Cognitive Psychology 107 (2018) 22–43

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

# Structure-function fit underlies the evaluation of teleological explanations

Emily G. Liquin*, Tania Lombrozo

*Department of Psychology, Princeton University, Princeton, NJ 08540 USA*

ARTICLE INFO

*Keywords:*
Explanation
Teleology
Causal reasoning

ABSTRACT

Teleological explanations, which appeal to a function or purpose (e.g., "kangaroos have long tails for balance"), seem to play a special role within the biological domain. We propose that such explanations are compelling because they are evaluated on the basis of a salient cue: structure-function fit, or the correspondence between a biological feature's form (e.g., tail length) and its function (e.g., balance). Across five studies with 852 participants in total, we find support for three predictions that follow from this proposal. First, we find that function information decreases reliance on mechanistic considerations when evaluating explanations (Experiments 1–3), indicating the presence of a salient, function-based cue. Second, we demonstrate that structure-function fit is the best candidate for this cue (Experiments 3–4). Third, we show that scientifically-unwarranted teleological explanations are more likely to be accepted under speeded and unspeeded conditions when they are high in structure-function fit (Experiment 5). Experiment 5 also finds that structure-function fit extends beyond biology to teleological explanations in other domains. Jointly, these studies provide a new account of how teleological explanations are evaluated and why they are often (but not universally) compelling.

## 1. Introduction

When the French naturalist Sonnerat described the peculiar primate that he encountered in Madagascar in the early 1780s, he emphasized its thin, elongated middle finger (Owen, 1863). We now know that this creature – the aye-aye – taps along logs to find grubs, and then uses its skinny finger to extract them. But local legends tell a different story. For example, the Sakalava people believe that the aye-aye enters houses through thatched roofs during the night to murder the sleeping humans within, using its elongated finger to cut the aortic vein of its victim (Goodman & Schütz, 2000).

What these evolutionary and traditional explanations for the aye-aye's elongated finger have in common is an appeal to *function*. They both support teleological explanations for a biological trait: that the aye-aye has a long finger to extract grubs, or that the aye-aye has a long finger to cut aortic veins. Such explanations, which appeal to a purpose, function, or goal, contrast with mechanistic explanations, which instead explain the presence of some trait by appeal to parts or causal processes, such as genes or natural selection.

While most agree that biological adaptations (such as the aye-aye's finger) can in fact be explained teleologically, mechanistically, or with both forms of explanation at once (e.g., Allen, 2009; Ayala, 2007), there's evidence that teleological explanations are favored in the biological domain. In explaining biological traits, both children and adults typically favor teleological explanations (e.g., that eyes exist "so people and animals can see") over mechanistic alternatives (e.g., that eyes exist "because bodies have special cells that

* Corresponding author.
*E-mail addresses:* eliquin@princeton.edu (E.G. Liquin), lombrozo@princeton.edu (T. Lombrozo).

https://doi.org/10.1016/j.cogpsych.2018.09.001
Accepted 28 September 2018
Available online 12 October 2018

combine to produce eyes"; Kelemen & DiYanni, 2005; Lombrozo, Kelemen, & Zaitchik, 2007). Moreover, across domains, there's evidence that teleological explanations may be less cognitively demanding: they are accepted more often under speeded conditions (Kelemen & Rosset, 2009; Kelemen, Rottman, & Seston, 2013) and by people who engage in less reflective thinking (Zemla, Steiner, & Sloman, 2016). These findings provide some empirical support for a claim by Richard Dawkins: that "we humans have purpose on the brain. We find it difficult to look at anything without wondering what it is 'for,' what the motive for it or the purpose behind it might be" (Dawkins, 1995, p. 81).

In the current paper, we explore one hypothesis for why teleological explanations play a special role in explaining biological adaptations. The motivating idea, suggested by Lombrozo et al. (2007), is that a good "fit" between a biological trait and some function provides a cue to the adequacy and quality of a teleological explanation. So, for example, the "fit" between an elongated finger and grub-extraction provides support for the idea that the aye-aye's finger evolved to extract grubs, and can thus be explained by appeal to this function. Of course, this inference is imperfect (and thus defeasible): noses have a good fit to the function of holding up glasses, but they did not evolve to do so, and thus cannot be explained by appeal to this function. So while good structure-function fit may offer support for a teleological explanation, additional reflection (and in some cases, additional causal knowledge) will sometimes reveal that a teleological explanation is in fact unwarranted.

Across five experiments, we evaluate our hypothesis that structure-function fit provides a basis for evaluating teleological explanations. Before turning to our experiments, however, we provide a brief review of past research to motivate the three specific predictions of this hypothesis that we go on to test.

### *1.1. Teleological versus mechanistic explanations*

The notion that people explain the world through different lenses has roots in Aristotle's four causes – the material cause, the formal cause, the efficient cause, and the final cause. The latter two correspond roughly to the distinction between mechanistic explanations and teleological explanations: the efficient cause is that which brings about the thing being explained, and the final cause provides the purpose for which a thing exists (Falcon, 2015). Dennett (1971) makes a related distinction in characterizing multiple "stances" that one can adopt in understanding and predicting a system's behavior: the physical stance allows interpretation in terms of causal mechanisms, while the design stance allows interpretation in terms of function and design. Keil (1994, 1995) similarly argues for multiple "modes of construal," available even to young infants, by which we interpret the world.

Research has demonstrated the psychological reality of these different "stances" or "modes" as reflected in different forms of explanation. *Mechanistic explanations*, which refer to the causal mechanism producing the thing to be explained, and *teleological explanations*, which refer to the function or purpose of the thing to be explained, have been shown to involve different types of reasoning and to produce different downstream cognitive effects. In particular, mechanistic and teleological explanations involve different causal commitments (Lombrozo & Carey, 2006), and encourage different criteria for causal ascription (Lombrozo, 2010). They also support different patterns of categorization (Lombrozo, 2009; see also Ahn, 1998) and generalization (Lombrozo & Gwynne, 2014), align with different goals (Vasilyeva, Wilkenfeld, & Lombrozo, 2017), and are differentially memorable across domains (Gwynne & Lombrozo, 2010).

Although teleological and mechanistic explanations have unique psychological profiles, they both seem to be treated as fundamentally causal. Within philosophy, Wright (1976) proposed that teleological explanations are implicitly causal, where the function they appeal to must have played a causal role in bringing about the thing being explained (see also Allen, 2009). As an example, we might explain why aye-ayes have a long, skinny finger by appeal to its role in extracting grubs. This is appropriate because long, skinny fingers were maintained and spread within the ancestral population of aye-ayes *because* this trait facilitated grub extraction (and thus led to differential reproductive success). Lombrozo and Carey (2006) examined whether people take this criterion into account when judging the acceptability of a teleological explanation. They found that people in fact rely on two criteria to judge the acceptability of a teleological explanation: first, the function must have played a causal role in bringing about the thing being explained (consistent with Wright's proposal) and second, the process by which the function did so must be perceived as generalizable.

### *1.2. The allure of teleological explanations*

As the quote by Dawkins suggests, teleological explanations may hold some special, intuitive appeal. This perspective has been advocated most forcefully by Kelemen and colleagues, who have developed a position called "promiscuous teleology." Descriptively, the claim is that teleological explanations are not restricted to particular circumscribed domains, but instead are applied and favored "promiscuously" across domains. Much of this work uses endorsement of "scientifically unwarranted" teleological explanations – such as "rocks are pointy so that animals won't sit on them" – to measure teleological bias. Children across several cultures robustly endorse these unwarranted teleological explanations (Kelemen, 1999c, 1999d, 2003; Schachner, Zhu, Li, & Kelemen, 2017), and they are also quite prevalent among adults in other cultures who have not been exposed to western education (Casler & Kelemen, 2008; Sánchez Tapia et al., 2016).

At a processing level, promiscuous teleology could be underwritten by a "developmentally persistent cognitive default" to interpret the world in terms of functions (Kelemen et al., 2013, p. 1075). One source of evidence for this idea is that under speeded conditions, adults make the mistake of accepting unwarranted teleological explanations rather than (for instance) rejecting warranted teleological explanations (Kelemen & Rosset, 2009; Rottman et al., 2017). Even professional physical scientists exhibit this pattern of errors (Kelemen et al., 2013). Endorsement of unwarranted teleological explanations is also more widespread among people who engage in less reflective reasoning (Zemla et al., 2016), which suggests that teleological explanations result from less effortful cognitive processing, and may reflect a more intuitive mode of thought.

On Kelemen's view (see also Kelemen, 1999b, 1999a), children initially come to understand teleology in terms of human intentional action. This understanding is quickly overextended, such that all entities with which humans interact in a goal-directed manner are construed as designed for that goal (e.g., "mountains are tall because they're for climbing"). Nonetheless, some selectivity in the scope of teleological explanations can be achieved as causal beliefs about the world are revised. In particular, the selectivity found in adulthood could relate to causal beliefs about Mother Nature or intelligent design (Diesendruck & Haber, 2009; Järnefelt, Canfield, & Kelemen, 2015; Kelemen & DiYanni, 2005): these beliefs have been linked to individual differences in endorsement of unwarranted teleological explanations in both children and adults, presumably because those who believe in the existence of an agent driving the forces of nature have a basis upon which to attribute functions and purposes to non-living natural entities. Additionally, different causal beliefs in different domains could explain variation across domains in teleological preferences.

In contrast to "promiscuous teleology" is a position sometimes referred to as "selective teleology" (also referred to as "discriminative teleology" by Greif, Kemler Nelson, Keil, & Gutierrez, 2006), according to which the scope of teleological thinking is more restrictive (Kelemen, 1999b, 1999a, 1999d, 1999c; Sánchez Tapia et al., 2016). Supporting a more circumscribed role for teleological explanations are the findings that adults are quick to reject "bad" teleological explanations (e.g., "animals grow ears because they need to smell things"; Kelemen & Rosset, 2009; Kelemen et al., 2013), and that young children restrict their *questions* about functions to the proper domains (Greif et al., 2006). For example, children will ask what a long tail "is for," but not what a kangaroo "is for."

Recent work paints an even more complex picture of the selectivity of teleological reasoning: distinct differences in teleological reasoning can be found across ages, cultures, and domains (Sánchez Tapia et al., 2016). For instance, teleological explanations for biological traits and actions are given more frequently by adults than by children, and by US participants than by participants from a Quechua-speaking region in Peru. What is consistent across both promiscuous and selective teleological positions is the idea that biological adaptations have a special connection to teleological thinking, and that this special connection persists into adulthood. What is less clear, however, is why this special connection exists, and how teleological explanations might be evaluated even *within* this domain. For instance, why is "animals grow fur because they need to smell things" rejected as a bad explanation, while other teleological explanations for biological features are (unreflectively) accepted? In the section that follows, we develop one proposal.

### *1.3. Structure-function fit and teleological explanation*

Following Lombrozo et al. (2007), we propose that the evaluation of teleological explanations is supported by a salient, defeasible cue: a good fit between the structure of a biological trait and some plausible function. We propose that this cue provides (provisional) evidence that a teleological explanation is warranted, and additionally affects the perceived quality of that explanation. So, for example, the claim that "animals grow fur because they need to smell things" can be readily rejected because the fit between having fur and smelling is poor – in other words, the "structure" of fur (thin, densely-packed, soft hairs) does not lend itself particularly well to the function of smelling. On the other hand, the claim that "trees produce oxygen so that animals can breathe" is in fact false, but might require further effort or reflection to reject, since the fit between oxygen production and respiration is more compelling. Finally, the explanation that "butterflies have wings for flying" might be judged more intuitively compelling than the claim that "butterflies have wings for thermal regulation" because the former is more likely to be perceived as high in structure-function fit (whether or not it provides the more accurate evolutionary explanation).

Structure-function fit can be a reliable basis for evaluating a teleological explanation, as a good fit between a structure and a function is not likely to be accidental. Wings are aerodynamic precisely *because* they were selected for their aerodynamic properties. However, even when a good fit does not arise by chance, it could arise for the wrong reason: not because the function shaped the structure, but because the structure shaped the function (as is the case with the structure of "the nose" and the function "holding up glasses"). So while structure-function fit may be a cue with some reliability, it is important to emphasize that it is fallible and *defeasible*. That is, the inference from good structure-function fit to the adequacy of a given teleological explanation can be defeated by additional information to the effect that the proper causal structure to support that teleological explanation is not present. In this way, an explanation with high structure-function fit (e.g., "earthworms tunnel underground in order to aerate the soil": the structure of underground tunnels particularly supports the function of soil aeration) can still be rejected as a poor explanation, given knowledge about the true mechanism by which the explanandum came to be (e.g., earthworms adapted to tunnel underground as a way to find food and to avoid the sun, and the function of soil aeration is merely a fortuitous consequence). Similarly, an explanation could be accepted despite being low in initial perceived structure-function fit: for example, the fit between peacocks' fanning of their feathers and the function of mating may not be immediately apparent, yet the explanation "peacocks fan their feathers to attract mates" could be accepted as a perfectly good explanation, given some causal knowledge about how this courtship ritual spread through the ancestral population of peacocks.

Importantly, structure-function fit is not the only possible or even plausible basis for the evaluation of teleological explanations. Evaluation could instead proceed solely by considering the domain of the explanandum (e.g., teleological explanations for biological traits and artifacts are good, teleological explanations for natural phenomena are bad) or by considering the structure or function alone. For example, a teleological explanation appealing to a function that is clearly adaptive for reproductive success could be judged better than an explanation appealing to a less "important" function, or to one that is less inferentially rich (Lombrozo & Rehder, 2012). While structure-function fit arguably requires a more complex evaluation in that it is inherently relational, it still falls short of a complete explanatory analysis. A more thorough method for the evaluation of teleological explanations would be to determine the presence or absence of the relevant causal etiology itself (and perhaps the importance or causal power of the function within the relevant causal system). Given that this more exhaustive alternative is rarely achievable, structure-function fit could provide a relatively easy but nonetheless useful basis for evaluation.

An additional reason to favor the proposal that people evaluate teleological explanations (at least in part) on the basis of structure-function fit is the fact that fit can plausibly be computed from quite basic cognitive and perceptual processes. In fact, there's evidence that children are able to recognize the relation between structure and function as early as eighteen months (Madole, Oakes, & Cohen, 1993) and make inferences on the basis of structure-function fit as early as two years of age (McCarrell & Callanan, 1995). McCarrell and Callanan (1995) suggest that "form-function correspondences" are "fundamental components of very young children's belief systems," where such correspondences are evaluated from perceptual information. For example, if one sees several times that round things roll under certain enabling conditions (e.g., when pushed or when placed on an incline), then one might infer that the structure of roundness (which can be assessed perceptually) has good fit with the function of rolling. These basic judgments of form-function correspondence could be the basis for evaluations of structure-function fit that become increasingly elaborated over the course of development.

### *1.4. The role of mechanisms in explanatory reasoning*

Structure-function fit may be a useful cue because it suggests a non-accidental relationship that could in fact have arisen from the very kinds of causal processes that support teleological explanations, such as natural selection or intentional design (Lombrozo & Carey, 2006). An explanation can thus be evaluated without detailed knowledge of the causal history of the entity being explained, or detailed knowledge of the causal mechanisms involved. In fact, one suggestion is that teleological reasoning may be useful precisely because it reduces the need to engage in highly mechanism-dependent reasoning. This is nicely illustrated by a quote from Daniel Dennett, who introduces the following example in motivating the design stance:

> Suppose I categorize a novel object as an alarm clock: I can quickly reason that if I depress a few buttons just so, then some hours later the alarm clock will make a loud noise. I don't need to work out the specific physical laws that explain this marvelous regularity; I simply assume that it has a particular design—the design we call an alarm clock—and that it will function properly, as designed. (Dennett, 2009, pp. 2–3)

In other words, reasoning about the *function* of a system allows us to understand the system, intervene on the system, and predict the working of the system to a sufficient degree, without requiring detailed knowledge of the mechanisms by which it operates or the process by which it was designed.

A few lines of evidence support the idea that teleological explanations decrease what we will call "mechanism-dependent reasoning" (see also Lombrozo & Wilkenfeld, in press). Lombrozo (2010) found that people evaluated causal claims differently when the candidate cause brought about the effect because it intended to do so or was fulfilling its function: in these cases, participants' judgments were less sensitive to whether the chain of events that mediated the relationship between the candidate cause and effect involved a direct transmission of force versus a more indirect process of double prevention. This suggests that when reasoning "functionally," causal evaluations were less dependent on the particulars of the mechanisms involved.

Another piece of evidence that reasoning in terms of functions might bypass mechanistic evaluation comes from Alter, Oppenheimer, and Zemla (2010), who studied how the illusion of explanatory depth (our tendency to overestimate the depth of our mechanistic knowledge) is affected by taking a more concrete or abstract construal of the world. While their manipulations were framed in terms of construal level theory (see Trope & Liberman, 2010 for a review), several of these manipulations primed participants to reason either in terms of mechanisms (lower level of construal) or in terms of functions (higher level of construal). They found that priming participants to reason in terms of functions increased the illusion of explanatory depth, indicating that function-based reasoning *decreased* the calibration of people's mechanistic understanding. This experiment supports one of the original explanations for the illusion of explanatory depth proposed by Rozenblit and Keil (2002): that people confuse an understanding of higher levels of analysis (i.e., functions) with an understanding of lower levels of analysis (i.e., mechanisms).

In sum, previous evidence suggests that function information can support a basis for reasoning that is largely independent of mechanisms. However, this claim has not been tested directly, and the psychological processes by which function information affects mechanism dependence remain unclear. We propose that these effects result, at least in part, from the use of structure-function fit as an intuitive but defeasible cue to the adequacy and quality of a teleological explanation, which can obviate the need to evaluate additional mechanism information. As a result, mechanism-dependent reasoning will be decreased when explanatory function information is available. In the following section, we detail our plan to test three predictions derived from this proposal.

### *1.5. Overview of present research*

We propose that teleological explanations can be intuitively compelling (at least in part) because their evaluation relies on structure-function fit, a salient cue that is relatively easy to evaluate and abstracted from mechanistic detail. This proposal explains previous results quite well. First, if teleological explanations are evaluated on the basis of structure-function fit (rather than detailed mechanistic analysis), this could explain the tendency of adults to initially endorse such explanations even when they are unwarranted, provided the structure seems to fit the function. Second, even young children are well-equipped to evaluate form-function correspondences, so structure-function fit potentially explains their evaluation of teleological explanations as well. Finally, structure-function fit can also explain variation in the acceptability of teleological explanations across items *within* a single domain.

Our proposal makes three predictions. First, we predict that because structure-function fit provides a ready basis for evaluation that is only weakly dependent on details of the underlying causal mechanisms involved, the presence of function information should decrease people's sensitivity to mechanistic considerations in evaluating explanations. In Experiments 1, 2, and 3 we find support for this prediction: when explanations for biological adaptations include functions, people are less sensitive to differences in mechanistic detail.

Second, we predict that the perceived quality of an explanation should closely track structure-function fit judgments, even for explanations that also include mechanism information. In Experiments 3 and 4, we compare perceived structure-function fit with a variety of plausible alternatives, and we find that structure-function fit ratings are the best predictor of explanation quality ratings.

Third, we predict that structure-function fit is both easy to evaluate and subject to error. If this is the case, structure-function fit should predict the kinds of errors in teleological reasoning previously reported by Kelemen and colleagues (Kelemen & Rosset, 2009; Kelemen et al., 2013). In Experiment 5, we use perceived structure-function fit to predict endorsement of unwarranted teleological explanations under both speeded and unspeeded conditions. We demonstrate that structure-function fit strongly predicts the over-extension of teleological reasoning in both conditions.

## 2. Experiment 1

In Experiment 1, we test our first prediction: that when an explanation for a biological trait contains explanatory function information, people will evaluate the explanation with less sensitivity to considerations about the mechanism(s) by which the trait arose. This prediction is motivated by our hypothesis that teleological explanations for biological traits are evaluated by considering structure-function fit, which abstracts away from detailed causal etiology. Before considering structure-function fit directly, however, as we do in Experiments 3–5, we aim to demonstrate the existence of *some* compelling cue that supports the evaluation of explanations that contain explanatory function information and that consequently diminishes sensitivity to mechanistic considerations. To do so, we present participants with explanations that vary in mechanistic detail, and that do or do not contain function information. If function information is not included, we would expect explanations with higher levels of mechanistic detail to be rated more satisfying than explanations with lower levels of mechanistic detail. On the other hand, when function information is present, we expect any effect of mechanistic detail to be attenuated or eliminated.

### 2.1. Method

#### 2.1.1. Participants

Participants in all experiments were recruited from Amazon Mechanical Turk and paid at the rate of $7.50 per hour (i.e., $0.25 for a 2-minute experiment). All participants provided informed consent under a protocol approved by the Institutional Review Board at University of California, Berkeley. Participation in all experiments was restricted to users in the United States with a minimum approval rating of 95% on at least 50 previous Mechanical Turk tasks. The sample size for Experiment 1 was determined by power analysis to have an 80% chance of detecting an effect, using effect sizes ranging from $\eta_p^2 = 0.03$ to $\eta_p^2 = 0.43$ from pilot data. Sample sizes for the remaining experiments were calculated so that at least 40 to 50 participants would be in each condition after exclusions, unless otherwise noted.

For Experiment 1, the participants were 201 adults (89 males, 112 females) ranging from 18 to 74 years of age ($M_{age} = 34$, $SD = 11$). Forty-two additional participants were excluded for expressing nonbelief in evolution by natural selection (explained below), and 107 additional participants were excluded for failing to pass two attention checks (explained below).

#### 2.1.2. Materials and procedure

The task consisted of a two-minute survey in which participants were asked to rate the quality of a single explanation for a biological feature. The explanation involved a feature of one of four fictitious animals: the speckled feather pattern of spirks, the large paws of yugrens, the sharp teeth of pomahs, or the thick fur of tellops. The explanation that was offered included one of three levels of mechanistic detail (*low, medium,* or *high*), and it either included or did not include a function (*present* or *absent*). Participants were randomly assigned to one of the 24 possible explanations that resulted from crossing *item* (4) with *mechanistic detail* (3) and with *function presence* (2).

Each mechanistic explanation described the process of evolution by natural selection, but at varying levels of detail.[1] A sample explanation at each level of mechanistic detail is included in Table 1. Participants rated on a 7-point scale how satisfying they found the explanation assuming that it is true, from "extremely dissatisfying" to "extremely satisfying," with the midpoint of the scale marked as "neither satisfying nor dissatisfying."

[1] One may question (as one reviewer did) why we chose to use evolutionary explanations rather than ontogenetic explanations (appealing, for example, to genetics), especially given that function information about a biological trait in some ways *implies* the evolutionary mechanism of natural selection. This implication is a feature insofar as our teleological and mechanistic explanations are more equivalent in their entailments, but we also recognize the potential concern – and we address this concern in the discussion of Experiment 2 (Section 3.3). We chose to use evolutionary mechanisms because they provided the greatest content match across explanations. That is, evolutionary explanations allowed us to present participants with *identical* mechanism information across conditions, but at varying levels of detail. Further, participants were in a better position to evaluate the adequacy of the provided mechanism information, as evolution by natural selection is a well-known (though often misunderstood) mechanism.

**Table 1**
Experiment 1 Sample Stimuli.

| Yugrens are a kind of animal with large paws. Suppose someone asks the following question, and the answer below is offered: **Question:** Why do yugrens have large paws? | |
|---|---|
| Mechanistic Detail | Explanation |
| *Low* | **Answer:** Because of the process of evolution by natural selection*: large paws are better for balancing on rocky ground.* |
| *Medium* | **Answer:** Because of the process of evolution by natural selection. In the past, yugrens varied in paw size, and those that had larger paws were more likely to reproduce successfully*, as large paws are better for balancing on rocky ground.*[†] |
| *High* | **Answer:** Because of the process of evolution by natural selection. In the past, yugrens varied in paw size, and those that had larger paws were more likely to reproduce successfully*, as large paws are better for balancing on rocky ground.*[†] Because large paws are in part heritable (due to genetic factors), those yugrens that most successfully reproduced tended to generate offspring with large paws, thus increasing the proportion of the population with large paws. |

[†] Function information, *in italics*, was presented only in the *function present* condition.

After rating explanation quality, participants indicated on a 5-point scale their agreement with five statements about the origin of living things. As religious and scientific beliefs have been previously tied to teleological explanatory preferences (Kelemen & DiYanni, 2005; Kelemen & Rosset, 2009), these questions were included as a potential source of individual differences in satisfaction with *function present* explanations. Participants were asked whether they agree that "evolution by natural selection explains the origin of nonhuman living things," that "evolution by natural selection explains the origin of humans," that "God created the first living things," and that "God designed the current properties and traits of living organisms." Additionally, participants were asked whether they agreed with the statement "Nature is a powerful being," which has been found by Kelemen and colleagues (Järnefelt et al., 2015; Kelemen et al., 2013) to relate to acceptance of unwarranted teleological explanations.

Agreement with the first evolution statement was also used as a basis for excluding participants who could not be expected to accept the evolutionary explanations used in our stimuli, and to thereby provide the intended ratings of explanatory satisfaction. Specifically, we excluded participants who indicated "somewhat disagree" or "strongly disagree" in response to the statement. The qualification "nonhuman" was included to avoid unnecessary exclusions: prior research finds that some people accept evolution by natural selection for plants and for non-human animals but not for humans (Ranney & Thanukos, 2011), and none of our stimuli involved humans.

One attention check, placed within the series of questions about participants' beliefs, asked participants to select the choice "strongly disagree." A second attention check, which was presented after the belief ratings, read as follows:

> Regardless of education, some people may feel that they understand the process of evolution well, and others may feel that they don't understand the process of evolution at all. Where do you fall on the following scale? To verify your attention, please ignore the previous text and select the third option.
> 1. I do not understand the process of evolution at all.
> 2. I slightly understand the process of evolution.
> 3. I somewhat understand the process of evolution.
> 4. I mostly understand the process of evolution.
> 5. I completely understand the process of evolution.

Participants were excluded for failing either one or both of these attention checks, which were designed to catch participants who were not reading the items and instructions carefully.

Finally, participants indicated the extent of their previous education in evolutionary biology and provided their age and gender.

### 2.2. Results

To control for possible differences across our stimulus items, all analyses (across all experiments, unless otherwise noted) included random intercepts for item. All main effects and interactions were computed using marginal sums of squares over mixed-effects models. To examine the effect of mechanistic detail (*low*, *medium*, *high*) and function (*present*, *absent*) on explanatory satisfaction ratings, we conducted a 2 × 3 ANOVA. This analysis revealed a main effect of mechanistic detail, $F(2, 192) = 6.40$, $p = .002$, a main effect of function, $F(1, 192) = 5.08$, $p = .03$, and a significant interaction, $F(2, 192) = 6.88$, $p = .001$ (see Fig. 1). To unpack this interaction, we conducted two one-way ANOVAs, one at each level of function presence, with mechanistic detail as a between-subjects variable and explanatory satisfaction ratings as the dependent variable. When a function was *absent*, the effect of mechanistic detail on explanatory satisfaction was highly significant, $F(2, 95) = 8.93$, $p < .001$. When a function was *present*, this effect failed to reach significance, $F(2, 95) = 2.40$, $p = .10$.

These results indicate that participants found explanations with greater mechanistic detail to be more satisfying than explanations with minimal mechanistic detail. Additionally, participants found explanations that contained function information to be more satisfying than explanations that did not. However, in accordance with our hypothesis that function information decreases reliance on mechanism information in evaluating an explanation, the effect of mechanistic detail on explanation satisfaction was attenuated when a function was present, relative to when a function was absent.

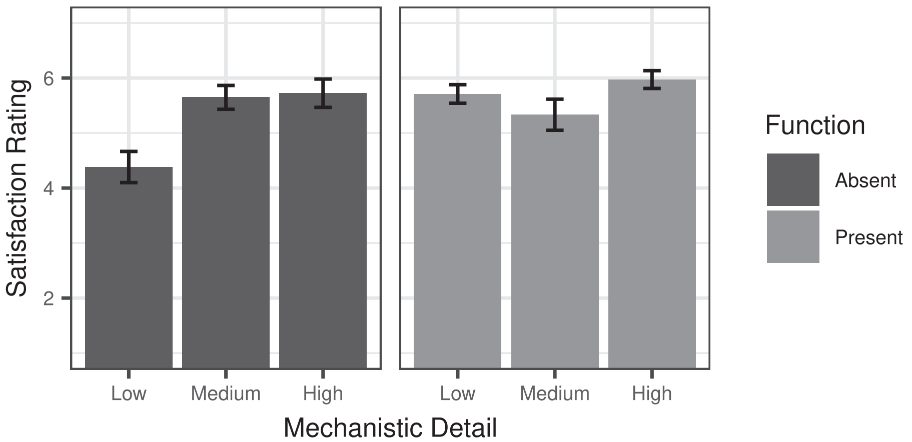


**Fig. 1.** Experiment 1 mean explanation satisfaction ratings as a function of mechanistic detail and function presence. Error bars ± 1 SEM.

Using a Bonferroni correction for multiple comparisons, we ran six regression analyses predicting explanatory satisfaction in the *function present* condition from the five belief measures and the self-report education measure. None of these regression models reached significance.

### *2.3. Discussion*

Experiment 1 provides preliminary evidence that when explanations can be evaluated on the basis of explanatory function information, participants become less sensitive to an explanation's mechanistic components. Specifically, the presence of function information decreased the effect of mechanistic detail on explanatory satisfaction. We refer to this as *the attenuation effect.* This effect supports our initial hypothesis that there is some mechanism-independent cue by which teleological explanations are evaluated. However, it is also possible that our results instead reflect a ceiling effect in explanation ratings. Explanatory satisfaction ratings were quite high in all conditions, and mean ratings were near 6 out of 7 when a function was present. It is therefore possible that mechanism information did not influence satisfaction ratings under these conditions simply because participants were highly satisfied with all functional explanations. Experiment 2 tests our hypothesis with an alternative method that sidesteps this concern.

## 3. Experiment 2

In Experiment 2, we adopt a new method to test the hypothesis that when explanatory function information is available, explanations are preferentially evaluated on this basis, resulting in less sensitivity to mechanistic detail. Instead of rating a single explanation, participants are asked to *compare* two explanations with varying levels of mechanistic detail, both of which either include or do not include a function. If the results of Experiment 1 reflect a genuine decrease in participants' reliance on mechanistic information once function information is available, we would expect the attenuation effect to manifest as a less robust preference for the more mechanistically detailed explanation when functions are present, relative to when they are absent. Alternatively, if the results of Experiment 1 reflect a ceiling effect, we would expect this comparative explanation evaluation to reveal that more mechanistically detailed explanations are rated as better than less mechanistically detailed explanations even when functions are present, and for this to occur to a similar extent when functions are absent.

Experiment 2 additionally tests a deflationary account of our results: it is possible that adding *any* additional information to an explanation results in an attenuation effect, whether or not the additional information refers to functions. To test this account, we include a control condition in which both explanations contain additional information, but where this information is not about functions.

Finally, we test two hypotheses about what drives the attenuation effect. The first hypothesis is that the attenuation effect is driven by a shift in the primary criteria used to evaluate an explanation, with participants relying on structure-function fit (or some other aspect of the provided function that is independent of mechanistic detail) when functions are available for evaluation. The second hypothesis is that the effect is driven by the mere indication that a function is present, without further evaluation of structure-function fit or some other aspect of the function's content. To differentiate these hypotheses, we can investigate whether function information attenuates the role of mechanism information even when the function remains unspecified (e.g., "thick fur serves *an important function* for tellops"). On the first hypothesis, the attenuation effect should be smaller or absent when functions are unspecified. On the second hypothesis, the effect should remain.

### 3.1. Method

#### 3.1.1. Participants

One hundred seventy-eight adults, recruited from Amazon Mechanical Turk, participated in Experiment 2. Participants ranged in age from 18 to 69 ($M_{age} = 34$, $SD = 12$) and were 80 males, 96 females, and two other/prefer not to specify. Twenty-four additional participants were excluded for nonbelief in nonhuman evolution by natural selection, and 68 additional participants were excluded for failing to pass the same two attention checks used in Experiment 1.

#### 3.1.2. Materials and procedure

The task consisted of a 2.5-minute survey which employed a procedure similar to that of Experiment 1. On a single screen, participants were shown two explanations that two different people (Anne and Beth) offered for the existence of different biological traits. One explanation always contained greater mechanistic detail than the other, where these corresponded to the "medium" and "low" levels of mechanistic detail from Experiment 1. Participants were randomly assigned to the *specified function present, unspecified function present, function absent,* or *control* condition. For participants in the *specified function present* condition, both explanations also included function information. For participants in the *unspecified function present* condition, both explanations included information about an unspecified function (e.g., "thick fur serves an important function for tellops"). For participants in the *function absent* condition, no additional information was included. For participants in the *control* condition, additional non-functional information was included in each explanation. Specifically, control explanations included an additional statement to the effect that differential reproduction led to a gradual change in the population over time. A sample set of explanations from the *specified function present* condition is included below, with function information indicated in italics:

> Tellops are a kind of animal with thick fur, and pomahs are a kind of animal with sharp teeth. Suppose someone asks Anne the following question, and she offers the answer below:
> **Question:** Why do tellops have thick fur?.
> **Anne's Answer:** Because of the process of evolution by natural selection*: thick fur is better for keeping tellops' bodies warm.*
>
> Suppose someone asks Beth the following question, and she offers the answer below:
> **Question:** Why do pomahs have sharp teeth?
> **Beth's Answer:** Because of the process of evolution by natural selection. In the past, pomahs varied in tooth sharpness, and those that had sharper teeth were more likely to reproduce successfully*, as sharp teeth are better for defending against predators.*

Participants were asked to rate whose explanation was more satisfying, assuming that both explanations were true. They responded on a 7-point scale that ranged from "Anne's answer is much more satisfying" to "Beth's answer is much more satisfying," with the midpoint marked as "Both explanations are equally satisfying." The remaining scale points were unmarked. In order to encourage a holistic comparison of the explanations, the answers that Anne and Beth gave were to two different questions. Explanations for the features of yugrens and spirks always appeared together, and explanations for the features of pomahs and tellops always appeared together. The order of the animals in each pair and the animal that had higher mechanistic detail were independently counterbalanced across participants. Participants were randomly assigned to one of the combinations resulting from condition (4: *specified function present*, *unspecified function present*, *function absent*, or *control*) and stimulus set (2). A between-subjects design was used because all explanations were fairly similar (as all appealed to the mechanism of natural selection), and thus sequential ratings may not have been serially independent.

After comparing the two explanations, participants answered the same belief questions and demographic questions as in Experiment 1.

### 3.2. Results

Participants' comparison ratings were coded so that a score of 3 corresponded to strong preference for the more mechanistically detailed explanation, a score of −3 corresponded to strong preference for the less mechanistically detailed explanation, and a score of 0 corresponded to indifference between the two explanations. The average comparison rating across all participants was 1.53 ($SD = 1.62$), which was significantly different from zero, $t(176) = 5.81$, $p < .001$, indicating an average preference for the more mechanistically detailed explanation.

To test the prediction that a specified function would decrease preference for the more mechanistically detailed explanation, we ran a one-way ANOVA with condition (*specified function present*, *unspecified function present*, *function absent*, *control*) as a between-subjects variable and explanation comparison ratings as the dependent variable (see Fig. 2). This analysis revealed a significant effect of condition, $F(3, 173) = 12.98$, $p < .001$. Further inspection of the corresponding regression model coefficients (using treatment coding) revealed a significant difference between the *specified function present* condition and all three remaining conditions; *unspecified function present* condition, $b = 0.70$, 95% CI [0.08, 1.33], $t(173) = 2.22$, $p = .03$; *function absent* condition, $b = 1.80$, 95% CI [1.18, 2.41], $t(173) = 5.79$, $p < .001$; *control* condition, $b = 1.42$, 95% CI [0.80, 2.05], $t(173) = 4.49$, $p < .001$. There was also a significant difference in comparison ratings between the *unspecified function present* condition and the *function absent* condition, $b = -1.09$, 95% CI [−1.70, −0.50], $t(173) = -3.58$, $p < .001$. However, explanation comparison ratings were not significantly different between the *function absent* condition and the *control* condition.

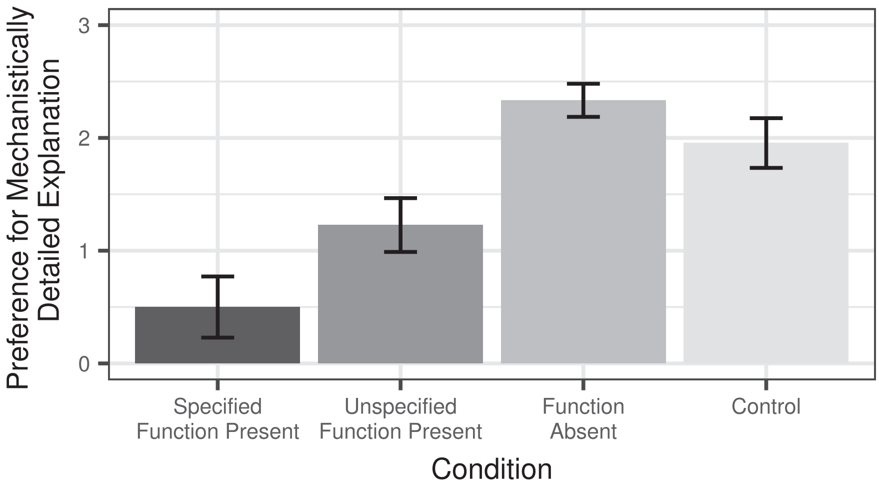


**Fig. 2.** Experiment 2 mean explanation comparison ratings as a function of condition. Error bars ± 1 SEM.

These results indicate that when functions were present, participants relied *less* on mechanistic detail to make explanatory judgments, exhibiting a weaker preference for the more mechanistically detailed explanation. Moreover, the attenuation effect appears to be at least somewhat specific to function information: the addition of information that is *not* relevant to function (in the control condition) did not decrease preference for the more mechanistically detailed explanation. The presence of unspecified function information *did* produce a significant attenuation effect, but to a lesser extent than full function information.

Using a Bonferroni correction for multiple comparisons, we ran six regression analyses predicting explanation comparison ratings in the *specified function present* and *unspecified function present* conditions from the five belief measures and the self-report education measure. None of these models reached significance.

### *3.3. Discussion*

These results offer a conceptual replication of our key finding from Experiment 1: explanatory function information attenuates the role of mechanistic detail in explanation evaluation. Moreover, the results rule out two alternative explanations. First, by using a comparative rating rather than an absolute rating, we showed that the attenuation effect is not an artifact of ceiling effects. Second, we demonstrated that an attenuation effect is not elicited by introducing *any* extra piece of information. Instead, this effect seems to be at least somewhat specific to function information, supporting the idea that teleological explanations support some basis for explanation evaluation that is more compelling than and partially independent from mechanistic detail.

The results from the *unspecified function* condition are particularly intriguing. First, the results support the idea that function information plays a special role in explanation evaluation. In both the *unspecified function* condition and in the *control* condition, participants compared pairs of explanations with *the very same text* included in both (either the claim about an important function, or the elaboration about differential reproduction). The fact that the former produced an attenuation effect – even though the explanations differed in mechanistic detail and not in their functional content – speaks to the power of functions.

Second, the fact that an unspecified function produced a weaker attenuation effect than a specified function is consistent with the idea that participants evaluated explanations with a specified function on the basis of structure-function fit. But if this is the case, it is not entirely clear why an unspecified function would generate any attenuation effect at all: only the content of the structure was available for evaluation. The results from Experiment 1–2 therefore leave us with several open questions that we take up in Experiments 3–4. First, is it the case that participants in fact evaluate structure-function fit, and not isolated aspects of the structure or function, such as its perceived importance? We consider this question in Experiment 3. Second, how, if at all, do participants evaluate structure-function fit when a function is unspecified? We take up this question in Experiment 4.

Before turning to these experiments, however, we consider an alternative interpretation of our results. It is possible that the decreased preference for an explanation with higher mechanistic detail in the presence of additional function information reflects an *incorporation effect* rather than an *attenuation effect*. That is, the function information might incorporate or imply the mechanism information, such that the presence of the described evolutionary mechanism is simply assumed when a function (specified or unspecified) is present. Once the proper mechanism is assumed, additional detail expounding upon the operation of this mechanism has little effect on explanatory satisfaction. In this way, the presence of function information does not *decrease* the consideration of mechanism information, but instead *incorporates* mechanism information into the function specification itself.

To test this account, we conducted an additional experiment (full details in the supplementary material, Experiment S1), replicating the *function absent* and *specified function present* conditions from Experiment 2. In contrast to the previous experiment, participants read explanations of varying mechanistic detail that appealed to *proximate* causes of the development of a trait (e.g., “Because yugrens have the YN5 gene, which leads to the creation of the protein keinopin and the enzyme meinase. Meinase modulates the expression of the protein keinopin, and this causes large paws”). In this case, added function information (e.g., “Large paws, in turn, are better for balancing on rocky ground") does not imply or incorporate the provided mechanism. Thus, the *attenuation* hypothesis predicts a replication of the decreased preference for the more mechanistically detailed explanation, while the *incorporation* hypothesis predicts no effect of function information on preference for a more detailed proximate-cause explanation. Though participants displayed a slightly weaker baseline preference for the more mechanistically detailed explanation (when function information was absent) than in Experiment 2 ($M = 0.92$, $SD = 1.92$), there was nonetheless a significant attenuating effect of function information, $t(174) = 3.13$, $p = .002$. When function information was present, participants displayed virtually no preference for the more mechanistically detailed proximate-mechanism explanation ($M = 0.02$, $SD = 1.80$). This strongly suggests that function information has an *attenuating* rather than an *incorporating* effect.

## 4. Experiment 3

In Experiment 3, we investigate whether structure-function fit can explain the explanatory satisfaction judgments people make when both function and mechanism information are present. We do so by soliciting participants' own evaluations of structure-function fit and using them to predict variation in explanation ratings. We also solicit judgments for a variety of alternative bases for evaluating a teleological explanation to investigate whether structure-function fit provides the best explanation for people's judgments. Finally, we include an internal replication of our key results from Experiment 2, showing that the presence of a specified function decreases the extent to which participants rely on mechanism information when evaluating explanations.

### 4.1. Method

#### 4.1.1. Participants

One hundred sixty-one adults participated in Experiment 3. Participants ranged in age from 18 to 73 ($M_{age} = 34$, $SD = 11$) and were 83 males, 77 females, and one other/prefer not to specify. Thirty-three additional participants were excluded for nonbelief in nonhuman evolution by natural selection, and 63 additional participants were excluded for failing to pass the same two attention checks used in the previous experiments.

#### 4.1.2. Materials and procedure

The task consisted of a 3.5-minute survey which employed a procedure nearly identical to that of Experiment 2. As in the previous experiment, participants read and compared either two mechanistic explanations of varying detail or two mechanistic explanations of varying detail that were augmented with function information. For those who received function information, we wanted to ensure adequate variation in structure-function fit, so in some cases similar fits were paired (*high-high* or *low-low*) and in some cases different fits were paired (*low-high* or *high-low*). The high and low structure-function fit stimuli are included in Table 2, and were classified as *low* or *high* based on pilot testing using the structure-function fit prompt described below. This design resulted in 5 conditions to which participants were randomly assigned: one *function absent* condition and four *function present* conditions (*high-high*, *low-low*, *high-low*, and *low-high*). The *function present* conditions were collapsed for all analyses, as our primary question was whether perceived structure-function fit (as rated by individual participants) corresponded to participants' explanation comparison ratings.

After completing the comparative rating of the two explanations, participants in the four *function present* conditions rated the structure-function fit of the features and functions described in the explanations, guided by the following prompt:

> Some biological traits have a good *"fit"* to their particular function, in that the physical structure of the trait lends itself particularly well to the performance of the function.
>
> For example:
> For giraffes, the trait "having a long neck" has a good fit with the function of reaching high sources of food.
> For humans, having a hand with an opposable thumb has a good fit with the function of grasping.
>
> *To what extent to you think yugrens' large paws have good fit with the function of balancing on rocky ground?*

Participants rated fit on a 7-point scale for both features that had been explained.

Additionally, participants answered a series of other questions about the features and functions that had been explained: feature goodness: "To what extent do you think it is good to have large paws?"; function goodness: "To what extent do you think it is good to

**Table 2**
Stimuli used in Experiments 3–4.

| Animal | Rich Feature (Exp. 3) | Blank Feature (Exp. 4) | High-Fit Function (Exp. 3 & 4) | Low-Fit Function (Exp. 3 & 4) |
|---|---|---|---|---|
| Spirk | Speckled feather pattern | Dorastic mintels | Blending in with the surroundings | Communicating with other spirks |
| Yugren | Large paws | Gleebling loozoids | Balancing on rocky ground | Detecting when predators are nearby |
| Tellop | Thick fur | Nalporing shizules | Keeping tellops' bodies warm | Spotting prey in dim light |
| Pomah | Sharp teeth | Zettific warotels | Defending against predators | Swimming in rivers with a strong current |

balance on rocky ground?"; reproductive importance: "To what extent do you think balancing on rocky ground is important for successful reproduction?"; survival importance: "To what extent do you think balancing on rocky ground is important for survival?"; and mating importance: "To what extent do you think balancing on rocky ground is important for attracting mates?". Answers to all questions were provided on 7-point scales. These additional questions ("additional cue measures") were included both to test whether perceived structure-function fit could be reduced to one of these simpler measures and also to separate the potential effects of structure-function fit on explanatory satisfaction from other plausible cues to a good explanation.[2]

### 4.2. Results

Participants' comparison ratings were coded so that a score of 3 corresponded to strong preference for the more mechanistically detailed explanation and a score of −3 corresponded to strong preference for the less mechanistically detailed explanation. The average comparison rating was significantly greater than zero, $M = 1.16$, $SD = 1.75$, $t(159) = 6.51$, $p < .001$, indicating a preference for the more mechanistically detailed explanation.

In order to assess whether comparison ratings differed across the function *present* versus *absent* conditions, a mixed-effects model was fit to the data, with random intercepts for stimulus item. The analysis revealed that function presence was a significant predictor of explanation comparison ratings, $b = 1.19$, 95% CI [0.57, 1.81], $t(158) = 3.77$, $p < .001$ (see Fig. 3). These results indicate that the presence of a function decreased participants' reliance on mechanistic detail in judging explanatory satisfaction, replicating our results from Experiments 1–2.

We next analyzed how structure-function fit ratings and the additional cue measure ratings predicted explanation preference. Rated structure-function fit for the less mechanistically detailed explanation was subtracted from rated structure-function fit for the more mechanistically detailed explanation. This difference score thus describes the perceived structure-function fit in favor of the more mechanistically detailed explanation, with scores ranging from −6 to 6. We calculated a difference score for the five additional cue measures (feature goodness, function goodness, reproductive importance, survival importance, and mating importance) in the same way.

Table 3 shows the correlations between the difference scores for each of the cue measures and structure-function fit. While structure-function fit correlated with most of these alternative cue measures, the correlations are modest in size. We conducted a regression analysis predicting structure-function fit ratings using these five measures as regressors, with a random intercept for stimulus item. Using Nakagawa and Schielzeth's (2013) measure of $R^2$ for mixed-effects models, we obtained a marginal $R^2$ value of 0.41. Therefore, while the fixed effects (the five cue measures) explain a reasonable amount of variance in structure-function fit ratings, structure-function fit ratings cannot be completely reduced to simpler judgments of function or feature goodness.

Next, we tested the hypothesis that structure-function fit ratings predict explanatory preference. The structure-function fit difference score was used as a predictor to estimate explanation comparison ratings in the subset of participants who read explanations with function information (controlling for differences in stimulus items with a random intercept). This analysis revealed a significant effect of structure-function fit, $F(1, 121) = 36.06$, $p < .001$ (see Fig. 4). Each one-point increase in structure-function fit favoring the more mechanistically detailed explanation predicted a 0.28-unit increase in preference for this explanation. This model explained 23% of the variance in explanation comparison ratings.

In order to test whether structure-function fit ratings were the *best* predictor of explanatory preference, we conducted a dominance analysis (Azen & Budescu, 2003), a method for determining the importance of predictors in multiple regression. If one predictor has *complete dominance* over another, this indicates that the dominant predictor makes a greater contribution in variance explained than the other predictor in all possible subset models of the full model. Therefore, if this analysis demonstrates that structure-function fit completely dominates all other cue measures in a regression model predicting explanation comparison ratings, we would have strong evidence that of the measures tested, structure-function fit is the most important determinant of explanatory satisfaction. A standard linear regression model (with no random intercept for stimulus item) using all six measures as predictors (structure-function fit, feature goodness, function goodness, reproductive importance, survival importance, and mating importance) revealed that structure-function fit had complete dominance over all other variables, suggesting that structure-function fit is the most important predictor of explanatory satisfaction.[3]

Finally, to test whether education or religious/scientific beliefs predicted explanation comparison ratings in the subset of participants who viewed function information, we ran six regression analyses (with a random intercept for item) for each of these six

[2] A preliminary version of this experiment was conducted without these additional cue measures. Three hundred seventeen adults participated (excluding an additional 183 participants) on Mechanical Turk. All reported results regarding explanation comparison ratings and structure-function fit ratings were replicated in this experiment. These additional questions regarding perceived benefits of the feature and function were included to rule out alternative explanations for our findings from this preliminary experiment.

[3] We also performed a bootstrap analysis with 1000 samples to estimate the strength of this result. The outcome of this analysis, called *reproducibility* by Azen and Budescu (2003), describes the percentage of bootstrap samples in which the result obtained in the sample holds. Therefore, for each pair of predictors, we can estimate the frequency with which structure-function fit will have complete dominance in the population. Structure-function fit completely dominated feature goodness in 66% of bootstrap samples, function goodness in 92% of bootstrap samples, function's importance for reproduction in 75% of bootstrap samples, function's importance for survival in 93% of bootstrap samples, and function's importance for attracting mates in 94% of bootstrap samples. These results strongly suggest that structure-function fit is a critical predictor of teleological explanatory satisfaction, and is a better predictor than the other measures included in this experiment (with feature goodness being the closest competitor).

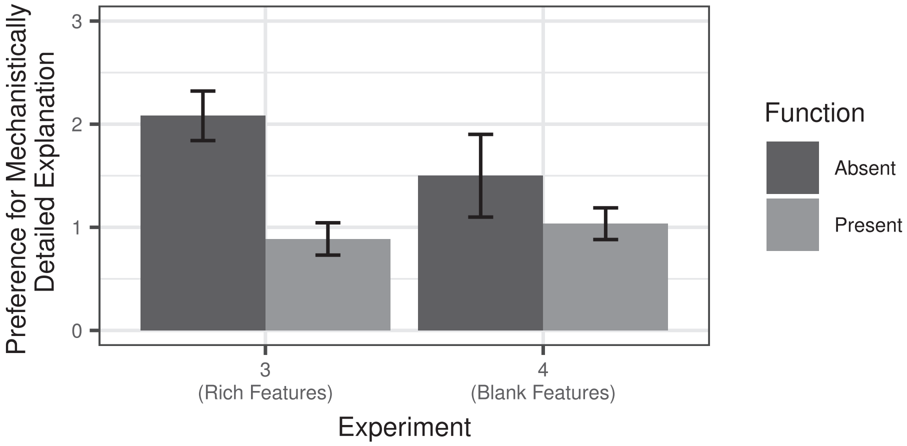


**Fig. 3.** Experiment 3 and 4 mean explanation comparison ratings when a function was absent versus present. Error bars ± 1 SEM.

**Table 3**
Correlations between structure-function fit and other measures of function/feature goodness.

| Measure | Experiment 3 | | Experiment 4 | |
|---|---|---|---|---|
| | Correlation with structure-function fit | Correlation with explanation comparison ratings | Correlation with structure-function fit | Correlation with explanation comparison ratings |
| Feature goodness | 0.41*** | 0.35*** | 0.16 | 0.16 |
| Function goodness | 0.50*** | 0.19* | 0.10 | 0.23* |
| Reproductive importance | 0.43*** | 0.35*** | 0.02 | 0.21* |
| Survival importance | 0.40*** | 0.28** | 0.08 | 0.22* |
| Mating importance | 0.21* | 0.19* | −0.03 | 0.06 |
| Structure-function fit | – | 0.48*** | – | 0.31*** |

*** p < .001, ** p < .01, * p < .05.

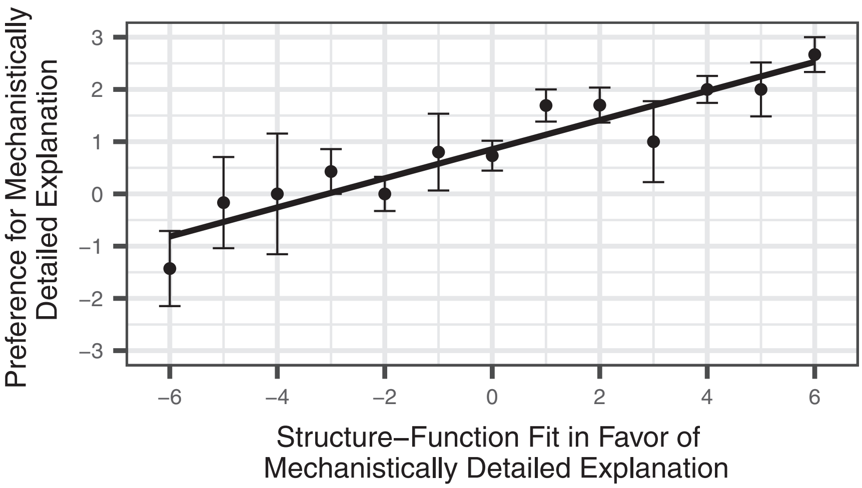


**Fig. 4.** Experiment 3 relation between structure-function fit ratings and explanatory preference. The mean explanation comparison rating at each level of structure-function fit in favor of the more mechanistically detailed explanation is plotted. Error bars ± 1 SEM.

measures using a Bonferroni correction for multiple comparisons. The only model that reached significance predicted comparison ratings from the belief that evolution by natural selection explains the origin of non-human species, $b = 0.58$, 95% CI [0.16, 0.99], $t(121) = 2.73$, $p_{Bonferroni} = 0.04$. This model revealed that greater belief in evolution increased the extent to which the more mechanistically detailed explanation was favored. Because this result was not found consistently across studies, we hesitate to draw strong conclusions from it.

### 4.3. Discussion

The findings from Experiment 3 suggest that structure-function fit does in fact play a large role in evaluating explanations that contain both functional and mechanistic content, and that structure-function fit cannot be reduced to nor dominated by a variety of alternative measures of the quality of individual features or functions. Additionally, we replicated the results of Experiment 2, providing further evidence that function information attenuates people's preference for a more mechanistically detailed explanation. However, another question from Experiment 2 remains: Why does an attenuation effect arise when a function is *unspecified?* We turn to this question in Experiment 4.

## 5. Experiment 4

Experiment 3 revealed that structure-function fit plays an important role in evaluating teleological explanations. However, this account of explanation evaluation is problematic when considering the results of Experiment 2: the mere suggestion that an important function exists (i.e., an unspecified function) was enough to produce a small attenuation effect. There are two ways to account for this result. One possibility is that participants evaluated structure-function fit despite limited information, perhaps on the basis of somewhat idiosyncratic inferences to "fill in" the information that was missing.[4] A second possibility is that in these cases, participants evaluated the explanations on some basis *other* than structure-function fit, such as a characteristic of the structure in isolation.

We can test these alternatives by presenting participants with a version of Experiment 3, but in which all explanations involve "blank" features (e.g., dorastic mintels). If participants make (potentially idiosyncratic) inferences about what the features are and use this as a basis for evaluating structure-function fit, then we would expect the results of Experiment 4 to mirror those of Experiment 3, though perhaps more weakly, and with greater variation across participants. Most critically, we would expect to find that perceived structure-function fit continues to be the strongest predictor of explanation quality. Alternatively, if participants evaluate explanations that lack a meaningful structure on the basis of some alternative cue, then we would expect to find that perceived structure-function fit is no longer a good predictor of explanation judgments, and that some alternative cue (perhaps based only on the specified function) dominates. We return to these competing possibilities in our discussion of Experiment 4.

### 5.1. Method

#### 5.1.1. Participants

Participants in Experiment 4 were 141 adults, with 63 males and 77 females ranging in age from 18 to 75 ($M_{age} = 35$, $SD = 13$). Twenty-six additional participants were excluded for nonbelief in nonhuman evolution by natural selection, and 78 additional participants were excluded for failing to pass the same two attention checks used in the previous experiments.

#### 5.1.2. Materials and procedure

The task used in Experiment 4 was identical to the task used in Experiment 3. However, all features from Experiment 4 were replaced with "blank" features (see Table 2).[5]

### 5.2. Results

Participants' comparison ratings were coded as in Experiment 3. The average comparison rating was significantly greater than zero, $M = 1.12$, $SD = 1.73$, $t(139) = 7.69$, $p < .001$, indicating a preference for the more mechanistically detailed explanation.

To assess whether comparison ratings differed as a result of function presence (*present*, *absent*), we performed a regression analysis with explanation comparison ratings as the dependent measure. Again, random intercepts for stimulus item were included. Function presence was not a significant predictor of comparison ratings, $t(138) = -1.24$, $p = .22$ (see Fig. 3). These results indicate that function presence did *not* have an effect on preference for the more mechanistically detailed explanation when features were blank.[6]

[4] In an experiment not reported here, 178 participants (and an additional 95 excluded) completed a procedure similar to that of Experiment 3. In addition to the function manipulation (*present-unspecified* vs. *absent*), participants were randomly assigned to read explanations for "*blank*" features (e.g., dorastic mintels) or "*rich*" features (e.g., a speckled feather pattern). Additionally, participants in all conditions rated their confidence that they could guess the unspecified function. A regression analysis on comparison ratings revealed a marginally significant effect of function presence, $F(1, 172) = 2.72$, $p = .10$, replicating the results of Experiments 2 and 3. Additionally, there was a marginally significant effect of confidence that the function could be guessed, $F(1, 172) = 3.40$, $p = .07$, with greater confidence associated with a smaller preference for the mechanistically detailed explanation. These results suggest that participants may infer an actual function when the function is unspecified, but this may not fully explain the effect of an unspecified function on explanation evaluation.

[5] Experiments 3 and 4 were in fact run in parallel, with participants randomly assigned to conditions across both experiments. We report them separately for easier exposition.

[6] Notably, a 2 × 2 ANOVA (with random intercepts for stimulus item) predicting comparison ratings from function (present vs. absent) and experiment (3 vs. 4) revealed *no* significant interaction, $F(1, 297) = 2.22$, $p = .14$. The overall effect of function was significant, $F(1, 297) = 11.52$, $p < .001$. These results suggest that there *may* in fact be an attenuation effect in the blank feature case, although we did not detect such an effect in the experiment-specific analysis. The analysis does suggest that we need to be cautious in drawing different conclusions from the presence of a significant attenuation effect in Experiment 3, and the absence of one in Experiment 4.

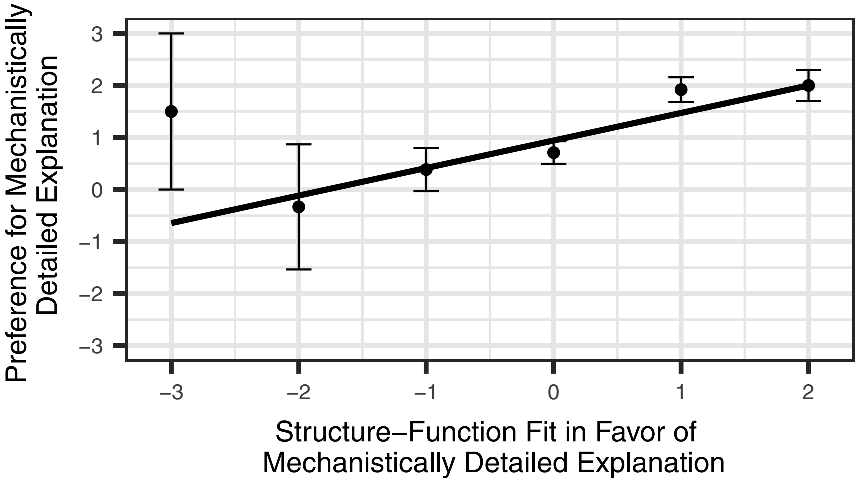


**Fig. 5.** Experiment 4 relation between structure-function fit ratings and explanatory preference for blank features. The mean explanation comparison rating at each level of structure-function fit in favor of the more mechanistically detailed explanation is plotted. Error bars ± 1 SEM.

Next, we analyzed how structure-function fit ratings and the alternative cue measure ratings predicted explanation preference. Difference scores for the six measures (structure-function fit, feature goodness, function goodness, reproductive importance, survival importance, and mating importance) in favor of the more mechanistically detailed explanation were calculated as in Experiment 3. Table 3 shows the correlations between each of these measures and structure-function fit. These correlations were relatively small, and none of them reached significance. We fit a mixed-effects model predicting structure-function fit ratings using the five alternative cue measures, including a random intercept for stimulus item. Again using Nakagawa and Schielzeth's (2013) $R^2$ measure for linear mixed models, we obtained a marginal $R^2$ of 0.03, which suggests that structure-function fit is largely unrelated to these other cue measures when rating blank-feature items.

Next, we analyzed whether structure-function fit ratings predicted explanation comparison ratings. Using data from the subset of participants who read function information, we fit a mixed-effects model predicting explanation comparison ratings from structure-function fit difference scores. This analysis revealed a significant main effect of structure-function fit, $F(1, 112) = 12.45$, $p < .001$ (see Fig. 5). Each one-point increase in structure-function fit favoring the more mechanistically detailed explanation predicted a 0.52-unit increase in preference for this explanation. This model explained 11% of the variance in explanation comparison ratings. Additionally, the dominance analysis performed in Experiment 3 was repeated with the Experiment 4 sample. All results were replicated: structure-function fit had complete dominance over the other five cue measures.[7]

Thus, while the effect of function presence on preference for mechanistic detail was not replicated with blank-feature items, there was still a strong relationship between structure-function fit ratings and explanation comparison ratings, and structure-function fit was a better predictor than the five tested alternatives. This supports the idea that despite limited information, participants were able to evaluate structure-function fit, and that structure-function fit in turn informed explanation judgments. If participants accomplished this by making somewhat idiosyncratic inferences about structure-function fit in the case of blank features (e.g., "if dorastic mintels support camouflage, they must be some patterned body part"), then we would also expect little consistency in the way structure-function fit ratings relate to explanation comparison ratings within items in Experiment 4, relative to Experiment 3. In other words, if there is no objective basis upon which participants made their structure-function fit ratings (i.e., there is no "true" value describing the degree to which dorastic mintels fit the function of camouflage), we would predict that the association between structure-function fit ratings and explanation comparison ratings would be considerably weakened if the structure-function fit data were shuffled within each specific item evaluated.

To test this claim, we designed the following analysis, which was conducted independently on data from both Experiments 3 and 4. First, the structure-function fit data were split into subsets on the basis of the feature and function evaluated. Next, structure-function fit ratings were randomly shuffled within these subsets. The difference score between structure-function fit ratings (in favor of the more mechanistically detailed explanation) was recalculated for each participant using the new shuffled ratings. Finally, a regression analysis was performed, estimating explanation comparison ratings with this difference score as a predictor. This procedure was repeated 1000 times for both Experiments 3 and 4. If participants made *systematic* judgments that tracked "true" structure-function fit in some way, the shuffled data should, on average, retain the relation between structure-function fit ratings and explanatory satisfaction. If, however participants made *idiosyncratic* structure-function fit ratings, perhaps by inferring unique contents for blank features, then we would expect no relation on average between structure-function fit ratings and explanatory preference in the shuffled data.

[7] The complete dominance of structure-function fit over each other rating was also replicated in 71% to 89% (varying by cue measure) of 1000 bootstrap samples.

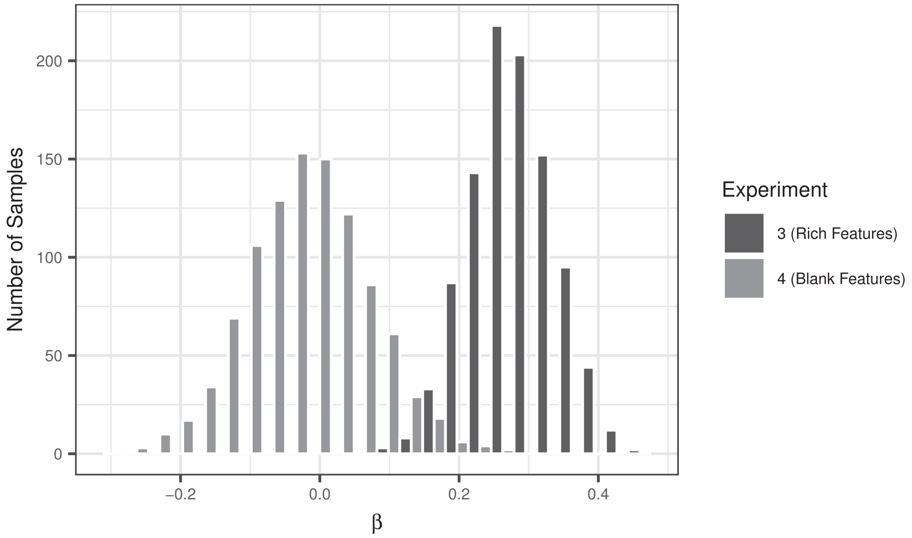


**Fig. 6.** Results of Experiment 3 and 4 structure-function fit shuffling analysis. Histograms display standardized $\beta$ coefficients describing the relation between explanation comparison ratings and structure-function fit ratings, obtained from 1000 samples in which structure-function fit was randomly shuffled.

This analysis revealed a positive relation between structure-function fit ratings and explanatory preference for the data from Experiment 3 (Mean $\beta = 0.28$, SD = 0.06), indicating that participants rated structure-function fit within the specific animal and function being evaluated in a systematic way. For the data from Experiment 4, however, the average relation between structure-function fit ratings and explanatory preference was close to zero (Mean $\beta = -0.02$, SD = 0.09), indicating that participants in this condition made idiosyncratic judgments of structure-function fit that were not determined by the feature and function being evaluated in an invariant way (see Fig. 6). In other words, the relation between structure-function fit and explanatory preference in the case of rich features seems to reflect properties of the items, while the same relation in the case of blank features seems to reflect some judgment that relates systematically to an individual's explanation ratings, but not to a feature of the stimuli that was invariant across participants.

Finally, using a Bonferroni correction for multiple comparisons, we ran six regression analyses on the Experiment 4 data predicting comparison ratings from the five belief measures and the self-report education measure. None of these models were significant.

### 5.3. Discussion

The findings from Experiments 3 and 4 support the idea that when evaluating explanations for biological adaptations that contain function information and vary in mechanistic detail, participants evaluate the explanations by assessing structure-function fit. When both the structure (feature) and function were specified, as in Experiment 3, this evaluation was systematic across participants, and resulted in an attenuation effect. Even when a feature was "blank," however, as in Experiment 4, participants seemed to evaluate structure-function fit (and not merely a property of the function), as structure-function fit continued to dominate other predictors. While the relation between rated structure-function fit and explanatory satisfaction seems somewhat mysterious in this case, we were able to explain this finding with additional analyses, demonstrating that structure-function fit ratings involving blank features were more idiosyncratic, potentially reflecting the idiosyncratic ways in which individual participants "filled in" the missing feature. These results may also explain our results from Experiment 2: the modest but significant attenuation effect that arose when unspecified function information was provided may have been a result of inferences about structure-function fit.

## 6. Experiment 5

Thus far, we have tested two of our main predictions: that teleological explanations are evaluated on the basis of structure-function fit, and that this leads to an attenuation effect whereby differences in mechanistic detail are given less weight in evaluating an explanation when function information is also available. Having tested these predictions, we consider whether structure-function fit can additionally explain existing evidence for "promiscuous teleology." In previous research, Kelemen and colleagues have found that adults under speeded conditions are especially likely to make the error of accepting unwarranted teleological explanations as true (Kelemen & Rosset, 2009; Kelemen, et al., 2013; Rottman et al., 2017). This has been taken to support the idea that humans have a default bias to construe the world in terms of function or purpose. Our results suggest an alternative or additional reason for the erroneous endorsement of some unwarranted teleological explanations: that they are judged high in structure-function fit, and therefore found satisfying (at least on first pass). If people rely on structure-function fit for the evaluation of teleological explanations, we might expect that the explanations that are judged to be highest in structure-function fit are also most likely to be incorrectly endorsed. In other words, we should be able to predict participants' acceptance of an unwarranted teleological explanation (e.g., "Trees produce oxygen so that animals can breathe") by their ratings of structure-function fit (e.g., their ratings of how well trees' production of oxygen "fits" facilitation of animals' breathing).

Further, we explore whether the relation between structure-function fit and unwarranted teleological explanation endorsement differs in speeded and unspeeded conditions. One possibility is that structure-function fit strongly influences explanation evaluation when time is limited, but that when time pressure is eliminated, people can draw on additional sources of information (such as knowledge of causal etiology) to determine whether a teleological explanation is in fact warranted. Under this account, we would expect the relation between structure-function fit and explanation endorsement to be *stronger* under speeded conditions than under unspeeded conditions. However, it is also possible that evaluating structure-function fit is itself somewhat effortful, in which case we might expect the relation between structure-function fit and explanation endorsement to be *weaker* under speeded than unspeeded conditions. This is also what we would expect if responding under speeded conditions simply introduces unsystematic noise. Finally, it could be that structure-function fit guides explanatory judgments to the same extent under speeded and unspeeded conditions, in which case we would not expect differences in the relationship between structure-function fit and explanation ratings across conditions. In Experiment 5, we test these accounts.

### 6.1. Method

#### 6.1.1. Participants

A power analysis indicated that at least 150 participants were needed to detect a small to moderate interaction between structure-function fit and condition with 80% power. Participants in Experiment 5 were 171 adults (80 female, 90 male, and one other/prefer not to specify), ranging from 19 to 72 years of age ($M_{age}$ = 36, $SD$ = 11). We used several methods to screen out participants who were not paying attention or not reading instructions carefully.[8] A reading check and a memory check, described below, excluded an additional 54 participants. An additional two participants were excluded because they did not respond to at least 75% of items in the speeded/unspeeded portion of the study.

#### 6.1.2. Materials & procedure

Experiment 5 consisted of a 15-minute survey. The first part of Experiment 5 was a replication of Kelemen et al. (2013), in which participants were presented with explanations, one at a time, to evaluate as "true" or "false." In the *speeded* condition ($N$ = 77), participants were told that they would have about 4 s to respond to each statement (actual time limit was 4.2 s) and were instructed to try to respond to each statement before time ran out. In the *unspeeded* condition ($N$ = 94), participants were instructed to respond to each statement after careful consideration, and they were given no time limit.[9] Stimuli were 100 one-sentence explanations: 30 "test sentences" and 70 "control sentences." Test sentences were scientifically unwarranted teleological explanations for natural phenomena (14 biological and 16 non-biological). Control sentences were of four types: 20 true mechanistic[10] explanations (8 biological and 12 non-biological), 10 true teleological explanations (all non-biological), 30 false mechanistic explanations (12 biological and 17 non-biological), and 10 false teleological explanations (6 biological and 4 non-biological). While our primary interest was in unwarranted biological explanations, we also prioritized similarity to Kelemen et al.'s (2013) experimental materials – both to test the generalizability of the predicted role of structure-function fit outside of the biological domain, and to minimize the potential for alternative explanations of any discrepancies in our findings.[11]

For the explanation acceptance task, the only changes in our procedure from Kelemen et al. were as follows. First, to shorten the length of the experiment, participants were given only four practice explanations before the task began (rather than the 20 used by Kelemen et al.). Second, participants were given no breaks during the speeded explanation task (as opposed to one three-second break after each block of ten questions). Finally, the maximum time allotted per item was 4.2 s (rather than 3.2 s), to allow for potential differences in reading speed and attention between Amazon Mechanical Turk workers (our sample) and adults tested in a laboratory setting (Kelemen et al.'s sample).

After the speeded explanation task, participants were given a memory test, which included four of the explanations they had seen previously and eight new explanations, presented in a random order. Participants were asked to identify which explanations they had read during the study. Participants who did not correctly identify at least 10 out of the 12 items as familiar or unfamiliar were excluded from further analysis.

[8] We did not exclude participants for nonbelief in evolution by natural selection in this study, as none of the test items were adaptationist explanations for biological traits. However, excluding these participants (18 additional participants) does not affect the pattern of results reported here.

[9] Though there were more participants in the unspeeded condition than the speeded condition, there was no significant difference in exclusion across these conditions, $\chi^2(1) = 0.29$, $p = .59$.

[10] In Kelemen et al.'s (2013) paper, the mechanistic explanations were referred to as "causal" explanations. We have revised the nomenclature to be consistent with our own usage, as we consider teleological and mechanistic explanations to both be sub-types of causal explanation.

[11] In Kelemen et al. (2013), three test items were excluded from analysis due to a lack of consensus on whether they were, in fact, unwarranted. These statements are included in our analyses. However, excluding these statements does not affect the pattern of results reported here.

After this memory test, participants rated structure-function fit on a 7-point scale for the explanandum and function of all test items from the speeded/unspeeded explanation acceptance task.[12] Participants were given the same prompt to guide their structure-function fit ratings as in Experiments 3 and 4 (with "entities, actions, or traits" replacing the phrase "biological traits"). Items were presented as follows: *"To what extent do you think [explanandum] has good fit with the function of [function]?"* (e.g., "To what extent do you think the Earth's ozone layer has good fit with the function of protecting the Earth from UV light?"). Items were presented in a random order.

Finally, participants completed the same belief and demographic questions as in all previous experiments. As previously described, a reading check was embedded within the belief ratings, which instructed participants to select "strongly disagree." Participants who did not select the correct option were excluded.

## 6.2. Results

### 6.2.1. Comparison with previous studies

Overall, participants incorrectly endorsed an average of 52% (SD = 22%) of unwarranted teleological explanations. This is comparable to the results of Kelemen et al. (2013), in which 47% (SD = 22%) of unwarranted teleological explanations were endorsed by community participants.

To test the effect of speeding on the endorsement of teleological explanations, we fit a generalized linear mixed model to participant endorsement of unwarranted teleological explanations (test items). Condition (*speeded, unspeeded*) was a fixed effect in the model, and random intercepts were included for participant and item. This analysis revealed a significant effect of condition, $b = 0.51$, 95% CI [0.10, 0.92], $z = 2.45$, $p = .01$. Controlling for participant and item, speeded participants were 166% more likely to endorse an unwarranted teleological explanation relative to unspeeded participants. This replicates the findings of Kelemen and colleagues (Kelemen & Rosset, 2009; Kelemen et al., 2013; Rottman et al., 2017).

Unlike previously-published findings, however, the effect of speeding on endorsement of *test items* was comparable to the effect of speeding on endorsement of *control items*. This was tested by a similar analysis to that described above. In addition to condition, the model included item type (*test, control*) as a fixed effect. Item accuracy (*correct, incorrect)* for all items was the dependent measure. We used a likelihood ratio test to compare a model including the interaction between item type and condition to a model excluding this interaction term. The likelihood ratio test was not significant, $\chi^2(1) = 0.09$, $p = .76$. This suggests that speeding generally decreased accuracy, and did not necessarily do so selectively for unwarranted teleological explanations. In general, performance on control items was worse than in previous studies: participants in the present research erred on 10% (SD = 10%) of control items in the unspeeded condition and 20% (SD = 15%) of control items in the speeded condition. In contrast, participants in Kelemen et al. (2013) erred on 7% (SD = 5%) of control sentences (averaged across both conditions and all participant groups). This increased error rate could be due to the smaller number of practice problems and lack of breaks during the task, or to characteristics of the population sampled (all MTurk workers could be working quickly to maximize the amount of money they can earn in a limited amount of time – thus even the unspeeded condition could be somewhat speeded in an MTurk setting). Kelemen et al. also employed different exclusion criteria. Notably, they excluded participants who weren't accurate on at least 80% of control items, which could have driven down the error rate in the reported sample.

### 6.2.2. Structure-function fit

To test whether structure-function fit ratings predicted acceptance of unwarranted teleological explanations, we fit a generalized linear mixed model, with structure-function fit rating and condition (*speeded*, *unspeeded*) as fixed effects, again with random intercepts for participant and item. Statement endorsement for test items (*true, false*) was the dependent measure. We used a likelihood ratio test to test a model including the interaction between condition and structure-function fit against a model excluding this interaction: this interaction was significant, $\chi^2(1) = 17.39$, $p < .001$ (see Fig. 7). Next, we investigated the relationship between structure-function fit and test item endorsement in each condition alone. In the *speeded* condition, each one-point increase in structure-function fit predicted a 163% increase in the probability of endorsing an unwarranted teleological explanation, $b = 0.49$, 95% CI [0.40, 0.57], $z = 11.33$, $p < .001$. In the *unspeeded* condition, each one-point increase in structure-function fit predicted a 206% increase in the probability of endorsing an unwarranted teleological explanation, $b = 0.72$, 95% CI [0.63, 0.82], $z = 15.48$, $p < .001$. These results suggest that structure-function fit affects the perceived quality of explanations across both speeded and unspeeded conditions, but is a stronger predictor under unspeeded conditions – perhaps due to increased noise in speeded participants' responses, or because structure-function fit is somewhat effortful to evaluate.

Next, we tested the generality of structure-function fit as a predictor of explanation endorsement across biological and non-biological domains. We again fit a generalized linear mixed model predicting explanation endorsement, with structure-function fit rating and domain (*biological, non-biological*) as fixed effects, and with random intercepts for participant and item. We compared a model including the interaction between domain and structure-function fit to a model excluding this interaction term. A likelihood ratio test revealed that this interaction was not significant, $\chi^2(1) = 1.67$, $p = 0.20$, providing no evidence that structure-function fit is differentially predictive of explanation endorsement across domains.

[12] In a preliminary experiment, we also asked for structure-function fit ratings for all teleological *control* items. These results, reported in the supplementary material (Experiment S2), hint at the generality of structure-function fit as a predictor of explanation acceptance. Across all domains and item types, structure-function fit significantly predicted explanation endorsement.

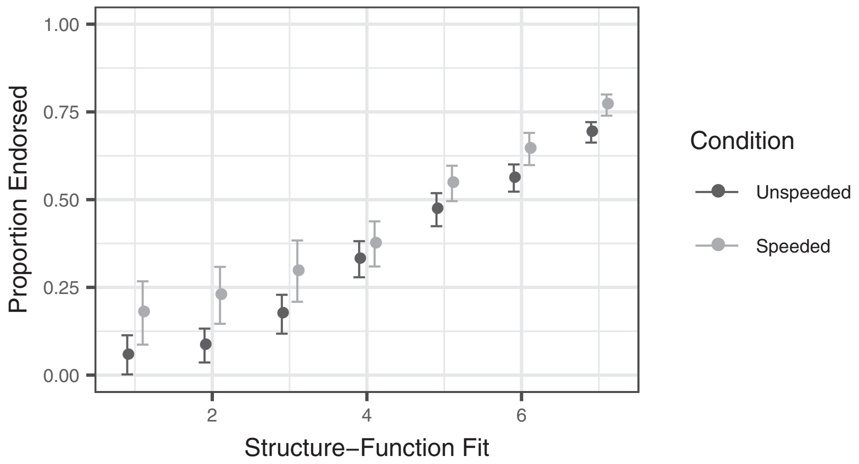


**Fig. 7.** Proportion of test items (unwarranted teleological explanations) in Experiment 5 endorsed at each level of rated structure-function fit. Error bars = bootstrap 95% CI.

### 6.2.3. Belief measures

Finally, to test whether education or religious/scientific beliefs predicted the probability of accepting unwarranted teleological explanations, we ran six generalized mixed models, predicting endorsement of unwarranted teleological explanations with each education/belief measure as the fixed effect. Each model also included random intercepts for participant and item. Using a Bonferroni correction for multiple comparisons, two of these models revealed a significant effect: endorsement of the statement "Nature is a powerful being" significantly predicted endorsement of unwarranted teleological explanations, $b = 0.41$, 95% CI [0.25, 0.58], $z = 4.85$, $p_{Bonferroni} < 0.001$, as did endorsement of the statement "God designed the current properties and traits of living organisms," $b = 0.19$, 95% CI [0.06, 0.32], $z = 2.89$, $p_{Bonferroni} = 0.02$. This is consistent with Kelemen's findings that belief in God and in Mother Nature predicted higher endorsement of unwarranted teleological explanations (Järnefelt et al., 2015; Kelemen et al., 2013).

## 6.3. Discussion

These results suggest that the role of structure-function fit in evaluating teleological explanations partially explains why participants often endorse such explanations even when they are scientifically unwarranted. In both speeded and unspeeded contexts, structure-function fit ratings were a strong predictor of endorsement of unwarranted teleological explanations. However, structure-function fit ratings were a stronger predictor in the *unspeeded* condition relative to the *speeded* condition. This suggests that evaluating structure-function fit could itself be an effortful process, and argues against the idea that structure-function fit provides a fast, initial basis for evaluation that can be subsequently revised by recruiting additional considerations of etiology or mechanistic detail. That said, we caution against rejecting this process-level account on the basis of these findings alone. The weaker relationship between structure-function fit and explanation acceptance in the speeded condition could be due to increased noise, rather than decreased reliance on structure-function fit. Additionally, the strength of the relationship between structure-function fit and unspeeded explanation acceptance does not necessarily imply that other cues to an explanation's adequacy (such as causal knowledge) are *not* being recruited. Future research is needed to address these questions.

One potential deflationary account of these results should be considered. Our results from Experiments 3–5 do not rule out the possibility that structure-function fit is not a *cause* of participants' judgments of explanation quality, but instead a *consequence* of these judgments. In other words, participants might first evaluate the quality of a teleological explanation, and based on this evaluation, generate an assessment of structure-function fit. One reason to question this alternative is because assessments of fit were deliberately solicited without any directional information about what was putatively explaining what (i.e., participants read a description of the feature and a description of a function, but not the entire explanation, when rating structure-function fit). Thus we would expect high ratings for "fit" if participants were asked about the fit between "the shape of the nose" and "holding up glasses," no matter that they would (presumably) reject the explanation that the nose is shaped as it is *because* it is for holding up glasses. An example from our stimuli is "Particles collide in order to produce chemical reactions" – in fact, chemical reactions occur *because* particles collide, not vice versa. Nonetheless, the "fit" between particle collision and chemical reaction is judged to be quite high ($M = 5.97$ in the present study). A second reason to favor the explanation that structure-function fit underlies teleological explanation judgments, rather than the reverse, comes from a feature of the data from Experiment 5: speeding affected the distribution of responses for explanation judgments, but not structure-function fit responses.[13] This is perhaps unsurprising, given that the latter judgment was not made under

[13] A model predicting structure-function fit ratings on the basis of condition made no significant improvement over a model containing only random intercepts for participant and item, $\chi^2(1) = 0.0002$, $p = .99$, suggesting that speeding the explanation acceptance task did not affect subsequent structure-function fit ratings.

time constraints, but nonetheless suggests that the experimental manipulation of explanation judgments did not itself result in an effect on structure-function fit judgments, as we would have expected if participants were responding to structure-function fit by considering the teleological explanations corresponding to each rating. (We consider additional conceptual arguments against this deflationary account in the General Discussion.)

## 7. General discussion

Many philosophers and biologists agree that biological adaptations support teleological explanations. Moreover, there is evidence that laypeople not only accept such explanations, but often favor them, sometimes applying them well beyond their scientifically-sanctioned scope. Why is this the case? We have proposed that teleological explanations are often compelling because their evaluation relies on structure-function fit, a cue to a teleological explanation's adequacy and quality that is largely independent of mechanistic considerations. Our experimental evidence supports the three predictions derived from this proposal. First, consistent with the idea that structure-function fit is easy to evaluate and largely independent of underlying mechanisms, we find that the presence of function information decreases people's sensitivity to mechanistic considerations in evaluating explanations (Experiments 1, 2, and 3). Second, we find that perceived structure-function fit predicts judgments of explanation quality, and that it does so more effectively than other plausible cues (Experiments 3 and 4). Third, we find that structure-function fit judgments underlie errors in teleological reasoning under speeded and unspeeded conditions (Experiment 5). Together, these studies provide empirical evidence for a hypothesis that has previously gone untested: that teleological explanations are evaluated on the basis of structure-function fit, and that this defeasible cue at least partially explains their intuitive appeal.

While these findings are not inconsistent with previous work by Kelemen and colleagues (e.g., Kelemen et al., 2013), they support an alternative framework for thinking about the allure of teleological explanations. Rather than claiming that teleological explanations are privileged because they reflect a "cognitive default," we suggest that teleological explanations can be intuitively compelling by merit of the method of evaluation they support: the use of structure-function fit. This evaluation process can potentially account for previous results on adults' acceptance of unwarranted teleological explanations under speeded conditions (Kelemen & Rosset, 2009; Kelemen et al., 2013; Rottman et al., 2017), when little reflection is involved (Zemla et al., 2016), and under cognitive impairment (Lombrozo et al., 2007). It can also potentially explain children's preference for teleological explanations across domains (Kelemen, 1999d, 1999c, 2003), and has the virtue of explaining why teleological explanations can vary in quality even within a given domain.

We have also provided evidence that explanatory function information can decrease mechanism-dependent reasoning. Specifically, the presence of function information attenuates the extent to which mechanistic detail is taken into account in evaluating explanations. It's worth considering just how far this mechanism-independence extends. Our introductory example from Dennett suggests that bypassing mechanisms is a virtue: design thinking can be powerful in no small part because it allows us to engage in explanation and prediction without a detailed understanding of relevant physical laws or the exact processes by which something came about. On the other hand, teleological explanations are only warranted when certain mechanistic considerations obtain. For instance, we explain the shape of glasses by appeal to noses, and not the reverse, because of the historical causal processes involved. It is therefore safe to assume that mechanistic considerations play an important role in blocking the inference from good structure-function fit to the acceptability of a teleological explanation. It's also quite likely that some mechanistic thinking is required to evaluate structure-function fit – for instance, to appreciate that a metal weight has moderate fit with the function of holding down paper while a feather does not. Our suggestion is therefore that teleological thinking can support a modest but convenient reduction in mechanism-dependence; not that teleological thinking is entirely independent of mechanistic considerations.

One notable feature of our account is that it explains a very complex judgment (explanation quality) in terms of a component that we have reason to believe is much more basic (structure-function fit). Indeed, young children and other species are able to evaluate rudimentary forms of structure-function fit. Children make inferences on the basis of structure-function fit as early as two years of age (McCarrell & Callanan, 1995) and attend to the relation between structure and function as early as eighteen months (Madole et al., 1993). These abilities manifest while children are still developing the tendency to ask for and provide explanations (Hickling & Wellman, 2001). Additionally, there is evidence that some primates (Manrique, Gross, & Call, 2010; Visalberghi et al., 2009) and crows (Chappell & Kacelnik, 2002) can select a tool on the basis of whether the tool's physical properties will allow them to achieve a desired goal. Thus, while verbal explanations are uniquely human, structure-function fit judgments seem to have more ancient roots. Importantly, this feature of our account addresses a potential concern about the relation between structure-function fit and teleological explanatory satisfaction. As judgments of structure-function fit are ontogenetically and phylogenetically more basic than explanation judgments, it seems unlikely that our results reflect an inference from explanation quality to structure-function fit, rather than the reverse.

To further test our account, future work should investigate the role of structure-function fit in judgments other than the explicit evaluation of explanations. Structure-function fit could serve as a cue not only to the adequacy of teleological explanations, but also, more generally, to the suitability of adopting a "design stance." If this is the case, we would expect that the presence of function information might have broad effects. For example, judgments of structure-function fit could affect how people learn about the functions and mechanisms of novel artifacts or biological systems. Consistent with this idea, recent research demonstrates that teleological preferences (i.e., the tendency to believe that natural phenomena are goal-directed) predict how well students learn the process of evolution by natural selection (Barnes, Evans, Hazel, Brownell, & Nesse, 2017), with a stronger teleological preference associated with greater misunderstanding. Additionally, and as mentioned in the introduction, adopting a functional construal can exacerbate the illusion of explanatory depth, leading to a greater overestimation of mechanistic understanding (Alter et al., 2010). In cases like these, where mechanistic understanding is important, interventions that prevent good structure-function fit from cuing a functional construal or design stance could have beneficial consequences for learning.

Many additional questions remain open for future research. First, while Experiment 5 (and Experiment S2 in the supplementary material) hint at the generality of these findings beyond the biological domain, it remains to be seen what role structure-function fit plays in other domains. For instance, one can ask whether teleological explanations for human-made artifacts (e.g., "chairs are flat so you can sit on them") and intentional behavior (e.g., "Alice raised her voice to get the classroom's attention") also support evaluation by structure-function fit over other possible cues, and whether this evaluation produces a mechanism attenuation effect similar to that reported here. Additionally, and relevant to the interpretation of Experiment 5, it remains to be seen whether or not structure-function fit is a *heuristic cue* (a less effortful but error-prone method of evaluation that can sometimes be substituted by a more difficult evaluation of causal etiology).

Also, it is worth considering what other cues might influence explanatory satisfaction. Across all of our experiments, structure-function fit ratings accounted for a significant but modest amount of variance in explanation quality judgments. While Experiments 3 and 4 considered several alternatives to structure-function fit (e.g., the function's importance for survival), future work might extend this investigation to identify other important cues. More generally, we know that explanation quality judgments are based on a host of factors: simplicity (Bonawitz & Lombrozo, 2012; Lombrozo, 2007; Pacer & Lombrozo, 2017), scope (Johnson, Johnston, Toig, & Keil, 2014; Preston & Epley, 2005), explanatory power (Schupbach, 2011), and more (Zemla, Sloman, Bechlivanidis, & Lagnado, 2017). It remains unclear how judgments of structure-function fit are integrated with these additional factors.

Relatedly, very little is known about the bases for evaluating *mechanistic* explanations. While the present research focused on mechanistic detail as a factor that could affect the perceived quality of a mechanistic explanation, several other factors are possible. For example, Hopkins, Weisberg, and Taylor (2016) demonstrated that reductive information – referring to fundamental parts and processes that underlie the phenomenon being explained – increases the perceived quality of a mechanistic explanation. Moreover, Frazier, Gelman, and Wellman (2016) found that in some cases, additional mechanistic detail did not improve the perceived quality of an explanation. Future research that explores cues to mechanistic explanatory satisfaction could also shed light on the following question: is there an asymmetry in explanatory preferences, such that teleological explanations are *privileged over* mechanistic explanations? Or do various cues to an explanation's adequacy determine whether mechanistic or teleological explanations are privileged in a given situation? While proponents of "promiscuous teleology" often present teleological explanations as privileged (e.g., Kelemen, 1999d), a deeper understanding of how mechanistic and teleological explanations are evaluated could support a more nuanced claim.

Another important question is when and how structure-function fit is overridden. As emphasized already, structure-function fit is a *defeasible* cue to the acceptability and quality of a teleological explanation. This is most obvious in our results from Experiment 5 – while structure-function fit predicted explanation acceptance for unwarranted teleological explanations, this cue was often overridden. When and how is this defeasible inference defeated? As we argue above, structure-function fit may be defeated by background beliefs about causal mechanisms (e.g., that glasses were designed to fit on noses, rather than vice versa), but this claim remains open for future study.

Finally, this study may be limited by our exclusive use of participants from Amazon Mechanical Turk within the United States. These research participants may be more demographically diverse than typical college student samples (Berinsky, Huber, & Lenz, 2012; Buhrmester, Kwang, & Gosling, 2011), but they certainly do not represent all populations of interest. Previous research has shown marked differences in teleological preferences across cultures and across ages (e.g., Casler & Kelemen, 2008; Sánchez Tapia et al., 2016). In order to explain these differences, future research should consider how teleological explanations are evaluated in non-Western cultures, as well as in children.

Luckily for the Sakalava people, the aye-aye's long finger does *not* serve the function of cutting the aortic vein of its victim. However, their folk beliefs about the natural world reveal an important truth about the role of functions in our understanding of the biological domain. The research reported here reveals that functions are not only important – they also have distinct consequences for the evaluation of explanations for the natural world. This provides a new framework for the study of teleological explanation, and more broadly, sheds light on how humans make sense of the complex world around them.

## 8. Declarations of interest

None.


## Acknowledgements

We would like to thank members of the Concepts and Cognition Lab at UC Berkeley for their useful comments on experimental design. EL presented some of the results reported here at the 2017 Meeting of the Society for Philosophy and Psychology, and we are grateful to that audience for their discussion and feedback. This work was supported by a James S. McDonnell Foundation Scholar Award in Understanding Human Cognition to TL and an NSF Graduate Research Fellowship to EL [grant numbers DGE 1752814 and DGE 1106400]. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the McDonnell Foundation or the National Science Foundation.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cogpsych.2018.09.001.

## References


Ahn, W. (1998). Why are different features central for natural kinds and artifacts?: The role of causal status in determining feature centrality. *Cognition, 69*(2), 135–178. https://doi.org/10.1016/S0010-0277(98)00063-8.

Allen, C. (2009). Teleological notions in biology. In E. N. Zalta (Ed.), Stanford Encyclopedia of Philosophy (Winter 2009). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/win2009/entries/teleology-biology/.

Alter, A. L., Oppenheimer, D. M., & Zemla, J. C. (2010). Missing the trees for the forest: A construal level account of the illusion of explanatory depth. *Journal of Personality and Social Psychology, 99*(3), 436–451. https://doi.org/10.1037/a0020218.

Ayala, F. J. (2007). Darwin's greatest discovery: Design without designer. *Proceedings of the National Academy of Sciences, 104*(suppl 1), 8567–8573. https://doi.org/10.1073/pnas.0701072104.

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods, 8*(2), 129. https://doi.org/10.1037/1082-989X.8.2.129.

Barnes, M. E., Evans, E. M., Hazel, A., Brownell, S. E., & Nesse, R. M. (2017). Teleological reasoning, not acceptance of evolution, impacts students' ability to learn natural selection. *Evolution: Education and Outreach, 10*(1), 7. https://doi.org/10.1186/s12052-017-0070-6.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20*(3), 351–368. https://doi.org/10.1093/pan/mpr057.

Bonawitz, E. B., & Lombrozo, T. (2012). Occam's rattle: Children's use of simplicity and probability to constrain inference. *Developmental Psychology, 48*(4), 1156–1164. https://doi.org/10.1037/a0026471.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1), 3–5. https://doi.org/10.1177/1745691610393980.

Casler, K., & Kelemen, D. (2008). Developmental continuity in teleo-functional explanation: Reasoning about nature among Romanian Romani adults. *Journal of Cognition and Development, 9*(3), 340–362. https://doi.org/10.1080/15248370802248556.

Chappell, J., & Kacelnik, A. (2002). Tool selectivity in a non-primate, the New Caledonian crow (Corvus moneduloides). *Animal Cognition, 5*(2), 71–78. https://doi.org/10.1007/s10071-002-0130-2.

Dawkins, R. (1995). God's utility function. *Scientific American,* 80–85.

Dennett, D. C. (1971). Intentional systems. *The Journal of Philosophy, 68*(4), 87. https://doi.org/10.2307/2025382.

Dennett, D. C. (2009). Intentional systems theory. In A. Beckermann, B. P. McLaughlin, & S. Walter (Eds.). *The Oxford Handbook of Philosophy of Mind* (pp. 339–350). Oxford, UK: Oxford University Press.

Diesendruck, G., & Haber, L. (2009). God's categories: The effect of religiosity on children's teleological and essentialist beliefs about categories. *Cognition, 110*(1), 100–114. https://doi.org/10.1016/j.cognition.2008.11.001.

Falcon, A. (2015). Aristotle on causality. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Spring 2015). Retrieved from https://plato.stanford.edu/archives/spr2015/entries/aristotle-causality/.

Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2016). Young children prefer and remember satisfying explanations. *Journal of Cognition and Development, 17*(5), 718–736. https://doi.org/10.1080/15248372.2015.1098649.

Goodman, S., & Schütz, H. (2000). The lemurs of the Northeastern slopes of the Réserve Spéciale de Manongarivo. *Lemur News, 5*, 30–33.

Greif, M. L., Kemler Nelson, D. G., Keil, F. C., & Gutierrez, F. (2006). What do children want to know about animals and artifacts?: Domain-specific requests for information. *Psychological Science, 17*(6), 455–459. https://doi.org/10.1111/j.1467-9280.2006.01727.x.

Gwynne, N. Z., & Lombrozo, T. (2010). *The cultural transmission of explanations: Evidence that teleological explanations are preferentially remembered* (pp. 1301–1306). Austin, TX: Cognitive Science Society.

Hickling, A. K., & Wellman, H. M. (2001). The emergence of children's causal explanations and theories: Evidence from everyday conversation. *Developmental Psychology, 37*(5), 668. https://doi.org/10.1037/0012-1649.37.5.668.

Hopkins, E. J., Weisberg, D. S., & Taylor, J. C. V. (2016). The seductive allure is a reductive allure: People prefer scientific explanations that contain logically irrelevant reductive information. *Cognition, 155*, 67–76. https://doi.org/10.1016/j.cognition.2016.06.011.

Järnefelt, E., Canfield, C. F., & Kelemen, D. (2015). The divided mind of a disbeliever: Intuitive beliefs about nature as purposefully created among different groups of non-religious adults. *Cognition, 140*, 72–88. https://doi.org/10.1016/j.cognition.2015.02.005.

Johnson, S., Johnston, A., Toig, A., & Keil, F. (2014). Explanatory scope informs causal strength inferences. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.). *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 707–712). Austin, TX: Cognitive Science Society.

Keil, F. C. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In L. Hirschfeld, & S. Gelman (Eds.). *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp. 234–254). Cambridge: Cambridge University Press.

Keil, F. C. (1995). The growth of causal understandings of natural kinds. In D. Sperber, D. Premack, & A. J. Premack (Eds.). *Symposia of the Fyssen Foundation. Causal cognition: A multidisciplinary debate* (pp. 234–267). New York: Clarendon: Press/Oxford University Press.

Kelemen, D. (1999a). 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, UK: Oxford University Press.

Kelemen, D. (1999b). Function, goals and intention: Children's teleological reasoning about objects. *Trends in Cognitive Sciences, 3*(12), 461–468. https://doi.org/10.1016/S1364-6613(99)01402-3.

Kelemen, D. (1999c). The scope of teleological thinking in preschool children. *Cognition, 70*(3), 241–272. https://doi.org/10.1016/S0010-0277(99)00010-4.

Kelemen, D. (1999d). Why are rocks pointy? Children's preference for teleological explanations of the natural world. *Developmental Psychology, 35*(6), 1440–1452. https://doi.org/10.1037/0012-1649.35.6.1440.

Kelemen, D. (2003). British and American children's preferences for teleo-functional explanations of the natural world. *Cognition, 88*(2), 201–221. https://doi.org/10.1016/S0010-0277(03)00024-6.

Kelemen, D., & DiYanni, C. (2005). Intuitions about origins: Purpose and intelligent design in children's reasoning about nature. *Journal of Cognition and Development, 6*(1), 3–31. https://doi.org/10.1207/s15327647jcd0601_2.

Kelemen, D., & Rosset, E. (2009). The human function compunction: Teleological explanation in adults. *Cognition, 111*(1), 138–143. https://doi.org/10.1016/j.cognition.2009.01.001.

Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General, 142*(4), 1074. https://doi.org/10.1037/a0030399.

Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology, 55*(3), 232–257. https://doi.org/10.1016/j.cogpsych.2006.09.006.

Lombrozo, T. (2009). Explanation and categorization: How "why?" informs "what?". *Cognition, 110*(2), 248–253. https://doi.org/10.1016/j.cognition.2008.10.007.

Lombrozo, T. (2010). Causal–explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology, 61*(4), 303–332. https://doi.org/10.1016/j.cogpsych.2010.05.002.

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition, 99*(2), 167–204. https://doi.org/10.1016/j.cognition.2004.12.009.

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

*Neuroscience, 8*, 1–12. https://doi.org/10.3389/fnhum.2014.00700.
Lombrozo, T., Kelemen, D., & Zaitchik, D. (2007). Inferring design: Evidence of a preference for teleological explanations in patients with Alzheimer's disease. *Psychological Science, 18*(11), 999–1006. https://doi.org/10.1111/j.1467-9280.2007.02015.x.
Lombrozo, T., & Rehder, B. (2012). Functions in biological kind classification. *Cognitive Psychology, 65*, 457–485. https://doi.org/10.1016/j.cogpsych.2012.06.002.
Lombrozo, T., & Wilkenfeld, D. (in press). Mechanistic versus functional understanding. In S. R. Grimm (Ed.), Varieties of Understanding: New Perspectives from Philosophy, Psychology, and Theology. New York, NY: Oxford University Press.
Madole, K. L., Oakes, L. M., & Cohen, L. B. (1993). Developmental changes in infants' attention to function and form-function correlations. *Cognitive Development, 8*(2), 189–209. https://doi.org/10.1016/0885-2014(93)90014-V.
Manrique, H. M., Gross, A. N.-M., & Call, J. (2010). Great apes select tools on the basis of their rigidity. *Journal of Experimental Psychology: Animal Behavior Processes, 36*(4), 409. https://doi.org/10.1037/a0019296.
McCarrell, N. S., & Callanan, M. A. (1995). Form-function correspondences in children's inference. *Child Development, 66*(2), 532. https://doi.org/10.2307/1131595.
Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x.
Owen, R. (1863). *Monograph on the Aye-aye (Chiromys Madagascariensis, Cuvier).* Taylor & Francis.
Pacer, M., & Lombrozo, T. (2017). Ockham's razor cuts to the root: Simplicity in causal explanation. *Journal of Experimental Psychology: General, 146*(12), 1761–1780. https://doi.org/10.1037/xge0000318.
Preston, J., & Epley, N. (2005). Explanations versus applications: The explanatory power of valuable beliefs. *Psychological Science, 16*(10), 826–832. https://doi.org/10.1111/j.1467-9280.2005.01621.x.
Ranney, M. A., & Thanukos, A. (2011). Accepting evolution or creation in people, critters, plants, and classrooms: The maelstrom of American cognition about biological change. In R. S. Taylor, & M. Ferrari (Eds.). *Epistemology and science education: Understanding the evolution vs. intelligent design controversy* (pp. 143–172). New York: Routledge.
Rottman, J., Zhu, L., Wang, W., Seston Schillaci, R., Clark, K. J., & Kelemen, D. (2017). Cultural influences on the teleological stance: Evidence from China. *Religion, Brain & Behavior, 7*(1), 17–26. https://doi.org/10.1080/2153599X.2015.1118402.
Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science, 26*(5), 521–562. https://doi.org/10.1207/s15516709cog2605_1.
Sánchez Tapia, I., Gelman, S. A., Hollander, M. A., Manczak, E. M., Mannheim, B., & Escalante, C. (2016). Development of teleological explanations in Peruvian Quechua-speaking and US English-speaking preschoolers and adults. *Child Development, 87*(3), 747–758. https://doi.org/10.1111/cdev.12497.
Schachner, A., Zhu, L., Li, J., & Kelemen, D. (2017). Is the bias for function-based explanations culturally universal? Children from China endorse teleological explanations of natural phenomena. *Journal of Experimental Child Psychology, 157*, 29–48. https://doi.org/10.1016/j.jecp.2016.12.006.
Schupbach, J. N. (2011). Comparing probabilistic measures of explanatory power. *Philosophy of Science, 78*(5), 813–829. https://doi.org/10.1086/662278.
Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440–463. https://doi.org/10.1037/a0018963.
Vasilyeva, N., Wilkenfeld, D., & Lombrozo, T. (2017). Contextual utility affects the perceived quality of explanations. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-017-1275-y.
Visalberghi, E., Addessi, E., Truppa, V., Spagnoletti, N., Ottoni, E., Izar, P., & Fragaszy, D. (2009). Selection of effective stone tools by wild bearded capuchin monkeys. *Current Biology, 19*(3), 213–217. https://doi.org/10.1016/j.cub.2008.11.064.
Wright, L. (1976). *Teleological explanations: An etiological analysis of goals and functions.* Berkeley, CA: University of California Press.
Zemla, J. C., Sloman, S., Bechlivanidis, C., & Lagnado, D. A. (2017). Evaluating everyday explanations. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-017-1258-z.
Zemla, J. C., Steiner, S. M., & Sloman, S. (2016). Analytical thinking predicts less teleological reasoning and religious belief. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.). *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1217–1222). Austin, TX: Cognitive Science Society.