COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 46 (2022) e13202

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13202

Are we Teleologically Essentialist?

Sehrang Joo, Sami R. Yousif

Department of Psychology, Yale University

Received 18 October 2021; received in revised form 22 August 2022; accepted 24 August 2022

Abstract

People may conceptualize certain categories as held together by a category-specific "essence"—some unobservable, critical feature that causes the external features of a category to emerge. But what is the nature of this essence? Recently, Rose and Nichols have argued that something's essence is fundamentally its telos or purpose. However, Neufeld has challenged this work on theoretical grounds, arguing that these effects arise only because people *infer* an underlying internal change when reasoning about a change in telos. In Neufeld's view, it is the underlying internal cause, and not the telos itself, that serves as an essence (consistent with classic views of scientific essentialism). Here, we ask: Is teleology the primary force behind psychological essentialism? We begin by successfully replicating Rose and Nichols' key findings in support of teleological essentialism. In two further experiments, however, we demonstrate that teleology may not be the central way that people understand the essences of living things. We show that internal changes matter at least as much as changes in teleology. These findings suggest that while teleology may be one important cue to category membership and the essences of living things, it may be premature to say that we are "teleologically essentialist."

Keywords: Psychological essentialism; Teleological essentialism; Scientific essentialism; Teleology

1. Introduction

What makes a bee a bee, and not a spider? What makes a hummingbird a hummingbird, and not a vulture? People are typically thought to represent categories like "bee" and "hummingbird" as held together by a sort of category-specific "essence"—an "underlying reality" that gives rise to external features (e.g., Gelman, 2004; Kelemen & Carey, 2007). This essence need not be a tangible, identifiable thing; some views posit the existence of a sort of "placeholder" essence (i.e., people can represent something as *having* an essence without yet fully specifying *what* that essence is; Gelman, 2003, 2004). There are many ways to 'fill

Correspondence should be sent to Sehrang Joo, Departement of Psychology, Princeton University, Princeton, NJ 08540, USA. E-mail: sehrang.joo@princeton.edu

in' this placeholder. One common way we may do so, for instance, is via a growing scientific understanding of the organism's *insides* (e.g., Gelman, 2003; Gelman & Heyman, 1999; Gelman & Markman, 1986, 1987; Inagaki & Hatano, 2002; Newman & Keil, 2008). For example, something's DNA may loosely be considered its essence. In contrast to this view, Rose and Nichols (2019, 2020) propose a radically different idea: the notion of *teleological essentialism*, a view in which it is instead an organism's purpose that defines its essence.

1.1. Teleological essentialism

According to teleological essentialism, what makes a bee a bee is not its black and yellow coloring or its buzzing, or even its DNA—but its *purpose* (e.g., one might think a bee's purpose is to pollinate flowers or to make honey). The *telos* of the bee, whatever it may be, is what ultimately gives rise to its other features—including its internal makeup, external features, and other behaviors. In other words, teleological essentialism shifts the focus away from people's intuitions about animals' scientific nature, and instead recharacterizes psychological essentialism around an understanding of animals' purposes.

Teleological essentialism has the potential to explain a wide range of laypeople's judgments. For instance, teleological essentialism may explain why people say that tea (which they say is 91% H_2O) is not water, although a swimming pool (which they say is 82.6% H_2O) is water (Malt, 1994)—because the latter serves the function of water, whereas the former does not. Similarly, it may explain why teleological judgments are so widely influential on other sorts of judgments. For instance, consider the relevance of teleological information to people's composition judgments (i.e., whether some new object is created from the combination of smaller, existing objects). While gluing any random two objects together (e.g., a book and a stick) is typically seen as insufficient for truly creating a new object, Rose and Schaffer (2017) show that if the combination has some new purpose (e.g., as a fly swatter), people are more likely to agree that a new object has been composed. If the essence of an object is its telos, then this influence is easy to understand: Once there is a new essence (i.e., new telos), then there is a new object. In this way, teleological essentialism may help bring together existing work demonstrating the influence of teleological information also in how people understand persistence (i.e., whether something is the same or a new object after a major transformation; Rose, Schaffer, & Tobia, 2018) or evaluate explanations (Joo et al., 2022; Kelemen, 1999; Kelemen & Rosset, 2009; Kelemen, Rottman, & Seston, 2013).

Teleological essentialism is clearly an intriguing possibility, and it may also be the right way to capture people's essentialist intuitions. Consider, for example, this case from Rose and Nichols (2019):

> Some very talented and skilled scientists, Suzy and Andy, decide that they are going to perform a special operation on a bee. They removed its wings and antennae, lengthened its legs, and added a new pair of legs. They also inserted into the back of it something for making webs and trained the animal so that it would eat insects. After running some tests, they found that the thing after the special operation didn't pollinate flowers or

make honey. Instead, it only spun webs to catch insects and eat them. (Experiment 1 in Rose & Nichols, 2019)

In this scenario, people indicate that the bee has become a spider (on a scale from 1 = it is definitely a bee, to 7 = it is definitely a spider, the mean score reported is 5.73). In a contrasting condition in which the bee's telos is preserved rather than changed, people indicate that the bee remains a bee (on the same scale, the mean score reported is 2.57). This basic phenomenon holds true in several cases, including when the scientists directly change the original bee's features (as shown here), operate on the creature's "insides," or conduct an experiment placing a baby bee in a spider's cage. Rose and Nichols (2020) extended this finding, showing similar effects not only for animals, but also for artifacts and nonliving natural kinds. On the basis of these findings, Rose and Nichols conclude that, "…teleological essentialism is general, [meaning] that essence is represented by a kind of telos for a broad range of categories" (p. 16).

1.2. Against teleological essentialism

Recent work, however, challenges teleological essentialism on theoretical grounds. Neufeld (2021a) claims that the results observed by Rose and Nichols (2019, 2020) can be explained by a canonical view of scientific essentialism. Specifically, Neufeld argues that people may use information about a change in something's telos as a way to infer a change in its underlying scientific essence.

Consider, for instance, this scenario from Hampton, Estes, and Simmons (2007): A bird falls out of the sky into a vat of hazardous chemicals, and it emerges from the chemicals looking more like a spider than a bird. The critical finding here, according to Neufeld, is not that people view the creature after its transformation as a spider, but that they do so because of some deeper inference about the insides of the creature. That is, participants reasoned that some internal change must have given rise to the observed physical or behavioral changes, and this inferred internal change was what led them to classify the creature as a spider. Perhaps the animals' DNA was altered, for example—in which case scientific essentialism *should* predict a change in category membership. In other words, on this view, only the *internal* (or scientific) essence is a true cause of the animals' external features, and so people infer that any change in such features is the result of a change in essence (see also Rehder & Hastie, 2001). This is consistent with classic work (Keil, 1989), in which changes attributed to external causes do *not* result in changes to category membership, ostensibly because it is clear that no essence has been altered.

Neufeld argues that these inferences are also present in the experiments from Rose and Nichols (2019). After a bee gains the telos of a spider (i.e., it starts spinning webs and catching insects instead of pollinating flowers and making honey), Rose and Nichols find that people think the creature after this change is a spider. In Neufeld's view, the description of the creature's telos after the experiment may still lead to inferences about further, perhaps deeper, internal changes. In other words, when participants are told that the creature spins webs like a spider, they assume that key "insides" have changed to be like those of a spider—

and when told that the creature continues to pollinate flowers, they assume that these key "insides" are still more like those of a bee. In other words, Rose and Nichols's (2019, 2020) vignettes may be ambiguous with respect to the causal nature of the observed teleological changes.

1.3. Current study

The goal of the experiments reported here is to test whether teleological essentialism holds up empirically in a range of scenarios. Our aim is to provide stronger test cases of teleological essentialism: Does teleology *uniquely* or even *primarily* drive people's essentialist intuitions? On the one hand, it is possible that teleological essentialism is robust, and that people do in fact categorize objects based on their teleology—over and above any inferences about underlying causes. On the other hand, it is possible that while people's understanding of animals' essences may be affected by teleological information, their intuitions are primarily the result of inferring deeper, internal changes.

We investigate people's essentialist intuitions by examining their categorization judgments. What kinds of changes (e.g., to its telos vs. to its DNA) determine what category something belongs to? We begin (Experiments 1a–c) by replicating three key experiments from Rose and Nichols (2019). In line with their original results, we show that teleology seems to play some role in people's categorization of animals. Next (Experiment 2), we directly contrast teleological changes with changes to deeper "insides." We find that people's categorization judgments are driven at least as much as by their understanding of animals insides as they are by their understanding of animals' tele. Finally (Experiment 3), we contrast teleological changes with insides changes in separate scenarios. We present participants with simplified vignettes in an effort to minimize the extent to which they may infer any additional changes, and find that in some cases, insides changes may primarily drive people's categorization judgments.

2. Experiments 1a–c: Direct replications

Here, we aim to replicate Experiments 1–3 from Rose and Nichols (2019). We contrast cases where a creature's telos changed with cases in which it did not, and ask participants to evaluate "to what extent [they] think that the thing after the special operation is a bee or spider?"

2.1. Method: Experiment 1a

2.1.1. Participants

One hundred adult participants completed a survey online through Amazon Mechanical Turk, recruited through CloudResearch. The sample size was preregistered. All participants lived in the United States. Data from an additional 25 participants were collected but excluded for failing one or more comprehension checks (see Procedure section).

2.1.2. Procedure

Participants were given the following scenario:

> Some very talented and skilled scientists, Suzy and Andy, decide that they are going to perform a special operation on a bee. They removed its wings and antennae, lengthened its legs and added a new pair of legs. They also inserted into the back of it something for making webs and trained the animal so that it would eat insects.

On the same page, they were then shown two images, from before and after the special operation (the "before" image looked like a bee, and the "after" image looked like a spider). Participants were then told that the creature either changed or preserved its telos, or purpose, after the operation:

> [Telos changed] After running some tests, they found that the thing after the special operation *didn't* pollinate flowers or make honey. Instead, it *only* spun webs to catch insects and eat them.

> [Telos preserved] After running some tests, they found that the thing after the special operation *didn't* spin webs to catch insects and eat them. Instead, it *only* pollinated flowers and made honey.

Participants were asked two true/false comprehension questions. They were asked to evaluate a statement about the events of the scenario (e.g., "Suzy and Andy performed a special operation on a bee") and a statement about the teleology manipulation (e.g., "The thing after the special operation only spins webs to catch insects and eat them.") These comprehension questions, as well as all other details of the vignette, were exactly the same as those used in Rose and Nichols (2019) Experiment 1.

Participants were then asked "To what extent do you think that the thing after the special operation is a bee or a spider?" on a 1–7 scale, where 1 was "It is definitely a bee" and 7 was "It is definitely a spider."

Data, materials, and preregistration information for this experiment and all following can be found on the Open Science Framework (OSF) at https://osf.io/x5gf6/?view_only=a43050f1fa0f4802b0025b3a59beb56c.

2.2. Results: Experiment 1a

Results from Experiments 1a–c are shown in Fig. 1. Participants' categorization of the creature after its transformation depended on whether or not its telos had changed (independent samples t-test, $t(98) = 7.93$, $p < .001$, $d = 1.59$). Specifically, when the creature's telos changed to that of a spider, participants thought it was a spider ($M = 5.46$, $SD = 1.69$; one sample t-test, $\mu = 3.5$, $t(49) = 6.10$, $p < .001$, $d = 0.86$) and when the creatures' telos was still that of a bee, participants thought it was a bee ($M = 2.72$, $SD = 1.76$; one sample t-test, $\mu = 3.5$, $t(49) = 5.14$, $p < .001$, $d = 0.73$).

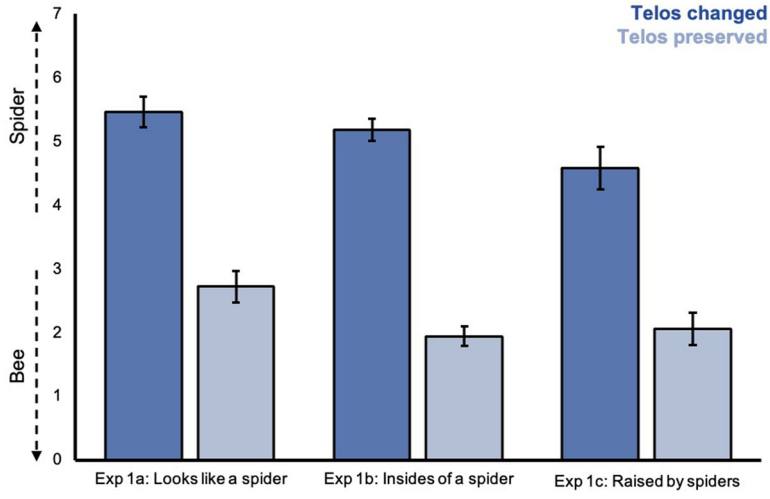

Fig. 1. Results of Experiments 1a–c, replicating Experiments 1–3 in Rose and Nichols (2019). In all experiments, participants thought the creature after its transformation was more like a spider when its telos changed than when it did not. In Experiment 1a, the creature ended up looking like a spider. In Experiment 1b, the creature's insides changed. In Experiment 1c, it grew up in a cage full of spiders. Error bars represent +/– 1 standard error.

2.3. 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 the previous experiment. Data from an additional 10 participants were excluded for failing one or more comprehension checks.

Participants were now presented with the vignette from Rose and Nichols (2019) Experiment 2:

> Some very talented and skilled scientists, Suzy and Andy, decide that they are going to perform a special operation on a bee. They decide to remove the insides of the bee and replace them with the insides from a spider.

They were shown the same image of a bee twice: once to represent the creature before the operation and once to represent it after the operation.

2.4. Results: Experiment 1b

Participants' categorization of the creature after its transformation depended on whether or not its telos had changed, independent samples t-test ($t(98) = 13.84$, $p < .001$, $d = 2.77$). Specifically, when the creature's telos changed to that of a spider, participants thought it was a spider ($M = 5.18$, $SD = 1.24$; one sample t-test, $\mu = 3.5$, $t(49) = 6.73$, $p < .001$,

$d = 0.95$), and when the creatures' telos was still that of a bee, participants thought it was a bee ($M = 1.94$, $SD = 1.10$; one sample t-test, $\mu = 3.5$, $t(49) = 13.30$, $p < .001$, $d = 0.188$).

2.5. Method: Experiment 1c

All elements of the experimental design were identical to those of previous experiments, 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 the previous experiment. Data from an additional seven participants were excluded for failing one or more comprehension checks.

Participants were now presented with the vignette from Rose and Nichols (2019) Experiment 3:

> Some very talented and skilled scientists, Suzy and Andy, decide that they are going to perform a special experiment with a newborn bee. After an egg hatches, they place the newborn bee in a cage full of spiders.

No images were shown because no images were shown in the original experiment.

2.6. Results: Experiment 1c

Participants' categorization of the creature after its transformation depended on whether or not its telos had changed, independent samples t-test ($t(98) = 6.04$, $p < .001$, $d = 1.21$). Specifically, when the creature's telos changed to that of a spider, participants thought it was a spider ($M = 4.58$, $SD = 2.33$; one sample t-test, $\mu = 3.5$, $t(49) = 3.23$, $p = .002$, $d = 1.97$), and when the creatures' telos was still that of a bee, participants thought it was a bee ($M = 2.06$, $SD = 1.81$; one sample t-test, $\mu = 3.5$, $t(49) = 5.62$, $p < .001$, $d = 1.14$).

2.7. Discussion: Experiments 1a–c

Here, we replicate the key findings of Rose and Nichols (2019). Consistent with teleological essentialism, people's classifications of the creature after its transformation were affected by whether or not its telos had changed.

3. Experiment 2: Teleological versus insides changes

The previous experiments are consistent with a view that people are teleologically essentialist. However, a strong view of teleological essentialism would predict not only that people's category judgments are affected by the creature's telos, but also that their judgments are affected *only* or at least *primarily* by the creature's telos. Here, we test this prediction directly. We contrast cases where a creature's telos did/did not change with cases where its "insides" (i.e., its organs and DNA) did/did not change to be more like a different animal, and ask participants how best to classify the creature after its transformation: as its original classification, as its possible new classification, or neither.

3.1. Method

3.1.1. Participants

In addition to the bee to spider transformation adapted from Rose and Nichols (2019), we added two additional scenarios in this experiment: a transformation from a hummingbird to a vulture (see also Rose & Nichols, 2020), and a transformation from a spider to a hummingbird. The full text of all vignettes is available on our OSF page. Across these scenarios, 1200 new participants completed the survey online, recruited directly from Amazon Mechanical Turk. This sample size was chosen for each vignette to double the number of participants per condition from Experiment 1 (since, unlike the previous experiments, this experiment was not a direct replication). The sample size was preregistered. All participants lived in the United States. Data from an additional 123 participants were collected but excluded for failing one or more comprehension checks (see Procedure section).

3.1.2. Procedure

In a 2 (Telos: changed/stayed the same) x 2 (Insides: changed/stayed the same) x 3 (Transformation: spider to bee, hummingbird to vulture, spider to hummingbird) between-subjects design, participants were randomly assigned to one vignette. Language in the vignettes was adapted from Rose and Nichols (2019) with minimal changes. Participants in all conditions were shown the following setup:

> Some very talented and skilled scientists, Suzy and Andy, decide that they are going to perform a special experiment on a bee (spider, hummingbird).

This was followed by information about whether the creature's telos changed and about whether its insides changed. For example, in the bee to spider transformation conditions, participants saw a version of the following:

> Before the special experiment, the bee pollinated flowers and made honey.
>
> [Telos Changed] After the special experiment, it no longer pollinated flowers and made honey. Instead, it spun webs to catch insects and eat them. In fact, Suzy and Andy found that it only ever acted like a typical spider.
>
> [Telos Same] After the special experiment, it still pollinated flowers and made honey. And it still didn't spin webs to catch insects and eat them. In fact, Suzy and Andy found that it only ever acted like a typical bee.
>
> [Insides Changed] Afterwards, Suzy and Andy also found that the creature's insides seemed completely different—all the organs and even the DNA itself were now much more like those of a spider than those of a bee. In fact, Suzy and Andy could not find any discernible differences between the creature's insides and the insides of a typical spider.

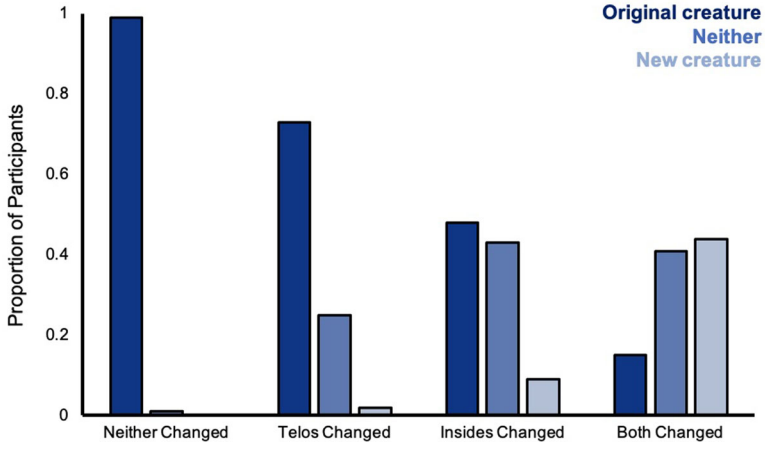

Fig. 2. Results from Experiment 2. Participants thought that a creature that changed neither its telos nor insides (i.e., a bee changing like a spider) after an experiment was still the same kind of creature. When *both* the creature's telos and its insides changed, participants tended to think it was now a different kind of animal. However, participants had mixed intuitions when considering cases where only the creature's telos *or* its insides changed, and a majority continued to classify it as its original kind.

> [Insides Same] Afterwards, Suzy and Andy also found that the creature's insides seemed like they hadn't changed at all—all the organs and even the DNA itself were still much more like those of a bee than those of a spider. In fact, Suzy and Andy could not find any discernible differences between the creature's insides and the insides of a typical bee.

In Rose and Nichols (2019), participants were always shown information about the creature's telos *last*. Here, we reverse this order in case this was one factor influencing participants' judgments. Participants were always shown information about the creature's telos first, followed by information about its insides.

Participants were then asked a forced-choice question about the best way to classify the creature after the special experiment: as a bee, a spider, or neither. We chose this discrete measure in order to ask this question in a commonsense way and to give participants clear options, rather than rely on their interpretations of a scale.

Finally, participants were asked two true/false comprehension questions. One question, about whether or not the creature's telos changed (i.e., "The thing after the special experiment only spins webs to catch and eat them"), was taken directly from Rose and Nichols' procedure. The second question was the analogous check of participants' understanding of whether or not the creature's insides changed (i.e., "The creature after the special experiment has organs that seem more like a spider's than like a bee's").

3.2. Results

Results from Experiment 2 are shown in Fig. 2. Data were analyzed using R with the nnet package, and were fit to multinomial logistic models. There was a significant main effect of

both a change in insides, $X^2(3, N = 1200) = 461.40$, $p < .001$, and a qualitatively smaller effect of a change in telos, $X^2(3, N = 1200) = 223.0256.12$, $p < .001$. The same was true when looking at the main effects of both telos and insides within each vignette, $ps < .001$.

Compared to when neither the creature's telos nor its insides changed, participants were *less* likely to classify the creature as still its original kind either when its telos changed (neither changed vs. telos changed: 99% vs. 74%; two-proportions z-tests, $zs > 3.95$, $ps < .001$) or when its insides changed (neither changed vs. telos changed: 99% vs. 52%; two-proportions z-tests, $zs > 7.39$, $ps < .001$). However, participants were also *more* likely to classify the creature as its new kind when both its telos and insides changed, compared to both when only its telos changed (both changed vs. telos changed: 45% vs. 4%; two-proportions z-tests, $zs > 5.42$, $ps < .001$) or only its insides changed (both changed vs. insides changed: 45% vs. 8%; two-proportions z-tests, $zs > 3.91$, $ps < .001$).

Participants were also overall more likely to classify the creature as "neither" when only its insides changed versus when only its telos changed (insides changed vs. telos changed: 40% vs. 23%; two-proportions z-test, $z = 4.67$, $p < .001$). This pattern was qualitatively true in all scenarios but was statistically significant in only the spider to hummingbird (two-proportions z-test, $z = 3.92$, $p < .001$) and hummingbird to vulture (two-proportions z-test, $z = 2.69$, $p = .007$) cases.

Finally, across all three transformations, there was a significant interaction between a change in telos and a change in insides, $X^2(3, N = 1200) = 26.13$, $p < .001$. In other words, participants were most likely to think that the creature was now a different kind of animal when *both* its telos and insides changed (vs. when either changed in isolation).

3.3. Discussion

While there was a main effect of telos changes on people's judgments, there was also (an equally large) main effect of changes to the creature's insides. When comparing only the conditions in which the creature's insides changed, the results of this experiment conceptually replicated the influence of teleology on people's understanding of animals' essences: As in Experiments 1a–c, people tended to think the creature was a bee when its telos did not change and to think it was a spider when it did. However, the 2×2 design in this experiment also demonstrated that the same could be said of changes to the creature's insides. When comparing only the conditions where the creature's telos changed, people tended to think the creature was a bee when its insides did not change and to think it was a spider when they did. In other words, these results are inconsistent with a strong view of teleological essentialism. People's judgments were clearly not driven *only* or even *primarily* by teleological information. Moreover, information about the insides seemed to be overall more informative of how the creature should be classified.

What might this mean for teleological essentialism more generally? On the one hand, it seems that people's judgments are affected *to some extent* by teleological information. On the other hand, trying to isolate isides and telos changes from each other demonstrated the difficulty in pulling these concepts fully apart. Many of our participants directly noted that it seemed unusual or even impossible both for the creature to act like some new kind of

animal without any discernible changes to its insides, and for it to act like the same kind of animal when its insides were indistinguishable from that of something else (all responses to this free-response question are available in full on our OSF page). It was challenging, in practice, to convince participants that no internal changes occurred when the creature changes to such an extent. Thus, the effects of teleology that we do observe may nevertheless be due to some other inference that participants are making.

4. Experiment 3: Teleological versus insides changes isolated

In the previous experiment, people's category judgments were equally influenced by both teleological and internal changes. Here, we (1) seek to conceptually replicate these results in a different paradigm, and (2) ask whether one factor may sometimes carry greater weight than another.

We began with one of Rose and Nichols' (2019) vignettes—one in which scientists experiment on a baby bee by placing it in a cage full of spiders—and examined cases where either the creature's telos changed (and it spins webs to catch insects) or its insides changed (and it has DNA like that of a spider). As in Experiment 2, we also adapted this vignette for a transformation from a hummingbird to a vulture and from a spider to a hummingbird.

Here, we made two key changes from Experiment 2. First, participants were only given information about *either* a telos change *or* an insides change. Second, we simplified the language of the vignettes by eliminating mention of "scientists" conducting a "special experiment." We removed this language because we think it may directly encourage erroneous inferences. For example, one might think that the scientists made other changes to the creature while setting up their experiment, and in this case, teleological changes may appear to have a bigger impact on participants' category judgments than they otherwise would. These changes resulted in a simpler version of a transformation between kinds (see also Gelman & Wellman, 1991).

4.1. Method

4.1.1. Participants

Six hundred new participants completed the survey online, recruited directly from Amazon Mechanical Turk. This sample size was chosen to be identical (within each condition) to the number of participants in each condition of the previous experiment. This experiment was also preregistered.

4.1.2. Procedure

Participants were randomly assigned to either the telos changed or the insides changed condition and to one of the three transformation scenarios in a 2 (Transformation: Telos Changed, Insides Changed) x 3 (Species type: spider to bee, hummingbird to vulture, spider to hummingbird) design. A full set of vignettes is available on our OSF page. In the bee to spider transformation conditions, for example, participants were shown one version of the following:

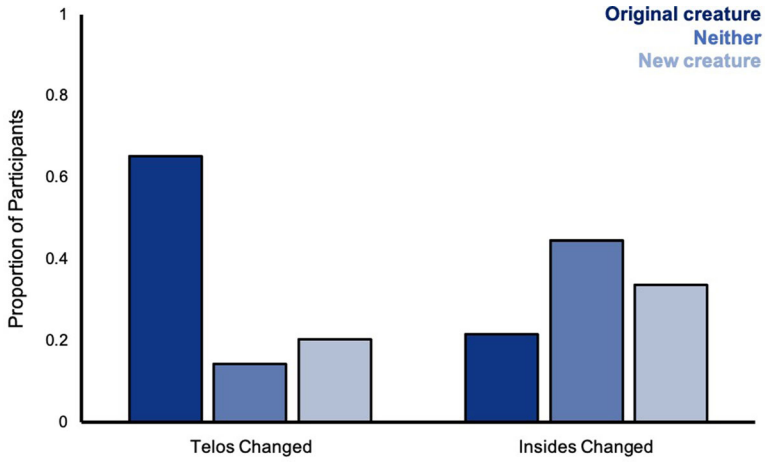

Fig. 3. Results from Experiment 3. Participants were shown a simplified version of a scenario in which a baby creature is placed in a cage of a different kind of animal. When its telos changed, they tended to think it was still the same creature as it was originally. But when its insides changed, they thought the creature had changed enough to be something different.

Suppose that the moment it hatches, a newborn bee is placed in a cage full of spiders.

> [Telos Changed] After two weeks, the creature doesn't pollinate flowers or make honey. Instead, it only spins webs to catch them and eat them.
>
> [Insides Changed] After two weeks, the creature doesn't seem to have the insides of a bee. Instead its insides seem completely different—all the organs and even the DNA itself are now much more like those of a spider than those of a bee.

Language for the telos changed condition was adapted from Rose and Nichols' language of telos change, consistent with prior experiments. Language for the insides changed condition was adapted from that of the previous experiment.

Participants were then asked "Which of the following best describes the animal?" and could choose between the original classification (i.e., "bee," "hummingbird," or "spider"), the new classification (i.e., "spider," "vulture," or "hummingbird"), and "neither." We set up this study in the simplest way possible: no images were included, all information and the question was on a single page, and no further questions were asked.

4.2. Results

Results from Experiment 3 are shown in Fig. 3. In general, participants thought that the creature had changed more when its insides had changed than when its telos had changed. Across all transformations, participants were less likely to classify the creature as its original

kind when its insides changed than when its telos changed (insides change vs. telos change: 22% vs. 65%; two-proportions z-test, $z = 10.79$, $p < .001$). This was also true in each transformation type individually (two-proportions z-tests, zs > 4.66, ps $< .001$). Participants were also more likely to classify the creature as now "neither" one category nor the other when its insides changed than when its telos changed (insides change vs. telos change: 45% vs. 14%; two-proportions z-test, $z = 8.15$, $p < .001$). This was also true in each transformation type individually (two-proportions z-tests, zs > 4.88, ps $< .001$).

Finally, participants were more likely, overall, to classify the creature as fully something different (e.g., a spider when it began as a bee) when its insides changed than when its telos changed (insides change vs. telos change: 34% vs. 20%; two-proportions z-test, $z = 3.67$, $p < .001$). While this pattern was qualitatively consistent in each transformation, it was not statistically significant within individual scenarios.

4.3. Discussion

In general, participants think that a creature is not its original kind once its insides have changed to resemble some other kind of animal. These results, along with those of Experiment 2, suggest that insides changes have at least as much bearing on essences as teleological changes. Here, in fact, changes to the animals' insides mattered *more* than teleological changes; when, for example, a bee's telos changed (without, presumably, any other changes), participants still thought it was a bee.

One possibility for why teleological changes did not impact participants' category judgments in this experiment (as opposed to prior experiments) is that we removed all language of "scientists" performing a "special experiment." Such language (present in previous experiments) may have led participants to make further inferences about underlying causes. Take this first sentence of the vignette from Experiment 3 and Rose and Nichols (2019):

> Some very talented and skilled scientists, Suzy and Andy, decide that they are going to perform a special operation on a bee.

Here, the reader is given several important pieces of information. First, they are told that Suzy and Andy are intentionally performing an operation. Second, they are told that this operation is "special" (although the meaning of this word is never explained). Third, they are told that Suzy and Andy are experts ("very talented and skilled"), meaning that whatever intention they have to perform an operation is likely not aimless. These seemingly innocuous details can meaningfully shape how a reader may interpret the information that follows. For example, Neufeld (2021a) argues that people will take a change in telos as evidence for a change in something's scientific essence. If so, this may explain why this additional information about scientists performing an experiment influences participants' intuitions about teleological changes: Such information may facilitate further inferences (i.e., about animals' insides) in the case of teleology but not in the case where participants are directly told about changes to the creature's insides.

5. General discussion

What is the best way to characterize psychological essentialism? While we replicate key findings in support of teleological essentialism (Experiments 1a–c), our results ultimately suggest that teleology is not the driving force of psychological essentialism (Experiments 2 and 3). In other words, teleological information may be one of many factors relevant to people's categorization judgments but seems not to itself constitute something's essence.

5.1. Scientific versus teleological essentialism?

A strong view of teleological essentialism would require that teleology is the primary way that people understand the essences of living things. In contrast to this view, however, the studies reported here results offer empirical support for the argument made by Neufeld (2021a): People's category judgments may be influenced by teleological information not because they are directly teleologically essentialist, but because they infer that deeper internal changes gave rise to this change in telos. In other words, while something's telos may not itself be an essence, people may use teleological information in order to *infer* something's scientific essence. This perspective can explain both why people are puzzled about cases in which only a creature's telos or only its insides changed (see Experiment 2), and why their intuitions about "insides" changes appear more stable than their intuitions about telos changes (see Experiment 3).

In line with the argument made by Neufeld (2021a), we find that the effects observed by Rose and Nichols (2019) may in fact be consistent with standard views of essentialism. This argument also has implications for scientific essentialism. Namely, even the information about animals "insides" provided in our experiments may serve as a cue to some deeper essence. Even something's DNA may not directly comprise its essence. Rather, these explicit internal features may simply be highly diagnostic for an even deeper causal essence.

The debate between teleological and scientific essentialism also raises questions about how people may categorize animals and other natural kinds even in cases not involving transformation (see also Lombrozo & Rehder, 2012). For instance, how might people characterize an animal when shown a photograph of a spider and told that the creature makes honey? How might they characterize a novel animal with the teleological or behavioral characteristics of one species but a closer DNA match with another?

Finally, regardless of the ultimate verdict on teleological essentialism, people's use of teleological information *at all* remains quite striking. In short, even if scientific essentialism ultimately best characterizes people's essentialist intuitions, it seems that something's telos may be a particularly salient cue to its internal essence. Something's telos may not be its essence *per se*, but learning about something's telos (or a change to it) may nevertheless be highly informative. We know that teleological thinking plays an important role in many different areas of cognition—influencing judgments of category membership (Lombrozo & Rehder, 2012; Matan & Carey, 2001), persistence (Rose, 2014), mereology (Rose & Schaffer, 2017), and explanation (Joo et al., 2022; Kelemen, 1999; Kelemen & Rosset, 2009; Kelemen et al., 2013). One way to understand this influence is to something's purpose as a simple,

but particularly useful, cue to many complex properties. For instance, learning that something like a microwave is for heating food can also convey that those pieces of metal make a single object, that heating food is a key characteristic of microwaves generally, and that this microwave exists because it was designed to serve that purpose. Similarly, learning something's telos could convey something about its deeper insides. A change in something's telos may be a heuristic through which we infer that its insides have changed to the point that it is now best calssified as some other kind of thing. In other words, even if people are not primarily teleologically essentialist, teleology seems to play some role in how people figure out objects' essences.

Teleological essentialism may also turn out to be the best way to describe the essences of things other than animals. Scientific essentialism is largely restricted to natural kinds (Gelman, 2003; Gelman & Heyman, 1999; Gelman & Markman, 1986, 1987; Inagaki & Hatano, 2002; Newman & Keil, 2008)—but people have also been found to ascribe essences to many other sorts of things, including artifacts (e.g., Barton & Komatsu, 1989; Puebla & Chaigneau, 2014), various social categories (e.g., Haslam & Levy, 2006; Haslam, Bastian, & Bissett, 2004; Morton, Postmes, Haslam, & Hornsey, 2009; Williams & Eberhardt, 2008), moral character (Heiphetz, 2019, 2020), and abstract concepts (e.g., Knobe, Prasada, & Newman, 2013; see also Neufeld, 2021b on differences between what constitutes an essence in these different domains). Recent work also suggests that there may be a general way of understanding essentialism across these cases, even though only some of them seem to be scientific or causal (Newman & Knobe, 2019, see also Neufeld, 2021b). A better understanding of psychological essentialism may also involve future research into boundary cases between domains. Take, for example, the characterization of a swimming pool (but not tea) as water (Malt, 1994). Is this case about a natural kind (water) or about a human-made artifact (a swimming pool)? Similarly, might people treat highly domesticated animals (e.g., livestock) in some ways both as an animal (i.e., in comparison to wild animals) and as an artifact (i.e., as something humans have engineered for their use)? A broader account of people's essentialist intuitions and the role of teleological thinking in those intuitions may involve a better understanding of these boundary cases between domains. Whether teleology plays some role in these other cases not typically explained by scientific essentialism, or in some more generalized understanding of essence, remains an open question.

5.2. Other (methodological) concerns

Our changes to Rose and Nichols' vignettes are simple. Yet, they have striking effects on people's judgments. These findings thus have methodological implications for not only the study of teleology and essentialism, but also for related bodies of work (e.g., on categories, kinds, etc.). We discuss three key takeaways.

First, participants respond differently to a continuous versus discrete response scheme, potentially resulting in findings that may be interpreted quite differently. For instance, the high rate of "neither" responses that we found with a discrete response scheme may suggest that participants do not necessarily treat the middle range of the scale as a true midpoint. With discrete options, however, it may be that participants could express clearly when they thought

the creature had changed too much to still be a bee but was still not quite a spider. Our point is not that one of these measures is more appropriate in this case or in general, but simply to bring attention to this subtle but potentially critical design detail.

Second, extraneous information can affect how participants understand the scenario—particularly when this information concerns intentional agents. We suggest that details (such as skilled scientists performing an experiment) do more than simply add context to a vignette. Instead, the presence of such information may meaningfully influence people's interpretation of the manipulation. In some of these experiments, for instance, it may have been that people inferred that the scientists made some internal change to the organism (even though internal changes were never explicitly mentioned in the vignette).

Finally, it may be important to reconsider what it means for something like a bee or a spider to have a telos in the first place. Following Rose and Nichols (2019, 2020), we treat characteristic behaviors of a bee or a spider as the telos of that kind. We take for granted (as Rose and Nichols did) that features like pollinating flowers and making honey or spinning webs and catching insects best characterize these animals' tele. Yet, one possibility is that animals may not truly have tele at all (consider, for instance, what may be the telos of a hippopotamus, or an octopus, or a beetle). Instead, it may be that the manipulations employed here simply reflect the fact that pollinating flowers and making honey are *prototypical features* of bees (whether or not those prototypical features constitute tele). If true, it would mean that even if we did observe strong effects in this paradigm, strong claims about teleological essentialism may still be unwarranted.

Another possibility is that even if animals *do* have tele, multiple criteria exist in principle for defining this telos. For instance, is an animal's telos the function it may serve for humans (similar to the sense in which the purpose of a hammer is to drive in nails)? Or is it something it does for a larger "organism," such as the ecosystem to which it belongs (similar to the sense in which the purpose of a giraffe's neck is to reach tall trees; see also ojalehto et al., 2013)? Or is it its realizable potential (similar to the sense in which an acorn's purpose may be to become an oak tree)? These different approaches may concretely affect how one studies something's telos: Depending on one's view, it may be that the purpose of a spider is not to spin webs or even to catch insects, but specifically to manage insect populations. Do changes to all of these *kinds* of tele result in similar influences on category membership? We would think not. Yet, the view put forth by Rose and Nichols does not specify when we should or should not expect such effects. To further the debate on teleolgical vs. scientific essentialism, it is vital that we first more clearly define what it means to have a telos in the first place.

5.3. Conclusion

Are people teleologically essentialist? The experiments reported here suggest that teleology may play some part in how people categorize and conceptualize natural kinds—but that information about animals' insides may still play the largest role. Thus, it may be that people may reason about objects' purposes in addition to reasoning directly about their insides, at least in part because identifying something's purpose may shed light on its internal parts.

While the findings of Rose and Nichols (2019, 2020) reveal yet another way that teleological thinking is central to human cognition, they appear to stop short of radically recharacterizing psychological essentialism.

Acknowledgments

For his unwavering support and generosity these past 5 years, we thank Frank Keil. For extensive conversations about all things teleology, we thank Josh Knobe. For a thought-provoking discussions and for inspiring this project, we thank Elonore Neufeld and David Rose. For helpful conversation and feedback, we thank Tobias Gerstenberg and all members of the Stanford Causality in Cognition Lab. And for many other helpful comments, we thank all members of the Yale Cognition and Development Lab.

Funding

This project was supported by a National Science Foundation Graduate Research Fellowship awarded to S. R. Yousif.

Open Research Badges

The experiments reported in this paper were formally pre-registered on the Open Science Framework at https://osf.io/x5gf6/?view_only=a43050f1fa0f4802b0025b3a59beb56c. In the results, we explicitly mark which analyses are confirmatory versus exploratory following this plan, and all derivations from the plan. All materials for the work presented in this paper (data, analysis scripts, and protocols) are available at https://osf.io/x5gf6/?view_only=a43050f1fa0f4802b0025b3a59beb56c.

References

Barton, M. E., & Komatsu, L. K. (1989). Defining features of natural kinds and artifacts. *Journal of Psycholinguistic Research*, *18*(5), 433–447.

Gelman, S. A. (2003). *The essential child*. Oxford University Press.

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, *89*, 404–409.

Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science*, *10*, 489–493.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*, 183–209.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, , *58*(6), 1532–1541.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, *383*, 213–244.

Hampton, J. A., Estes, Z., & Simmons, S. (2007). Metamorphosis: Essence, appearance, and behavior in the categorization of natural kinds. *Memory and Cognition*, *357*, 1785–1800.

Haslam, N., Bastian, B., & Bissett, M. (2004). Essentialist beliefs about personality and their implications. *Personality and Social Psychology Bulletin*, *3012*, 1661–1673.

Haslam, N., & Levy, S. (2006). Essentialist beliefs about homosexuality: Structure and implications for prejudice. *Personality and Social Psychology Bulletin*, *32*, 471–485.

Heiphetz, L. (2019). Moral essentialism and generosity among children and adults. *Journal of Experimental Psychology: General*, *14812*, 2077–2090.

Heiphetz, L. (2020). The development and consequences of moral essentialism. *Advances in Child Development and Behavior*, *59*, 165–194.

Inagaki, K., & Hatano, G. (2002). *Young children's naive thinking about the biological world*. Psychology Press.

Joo, S., Yousif, S. R., & Kell, F. C. (2022). Understanding "Why:" How Implicit Questions Shape Explanation Preferences. *Cognitive Science*, *46*(2), e13091.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. MIT Press.

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

Kelemen, D., & Carey, S. (2007). The essence of artifacts: Developing the design stance. In E., Margolis & S., Laurence (Eds.), *Creations of the mind: Theories of artifacts and their representation* (pp. 212–230). Oxford University Press.

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.

Knobe, J., Prasada, S., & Newman, G. (2013). Dual character concepts and the normative dimension of conceptual representation. *Cognition*, *127*, 242–257.

Lombrozo, T., & Rehder, B. (2012). Functions in biological kind classification. *Cognitive Psychology*, *654*, 457–485.

Malt, B. C. (1994). Water is not H_2O. *Cognitive Psychology*, *27*, 41–70.

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

Morton, T. A., Postmes, T., Haslam, S. A., & Hornsey, M. J. (2009). Theorizing gender in the face of social change: Is there anything essential about essentialism? *Journal of Personality and Social Psychology*, *96*, 653–664.

Neufeld, E. (2021a). Against teleological essentialism. *Cognitive Science*, *454*, e12961.

Neufeld, E. (2022). Psychological essentialism and the structure of concepts. *Philosophy Compass*, *96*(5), e12823.

Newman, G. E., & Keil, F. C. (2008). 'Where's the essence?': Developmental shifts in children's beliefs about the nature of essential features. *Child Development*, *79*, 1344–1356.

Newman, G. E., & Knobe, J. (2019). The essence of essentialism. *Mind & Language*, *345*, 585–605.

ojalehto, B., Waxman, S. R., & Douglas, L. M. (2013). Teleological reasoning about nature: Intentional design or relational perspectives? *Trends in Cognitive Sciences*, *17*, 166–171.

Puebla, G., & Chaigneau, S. E. (2014). Inference and coherence in causal-based artifact categorization. *Cognition*, *1301*, 50–65.

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General*, *1303*, 323.

Rose, D. (2014). Persistence through function preservation. *Synthese*, *192*, 97–146.

Rose, D., & Nichols, S. (2019). Teleological essentialism. *Cognitive Science*, *434*, e12725.

Rose, D., & Nichols, S. (2020). Teleological essentialism: Generalized. *Cognitive Science*, *443*, e12818.

Rose, D., & Schaffer, J. (2017). Folk mereology is teleological. *Noûs*, *51*, 238–270.

Rose, D., Schaffer, J., & Tobia, K. (2018). Folk teleology drives persistence judgments. *Synthese*, *19712*, 5491–5509.

Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology*, *94*, 1033–1047.