be exportable to novel cases. The causal requirements articulated by Wright's account follow from the need for intervention, which entails appreciating causal and not merely inferential relationships. However, explanations typically involve only a small fraction of the causes responsible for what is being explained; in explaining the cause of a forest fire one is likely to mention the unattended campfire and not the presence of oxygen, no matter that both are necessary causal factors. EFE's insistence on predictively useful information suggests which causes will be selected: those that are likely to subserve prediction by making a difference in future cases. Such causes must vary across contexts, covary with the effect of interest, and appear sufficiently often to warrant notice. A good way to identify such causes is to establish whether a particular process fits a familiar pattern, and to determine which kinds of roles are typical of the pattern (e.g. function, enabling-condition) but vary across contexts such that specifying what instantiates the role (e.g. a particular goal or oxygen) is informative.

Explanation for Export predicts our findings for teleological explanation as follows. Given that causal information is privileged as a result of the need for intervention, functions will only be legitimate explanations when they were also causes of what is being explained. However, having played a causal role is not sufficient: there must be reason to think the function is like the campfire rather than the oxygen. This additional requirement will be met when the function is predictively useful. A good cue for the predictive utility of a function is the extent to which the relevant process has previously occurred in a function-driven manner. If a causal process conforms to a function-driven causal pattern, then the function can be expected to make a difference in future cases. Thus teleological explanations should be restricted to cases where the function not only played a causal role, but did so via a causal process that conforms to a predictable pattern.

7.3. Relationship to previous work

Explanation for Export is related to recent accounts of explanation from philosophy (e.g. Strevens, 2004; Woodward, 2003), in structure if not motivation. In particular, EFE shares elements with Michael Strevens's Kairetic account of explanation, which requires explanations to cite causes that are "difference-makers" in the sense of having made a difference to whether or not what is being explained came about (Strevens, 2004). While several accounts of explanation emphasize difference-makers (e.g. Lewis, 1986), Strevens's account is unique for the manner in which difference-makers are identified. He proposes that difference-makers can be isolated by borrowing aspects of the criteria that make for good explanations according to unification accounts of explanation. In particular, he requires that the causal models invoked in explanations be as general as possible—that is, be satisfied by a maximum number of physically possible systems. As in EFE, the generality requirement provides a way to specify when a particular cause is explanatory. Unlike EFE, however, Strevens' and other philosophical accounts do not motivate the constraints on explanation by appeal to the psychological function or utility of explanation.

Consistent with the motivation for EFE, researchers in artificial intelligence, cognitive development, and education have suggested that the quality of an explanation depends on its future utility. Pearl (2001) writes that "the sense of 'comprehensibility' that accompanies an adequate explanation is a natural byproduct of the transportability of (and hence our familiarity with) the causal relationships used in the explanation" (p. 26). Pearl's notion of "transportability" is analogous to the idea that a good explanation should be exportable. Also from artificial intelligence, methods known as explanation-based learning and explanation-based generalization stem from the idea that explanation can serve as a mechanism for generalization to new cases (e.g. Lewis, 1988).

In developmental psychology and education, researchers have focused on the role of explanation in learning (e.g. Chi, DeLeeuw, Chiu, & LaVancher, 1994;

Neumann & Schwartz, 1998; Neumann & Schwartz, 2000). For example, in a paper provocatively titled "Explanation as Orgasm and the Drive for Causal Knowledge", Alison Gopnik (2000) defends the view that the satisfaction accompanying an explanation is the reward for exercising our theory-building capacities. As orgasm is to reproduction, so the satisfaction of explaining is to the causal representation of the world. EFE complements this view by suggesting the properties of explanations in virtue of which they may contribute to functions subserved by folk theories, like the prediction and control of the environment. While Gopnik's view suggests that explanations are satisfying because they play a role in theory formation, EFE spells out what that role is. As a result, the feeling of satisfaction accompanying an explanation can be distinguished from its contribution to theory-formation, though in practice the two may to be tightly coupled.

EFE explicitly states an idea running through these claims from AI and psychology, and presents a working hypothesis about the nature of explanation. If a view along these lines proves empirically adequate and amenable to more precise formulation, it would not only describe human explanatory intuitions, but also reveal their function: it would provide a teleological explanation of explanation.

7.4. Conclusions

Our findings suggest that the selective application of teleological explanation reflects deep underlying commitments about the causal structure of the world and the kinds of processes that typically occur within it. We have presented five experiments that support a theory of teleological explanation largely consistent with the causal theory of Wright (1976), but suggesting an additional role for the generality of a process. The findings motivate a broad hypothesis about explanation, Explanation for Export (EFE), which suggests that the function of explanation is to highlight information likely to subserve future prediction and intervention.

Acknowledgements

The authors would like to thank Elizabeth Allen, Tom Griffiths, and Michael Weisberg for helpful comments on earlier drafts; Elisabeth Camp for comments on a related paper; and Peter Godfrey–Smith for helpful discussions as well as for suggesting the name "Explanation for Export". This work was supported by an NDSEG graduate fellowship awarded to the first author. Conditions from Experiments 1, 3, and 4 were presented at the 25th annual conference of the Cognitive Science Society. These findings were also presented at the first annual joint conference of the Society for Philosophy and Psychology and the European Society for Philosophy and Psychology.

Appendix. Stimuli from Experiment 1

Below is the complete set of stimuli used in Experiment 1, with the causal history, domain, and version of each scenario labeled in brackets. The why-question and answers corresponding to each set of stimuli follow with the explanation type, also labeled in brackets.

[**History**: Design, **Domain**: Artifact, **Version**: Size]

Sally is an engineer designing satellite dishes. She wants them to be highly sensitivity, so she makes them with large quantities of metal to yield a large dish size.

[**History**: NICE, **Domain**: Artifact, **Version**: Size]

Sally is an engineer designing satellite dishes. The best-selling dishes are the largest ones, because they are the most sensitive. Although Sally doesn't realize that the large ones sell better because of their sensitivity, she decides to make all of her dishes large to meet demand. As a result she makes her dishes with large quantities of metal, yielding a large dish size.

[**History**: Accident, **Domain**: Artifact, **Version**: Size]

Sally is an engineer designing satellite dishes. When making new dishes she accidentally uses very large quantities of metal, yielding a large dish size–much larger than she intended. However, larger dishes are more sensitive, so she doesn't have any trouble selling them.

Why are Sally's satellite dishes large?

[Teleological] Because larger dishes are more sensitive.
[Intention-based] Because that's the way Sally wanted them.
[Mechanistic] Because Sally [accidentally] made them from large quantities of metal.
[True, irrelevant] Because technology is useful.
[Potentially false, irrelevant] Because Sally likes plastic.

[**History**: Design, **Domain**: Biological Part, **Version**: Size]

Sally is a genetic engineer trying to create cats who catch mice more effectively, so she produces cats who have large ears and hence hear mice from a farther distance. This allows the cats to find and catch mice more easily.

[**History**: NICE, **Domain**: Biological Part, **Version**: Size]

Sally is a genetic engineer who creates cats as household pets. Many of her clients buy cats in order to catch mice in their homes, so they tend to prefer the cats who have large ears, and hence hear and catch mice more effectively. Sally doesn't realize that large ears are popular because they help the cats catch mice, but she does notice that large ears are the common favorite. As a result she decides to genetically engineer only cats with large ears.

[**History**: Accident, **Domain**: Biological Part, **Version**: Size]

Sally is a genetic engineer who creates cats as household pets. She accidentally introduces a gene sequence that results in cats with large ears. However, cats with large ears are better at hearing and catching mice in homes, so her clients are pleased and the new cats are a hit.

Why do Sally's cats have large ears?

[Teleological] Because cats with larger ears catch mice more effectively.
[Intention-based] Because that's the way Sally wanted them.
[Mechanistic] Because Sally [accidentally] genetically modified them to be that size.

[True, irrelevant] Because mice have whiskers.
[Potentially false, irrelevant] Because Sally likes dogs.

[**History**: Design, **Domain**: Non-biological natural kind, **Version**: Size]
Sally owns a canyon park that attracts tourists who come to have picnics in the many caves. The best caves are those that are large enough to produce an audible echo, so she enlarges all the caves so that they produce louder echoes.
[**History**: NICE, **Domain**: Non-biological natural kind, **Version**: Size]
Sally owns a canyon park that attracts tourists who come to have picnics in the many caves. The best caves are those that are large enough to produce an audible echo, so tourists tend to prefer these. Sally doesn't realize that these are popular because of the echo, but she does notice that larger caves are everyone's favorites. As a result she decides to enlarge all the caves.
[**History**: Accident, **Domain**: Non-biological natural kind, **Version**: Size]
Sally owns a canyon park that attracts tourists who come to have picnics in the many caves. When reinforcing the caves to make them safer, Sally accidentally enlarges them, yielding only large caves. However, large caves are more popular because they produce audible echoes, so the tourists love the modified caves.

Why are Sally's caves large?
[Teleological] Because larger caves produce audible echoes.
[Intention-based] Because that's the way Sally wanted them.
[Mechanistic] Because Sally [accidentally] enlarged them.
[True, irrelevant] Because tourists go site seeing.
[Potentially false, irrelevant] Because Sally dislikes sandwiches.

[**History**: Design, **Domain**: Artifact, **Version**: Form]
Fred is a gardener designing tools to eliminate weeds. He wants to be able to remove weeds permanently by destroying their roots, so he builds a tool by sharpening the tip of a spade. This yields a pointy end for damaging roots as it digs.
[**History**: NICE, **Domain**: Artifact, **Version**: Form]
Fred is a gardener who sells spades with tips that he modifies into all sorts of shapes. His most popular tools are the spades with the pointiest tips, because they destroy the roots of weeds as they dig. Fred doesn't realize that they're popular because the pointiness destroys weed roots, but he does notice that the pointy tools sell best. As a result, Fred decides to manufacture only pointy tools, so he sharpens the tips of spades to make them pointy.
[**History**: Accident, **Domain**: Artifact, **Version**: Form]
Fred is a gardener who sells spades with tips that he modifies into all sorts of shapes. When modifying a new set of spade tips he accidentally sharpens them, making the tips pointy. However, the pointy tips damage roots as they're used for digging up weeds, so they're an instant hit.

Why are Fred's tools pointy?

[Teleological] Because the pointy end damages weed roots.
[Intention-based] Because that's the way Fred wanted them.
[Mechanistic] Because Fred [accidentally] sharpened the tips of the spades.
[True, irrelevant] Because weeding eliminates weeds.
[Potentially false, irrelevant] Because Fred doesn't like vegetables.

[**History**: Design, **Domain**: Biological part, **Version**: Form]
Fred is a genetic engineer who creates plants and animals that facilitate agriculture. He wants to help eliminate weeds, so he modifies a weed-eating gopher by giving it pointy claws. This way the gophers not only disturb weeds by digging and nibbling on them, but also damage the roots as they dig.
[**History**: NICE, **Domain**: Biological part, **Version**: Form]
Fred is a genetic engineer who creates plants and animals that facilitate agriculture, specializing in weed-eating gophers. His clients tend to purchase gophers with pointy claws, because they help destroy weeds not only by nibbling on them, but also by damaging the roots as they dig. Fred doesn't realize that pointy claws allow gophers to damage weed roots, but he does notice that the pointy claws are the most popular. As a result he decides to create all of his gophers with pointy claws.
[**History**: Accident, **Domain**: Biological part, **Version**: Form]
Fred is a genetic engineer who creates plants and animals that facilitate agriculture, specializing in weed-eating gophers. When creating new gophers he accidentally introduces a gene sequence that results in gophers with pointy claws. However, the pointy claws help damage weed roots as the gophers dig, so they're an instant hit.

Why do Fred's gophers have pointy claws?

[Teleological] Because the pointy claws damage weed roots.
[Intention-based] Because that's the way Fred wanted them.
[Mechanistic] Because Fred [accidentally] genetically modified the gophers to be like that.
[True, irrelevant] Because gophers aren't marsupials.
[Potentially false, irrelevant] Because Fred doesn't like vegetables.

[**History**: Design, **Domain**: Non-biological natural kind, **Version**: Form]
Fred landscapes gardens with rocks and other natural materials. He doesn't want birds to perch on the rocks, so he decided to make them pointy by breaking them into pointy shards. The pointiness effectively keeps birds from perching on the rocks.
[**History**: NICE, **Domain**: Non-biological natural kind, **Version**: Form]
Fred sells landscaping materials for gardens, especially rocks. Many of his clients don't like birds perching on the rocks they buy for their gardens, so they tend to buy the pointiest rocks, as the pointiness keeps birds off. Fred doesn't realize that the pointy rocks keep birds from perching on them, but he does notice that pointy rocks are the most popular. As a result he decides to break rocks into pointy shards and sell only pointy rocks.

[**History**: Accident, **Domain**: Non-biological natural kind, **Version**: Form]
Fred sells landscaping materials for gardens, especially rocks. When modifying his rocks he accidentally makes them very pointy on top. However, many of his clients prefer pointy rocks because they keep birds from perching on them, so the pointy rocks are a hit.

Why are Fred's rocks pointy?
[Teleological] Because the pointiness keeps birds off.
[Intention-based] Because that's the way Fred wanted them.
[Mechanistic] Because Fred [accidentally] broke rocks into pointy shards.
[True, irrelevant] Because birds have wings.
[Potentially false, irrelevant] Because Fred doesn't like vegetables.

[**History**: Design, **Domain**: Artifact, **Version**: Color]
Carl is worried that car drivers don't see people who are jogging at night, so he dyes normal hats with a glow-in-the-dark color to create "jogging hats" that are very bright and attention grabbing at night.
[**History**: NICE, **Domain**: Artifact, **Version**: Color]
Carl manufactures and sells hats at a sporting club. Most of his clients are joggers who worry that car drivers don't see them when they jog at night, so they buy the brightest colored hats Carl sells. Carl doesn't realize that they buy these hats because they're attention grabbing, but he does notice that bright hats sell well. As a result he decides to make a series of glow in the dark hats by dying them with glow-in-the-dark dye. The new hats are an instant hit.
[**History**: Accident, **Domain**: Artifact, **Version**: Color]
Carl manufactures and sells hats at a sporting club. When dying a new series of hats he accidentally uses glow-in-the-dark dye instead of the black dye he intended, resulting in glow-in-the-dark hats. However, most of his clients are joggers who worry that car drivers don't see them when they jog at night, so they prefer bright hats. The glow-in-the-dark hats are thus an instant hit.

Why are Carl's hats glow-in-the-dark?
[Teleological] Because they're attention grabbing at night.
[Intention-based] Because that's the way Carl wanted them.
[Mechanistic] Because Carl [accidentally] dyed them like that.
[True, irrelevant] Because hats are accessories.
[Potentially false, irrelevant] Because Carl is worried that joggers might develop knee-problems.

[**History**: Design, **Domain**: Biological Part, **Version**: Color]
Carl is worried that car drivers don't see people who are walking their dogs at night, so he genetically engineers dogs with glow-in-the-dark noses that are very bright and attention grabbing.
[**History**: NICE, **Domain**: Biological Part, **Version**: Color]
Carl is a genetic engineer who works on designer pets, specializing in dogs. Because

potential dog owners worry about not being seen by car drivers when walking their dogs at night, they tend to choose dogs with brightly colored noses. Carl doesn't realize that they choose such noses because they're attention grabbing, but he does notice that bright noses are popular. As a result he decides to engineer dogs with glow-in-the-dark noses. The new dogs are an instant hit.

[**History**: Accident, **Domain**: Biological Part, **Version**: Color]

Carl is a genetic engineer who works on designer pets, specializing in dogs. When modifying a new dog genome he accidentally inserts a sequence that results in dogs with glow-in-the-dark noses. However, because potential dog owners worry about being seen by car drivers when walking their dogs at night, they prefer dogs with bright noses that are attention grabbing. Thus the dogs with glow-in-the-dark noses are an instant hit.

Why do Carl's dogs have glow-in-the-dark noses?

[Teleological] Because they're attention grabbing at night.
[Intention-based] Because that's the way Carl wanted them.
[Mechanistic] Because Carl [accidentally] genetically modified them to be like that.
[True, irrelevant] Because dogs have an excellent sense of smell.
[Potentially false, irrelevant] Because Carl likes cats.

[**History**: Design, **Domain**: Non-biological natural kind, **Version**: Color]

Carl is worried that car drivers don't see people who are taking walks at night, so he creates a glow-in-the-dark isotope of a common element in dirt that can be used along dirt paths near roads. The result is glow-in-the-dark dirt paths that are very bright and attention grabbing.

[**History**: NICE, **Domain**: Non-biological natural kind, **Version**: Color]

Carl modifies materials to be used for outdoor trails. Because those who buy his materials worry that people taking walks at night won't be seen by car drivers, they tend to choose the brightest materials for walking trails. Carl doesn't realize that they choose such materials because they're attention grabbing, but he does notice that bright materials are most popular. As a result he decides to create a glow-in-the-dark isotope from one of the common materials in dirt, resulting in attention grabbing glow-in-the-dark dirt paths that are an instant hit.

[**History**: Accident, **Domain**: Non-biological natural kind, **Version**: Color]

Carl modifies materials to be used for outdoor trails. When modifying the chemical composition of dirt for walking trails, he accidentally creates a glow-in-the-dark isotope of one of the common elements, resulting in glow-in-the-dark dirt for trails. However, because city planners worry about pedestrians being seen by car drivers when taking walks at night, they tend to prefer the brightest materials. Thus the glow-in-the-dark dirt is an instant hit.

Why does Carl's dirt glow in the dark?

[Teleological] Because it's attention grabbing at night.
[Intention-based] Because that's the way Carl wanted it.

[Mechanistic] Because Carl [accidentally] created an isotope to make dirt like that.
[True, irrelevant] Because pedestrians have right-of-way.
[Potentially false, irrelevant] Because Carl hates chemistry.

References

Atran, S. (1994). Core domains versus scientific theories: Evidence from systematics and Itza-Maya folkbiology. In L. A. Hirschfeld, & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 316–340). Cambridge, England: Cambridge University Press.

Atran, S. (1995). Causal constraints on categories. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multi-disciplinary debate* (pp. 205–233). Oxford, England: Clarendon Press.

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

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

Casler, K., & Kelemen, D. (2003). *Teleological explanation and conceptual change: Reasoning about nature among Romanian Roma (Gypsy) adults. Poster presented at the meeting of the cognitive development society, Park City, UT*.

Chi, M. T. H., DeLeeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanation improves understanding. *Cognitive Science*, *13*, 145–182.

Cummins, R. (1975). Functional analysis. *The Journal of Philosophy*, *72*(20), 741–760.

Evans, E. M. (2000). Beyond scopes: Why creationism is here to stay. In K. Rosengren, C. Johnson, & P. Harris (Eds.), *Imagining the impossible: The development of magical, scientific and religious thinking in contemporary society* (pp. 305–333). Cambridge, England: Cambridge University Press.

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

Gelman, S. A., Coley, J. D., & Gottfried, G. M. (1994). Essentialist beliefs in children: The acquisition of concepts and theories. In L. A. Hirschfeld, & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 341–366). Cambridge, England: Cambridge University Press.

German, T., & Johnson, S. A. (2002). Function and the origins of the design stance. *Journal of Cognition and Development*, *3*, 279–300.

Gopnik, A. (2000). Explanation as orgasm and the drive for causal knowledge: The function, evolution, and phenomenology of the theory-formation system. In F. Keil, & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 299–324). Cambridge, MA: MIT Press.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, *15*, 135–175.

Keil, F. C. (1992). The origins of an autonomous biology. In M. R. Gunnar, & M. Maratsos, *Modularity and constraints in language and cognition*. *Minnesota symposium on child psychology* (Vol. 25) (pp. 103–138). Hillsdale, NJ: Earlbaum.

Keil, F. C. (1994). The birth and nurturance concepts by domains: The origins of concepts of living things. In L. A. Hirschfeld, & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 234–254). Cambridge, England: Cambridge University Press, 234–254.

Keil, F. C. (1995). The growth of causal understanding of natural kinds. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multi-disciplinary debate* (pp. 234–262). Oxford, England: Clarendon Press, 234–262.

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

Kelemen, D. (1999b). Why are rocks pointy? Children's preference for teleological explanations of the natural world. *Developmental Psychology*, *35*(6), 1440–1452.

Kelemen, D. (1999c). Function, goals and intention: Children's teleological reasoning about objects. *Trends in Cognitive Science*, *3*(12), 461–468.

Kelemen, D. (1999d). Beliefs about purpose: On the origins of teleological thought. In M. Corballis, & S. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 278–294). Oxford, England: Oxford University Press, 278–294.

Kelemen, D. (2003). British and American children's preferences for teleo-functional explanations of the natural world. *Cognition*, *88*(2), 201–221.

Kelemen, D. (2004). Are children "intuitive theists?": Reasoning about purpose and design in nature. *Psychological Science*, *15*(5), 295–301.

Kelemen, D., & Carey, S. (in press). The essence of artifacts: Developing the design stance. In S. Laurence, & E. Margolis (Eds.), *Creations of the mind: Theories of artifacts and their representation*. Oxford, England: Oxford University Press.

Kelemen, D., & DiYanni, C. (in press). Intuitions about origins: Purpose and intelligent design in children's reasoning about nature. *Journal of Cognition and Development*.

Lewis, C. (1988). Why and how to learn why: Analysis-based generalization of procedures. *Cognitive Science*, *12*, 211–256.

Lewis, D. (1986). *Causal explanation*, Vol. 2. Oxford, England: Oxford University Press.

Matan, A., & Carey, S. (2001). Developmental changes within the core of artifact concepts. *Cognition*, *78*, 1–26.

Medin, D.L., & Atran, S. (in press). The native mind: Biological categorization, reasoning and decision making in development across cultures. *Psychological Review*.

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

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

Neumann, Y., & Schwartz, B. B. (1998). Is self-explanation while solving problems helpful? The case of analogical problem-solving. *British Journal of Educational Psychology*, *68*, 15–24.

Neuman, Y., & Schwarz, B. B. (2000). Substituting one mystery for another: The role of self-explanations in solving algebra word-problems. *Learning and Instruction*, *10*, 203–220.

Pearl, J. (2001). *Causality: Models, reasoning, and inference*. Cambridge, England: Cambridge University Press.

Salmon, W. C. (1989). *Four decades of scientific explanation*. Minneapolis, MN: University of Minnesota Press.

Shtulman, A. (under review). *Evidence of incommensurability between naive and scientific theories of evoution*. Manuscript in preparation.

Slaughter, V., Jaakkola, K., & Carey, S. (1999). Constructing a coherent theory: Children's biological understanding of life and death. In M. Siegler, & C. Peterson (Eds.), *Children's understanding of biology and health* (pp. 71–98). Cambridge: Cambridge University Press.

Strevens, M. (in preparation). *The Kairetic Account of Explanation*. Book manuscript in preparation.

Strevens, M. (2004). The causal and unification accounts of explanation unified—causally. *Noûs*, *38*, 154–176.

Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, *43*, 337–375.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford, England: Oxford University Press.

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