

Understanding 'Why': How implicit questions shape explanation preferences

Sehrang Joo, Sami R. Yousif, and Frank C. Keil

Version: Dec 2021; Accepted to *Cognitive Science*

Abstract

Adults and children ‘promiscuously’ endorse teleological answers to ‘why’ questions—a tendency linked to arguments that humans are intuitively theistic and naturally unscientific. But how do people arrive at an endorsement of a teleological answer? Here, we show that the endorsement of teleological answers need not imply unscientific reasoning (n=880). A series of experiments shows that (a) ‘why’ questions can be understood as a query for one of two distinct kinds of information, and (b) these ‘implicit questions’ can explain adults’ answer preferences without appeal to unscientific worldviews. As a strong test of this view, we show that people endorse teleological answers that can answer relevant (implicit) questions about something’s purpose, even when those answers are explicitly non-causal. Thus, we argue that endorsement of teleological answers does not necessarily equate to endorsement of teleological ‘explanations’: Instead, explanation preferences may simply be an indication of people’s pragmatic expectations about the questions that others ask. This view reframes how we should think about explanation preferences in general, while also offering practical insight into the pragmatics of question-asking.

Keywords: Explanation; Teleology; Mechanism; Why questions

1 Introduction

*"I wonder why things have to change," murmured Piglet.
Pooh thought for awhile, then said, "It gives them a chance to get better."*

- David Benedictus,
Winnie-the-Pooh: Return to the Hundred Acre Wood

Imagine a child asks you a question like, “Why do things change?” Like Pooh, you might sometimes be tempted to give an answer similar to this one: “It’s so that they have a chance to get better.” However, to many scientists and philosophers who study explanation, that answer is deeply flawed: It tries to explain a part of the natural world via appeal to a *telos* that played no causal role in its origin. Even if changes in nature do have this *telos* or purpose (i.e., to allow things to get better), this purpose clearly doesn’t seem to explain what *causes* change. Yet, in many ways, this answer feels perfectly natural—of the sort we encounter and provide all the time. Here, we aim to understand why it is that ostensibly flawed explanations like this one *feel* like satisfying answers (and are frequently endorsed by adults and children alike).

This answer about change offers a *teleological explanation*—one that explains by appeal to purpose. Both children and adults appreciate such explanations, preferring them over comparable non-teleological explanations in a variety of domains (see e.g., Banerjee & Bloom, 2015; Kelemen, 1999; Lewry et al., 2020; Schachner et al., 2016). Endorsement of teleological explanations in some domains is clearly reasonable: There is a straightforward sense in which we can say that hammers are explained by their purpose to drive nails. However, it makes less sense to say that some change in nature can be explained by its purpose. Most changes in nature occur by naturally, or by accident—without some end state in mind. Yet even adults are drawn to such teleological explanations (e.g., Kelemen & Rosset, 2009; Kelemen et al., 2013), raising questions about how humans naturally understand the world (and explanations about it).

For example, teleological explanation preferences have been linked to theories like ‘intuitive theism’ (i.e., the idea that children are endowed with a natural belief in a designer, or God(s); Kelemen, 2004), as well as the ‘agentive worldview’ (i.e., the idea that humans naturally over-ascribe agency in virtually all domains of cognition; Rose & Schaffer, 2017; Rose, 2020). On such views, people ascribe teleological explanations for natural kinds because they *do* perceive those natural kinds as specifically resulting from some broader purpose. Rocks and trees for example, may have been designed by God to serve humans and animals. Thus, these views regard the endorsement of teleological explanations (in some domains) as inherently unscientific—revealing a deep-set conflict between our intuitive worldview and modern science (e.g., Allen & Bekoff 1995; but see also, Lombrozo & Carey, 2006). As such, understanding these explanation preferences and their origins is key to understanding not only how we are inclined to *explain* the world, but also how we come to make sense of it in the first place.

1.1 Understanding ‘Why’

Most prior work has (reasonably) focused on explanations; here, we focus on the *questions* that prompt them. Prior work often assumes that ‘why’ questions are ambiguous, and researchers have cited this ambiguity as a reason to use ‘why’ questions to examine people’s explanation preferences: Kelemen and DiYanni (2009) write, for example “...research therefore use[s] ‘why’ questions... because they can be interpreted as requests for causal or teleo-functional explanation depending on the listener’s construal of the phenomenon under question.” At the same time, this work also assumes that actually answering this ambiguous question with an explanation must involve a causal account—thus interpreting teleological answers as a kind of (sometimes unscientific) *causal* explanation (see, e.g., Kelemen, 1999; Schachner et al., 2016; Lombrozo et al., 2007; Lombrozo & Gwynne). We ask: (1) Are ‘why’ questions truly ambiguous, and (2) is it fair to assume that a ‘correct’ answer to a ‘why’ question must provide a causal explanation?

Consider the implications of this first question. Figure 1 shows an innocuous ‘why’ question, two possible ways the question could be rephrased, and two possible ways the question could be answered. The red line shows how we typically think of the endorsement of teleological explanations—as stemming directly from ‘why’ questions. But suppose that “Why?” is *not* ambiguous. This figure shows that if people understand the initial ‘why’ question as specifically implying a ‘purpose’ question, it would not be strange in any way to prefer the teleological answer (as that answer is no longer meaning to causally *explain*). People may, for instance, understand the question “Why do things change?” as a question about the purpose of change. And in that case, a teleological answer is a correct response to this purpose question. Such an answer is not only adequate—but in fact the only correct way to respond to the implied question.

To linger on the relevant distinction for a moment: Our goal is not to argue that teleological explanations are never causal—but to demonstrate that *whether or not they are appealing* is a distinct question from *whether or not they are causal*. Some teleological explanations are straightforwardly causal explanations because they involve purposes that were designed by humans (e.g., everyone would agree that a hammer’s features are actually a result of its intended purpose of driving in nails). Other teleological explanations typically seem obviously non-causal (at least without referring to supernatural forces or beings; e.g., a rock’s properties are usually not thought to result from some purpose it happens to serve for humans; see also, Kelemen, 2004). Yet other teleological explanations seem to be somewhere between these extremes (e.g., teleology in biology, see Lombrozo Carey, 2006; c.f., Allen Bekoff, 1995). Here, we mean to make no strong arguments about the causal status of these explanations or about teleological explanations in general. Instead, our claim is simply that the basic way that people evaluate explanations may be separable from how they evaluate causality.¹

1.2 Current Study

The majority of explanation studies ask people to evaluate different explanations that answer the same ‘why’ question. Here, we break down questions and answers separately and investigate people’s assumptions about both (see Figure 2).

First (Experiments 1a-b), we ask people to interpret ambiguous ‘why’ questions, and find that adults do have domain-specific expectations about such questions (mirroring known explanation preferences). Then

¹To say this in a different way, consider the following analogy: Something being amoral (i.e., someone being tall) is clearly distinct from something being immoral (i.e., someone having murderous intent). To use similar terms, we might say that our paper argues that explanation preferences are primarily ‘acausal.’ Causality may sometimes be irrelevant to how people choose to endorse an explanation, even in response to an explanatory ‘why’ question.

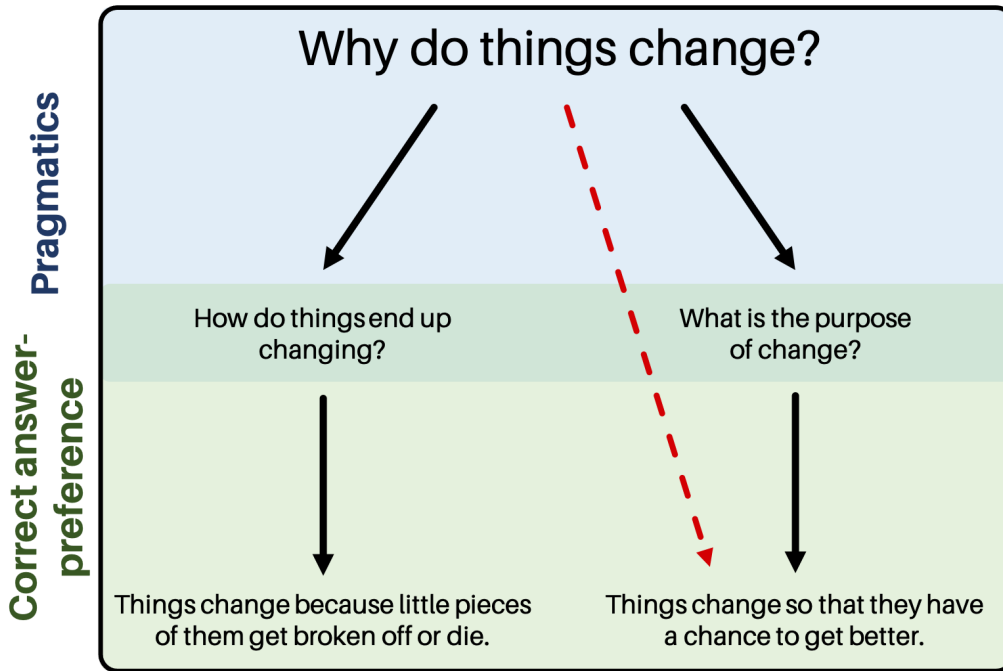


Figure 1: A depiction of how the pragmatics of ‘why’ questions influences the answers people prefer in response to those questions. Endorsement of teleological explanations can only be unscientific if we imagine a direct link between ‘why’ questions and teleological answers (dashed red line). However, if we imagine that people understand certain ‘why’ questions as implying a ‘purpose’ question (right side), then people’s answer preferences are understandable.

(Experiments 2a-b), we have people play a ‘jeopardy’ game in which they must generate questions from answers. Using this approach, we show that people map specific answers (e.g., teleological vs causal) on to specific implicit questions (‘purpose’ vs ‘how’) to. In one final set of experiments (Experiments 3a-b), we directly address a key prediction of this view: that people will endorse teleological answers that cannot be straightforwardly causal.

2 Experiments 1a-b: What do people really want to know?

Are ‘why’ questions (pragmatically) ambiguous? Here, we showed participants ambiguous ‘why’ questions and asked them what the agent asking the question “really wanted to know.”

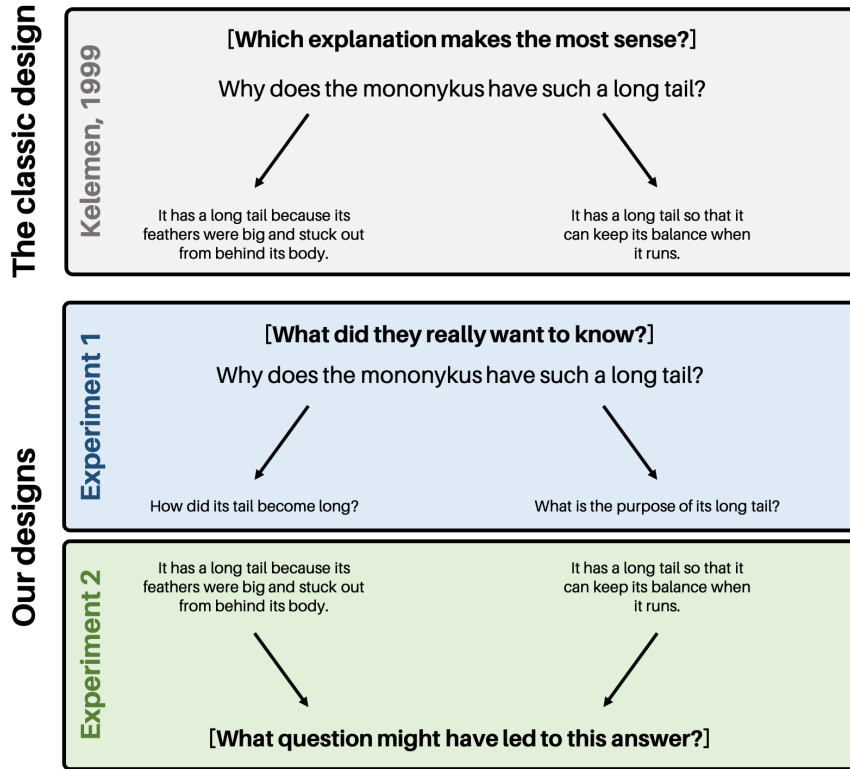


Figure 2: A depiction of our paradigms in comparison to prior work. While our stimuli are closely modeled after those used by Kelemen (1999), we use novel paradigms to examine underlying assumptions in people’s explanation preferences. Instead of asking about explanation preferences directly, we break down the classic design into its component parts—asking (1) how people interpret ‘why’ questions (Experiment 1) and (2) how they understand the link between specific explanations and implicit questions (Experiment 2).

2.1 Method: Experiment 1a

2.1.1 Participants

One hundred adult participants completed a survey online through Amazon Mechanical Turk. The sample size was chosen on the basis of independent pilot data and was preregistered. All participants lived in the United States. Data from an additional 39 participants were excluded for failure to follow directions or failure to generate interpretable questions (all data from excluded participants is also available on our OSF page).

2.1.2 Stimuli

Participants were shown an ambiguous ‘why’ question and simply asked what the agent asking the question “really want[ed] to know.” Total material consisted of twelve ‘why’ questions, four each in the domains of animals, non-living natural kinds (NLNK), and artifacts (see Table 1).

These stimuli (as well as the explanations in Experiments 2a-b) are based on those from Kelemen (1999) with a few modifications, including: (1) One additional domain (artifacts) was added; (2) The specific language of the stimuli themselves were changed to be in the present tense and refer to singular items; (3) Superfluous language was removed since only adults were tested; and (4) Stimuli were presented in a fully randomized order (rather than in pairs). Because these questions and explanations were designed to study children (or modeled after those original stimuli), they are linguistically simple. However, it is important for our view that we use stimuli that are representative of this field of study.

2.1.3 Procedure

For each item, participants were presented with an ambiguous ‘why’ question and were asked to list either all of the questions that the agent “may want the answer to.” They were specifically instructed to provide at least two questions and were able to list up to five. Participants were prompted to think of “more specific questions” and told that “it may help to think about questions that don’t contain the word ‘why’.” They were first shown a training item in which they were instructed to interpret the question “Why is the sky blue?”. Participants’ freely generated questions were then coded into ‘how’, ‘purpose’, and ‘other’ categories of questions. A question was coded as a ‘purpose’ question if it contained the words “purpose”, “for”, “use”, “function”, or “advantage.” A question was coded as a ‘how’ question if it contained the words “how”, “cause”, “made”, and “led.” Questions that asked “how come” or “the reason that” something was the way it was were treated as functional synonyms of ‘why’ questions and were thereby excluded from further analysis. All criteria for coding questions were preregistered. (There may be other questions that could have been categorized in further ways and/or different criteria for sorting between ‘how’ and ‘purpose’ questions. Our method in sorting questions was not exhaustive of all the possibilities nor a perfect system; however, (1) All of our data are publicly available, and we invite readers to analyze them in other ways as they wish, and (2) We are confident in the qualitative pattern we present regardless of the specific criteria used to code questions.)

Participants were assigned to one of two between-subjects conditions. Participants in each condition saw six ‘why’ questions, two from each domain (animals, NLNK, artifacts), presented in a random order. Pilot testing revealed that a fully within-subjects design resulted in an overly long task, given that each item requires participants to generate and type in at least two questions. Thus the purpose of these between-subjects conditions was simply to reduce task load; there were no meaningful differences between conditions.

Data, materials, and preregistration information for this experiment and all following can be found on the Open Science Framework (OSF) at https://osf.io/xtdu9/?view_only=acea58b086994c00a79f86072e798820.

Table 1: “Why” questions and explanations. The far left column contains the three domains from which we draw our stimuli (animals, non-living natural kinds, and artifacts). The middle columns list items in each domain along with the ‘why’ question asked about one of its features. The far right column presents two explanations for each item. These explanations are meant to answer the ‘why’ questions that they follow. Of each pair, the top explanation is a mechanistic explanation, and the bottom explanation is a teleological explanation. The animal and NLNK items are adopted with minor adjustments from prior work (Kelemen, 1999). The ‘why’ questions are investigated in Experiments 1a-b and the explanations are investigated in Experiments 2a-b.

	'Why' Question	Explanations
<i>Animals</i>	Why does the mononykus have such a long tail?	The mononykus has a long tail because its feathers were big and stuck out from behind its body. The mononykus has a long tail so that it can keep its balance when it runs.
	Why does the cryptoclidus have such a long neck?	The cryptoclidus has a long neck because the stuff inside got all stretched out and curved. The cryptoclidus has a long neck so that it can grab at fish and feed on them.
	Why does the macrauchenia have such a big snout?	The macrauchenia has a big snout because its face muscles and bones pulled down and got longer. The macrauchenia has a big snout so that it can pull down leaves from trees and eat them.
	Why does the moeritherium have such flat feet?	The moeritherium has flat feet because its toe bones got shortened and all smoothed out. The moeritherium has flat feet so that it can stand on wet ground without slipping.
<i>NLNK</i>	Why are the rocks so pointy?	The rocks are pointy because little bits of stuff piled up on top of one another over a long time. The rocks are pointy so that animals won't sit on them and smash them.
	Why is the pond so still?	The pond is still because no moving water ever ran into it. The pond is still so that it won't ever spill and lose all its water.
	Why is the sand so grainy?	The sand is grainy because bits of shells got broken and mixed up making it that way. The sand is grainy so that it won't get blown away and scattered by the wind.
	Why are the stones so green?	The stones are green because lots of colored stuff mixed together to make them that way. The stones are green so that they can't be seen in the grass and no one will pick them up and take them.
<i>Artifacts</i>	Why is the baking tool full of holes?	The baking tool is full of holes because it was cut with something sharp. The baking tool is full of holes so that it can hook onto small parts on a hot oven rack.
	Why does the glassblowing tool have empty loops?	The glassblowing tool has empty loops because its metal was bent into shape. The glassblowing tool has empty loops so that it can make circles in melted glass into spheres.
	Why is the roller covered with ridges?	The roller is covered with ridges because some of the wood was carved out. The roller is covered with ridges so that it can press against even little muscles that are hurting.
	Why does the ostrich pillow have a big hole?	The ostrich pillow has a big hole because the fabric wasn't sewn shut on one side. The ostrich pillow has a big hole so that it can fit someone's whole head inside when they sleep.

2.2 Method: Experiment 1b

All elements of the experimental design were identical to those of Experiment 1a, except as stated below. One hundred new participants completed the survey online through Amazon Mechanical Turk. This sample size was chosen to be identical to that in Experiment 1a.

Even if participants in Experiments 1a themselves generate a specific pragmatic questions for ‘why’ questions, we might wonder if this specific interpretation is truly the main way that ‘why’ questions are pragmatically interpreted—or if, instead, one question is simply what first comes to mind. Here, therefore, we offered participants a forced-choice. Participants were asked to choose between a ‘how’ question (e.g., “How did the mononykus’ tail become long?”) and ‘purpose’ question (e.g., “What is the purpose of the mononykus’ long tail?”; a full list of these questions is available on our OSF page under Table S1). They were asked which question the agent asking an ambiguous ‘why’ question *really* wanted to know.

2.3 Results: Experiment 1a

Results for Experiments 1a-b are shown in Figure 3. Participants’ questions were not evenly distributed across categories (‘how’, ‘purpose’ and ‘other’) when asked to generate questions from an ambiguous ‘why’ question, $\chi^2(5, N = 566) = 64.50, p < .001$. We further broke down these results by simply counting the total number of ‘how’ vs. ‘purpose’ questions that were generated within each domain. We analyzed participants’ likelihood of generating ‘how’ vs. ‘purpose’ questions by treating counts of participants’ questions within each domain as an average. A total number of ‘how’ questions and ‘purpose’ questions was generated for each participant within each domain. Participants were more likely to generate ‘how’ questions ($M=1.29, SD=1.17$) than ‘purpose’ questions ($M=.45, SD=.72$) when asked about non-living natural kinds ($t(99)=6.62, p < .001, d=.66$). They were also more likely to generate ‘purpose’ questions ($M=1.44, SD=1.23$) than ‘how’ questions ($M=.74, SD=.86$) for artifacts ($t(99)=4.49, p < .001, d=.45$). However, there were no significant differences between numbers of ‘how’ and ‘purpose’ questions generated about animals ($t(99)=1.19, p=.24$).²

As a secondary approach, we also analyzed questions by counting individual participants who generated all ‘purpose’ or all ‘how’ questions. In the first approach, one hypothetical participant generating one ‘how’ question and four ‘purpose’ questions would have skewed the question count toward ‘purpose’ questions

²An anonymous reviewer asked if these effects are stable if between-stimulus and between-participant differences are taken into account as randomized effects. Here and in subsequent experiments, we also fit the data to linear mixed-effects models (including stimulus item and participant as random effects where relevant). These models did not qualitatively change our results.

What kind of question is implied by an ambiguous ‘why’ question?

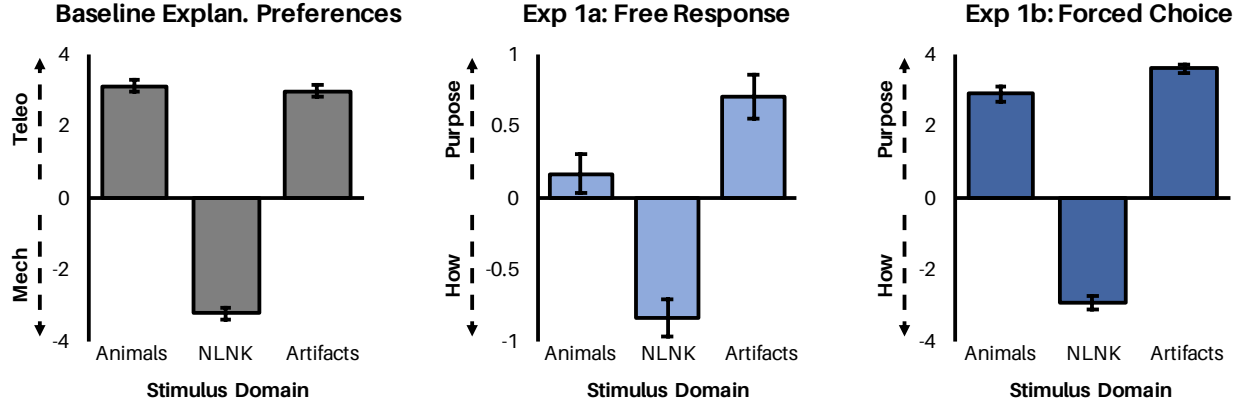


Figure 3: Results from experiments disambiguating ‘why’ questions. The x-axis represents chance performance, and explanation/question type is represented along the axis. In the leftmost graph, participants’ explanation preferences (found in a baseline experiment replicating Kelemen, 1999) are represented along the y-axis. In the right two graphs, participants’ coded (Experiment 1a) or scored (Experiment 1bc) responses are represented along the y-axis. In Experiment 1a, participants were asked to generate what question the agent asking the question really wanted to know. In Experiment 1b, participants were made to choose between a ‘how’ and a ‘purpose’ question in a forced-choice paradigm. Error bars represent ± 1 SE.

(even though they were indicating that the ‘why’ question was ambiguous); this more fine-grained approach allowed us to account for discrepancies in the number of questions generated by different participants. More participants generated only ‘purpose’ than only ‘how’ questions for artifacts ($\hat{p}=.70$), $p<.001$ (binomial test). The reverse was true of NLNK: More participants generated ‘how’ than ‘purpose’ questions ($\hat{p}=.81$), $p<.001$ (binomial test). However, for animals, participants were not more likely to generate only ‘purpose’ questions than ‘how’ questions ($\hat{p}=.58$), $p=.07$ (binomial test). We also conducted an even stricter test of our results to further clarify the relationship between participants’ preferred interpretation and the perceived ambiguity of the ‘why’ questions, by comparing the number of participants generating a single type of question to the number of those who ambiguously gave both ‘purpose’ and ‘how’ questions. For both animals and artifacts, participants were also more likely to generate only ‘purpose’ questions than to interpret ‘why’ ambiguously, animals: ($\hat{p}=.60$), $p=.03$ (binomial test); artifacts: $\hat{p}=.60$), $p<.001$ (binomial test), and for NLNK, (participants were more likely to generate only ‘how’ questions ($\hat{p}=.74$), $p<.001$ (binomial test).

(In Experiment 1a, we asked participants what an agent asking a ‘why’ question really wanted to know. Separately, we also ran a functionally identical study with a closely related question, asking participants what a ‘why’ question could “plausibly imply.” These results are qualitatively identical and are available in full on our OSF page.)

2.4 Results: Experiment 1b

When presented with ‘why’ questions about animals and artifacts, most participants thought the agent was really asking a ‘purpose’ question (animals: $\hat{p}=.86$; artifacts: $\hat{p}=.95$), $p<.001$ (binomial tests). In contrast, most participants chose the ‘how’ question for NLNK, ($\hat{p}=.87$), $p<.001$ (binomial test).

We also analyzed participants’ expectations by treating participants’ responses across items in the same domain as an average. Participants’ responses were scored as a 1 if they chose the ‘purpose’ question and a 0 if they chose the ‘how’ question. We then added these scores within each domain for each participant. We again found that participants thought that ‘why’ questions about animals ($M=2.88$, $SD=2.04$, $t(99)=14.14$, $p<.001$, $d=1.41$) and artifacts implied ‘purpose’ questions, ($M=3.58$, $SD=1.28$, $t(99)=27.96$, $p<.001$, $d=2.80$). For NLNK, participants instead thought that ‘why’ questions implied ‘how’ questions ($M=-2.94$, $SD=1.79$, $t(99)=16.47$, $p<.001$, $d=1.65$).

Finally, we also explicitly compared participants’ expectations between domains by using a repeated-measures ANOVA. There was a significant main effect of domain, $F(2,99)=434.21$, $p<.001$. Post-hoc tests demonstrated that participants’ were more likely to assume the ‘why’ question implied a ‘purpose’ question for animals than for NLNK, $t(99)=20.67$, $p<.001$, Bonferroni corrected, $d=2.07$. Similarly, participants were also more likely to assume the ‘why’ question implied a ‘purpose’ question for artifacts than for NLNK, $t(99)=25.87$, $p<.001$, Bonferroni corrected, $d=2.59$. While participants were also more likely to expect the ‘purpose’ questions for animals than for artifacts, these effects were much smaller, $t(99)=3.78$, $p<.001$, Bonferroni corrected, $d=.38$.

(In Experiment 1b, we asked participants to identify what somebody else asking a ‘why’ question really wanted to know. Separately, we also ran a functionally identical study with a closely related question, asking participants what they themselves would want to know. These results are qualitatively identical and are available in full on our OSF page.)

2.5 Discussion: Experiments 1a-b

In domains where adults prefer teleological explanations, they also think that an agent asking a ‘why’ question is intending to ask a ‘purpose’ question (and vice versa for mechanistic explanations and ‘how’ questions; see Figure 3). This was true both when participants were asked to rephrase the questions themselves and when they were made to choose between two possible interpretations—suggesting that these results capture people’s genuine pragmatic expectations.

In order to establish people’s explanation preferences for comparison, we also ran a baseline experiment replicating the classic paradigm (i.e., presenting people with ‘why’ questions and asking them to choose either the teleological or mechanistic explanation in response). (These results are also available in full on our OSF page). Participants not only had specific interpretations, unambiguous interpretations of ‘why’ questions, but these interpretations (in both Experiments 1a and 1b) also reflected their explanation preferences (tested in our baseline experiment and consistent with Kelemen, 1999).

In practice, this means that the endorsement of a teleological answer may not necessarily be an endorsement of a causal account; people may instead be evaluating these answers as responses to a *different* question—one that seeks information about something’s purpose, regardless of whether that purpose actually explains. This reframing implies that the endorsement of teleological answers is straightforward and reasonable, and need not demonstrate any fallacious or unscientific worldview (see Figure 1).

3 Experiments 2a-b: Jeopardy

Do people necessarily understand teleological answers as causally explanatory? Or might teleological answers offer a *different* sort of information? To find out, we have participants complete a ‘jeopardy’ game in which they must generate questions from answers (see Figure 2).

If people’s endorsement of teleological answers is indeed ‘unscientific’ (e.g., Kelemen, 2004; Rose, forthcoming), we should expect that people see teleological answers as responses to (implied) ‘how’ questions. Such responses would indicate that people understand teleological answers as a way to explain what caused the phenomenon in question. Suppose that instead, however, people see teleological answers as responses to (implied) ‘purpose’ questions. In this case, we shouldn’t consider the endorsement of such answers ‘unscientific’, because people are endorsing them as something other than causal explanations of how things came to exist. Thus, this jeopardy paradigm may offer us direct insight into how people are understanding the teleological answers they endorse.

3.1 Method: Experiment 2a

3.1.1 Participants

One hundred new adult participants completed a survey online through Amazon Mechanical Turk (data from 7 additional participants were collected but excluded for failing the training item; data from 5 additional participants were collected but excluded for failing to follow directions or produce interpretable questions;

see the Procedure section). This sample size was chosen to be identical to those in the previous experiments. This experiment was also preregistered.

3.1.2 Procedure

Total materials consisted of twenty-four explanations; twelve were mechanistic explanations and twelve were teleological explanations (see Table 1).

Participants first saw a training item that introduced the ‘jeopardy’ task. They were told that instead of being shown questions and answering them, they would be shown *answers* and be asked to generate the questions that led to those answers. As an example of the task, they were shown the answer “The book is over there, on top of the shelf but underneath the scarf” and asked to give a question that might have led to this answer. Participants who did not provide a question along the lines of “Where is the book?” were excluded and replaced for failing to understand the task (see the Participants section).

In each of the test items, participants were shown an image of the item and were told that somebody had provided an answer about that item (e.g., “Someone answered, ‘The mononykus has a long tail because its feathers were big and stuck out from behind its body.’”). See Table 1 for these explanations. They were asked, “What question might have led to this answer?” and were explicitly prompted not to use the word ‘why’.

If explanation preferences arise from pragmatic expectations, we should expect that participants expect mechanistic answers to have come from ‘how’ questions and teleological answers to have come from ‘purpose’ questions. Such a pattern of results would demonstrate a direct connection between explanation preferences and the implicit questions identified in the previous experiments.

Participants who either (1) systematically generated non-interpretable questions or (2) ignored the direction not to use the word ‘why’ were excluded and replaced (see the Participants section). Participants’ freely generated questions were then coded into ‘how’, ‘purpose’, and ‘other’ categories of questions using the same criteria as in Experiments 1a-b. All criteria were preregistered, and all of our data are publicly available.

3.2 Method: Experiment 2b

All elements of the experimental design were identical to those of Experiment 2a, except as stated below. One hundred new participants completed the survey online through Amazon Mechanical Turk (data from 1

additional participant was collected but excluded for failing the training item; see below). This sample size was chosen to be identical to those in the previous experiments. This experiment was also preregistered.

As in Experiment 1b, we wanted to replicate the free-response patterns of responses in a forced-choice task—in order to address the possibility that participants’ were generating the question that first came to mind but that they did not think specific answers could address only one or the other implicit question. Participants were thus now asked which of two questions led to the explanation they were given. One was a ‘how’ question (e.g., “How did the mononykus’ tail become long?”) and the other was a ‘purpose’ question (e.g., “What is the purpose of the mononykus’ long tail?”; a full list of these questions is available on our OSF page, under Table S1). The training item was also presented as a forced choice.

3.3 Results: Experiment 2a

The results for Experiments 2a and 2b are shown in Figure 4. Participants’ questions were not evenly distributed across categories of questions, $X^2(5, N = 923) = 166.37, p < .001$. The same was true when analyzing questions within each domain; whether thinking about answers regarding animals, ($X^2(5, N = 317) = 35.26, p < .001$), NLNK, ($X^2(5, N = 286) = 121.87, p < .001$), or artifacts, ($X^2(5, N = 320) = 116.58, p < .001$) participants generated different kinds of questions for different kinds of explanations.

We also analyzed participants’ likelihood of generating ‘how’ vs. ‘purpose’ questions (for both mechanistic and teleological explanations) by treating participants’ questions across items in the same domain as an average. ‘How’ questions were scored as a -1, ‘purpose’ questions were scored as a 1, and ‘other’ questions were scored as a 0. We then added these scores for each participant across all teleological explanations and, separately, across all mechanistic explanations. When presented with mechanistic explanations, participants were more likely to generate ‘how’ questions, ($M = -1.74, SD = 2.35, t(99) = 7.40, p < .001, d = .74$). Conversely, when presented with teleological explanations, participants were more likely to generate ‘purpose’ questions, ($M = 1.42, SD = 2.15, t(99) = 6.60, p < .001, d = .66$). The difference between participants’ questions in these two different explanation types was significant, $t(99) = 10.13, p < .001, d = 1.01$. This pattern was independently true within each domain, all $p < .001$, all $d > .5$.

3.4 Results: Experiment 2b

We analyzed the results of this experiment in the same way that we analyzed Experiment 2b. When presented with teleological explanations, a significant proportion of participants thought that the explanation resulted from a ‘purpose’ question, ($\hat{p} = .95$), $p < .001$ (binomial test). In contrast, a significant proportion of

What kind of question is implied by different answers? (Jeopardy Paradigm)

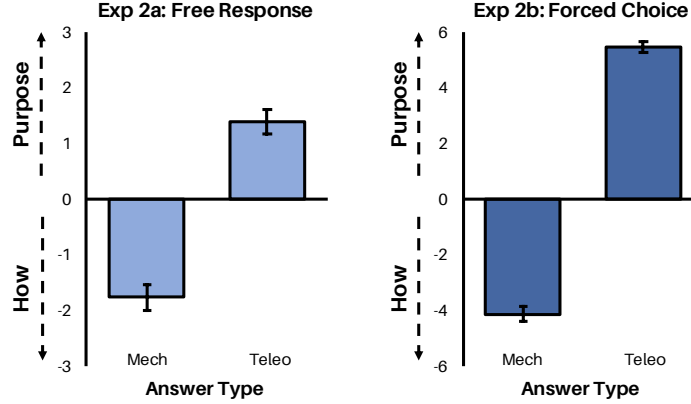


Figure 4: Results from ‘jeopardy’ experiments (Experiments 2a, 2b). The x-axis represents chance performance, and explanation type is represented along the axis. Participants’ coded (Experiment 2a) or scored (Experiment 2b) responses are represented along the y-axis. In both cases, participants thought that mechanistic explanations implied ‘how’ questions and that teleological explanations implied ‘purpose’ questions. Error bars represent ± 1 SE.

participants chose the ‘how’ question for mechanistic explanations, ($\hat{p}=.85$), $p<.001$ (binomial test). This pattern was independently true for each domain, all $\hat{p}>.75$, all $p<.001$.

We also analyzed participants’ likelihood of choosing the ‘how’ vs. ‘purpose’ questions (for both mechanistic and teleological explanations) by treating participants’ questions across items in the same domain as an average. ‘How’ questions were scored as a -1, and ‘purpose’ questions were scored as 1. We then added these scores for each participant across all teleological explanations and, separately, across all mechanistic explanations. When presented with mechanistic explanations, participants were more likely to choose the ‘how’ question, ($M=-4.16$, $SD=2.65$, $t(99)=15.72$ $p<.001$, $d=1.57$). Conversely, when presented with teleological explanations, participants were more likely to choose the ‘purpose’ question, ($M=5.44$, $SD=1.91$, $t(99)=28.50$, $p<.001$, $d=2.85$). The difference between participants’ choice of question in these two different explanation types was significant, $t(99)=25.60$, $p<.001$, $d=2.56$. This pattern was independently true within each domain, all $p<.001$, all $d>2.00$.

3.5 Discussion: Experiments 2a-b

Collectively, Experiments 2a-b show a direct link between implicit questions and teleological vs. mechanistic explanations. People thought that teleological answers were responding to ‘purpose’ questions and that mechanistic answers were responding to ‘how’ questions—regardless of whether they were generating the questions themselves or choosing one over the other. This relationship between kind of question and kind

of answer was also stable across domains. Not only do people have specific expectations about implicit questions in ‘why’ questions, but they then think that particular answers can address *only one* of these specific questions.

Why do people endorse teleological answers? We propose that people first have a specific expectation about a particular ‘why’ question (e.g., for an artifact, people likely assume that “Why?” is really asking a ‘purpose’ question). They are then faced with a choice of explanations, each of which can only answer one (implicit) question. They therefore choose an *answer* based on the implied *question* that they infer (see Figure 1). Thus, people may prefer teleological answers regardless of whether they are reasoning about them as being causally explanatory: A teleological answer may simply be a better answer for the (implied) question at hand.

4 Experiments 3a-b: Implications of understanding ‘why’

In introducing this paper, we suggested that people may sometimes endorse teleological answers for reasons other than their ability to provide a causal account. So far, we have shown that people do have pragmatic expectations about ‘why’ questions, such that one could plausibly answer a ‘why’ question without thinking about causal explanation. Here, we test a stronger prediction: that people will endorse teleological answers even when these purposes are explicitly non-causal. We presented participants with artifacts that were clearly designed for one purpose but were now being used for some other purpose. This allowed us to offer a teleological ‘explanation’ that referred to the artifact’s new purpose but could not actually *explain* the existence of its features. On the view that we are advancing, participants should still be drawn to the teleological answer that refers to the object’s new purpose—not because they think that purpose is causally explanatory, but because they think it can answer the implied ‘purpose’ question.

Potentially objectionable teleological explanations often involve a non-living natural entity (e.g., rocks). Here, however, we have opted to focus on artifacts. Using artifacts allowed us to manipulate cases both where purposes were clearly causal and when they were not. The goal of this study was to address a definitive case where people endorse explicitly non-causal teleological answers, regardless of domain. If people endorse non-causal teleological answers for artifacts, then we have reason to believe they may do so more generally—including in domains (like non-living natural kinds) where teleological answers are rarely causal.

4.1 Method: Experiment 3a

4.1.1 Participants

Two hundred and forty new adult participants completed a survey online through Amazon Mechanical Turk (data from 28 additional participants were collected but excluded for failing the comprehension check; see the Procedure section). This sample size was chosen to roughly match the number of participants per condition to prior experiments. This experiment was also preregistered.

4.1.2 Stimuli

Stimuli Each item consisted of a vignette about an artifact that had been designed for one purpose, but was now being used for a different purpose. These items were adapted from existing work (Joo, Yousif, & Knobe, 2021). Total materials consisted of four unique artifacts (straws / wind chimes, blankets / towels, ice cube trays / organizers, and washboards / percussion instruments). The order of the purposes was counterbalanced (e.g., for half of the subjects, one of the artifacts was designed as a straw but used as a wind chime, and vice versa). An example vignette is reproduced below:

These metal tubes were made as straws. That means that they were designed to move liquids from a cup to someone’s mouth.

But one day everyone in the town came together and agreed that the tubes shouldn’t be used in drinking liquids and that they should all instead be wind chimes. They decided that all of the metal tubes would be used only to make music when the wind blows.

Participants were then presented with a ‘why’ question (e.g., “Someone in the town asked, “Why are the metal tubes hollow?”) and were asked to choose between a mechanistic and a teleological explanation. Importantly, the teleological explanation always referred to the artifact’s new purpose (and not the purpose for which it was originally designed), meaning this purpose clearly could not be an explanation of the artifact’s features. A full list of vignettes and explanations is available on our OSF page, under Table S2.

4.1.3 Procedure

Participants were assigned to one of eight between-subjects conditions, such that each participant saw only one vignette. Four possible artifacts were presented, and the order of purposes was counterbalanced to produce eight conditions. Participants were shown the vignette along with its corresponding ‘why’ question. They were asked to choose an explanation a teleological vs. mechanistic explanation based on “which explanation makes the most sense.” This set up was designed to be closely matched with that in Kelemen (1999).

On a separate page after the task, participants were asked to provide an explanation of why they chose the answer that they did. These free-response answers were not part of any formal analyses, but we note that people most commonly referred to the importance of the item’s current use when choosing the teleological answer. All of these responses are available on our OSF page. Finally, participants were asked a true or false question about the artifact’s original design (e.g., “True or False: The metal tubes were originally designed as straws”); participants who failed this comprehension check were excluded and replaced (see the Participants section).

4.2 Method: Experiment 3b

All elements of the experimental design were identical to those of Experiment 3a, except as stated below. Two hundred and forty new adult participants completed a survey online through Amazon Mechanical Turk (data from 26 participants were collected but excluded for failing the comprehension check). Participants were now asked to choose between two teleological explanations—one which was the product of intentional design and one (also shown in Experiment 3a) could not be causally explanatory.

4.3 Results: Experiment 3a

We analyzed these data in the simplest way possible, by comparing the overall number of endorsements for the different answers types. We found that participants significantly preferred the teleological explanation over the mechanistic explanation, ($\hat{p}=.56$), $p=.03$ (binomial test) *despite* being shown vignettes that made it explicit that the teleological explanation in question was not a legitimate causal explanation. While this is a small effect, it is striking that people are still drawn to choosing teleological explanations at all. If participants were treating the ‘why’ questions in this experiment as questions seeking causal explanations, then they should have overwhelmingly chosen the mechanistic answer (since the teleological answer was obviously non-causal).

4.4 Results: Experiment 3b

When offered two teleological explanations, participants preferred the explanation that could offer a causal explanation (i.e., that referred to the artifact’s designed purpose) over the one that did not (i.e., that referred to the artifact’s new purpose); ($\hat{p}=.60$), $p<.001$ (binomial test). Note, however, that a significant portion (40%) were still drawn to the teleological explanation which could not offer a causal account.

4.5 Discussion: Experiments 3a-b

These experiments present straightforward instances where a possible teleological answer is clearly separable from a correct causal account of the features in question. Participants were systematically biased towards choosing a teleological explanation, even when that answer was clearly not a possible causal explanation. These results suggest that multiple factors are at play in people’s explanation preferences. On the one hand, people clearly understand causal accounts and will sometimes choose answers that can give causal explanations over those that cannot (Experiment 3b). Critically, though, the fact that people remain drawn towards a non-causal teleological explanation—even when a comparable teleological explanation is available—suggests that people’s ‘explanation’ preferences cannot be entirely explained by their understanding of causal explanation.

This does not mean that people’s preferences are irrational. Quite the opposite, these results imply that people’s endorsement of teleological answers in response to ‘why’ questions need not imply anything at all about how they understand the causality of teleological answers—or whether they endorse some broader theistic belief. Instead, as we show here, people likely understand these teleological answers as responses to implied ‘purpose’ questions.

5 General Discussion

What makes a good explanation? Most past research has addressed this question by focusing on the features of the explanations themselves. In contrast, our work has focused on the ways in which people interpret the *questions* preceding these explanations. We show that, over and above the explanatory virtues of a particular answer, people are drawn to those answers that address the relevant implicit question—and that, sometimes, this question has little to do with causal explanation.

Experiments 1a-b demonstrated that ‘why’ questions are *not* pragmatically ambiguous: People have clear, systematic expectations about implicit questions. Experiments 2a-b showed that people understand the relationship between certain kinds of ‘explanations’ (e.g., teleological vs. mechanistic) and certain kinds of questions (e.g., ‘purpose’ vs. ‘how’). Finally, Experiments 3a-b showed that people endorse teleological explanations that address relevant implicit questions but which cannot offer a causal explanation (i.e., because we explicitly offered an alternative causal account). Collectively, these results show that people endorse teleological answers for reasons distinct from their ability to provide a causal account.

5.1 Understanding ‘Why’

Most studies on explanation preferences depend on ‘why’ questions (e.g., Banerjee & Bloom, 2015; Kelemen & Rosset, 2009; Liquin & Lombrozo, 2018; Schachner et al., 2016). Thus, this new insight into ‘why’ questions may change how we understand known ‘explanation’ preferences—and how we interpret them.

Our proposal is simple: Adults understand ambiguous ‘why’ questions as one of two possible implied questions; preferences for teleological answers are primarily explained by an inference about an implied ‘purpose’ question. Thus, a teleological answer can be satisfying regardless of whether it provides causal account, insofar as it responds to a question that was not seeking such an account and does not require such an account in a satisfactory answer (see also Joo, Yousif, & Keil, 2021). In this case, a teleological answer is an appropriate answer to the inferred ‘purpose’ question—indeed, the only appropriate answer to such a question.

This reframing has important consequences for many of the theories built around people’s teleological preferences. For instance, much debate in both philosophy and psychology has centered around adults’ endorsement of teleological explanations about animals and whether these explanations require intelligent design (e.g., Lombrozo & Carey, 2006; Ruse, 2000). Our view would suggest that people’s preferences have more to do with implicit questions than any causal beliefs (whether those beliefs are explicitly about intentional design or about other kinds of causal influence). Even outside of the context of teleology and natural selection, this perspective thus offers a way to understand people’s (sometimes very broad) teleological preferences without needing to classify them as “scientifically unwarranted” (c.f., e.g., Kelemen et al., 2013; Rose, forthcoming). While teleological answers sometimes do provide a causal account, and people may be especially likely to endorse such answers in those cases, it does not mean that all endorsement of teleological answers implies causal reasoning. As a consequence of this view, we can understand why people sometimes endorse teleological explanations even in cases where teleology seems not to have any causal relevance at all; even in these cases, teleological answers may address the relevant implicit question.

5.2 What is teleology *for*, if not explanation?

If the appeal of teleological answers encompasses cases where teleology *isn’t* being treated as a straightforwardly causal answer, then how can we alternatively characterize people’s affinity for such information? One possibility is to take a generally broader view of explanation. On some existing views, what constitutes a valid explanation in everyday contexts may simply be anything that increases understanding of a given phenomenon in some helpful way (see, e.g., Wilson / Keil, 1998). Indeed, this stance is necessary to make

sense of certain kinds of explanations (e.g., mathematical or nomological explanations) which straightforwardly do not offer a causal account. Teleological explanations may understood in a similar way—as a kind of explanation that does not require an etiological causal account in order to be satisfactory.

Another possibility is to recharacterize people’s preferences entirely as a kind of *information* preference rather than as an *explanation* preference. It may be, for instance, that people simply have an affinity for learning about what things are for—perhaps because such information succinctly conveys how we can or should interact with these things (i.e., knowing that a hammer is for driving nails includes knowing to look for a hammer when needing to drive in nails and not at some other time).

Regardless, a new question arises about what makes teleology compelling, independent of its potential to causally explain. Several intriguing possibilities exist. First, teleology may be useful in contextualizing objects in a larger world. For instance, learning that an animal’s tail is good for keeping its balance may help place the tail in the context of the whole animal, and learning that a tree serves the function of removing carbon dioxide from the atmosphere may help illustrate a wider ecosystem or its relationship to humans (see ojalehto et al., 2013). In other words, regardless of its ability to *explain*, teleology may provide information about the relationship between different objects and some relevant community (see also Joo, Yousif, & Keil, 2021; Siegel & Callanan, 2007; Scheele, 2006). Second, teleological explanations may also help individuals make predictions about unfamiliar entities. Consider, for example, a child encountering a microwave for the first time. Learning that the microwave is for heating food tells them when it is likely to be used and how to interact with it meaningfully in the future (see also Lombrozo & Gwynne, 2014). Some kinds of teleological thinking may also help children and adults to make inferences about similar or the same object (i.e., the child may learn something about how to understand microwaves in general, see Korman & Khemlani, 2020). Finally, teleology may also help carve up the world into meaningful categories and objects. We know, for example, that teleology plays a key role in judgments of category membership (Matan & Carey, 2001; Rose & Nichols, 2019, 2020); object constitution (Rose & Schaffer, 2017); and persistence (Rose et al., 2018).

5.3 Conclusion

Here, we show that perhaps our simplest, most fundamental, and possibly most ubiquitous question, “Why?”, is not a single question at all. “Why?” refers to at least two distinct (and more specific) questions, each of which seeks a unique form of information. This simple insight may reshape our understanding of explanation preferences, raising questions about the broad theories on which such preferences are based. We argue that in order to better understand the nature of explanations, we must begin by understanding the questions that precede them.

Acknowledgments

For extensive conversations about all things teleology, we thank Josh Knobe. For pushing us to think harder about causation and explanation, we thank Tania Lombrozo. For thinking through ‘why’, ‘how’, ‘purpose’, and in-between questions with us, we thank Flora Zhang and Jessie Cheung. For helping to facilitate this collaboration in the first place, we thank Richard Ahl. And for many other helpful comments and conversation, we thank Nicole Betz, Katie Coyne, Alicia Mazzurra, Brynn Sherman, and all members of the Yale Cognition and Development Lab. This project was supported by a National Science Foundation Graduate Research Fellowship awarded to S. R. Yousif.

References

- Allen, C., & Bekoff, M. (1995). Function, Natural Design, and Animal Behavior: Philosophical and Ethological Considerations. In N.S. Thompson Editor (Eds.), *Perspectives in Ethology, Vol.1* (pp. 1-47). Plenum Press.
- Banerjee, K., & Bloom, P. (2015). "Everything Happens for a Reason": Children's Beliefs About Purpose in Life Events. *Child Development, 86*, 503–518.
- Joo, S., Yousif, S.R., & Knobe, J. (2021). .Teleology beyond explanation. *Mind & Language*
- Kelemen, D. (1999). Why are rocks pointy? Children's preference for teleological explanations of the natural world. *Developmental Psychology, 35*, 1440–1452.
- Kelemen, D. (2004). Are Children "Intuitive Theists"? Reasoning About Purpose and Design in Nature. *Psychological Science, 15*, 295–301.
- Kelemen, D., & DiYanni, C. (2005). Intuitions About Origins: Purpose and Intelligent Design in Children's Reasoning About Nature. *Journal of Cognition and Development, 6*, 3–31.
- Kelemen, D., & Rosset, E. (2009). The Human Function Compunction: Teleological Explanation in Adults. *Cognition, 111*, 138-143.
- 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*, 1074-1083.
- Korman, J. & Khemlani, S. (2020). Teleological generics. *Cognition, 200*, 104157.
- Lewry, C., Lombrozo, T., & Kelemen, D. (2020). People view humans as existing for purposes and condemn those who fail to fulfill them. *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*. Toronto: ON: Cognitive Science Society.
- Liquin, E. G., & Lombrozo, T. (2018). Structure-function fit underlies the evaluation of teleological explanations. *Cognitive Psychology, 107*, 22-43.
- Lombrozo, T., & Carey, S. (2006). Functional explanations and the function of explanation. *Cognition, 99*, 167-204.
- Lombrozo, T., & Gwynne, N. Z. (2014). Explanation and inference: mechanistic and functional explanations guide property generalization. *Frontiers in Human Neuroscience, 8*, 700.

- Matan, A., & Carey, S. (2001). Developmental changes within the core of artifact concepts. *Cognition*, 78, 1–26.
- oijalehto, b., Waxman, S.R., & Medlin L.M. (2013). Teleological Reasoning About Nature: Intentional Design or Relational Perspectives? *Trends in Cognitive Sciences*, 17, 166-171.
- Rose, D. (Forthcoming). Mentalizing Objects. In *Oxford Studies in Experimental Philosophy* 4. Oxford, England: Oxford University Press.
- Rose, D. & Nichols, S. (2019), Teleological Essentialism. *Cognitive Science*, 43, 12725.
- Rose, D. & Nichols, S. (2020), Teleological Essentialism: Generalized. *Cognitive Science*, 44, 12818.
- Rose, D., & Schaffer, J. (2017). Folk Mereology is Teleological. *Noûs*, 51, 238–270.
- Rose, D., Schaffer, J., & Tobia, K. (2020). Folk teleology drives persistence judgments. *Synthese*, 1-19.
- Ruse, M. (2000). Teleology: yesterday, today, and tomorrow? *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 31, 213–232.
- 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.
- Scheele, M. (2006). Function and use of technical artefacts: social conditions of function ascription. *Studies in History and Philosophy of Science*, 23-36.
- Siegel, D., & Callanan, M. (2007). Artifacts as conventional objects. *Journal of Cognition and Development*, 8, 183–203.
- Wilson, R. A., Keil, F. (1998). The shadow and shallows of explanation. *Minds and Machines: Journal for Artificial Intelligence, Philosophy and Cognitive Science*, 8, 137-159.