UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Teleological essentialism across development

Permalink
https://escholarship.org/uc/item/0z45p9zs

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Rose, David
Jaramillo, Sara
Nichols, Shaun
et al.

Publication Date
2022

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Teleological essentialism across development

David Rose, Sara Jaramillo, Shaun Nichols, Zachary Horne

Abstract

Do young children have a teleological conception of the essence of natural kinds? We tested this by examining how the preservation or alteration of an animal's purpose affected children's persistence judgments (N = 40, ages 4 - 12, M_{Age} = 7.04, 61% female). We found that even when surface-level features of an animal (e.g., a bee) were preserved, if the entity's purpose changed (e.g., the bee now spins webs), children were more likely to categorize the entity as a member of a different natural kind (e.g., a spider) and these effects were similar in magnitude to altering the surface-features of a natural kind. Our results suggest that we might view teleological properties as partially constitutive of the essence of natural kinds.

Keywords: teleology; essentialism; development; concepts

Introduction

Many categories cannot be characterized in terms of observable features alone (Gelman, 2003). Everyday categories—bumble bees, gold, Americans—are instead represented as having an essence which makes them what they are (Gelman, 2004, Hussak & Cimpian, 2019). *Essentialized* categories are those that people think of as having a "true nature," which people presume to be both constant across category members and responsible for the similarities of those members. How are these essences represented? A leading account is they are "placeholders": People believe "that there is some causal essence that holds a category together" even if they do not know what that essence is (Gelman, 2004). Placeholders might be elaborated with content that is, for example, scientific or social in nature, but some recent research suggests it is at least sometimes *teleological* – people may have an Aristotelian view of natural kinds where a category's essence is constituted by its teleology (Rose & Nichols, 2020; Haward et al., 2018). For example, people think an individual is no longer a member of a natural kind (e.g., a bee) when its purpose (e.g., making honey) is changed, even when its parts are unchanged (Rose & Nichols, 2019). These observations cohere with findings documenting the relationship between casual reasoning and categorization (Rehder & Kim, 2006) and expand on developmental research on teleological explanation, which indicates that children explain a wide range of phenomena in terms of the perceived purpose of entities (Kelemen, 1999). For instance, children think that, "Clouds are for raining", "Birds are for flying", and Tigers are "for eating and being looked at in the zoo"; things exist to satisfy their purpose (Kelemen, 1999). Together, this research raises the possibility that even for children a category's teleology may be constitutive of its essence.

Keil (1989) examined how kindergarteners (M_{Age} = 5.8), second (M_{Age} = 7.6), and fourth graders (M_{Age} = 9.9) categorized a raccoon that underwent a surgery that made it look like a skunk, finding kindergarteners thought that the animal was now a skunk, but fourth graders thought the animal was still a raccoon. Similar observations initially led researchers to hypothesize that young children's persistence judgments are primarily, if not entirely, determined by observable features of a category (Flavell, 1986). Contrary to this early work, several researchers have since argued that essentialism (Gelman, 2004) or essentialist-like thinking (Cimpian & Salomon, 2014; Cimpian & Steinberg, 2014) emerges earlier in development than initially hypothesized. Still, it's an open question whether young children attend to the teleology of a natural kind in transformation tasks. Do children behave as if a category's teleology is a key feature of its essence? We propose children operate with an Aristotelian view of category essences – the content of an essence is teleological. Thus, we expect that changing an entity's purpose will have an impact on children's beliefs about the persistence of natural kinds. While we expect changes in surface-level features to exert a strong influence on persistence judgments as well (Baldwin, 1992),

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.* ©2022 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

we predict judgments about teleological essences will nonetheless impact persistence judgments even when surface-level feature changes (or preservation) were inconsistent with the preservation or change of a kind's teleological essence. We tested this in a repeated-measure design with 40 children between the ages of 4 and 12.

Experiment

We tested our hypothesis by examining children's persistence judgments in a transformation task involving biological kinds (Keil, 1989), examining how changing (or preserving) purpose, surface features, or both impacted children's persistence judgments about transformation cases. The experimenter explained to children that they would hear some stories about two doctors who like to "work on" (Keil, 1989) animals to make them look different. For example, children were shown a black-and-white drawing of a bumble bee and told that the doctors were going to perform a special operation on it where its "wings and antennae were removed" and "legs lengthened," after which they were shown a picture of a spider. Children were asked to categorize the insect as a spider or bee before and after the operation, and were also asked about the *purpose* of bees and spiders (e.g., whether they are for making honey or for spinning webs and catching insects).

Although transformation tasks are more artificial compared to real instances of transformation, such as cases of metamorphosis (Menendez et al., 2022), using a transformation task allowed us to control and independently manipulate two key factors (surface features and perceived purpose) we expected to influence children's judgments about persistence. In contrast, real-life cases of metamorphosis always involve a surface change, making it difficult to resolve how each factor affects participants' persistence judgments. Transformation tasks explicitly pit the preservation or alteration of surface features against a removal of the purpose of an entity, providing a strong test of our hypothesis. If by using a classic, central test of essentialist thinking—transformation tasks—we find evidence that teleological considerations play a role in categorization judgments then that would suggest that children essentialize categories in terms of teleology.

Method

Participants We recruited 40 children between the ages of 4 and 12 (61% girls, M_{Age} = 7.04 years old) from the Children's Museum of Phoenix to participate in a short study. The museum requires that anyone under 18 years old be allowed to participate; three additional subjects older than our target age range of 4 – 12 years old participated, but their data were not analyzed. 17 additional children started the study but did not complete it.

Intended Sample Size and Analytical Robustness We intended to collect data from 60 participants after excluding participants who were either outside of the target age range or who could not complete the study. However, the Covid-19 pandemic prevented further in-person data collection starting in March, 2020. In light of this unforeseen problem with data collection, we took several measures to assess the robustness of our findings.

First, we report analyses of all participants who finished the study (N = 40), but additional analyses from only those participants who responded to every question (N = 35), or participants who completed at least one trial within our target age range (N = 54) are also reported. We focus our analyses on the sample of participants who completed at least one trial (N = 40), but supplementary analyses from only those participants who completed the entire study (N = 35) can be found on the Open Science Framework (OSF) (https://osf.io/dqjry/$view\_only = 87afd7377b174565ab38776fa6d69966$).

Second, throughout, we fit Bayesian rather than frequentist regression models, allowing us to assess the sensitivity of our analyses to more skeptical priors. This allows us to examine how strongly skeptical priors (those biased against evidence for our hypothesis) impact our inferences. We explain this procedure in the analytic strategy section below.

Materials and Procedure We created transformation cases using four vertebrae pairs and four invertebrate pairs that are familiar to children (see Wordbank database), and commonly appear in books for children. These pairs had distinct traits

so there were no co-occurring physical features or purposes. For example, an invertebrate with wings would not be paired with another winged invertebrate. Likewise, a vertebrae whose purpose is to eat bananas and swing from trees (i.e., a monkey) would not be paired with another vertebrae that fulfills the same or similar function. Each pair was presented with a stylized black-and-white image. These materials can be found on the OSF.

Our experimental design guaranteed vertebrae and invertebrate pairs appeared in every condition. For each pair, there were four possible versions, corresponding to each of the four conditions. Once that pair was presented, it would not appear again in another condition. Trials and conditions were counterbalanced and randomized.

Children completed eight trials (four conditions $\times$ two trials per condition) in a fully within-subjects design. Examples of each condition and the stimuli are located on OSF. The experimenter explained to children that they would hear some stories about two doctors who like to work on animals to make them look different. In each trial, children were shown a picture of an vertebrae or invertebrate and the experimenter described the transformation that occurred. Children were then asked to identify what the animal was before and after the special operation, as well as the animal's purpose after the special operation.

It's possible that children do not think animals have purposes in the first place (but see Kelemen, 1999), so we asked a series of comprehension questions to confirm children endorsed these beliefs. We describe these questions below.

Comprehension We included comprehension questions for each kind member to verify: First, children were familiar with the animals in each trial. Second, children correctly identified what the animal does. Third, and most importantly, children agreed what the animal is "for". Children answered 88% of comprehension questions correctly, and when selecting between two possible purposes for an animal, children selected the prototypical purpose 96% of the time.

While it's unlikely children universally think that, for instance, "the purpose is of a monkey is to eat bananas", performance on these comprehension questions provides prima facie evidence that children view the purposes we chose as central to their conceptions of the animals used in the study. Furthermore, our results are directly predicted by and consistent with prior work demonstrating how pervasive teleological thinking is in children and adults (Kelemen, 1999; Kelemen et al., 2013).

Analytic Strategy To examine the effects of purpose, surface features, and their interaction on children's persistence judgments, we performed Bayesian logistic mixed-effects modeling using the R package brms (Buerkner, 2017). Bayesian statistics provide a principled way to regularize parameter estimates (namely, by changing the prior distribution). Throughout, we set regularizing priors on the parameters in our models. For all exploratory modeling, we performed an approximation of leave-one-out cross-validation and then assessed the out-of-sample performance of a set of models based on the change in their expected log predictive density, Δelpd (Vehtari et al., 2017).

To formally examine the magnitude of the effect of purpose, we performed Bayesian model stacking (Yao et al., 2018) by fitting a series of Bayesian models varying the prior distributions over the Purpose parameter – that is, the primary coefficient of interest. To aid the reader in interpreting our data and the statistical power of our design, we also conducted a frequentist sensitivity analysis, which indicated that with our sample size we can detect a Cohen's d as small as d =.45 with 80% power for a paired t-test (for example, to test the difference between the reference group and the condition in which only the purpose of an entity is changed). This effect size is approximately a log odds ratio of .82 using the equation log odds ratio $= \frac{d\pi}{\sqrt{3}}$ (Sanchez, 2003).

Results

We predicted that when holding surface features constant, but changing an animal's purpose, we would find children's judgments about the persistence of the animal would be impacted. Likewise, we predicted that when holding an animal's pur-

pose fixed but radically changing the surface features, children's judgments about the persistence of the animal would be impacted. We did not know whether these factors would interact. We fit a Bayesian mixed-effects logistic model regressing persistence judgments on two factors Surface (Reference group = Surface features preserved); and Purpose (Reference group = Purpose preserved) and their interaction, and included random slopes of these predictors, and cross-classified random intercepts for Subject and Animal Pair. Raw data, code, and model outputs are available on the OSF.

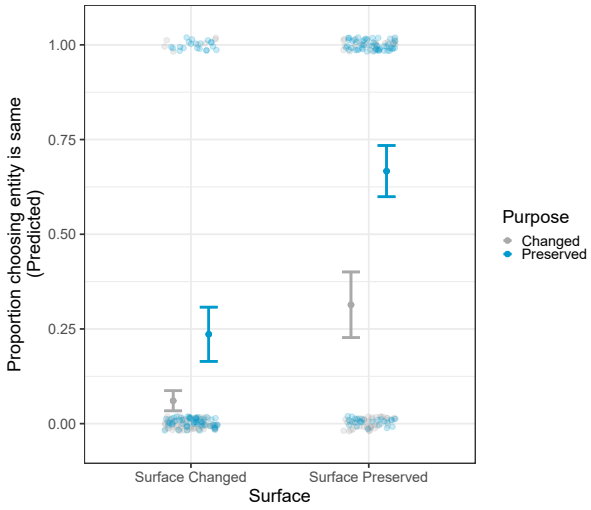

Figure 1: Conditional effects plot of children's responses across conditions. Error bars are ±1 standard error of the mean, and raw data is jittered for readability. A Bayesian logistic mixed-effects model (Reference group = Purpose Preserved, Surface Preserved) revealed that manipulating the purpose of an animal strongly predicted children's persistence judgments, $b_{\text{Purpose}} = -1.49$, 95% CI $(-2.31, -0.75)$, and the magnitude of this effect was similar to manipulating the surface features of the entity, $b_{\text{Surface}} = -1.90$, 95% CI $(-2.74, -1.09)$. These factors did not interact $b_{\text{Purpose} \times \text{Surface}} = -0.07$, 95% CI $(-0.53, 0.39)$.

We found that changing (or preserving) an animal's purpose affected children's judgments even when the surface features of the animal were entirely preserved or radically altered in the operation, Odds-Ratio = .22, 95% Credible Interval (.10, .48) (Fig. 1). Furthermore, the magnitude of the effect of removing the purpose of an animal was similar to that of removing its surface features Odds-Ratio = .15 , 95% CI (.06, .34), and these effects were not credibly different from each other, $b_{\text{Difference}} = -0.40$, 95% CI $(-1.60, 0.78)$. Across several model specifications and participant-inclusion criteria, we found children's judgments were credibly affected by the Purpose factor: $b_{\text{Purpose (Main model)}} = -1.49$, 95% CI $(-2.31, -0.75)$; $b_{\text{Purpose (Completed all trials)}} = -1.58$, 95% CI $(-2.43, -0.80)$; $b_{\text{Purpose (All participants)}} = -1.43$, 95% CI $(-2.20, -0.72)$; $b_{\text{Purpose (Model stacking)}} = -1.23$, 95% CI $(-2.14, -0.46)$

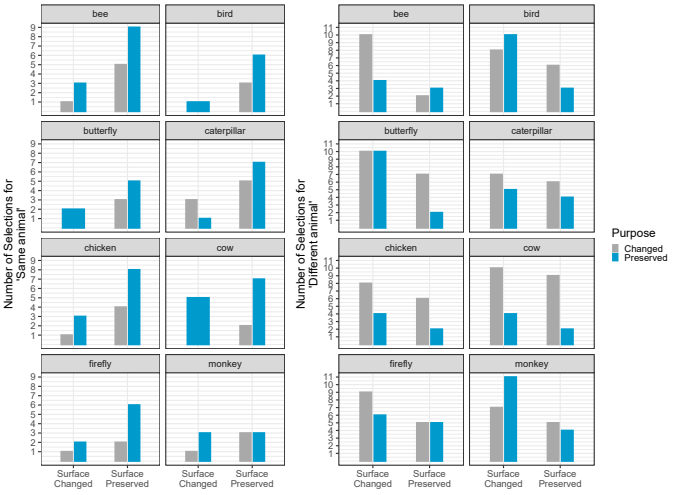

Figure 2: The frequency of children's responses across conditions for each vertebrae or invertebrate pair. In each trial, children decided whether an animal was the same animal (e.g., still a bee) or different animal (e.g., now a spider) across conditions. Note: the plot labels for each trial only list one animal for simplicity. The matching animal for each pair are in Tables S10-S17 of the Supplement.

Although we expected to see developmental shifts in the effect of purpose on children's re-

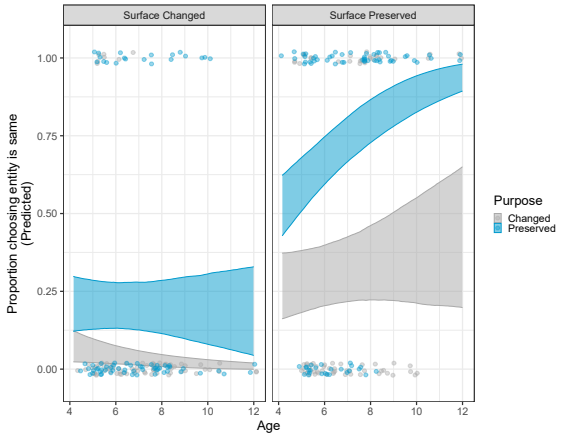

Figure 3: Conditional effects plot of children's responses by age, purpose, surface, and their interactions. Shaded regions represent ± 1 standard error of the mean. Raw data on axes are jittered for readability. Regardless of age, children's judgments were affected by manipulating the purpose and surface features of an animal. Age only predicted persistence judgments when both purpose and surface features of the entity were preserved across the operation (i.e., the reference group), $b_{\text{Age}} = 0.33$, 95% CI (0.12, 0.55).

sponses, we did not have specific hypotheses about the shape of these changes. Thus, we performed exploratory modeling to determine how different models including Age predicted observations our statistical model was not trained on. We found that interacting Age with Purpose trivially improved model fit relative to an model including Age, $\Delta elpd$ (Change in expected log predictive density) = 1.7, $SE = 1.1$, or without including Age at all, $\Delta elpd$ = 1.8, $SE = 1.6$ (Fig. 2). Thus, the magnitude of these differences suggests that a simpler, model including only Purpose and Surface best accounts for children's responses whilst minimizing unnecessary model complexity. Altogether, our results suggest purpose and surface features credibly impact children's judgments and the effects of these factors are similar across development.

Discussion

The idea that teleology partly constitutes categories can be traced back to Aristotle. For Aristotle, the essence is what defines the category. And, for at least many categories, the essence is given by the "telos" or its purpose. Adults appear to operate with an Aristotelian view of some category essences (Rose & Nichols, 2019; Rose & Nichols, 2020). Our findings here suggest that children might as well.

Using the same kinds of tests that are typically used to provide evidence of essentialist thinking—radical transformation tasks—we found that teleological considerations played an important role in persistence judgments. While some people have challenged the probative value of transformations tasks, among other things, with respect to teleological essentialism (see e.g., Neufield, 2021), the transformation test has been one of the main lines of evidence in favor of essentialism. To abandon that test is to undercut the case for essentialism more broadly. We take a more conservative approach here in affirming the value of transformation tasks as evidence for essentialism. But we do not mount an argument for that here and instead emphasize that using a standard test of essentialist thinking, we find that teleological considerations play a prominent role in children's categorization judgments. In particular, children thought the original animal changed category membership when its purpose was altered. Perhaps even more surprising is that even though changes in surface features are known to exert a dramatic impact on children's categorization judgements (e.g., Baldwin, 1992; Sloutsky, 2010), as we also observed, the purpose of an animal nonetheless exerted an influence of similar magnitude. Teleological considerations appear to play an important role in children's judgments of persistence across radical transformation. And they do so not just in the domain of artifacts (Kemler et al., 2000). Instead, teleological considerations seem to play a prominent role in categorization even when children consider the kinds of entities that have played a central role in probing essentialist thinking: natural kinds.

The Aristotelian tendency to essentialize cate-

gories in terms of their teleology emerged even for young children, with clear evidence that the judgments of children 5 years and older were impacted by manipulating the purpose factor. It seems, then, that the seeds of teleological essentialism are in place at a very early age. There are some signs that essentialist tendencies are in place by around roughly four years of age (see Sutherland & Cimpian, 2019; Rhodes & Mandalaywala, 2017; Gelman, 2003. But in earlier studies using radical transformation tasks, children under seven years old made categorization judgments based on *appearances* (Keil, 1989). By manipulating teleology, however, we found that substantially younger children exhibited essentialist tendencies on transformation tasks.

Our findings suggest that when children begin showing signs of essentializing in the context of radical transformation tasks, they already represent aspects of the content of essences in terms of teleology. However, this does not mean that essences are *never* represented in terms of placeholders. Indeed, any view on the content of essences is consistent with there being placeholders. But our findings suggest that when placeholders are elaborated, their content may be frequently elaborated in terms of their teleology. Although our findings suggest that children may indeed be operating with an Aristotelian conception of category essences, there are a range of further tests of essentialist thinking which could be used to examine our hypothesis. Our primary aim was to examine the teleology hypothesis by using a conservative task. But other tasks commonly used in research on essentialism could be used to examine our hypothesis in future research. For instance, switched-at-birth tasks (e.g., where an animal is raised in a community of different animals) provide an important test of essentialist thinking. Although our findings indicate that teleological considerations impact persistence judgments in radical transformation tasks, future research could investigate whether teleological considerations impact persistence judgments across a range of tests of essentialist thinking.

Most work on essentialism has not focused on the *content* of essences (but see Prasada, 2017), but understanding what and how category essences are given content is important for understanding essentialist representation. And doing so has important practical consequences. To consider just one example, if we do indeed represent categories in terms of teleology, then that suggests that we may have underestimated how difficult it is to get people to abandon that teleological way of thinking about the world (Kelemen et al., 2013). This is relevant for the prospect of reducing or eliminating essentialist thinking and also relevant to science education. One of the single advances of the scientific revolution was to move away from the teleological picture promoted by Aristotle and his followers. But that Aristotelian approach seems to be a deep-seated aspect of our worldview, not just about evolution, but about the categories themselves. Science education that exploits and builds on these tendencies might prove more effective than efforts aimed suppressing or bypassing them (Kelemen et al., 2013; Barnes et al., 2017).

References

Baldwin, D. A. (1992). Clarifying the role of shape in children's taxonomic assumption. *J Exp Child Psychol.*, *54*, 392–416.

Barnes, M. E., Evans, E. M., Hazel, A., Brownell, S. E., & Nesse, R. M. (2017). Teleological reasoning, not acceptance of evolution, impacts students' ability to learn natural selection. *Evolution: Education and Outreach*, *10*(1), 1–12.

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1), 1–28.

Cimpian, A., & Salomon, E. (2014). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences*, *37*(5), 461.

Cimpian, A., & Steinberg, O. D. (2014). The inherence heuristic across development: Systematic differences between children's and adults' explanations for everyday facts. *Cognitive Psychology*, *75*, 130–154.

Flavell, J. H. (1986). The development of children's knowledge about the appearance–reality distinc-

tion. *American Psychologist*, *41*(4), 418–425.

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

Gelman, S. (2004). Psychological essentialism in children. *Trends in Cognitive Science*, *8*(9), 404–409.

Haward, P., Wagner, L., Carey, S., & Prasada, S. (2018). The development of principled connections and kind representations. *Cognition*, *176*, 255–268.

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

Kelemen, D. (1999). The scope of teleological thinking in preschool children. *Cognition*, *70*(3), 241–272.

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.

Kemler Nelson, D. G., Russell, R., Duke, N., & Jones, K. (2000). Two-year-olds will name artifacts by their functions. *Child Development*, *71*(5), 1271–1288.

Menendez, D., Rosengren, K. S., & Alibali, M. W. (2022). Detailed bugs or bugging details? the influence of perceptual richness across elementary school years. *Journal of Experimental Child Psychology*, *213*, 105269.

Prasada, S. (2017). The scope of formal explanation. *Psychonomic Bulletin & Review*, *24*(5), 1478–1487.

Rehder, B., & Kim, S. (2006). How causal knowledge affects classification: A generative theory of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 659–683.

Rhodes, M., & Mandalaywala, T. M. (2017). The development and developmental consequences of social essentialism. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(4), e1437.

Rose, D., & Nichols, S. (2019). Teleological essentialism. *Cognitive Science*, *43*(4).

Rose, D., & Nichols, S. (2020). Teleological essentialism: Generalized. *Cognitive Science*, *44*(3), e12818.

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, *8*(4), 448.

Sloutsky, V. M. (2010). From perceptual categories to concepts: What develops? *Cognitive Science*, *34*(7), 1244–1286.

Sutherland, S. L., & Cimpian, A. (2019). Developmental evidence for a link between the inherence bias in explanation and psychological essentialism. *Journal of Experimental Child Psychology*, *177*, 265–281.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432.