Teleological essentialism in development

David Rose
Stanford University

Jadess Lowery
Stanford University

Siying Zhang
University of Washington

Ellen Markman
Stanford University

Abstract

Overcoming appearances in categorizing is an important intellectual achievement. But children can't help but rely on them, as classic shape similarity (Landau, Smith, & Jones, 1988) and perceptual transformation tasks (Keil, 1989) show. Scientific essences—placeholders for what gives something its identity that science fills out—might help but suggest a protracted development since scientific elaboration is challenging and rare (Gelman, 2003). We instead propose that children can more readily treat purposes as essences. Across three experiments, with 570 US based children, we find support for this. In contrast to being given only a label or inductively irrelevant properties, children categorize together dissimilar looking, familiar objects based on familiar purposes at the basic (Experiment 1) and superordinate level (Experiment 2). Children also used an object's purpose to judge its identity in the face of radical perceptual transformations (Experiment 3). That children can elaborate placeholders with purposes for biological kinds and artifacts suggests a reorientation of essentialism.

Keywords: categorization; teleology; essentialism; development; perceptual features.

Introduction

Linnaeus initially missed it. Brisson came close to capturing it. But Linnaeus returned to claim it: Whales aren't fish. Common sense makes a compelling case otherwise. Whales certainly look like most of the animals that we classify as fish. And for most of human history, the striking perceptual similarity of whales to fish has led common sense to prevail on this matter.

It's an important intellectual achievement that humans can categorize things in ways that go beyond similarities in their appearance—grouping things that look different into the same category and things that look alike into different categories—with important consequences. When perpetually dissimilar things are united into a single category that includes perceptually typical and atypical members, generalizations can change, an understanding of their properties can change, and advances in knowledge can be achieved. Such is the case in moving whales from the fish to mammal category.

Going beyond perceptual similarity in categorization isn't merely something that is relegated to the annals of scientific achievement. It's crucial even in ordinary classification. Consider the large variety of superordinate categories, from fruits to furniture, that people form. Yet members of these categories can have wildly different perceptual properties. It would be nearly impossible to form these superordinate categories if all we relied on was what things looked like. Even members of basic level categories, such as lamps, can look very different.

Yet perceptual features are immediately available for both adults and children. And they play an especially important role in early word learning and categorization (e.g., Baldwin, 1989; Booth, Waxman, & Huang, 2005; Diesendruck & Bloom, 2003; Gershkoff-Stowe & Smith, 2004; Graham & Diesendruck, 2010; Jara-Ettinger, Levy, Sakel, Huanca, & Gibson, 2022; Jones & Smith, 1993; Landau et al., 1988; Macario, 1991; Peretz-Lange & Kibbe, 2024; Samuelson & Smith, 2005; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). We don't deny that perceptual features are important in word learning, categorization, and identifying objects. Most of the time children and adults determine what something is by consulting its perceptual features. But solely relying on appearances poses serious limitations. How might we overcome them? We sometimes categorize by considering what makes something the kind of thing that it is, its essence (e.g., Medin & Ortony, 1989)—a thing's true nature which cannot be directly observed and gives an object its identity (Gelman & Ware, 2012).

Scientific essentialism

The view that we sometimes categorize things based on their essence grows out of work by Kripke (1980) and (Putnam, 1975) who maintained that, for natural kinds, science uncovers properties that make things what they are—what makes something a lemon isn't that it is yellow, tart, with a certain kind of peel, but that its has the genetic structure of lemons. This insight has been codified in a longstanding, influential view of essentialism in psychology—scientific essentialism (e.g., Gelman, 2003; Keil, 1989). Scientific essentialism acknowledges that laypeople rarely possess knowledge of the underlying features of things that science has discovered as being essential. Instead, on this view, scientific essences are typically represented in terms of placeholders that are restricted to natural and social kinds

(Atran, 1998; Gelman, 2003, 2013; Keil, 1994, 1989). And while there is a great deal of research on various aspects of scientific essentialism (see e.g., Gelman, 2004, for a review)—including that preschoolers can reason about the insides of things (e.g., Gelman & Wellman, 1991) and think that animals raised by different animals will grow up to be like their birth parents (e.g., Gelman & Wellman, 1991; Waxman, Medin, & Ross, 2007)—overcoming a reliance on perceptual features and categorizing instead based on scientific essences can be incredibly challenging. This is well documented in work showing that children heavily rely on shape similarity in word learning and categorization (e.g., Landau et al., 1988), and that children believe that an entity's identity changes after undergoing radical perceptual transformations (Keil, 1989). We turn to both cases next.

Shape similarity and radical perceptual transformations

Here we briefly discuss classic work documenting children's reliance on shape and perceptual appearances in transformation cases, beginning with work on shape similarity.

Shape similarity. Shape similarity plays a powerful role in early word learning and categorization. If, for instance, children are taught a new word in "dinosaur language", shown an apple and told it's a "dax", a shape match, such as a tennis ball, and a category match that is perceptually dissimilar, like a banana, when asked to find another dax, preschoolers select the shape match (e.g., Baldwin, 1992; Gentner & Namy, 1999; Imai, Gentner, & Uchida, 1994; Landau et al., 1988).

Labels can help children overcome a reliance on perceptual similarity (e.g., Cimpian & Markman, 2005; Gelman & Markman, 1986; Markman, 1989; Waxman, 1990) for basic-level categories, and help children generalize properties of perceptually dissimilar basic-level category members (e.g., Gelman & Markman, 1986, 1987). Labels are less effective at the superordinate level (e.g., Cimpian & Markman, 2005), and children struggle to generalize properties even when given the kind of scientific information about insides that scientific essentialism suggests is relevant to construing categories in terms of essences (Gelman & O'Reilly, 1988).

Radical perceptual transformations. If children are shown a raccoon that is made to look like a skunk, they show a protracted developmental pattern before they come to think that undergoing perceptual transformation doesn't actually make it a skunk (Keil, 1989). In fact, it isn't until children are in the 4^{th} grade—around 9 years of age—that they overcome a reliance on perceptual features in radical transformation tasks. Presumably children need to develop the rudimentary biological theory on which being a member of a kind isn't just a matter of looking like typical members of the kind, and come to view it as more important to have the appropriate hidden internal properties, its essence.

Our proposal

Overcoming perceptual similarity, especially shape, at both the basic and superordinate level in early word learning and categorization is difficult for young children. And determining whether something's identity persists across radical perceptual changes is even more challenging. Part of this is because perceptual information is readily available and highly salient, while placeholders for hidden, internal essences don't afford any particular kind of content to determine what the essence is such that it can be leveraged in determining category membership and identity. Scientific information can potentially provide this content, but it is rare that even adults come to elaborate placeholders in terms of scientific information (Gelman, 2003) and children struggle to make use of it (Gelman & O'Reilly, 1988).

We propose a different view of essentialism: teleological essentialism. We turn to this next.

Teleological essentialism. Teleological essentialism is the view that people tacitly regard essences in terms of a telos (Rose & Nichols, 2019, 2020). In other words, what something is for—its purpose—is treated as its essence. On this view, placeholders can be elaborated in terms of teleological information. Though it is unsurprising that artifacts are construed in terms of what they are for, teleological essentialism goes one step further, suggesting that even biological kinds can be essentialized in terms of teleology.

Rose and Nichols (2019) find that if a bee is made to look like a spider but preserves the purpose of bees, making honey, adults say it is still a bee. Even if it has all of its insides replaced with the insides of a spider. So long as it makes honey, it's a bee. And if baby bees are placed in a cage full of spiders, if they grow up and end up spinning webs, they are spiders. Moreover, if a bee is transformed to look like a spider, if it still makes honey, adults even predict that its offspring will be bees. Across a range of tests of essentialist thinking, from radical transformation, to switched-at-birth and offspring tasks, what something is for plays a role in determining what it is (Rose & Nichols, 2020). Since teleological considerations play a role in determining what something is when using standard tests of essentialist thinking, this suggests that purposes can serve as essences.

Adults elaborate essences in terms of purposes. Children might too. More specifically, what something is for might (a) be readily accessible to children (e.g., Butler & Markman, 2014; Deák, Ray, & Pick, 2002; Diesendruck & Bloom, 2003; Kelemen, 1999; Kemler Nelson, Russell, Duke, & Jones, 2000; Zhu, Goddu, Zhu, & Gopnik, 2024), (b) serve to elaborate placeholders, and (c) aid children in going beyond perceptual appearances when categorizing.

Our question

Our central question is whether in classic cases documenting the powerful role of perceptual appearances in categorizing—those involving striking shape similarities (e.g., Landau et al., 1988) and those involving complete and radical perceptual transformations (e.g., Keil, 1989)—teleological considerations lead children to overcome their reliance on appearances. The first two experiments focus on categorization tasks and the third study focuses on transformation tasks.

Experiments

For all results reported in this paper, we fit Bayesian logistic mixed effects models. We will refer to a statistical result of interest as "credible" when the 95% credible interval excludes 0, except when reporting odds ratios, which we will interpret as credible when the credible interval excludes 1.

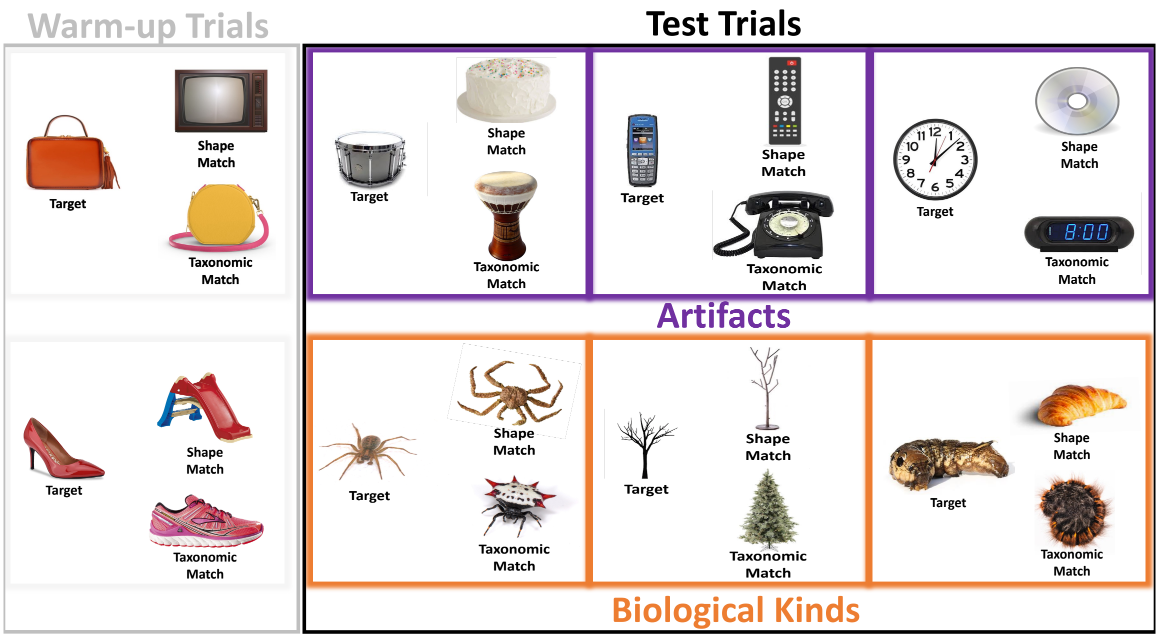

Figure 1. **Experiment 1 items**: Items for the warm-up (left box in gray) and test trials (right box in black). The test trials included three artifacts (purple boxes) and three biological kinds (orange boxes). Each trial included a target item and a shape and taxonomic match to the target item.

All materials, data, analyses, and links to pre-registrations are available here: `https://github.com/davdrose/teleological_essentialism_development`

Experiment 1: Shape and purpose at the basic level

The goal of this experiment was to determine whether giving children information about what things are for would lead them to be less likely to categorize objects at the basic level on the basis of their shape.

Methods

Participants. We recruited 180 participants through Lookit (Scott & Schulz, 2017) who were between the ages of 3 and 4 and met our pre-registered inclusion criteria (*gender*: 86 female, 93 male, 1 no response/other; *language*: 172 English, 8 no response/other). Each age group included 90 participants. Families were compensated $5 for their participation.

Materials. We created eight trials where six were test trials and two were warm-up trials. The test trials included three artifacts and three biological kinds. In all trials, children saw a target item, a shape and a taxonomic match to the target. The full set of materials is shown in Figure 1

Procedure. The experiment was programmed using Lookit (Scott & Schulz, 2017), and children were tested asynchronously. Participants were first introduced to a puppet named Maggie. They were told that Maggie is learning puppet language and were asked if they wanted to help teach her puppet language. Once they agreed, participants then

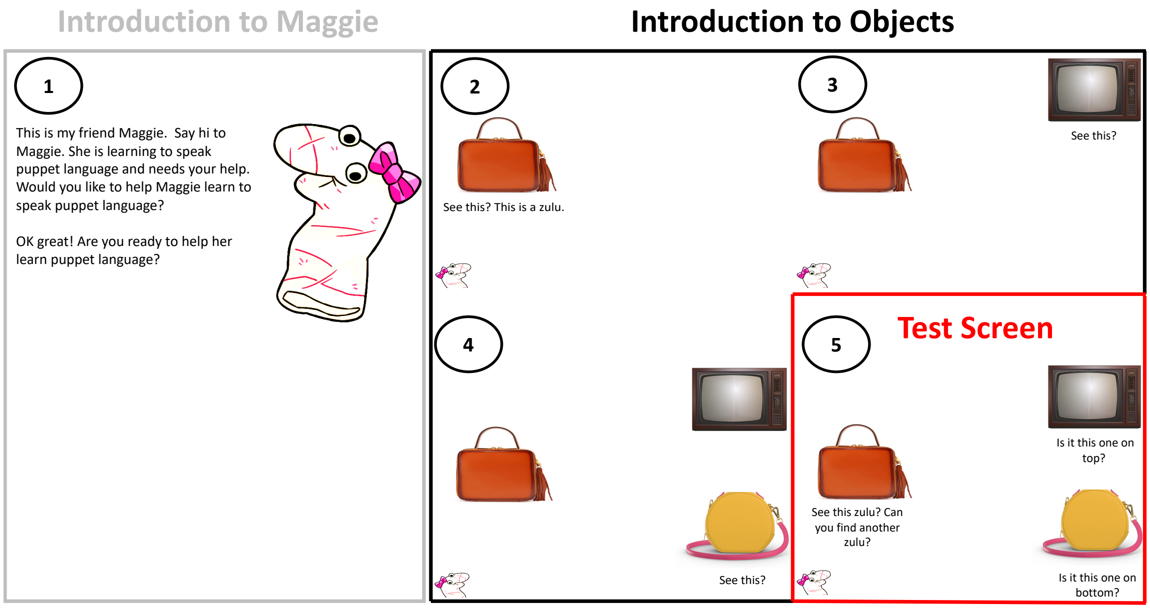

Figure 2. **Experiment 1 procedure**: Overview of experimental procedure. Children were (1) first introduced to Maggie, then (2) introduced to the target object, (3) shown a shape match, (4) a taxonomic match and (5) asked whether the shape or taxonomic match was a member of the target object's category. Text was shown on the slides to participants.

proceeded to the two warm-up trials. They were first shown, by itself, the target item, which was given a novel label (e.g., See this? This is a zulu.). Then participants saw the shape match and the taxonomic match. On the final screen, participants saw all of the items. The target item wiggled and participants heard, "See this zulu?". Then they were asked, "Can you find another zulu? Is it the one on the top or the one on the bottom?". The objects in the top and bottom position on the screen wiggled when they were referred to. Children made their responses out loud.

The test trials were presented in the same way as the warm-up trials. An example of the procedure is shown in Figure 2.

Design. Participants were randomly assigned to one of three conditions: a Label Only, Irrelevant or Purpose condition. Participants were also randomly assigned to one of three orders of the test items. These three orders were the same in each of our three conditions. Across the three orders, each test trial had the taxonomic match appear either below or above the shape match.

In the label only condition, a name was only given to the target item. The irrelevant condition was inspired by (Gelman & Markman, 1986, Study 3). The target item was given a label and inductively irrelevant information, such as that it gets wet when it rains. The taxonomic match was given the same inductively irrelevant information as the target. The shape match was given different irrelevant information. The purpose condition was similar to the irrelevant condition, except that the items were given purposes. An example of the kind of information provided in the different conditions is shown in Figure 3.

Figure 3. **Experiment 1 conditions**: Overview of experiment conditions. An example trial is shown for each condition: Label Only (blue), Irrelevant (red) and Purpose (green). In each, the standard, the "Zerp", is shown on the left and the shape match, the stool, and taxonomic match, drum, are shown on the top and bottom respectively. After seeing all of the objects, children were asked if they can find another Zerp and asked if it was the one on the top or bottom.

Results

Our pre-registered hypothesis was that children would be more inclined to engage in taxonomic categorization in the purpose condition in comparison to either the label only or irrelevant conditions. This is what we found (see Figure 4). Children were more inclined to categorize taxonomically when given a purpose (70%, 95% confidence interval (CI): [66%, 75%]) compared to when they were only given a label (39% [34%, 45%], odds ratio: 4.21 95% credible interval (CrI) [2.65, 6.26]) or inductively irrelevant properties (46% CI[41%, 51%], odds ratio: 3.19 CrI[2.04, 4.77]). We didn't make any predictions about whether the label only and inductively irrelevant property conditions would differ, but found that there was no credible difference in taxonomic responding between the label only and inductively irrelevant property conditions (odds ratio: .76 CrI[.49, 1.10]).

We also pre-registered the hypothesis that rates of taxonomic responding in the purpose condition would be greater than chance. This is what we found [(CrI[65%, 80%]]. Though we didn't make any predictions about whether rates of taxonomic responding in the label only and irrelevant condition would differ from chance, we found that they were credibly lower than chance in the label only condition (CrI[30%, 47%]) and not credibly different from chance in the irrelevant condition (CrI[36%, 55%]). Lastly, as a pre-registered exploratory analysis, we examined whether there were differences between artifacts and biological kinds. We found that these didn't credibly differ (0.16 CrI[- 0.10, 0.42]).

Discussion

When 3–4 year old children were given purposes for the objects, they no longer categorized based on shape: 70% of their responses were taxonomic. This isn't explained by a simple informational matching strategy since when given inductively irrelevant properties, only 46% of responses were taxonomic, which didn't differ from providing children with only a label, where 39% of responses were taxonomic.

Not only did providing purpose information lead children to overcome a reliance on

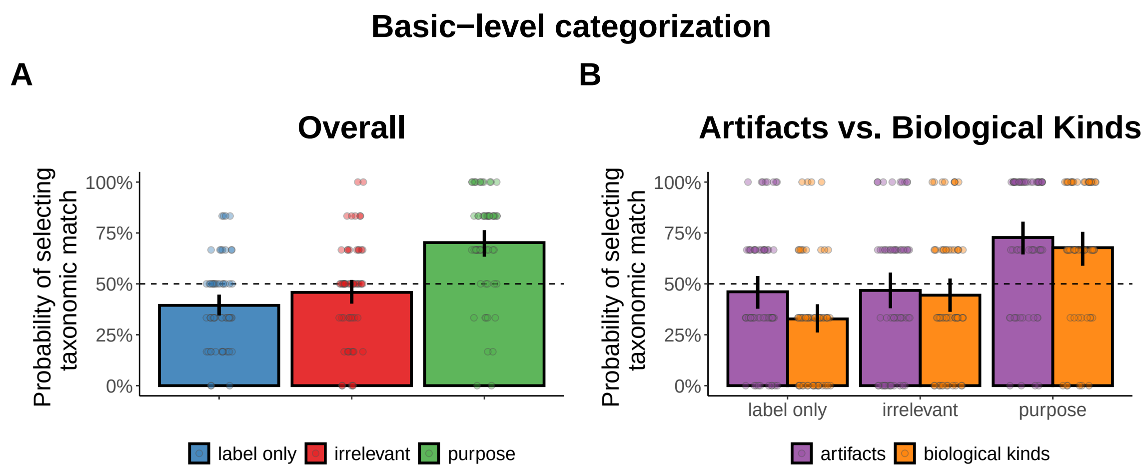

Figure 4. **Experiment 1 results**: (**A**) The mean probability, shown for each bar, of selecting the taxonomic match in each condition: label only (blue), irrelevant (red) and purpose (green). Black lines for each bar are bootstrapped 95% confidence intervals. Small dots are average taxonomic responses for individual participants. (**B**) The mean probability, shown for each bar, of selecting the taxonomic match for artifacts (purple) and biological kinds (orange) in each condition: label only, irrelevant and purpose. Black lines for each bar are bootstrapped 95% confidence intervals. Small dots are average taxonomic responses for individual participants for the artifact (purple) and biological kinds (orange).

shape, it was just as effective for biological kinds as for artifacts. We next ask whether the benefits of teleology in overcoming a reliance on shape extends to the superordinate level.

Experiment 2: Shape and function at the superordinate level

The goal of this experiment was to determine whether giving children information about what things are for would lead them to be less likely to categorize objects at the superordinate level on the basis of their shape.

Methods

Participants. We recruited 90 participants through Children Helping Science who were between the ages of 3 and 4 and met our pre-registered inclusion criteria (*gender*: 46 female, 38 male, 6 no response/other; *language*: 78 English, 12 no response/other). Each age group included 45 participants. Families were compensated $5 for their participation.

Materials. We created eight trials where six were test trials and two were warm-up trials. The test trials included three artifacts and three biological kinds. In all trials, children saw a target item, a shape and a taxonomic match to the target. The full set of materials is shown in Figure 5.

Procedure. The procedure was the same as in Experiment 1, except that children were tested synchronously over Zoom.

Design. The design was the same as in Experiment 1.

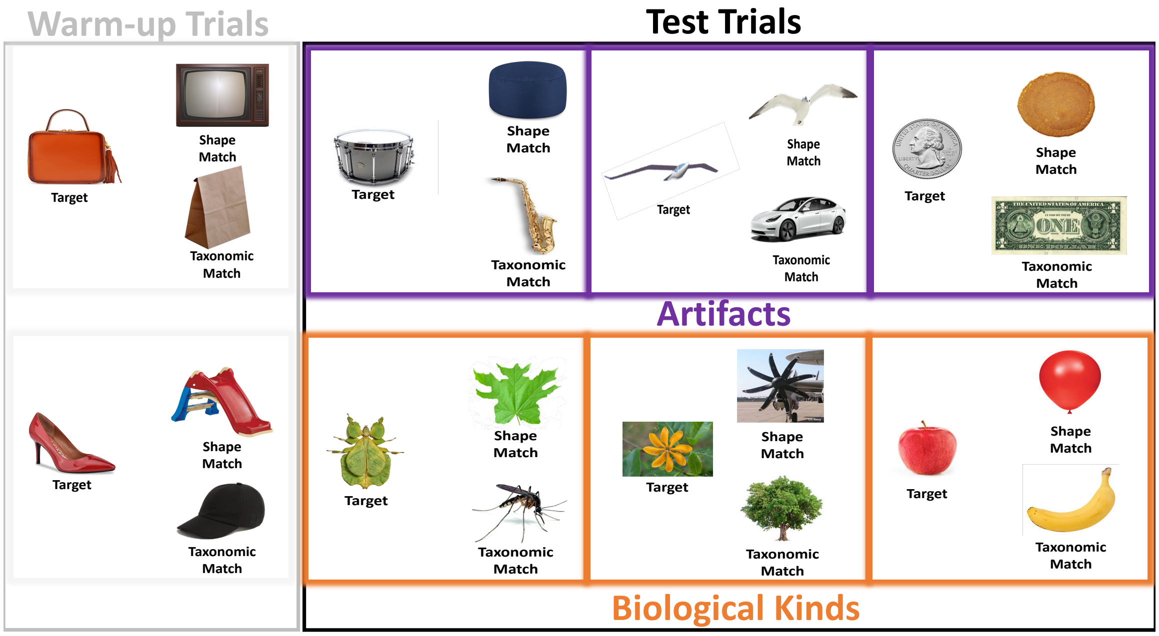

Figure 5. **Experiment 2 items**: Items for the warm-up (left box in gray) and test trials (right box in black). The test trials included three artifacts (purple boxes) and three biological kinds (orange boxes). Each trial included a target item and a shape and taxonomic match to the target item.

Results

Our pre-registered hypothesis was that children would be more inclined to engage in taxonomic categorization in the purpose condition in comparison to either the label only or irrelevant conditions. This is what we found (see Figure 6). Children were more inclined to categorize taxonomically when given a purpose (67% CI[60%, 74%]) compared to when they were only given a label (41% CI[34%, 49%], odds ratio: 3.41 CrI[1.58, 6.12]) or inductively irrelevant properties (36% CI[29%, 43%], odds ratio: 4.27 CrI[2.04, 7.92]). We also pre-registered that there would be no credible difference in taxonomic responding between the label only and inductively irrelevant property conditions. This is what we found (odds ratio: 1.28 CrI[.60, 2.28]).

We also pre-registered the hypothesis that rates of taxonomic responding in the purpose condition would be credibly greater than chance. This is what we found (CrI[57%, 81%]). Our pre-registered predictions for the label only and irrelevant conditions were that they would not credibly differ from chance. We found that they were credibly lower than chance in the irrelevant condition (CrI[21%, 47%]) but not credibly different from chance in the label only condition (CrI[27%, 54%]). Lastly, we preregistered that we would explore whether there were differences between artifacts and biological kinds. We found that there wasn't (-0.23 CrI[- 0.66, 0.20]).

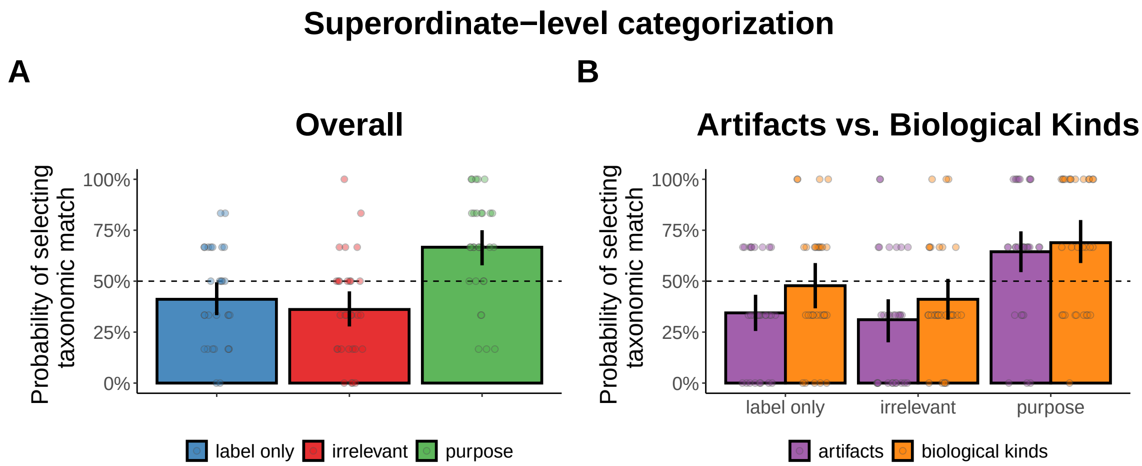

Figure 6. **Experiment 2 results**: (**A**) The mean probability, shown for each bar, of selecting the taxonomic match in each condition: label only (blue), irrelevant (red) and purpose (green). Black lines for each bar are bootstrapped 95% confidence intervals. Small dots are average taxonomic responses for individual participants. (**B**) The mean probability, shown for each bar, of selecting the taxonomic match for artifacts (purple) and biological kinds (orange) in each condition: label only, irrelevant and purpose. Black lines for each bar are bootstrapped 95% confidence intervals. Small dots are average taxonomic responses for individual participants for the artifact (purple) and biological kinds (orange).

Discussion

Even at the superordinate level, giving 3 - 4 year old children purposes for objects leads them to overcome a reliance on shape and categorize taxonomically. 67% of their responses were taxonomic. Again, this isn't explained by a simple informational matching strategy since when given inductively irrelevant properties, only 36% of responses were taxonomic. Providing only a label resulted in 41% of responses being taxonomic. Not only did providing purpose information lead children to overcome a reliance on shape, it was, again, equally as effective for biological kinds as for artifacts.

Overcoming a reliance on perceptual features in determining identity for things that undergo radical perceptual change, as in transformation tasks, is even more challenging than overcoming shape similarities in categorization. We ask, in our final experiment, whether knowing that things preserve their purpose across radical perceptual transformations might make children less likely to rely on appearances.

Experiment 3: Transformations

The goal of this experiment was to determine whether giving children information about what things are for would lead them to be less likely to categorize objects based on their appearance when they undergo radical changes in their perceptual features.

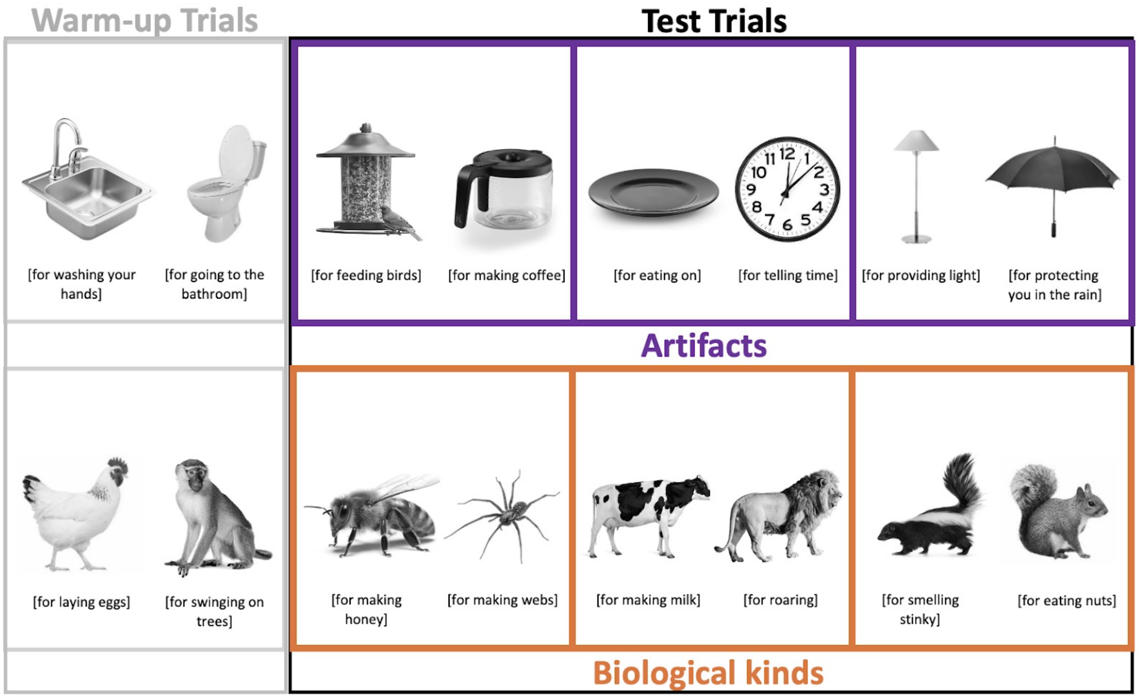

Figure 7. **Experiment 3 items**: Item pairs for the warm-up (left box in gray) and test trials (right box in black). The test trials included three artifact pairs (purple boxes) and three biological kind pairs (orange boxes). Information included in the brackets was given only in the purpose condition.

Methods

Participants. We recruited 300 participants through Lookit who were between the ages of 5 and 9 and met our pre-registered inclusion criteria (*gender*: 156 female, 143 males , 1 no response/other; *language*: 274 English, 26 no response/other). Each age group included 60 participants. Families were compensated $5 for their participation.

Materials. We created 8 trials involving transformations on four pairs of artifacts and four pairs on biological kinds. The full set of item pairs is shown in Figure 7. These were selected based on a two-part norming study aimed at determining which functions for different kinds were accessible to 4- to 6-year old children (see Appendix). We selected some of the top artifacts and animals whose purpose they children agreed on. Using those in the second part of our norming study, we gave 4–6 year olds 14 artifact pairs and 14 animal pairs and asked which had that purpose. Children were overwhelmingly inclined to identify the correct items when asked which of two had some particular purpose. From these, we selected four artifact and four biological kind pairs to be used here in Experiment 3.

Procedure. The experiment was programmed using Lookit (Scott & Schulz, 2017), and children were tested asynchronously. Each child was first introduced to Maggie and told that they, along with Maggie, would learn about Andy's projects. Next, the child was introduced to Andy and shown his workshop. Children were told that Andy has various supplies for his projects including crayons, paper, tools, and a toolbox, and that he will use his supplies to change things.

Children began with two warm-up trials. They were shown an object (e.g., a sink) and told that Andy was going to use this for one of his projects. Maggie then identifies the item (e.g., "That's a sink.") and then children are told that Andy will change the object with his supplies. They are then given a description of how the object is changed (e.g., made rounder, given a handle and lid) and then shown, on a new screen, a different object (e.g., a toilet) and told that the thing Andy changed with his supplies now looks like this. Maggie then says what the item looks like (e.g., "That looks like a toilet."). Children are then reminded that the thing Andy changed now looks like what is shown on the screen (e.g., a toilet) and asked what the object is ("Do you think this is a sink or a toilet?"). Children made their responses out loud.

The procedure was the same for warm-up and test trials. The warm-up and test trials were based on Keil's (1989) original transformation task, but the information given about the objects was different in the standard and purpose versions. In the purpose version, but not the standard version, children were told what each object was for. When the initial object (e.g., a sink) is introduced, Maggie says "That's a sink! And sinks are for washing your hands." After Andy changes it to look different (e.g., a toilet), Maggie says, "That looks like a toilet! And toilets are for going to the bathroom." Andy then says to Maggie that the object has the purpose of the original item (e.g., "This isn't for going to the bathroom. It is for washing your hands."). The standard version, as in Keil's (1989) original transformation task, only included information about the perceptual transformation.

The full procedure for purpose and standard versions are shown in Figure 8

Design. Children were assigned to the purpose version or standard version. For each condition, there were three orders where the order of the warm-up trials was fixed but the order of the test trials varied. Due to a coding error, one of our three orders of trials had test trials randomized for each participant. The other two trial orders were fixed for each participant.

Results

The results are shown in Figure 9. We first focus on the results comparing the standard and purpose versions that were predicted based on this manipulation. In line with our pre-registered prediction, in contrast to the standard version, we found that children in the purpose version of the transformation task were more likely to think the object retained its original identity despite the radical transformation for both artifacts (*purpose*: 51% CI[47%, 56%], *standard*: 9% CI[6%, 12%], odds ratio: 752 CrI[46, 5032]) and animals (*purpose*: 49% CI[45%, 55%], *standard*: 13% CI[10%, 16%], odds ratio: 500 CrI[18.9, 3795]). We also found, as predicted, that as children, regardless of whether they were in the purpose or standard version, got older, they were less likely to judge that the object changed categories (*artifacts*: 1.05 CrI[0.53, 1.68]; *biological kinds*: 1.00 CrI[0.43, 1.67]). However, we predicted that the difference between the purpose and standard versions in the extent to which children judged that the object retained its category membership after the transformation would increase with age, but we didn't find evidence of this (*artifacts*: -0.53 CrI[-1.64, 0.52]; *biological kinds*: -0.59 CrI[-1.75, 0.58]).

We also made specific pre-registered predictions for each condition. For the purpose version, we predicted that as children got older they would be more likely to judge that

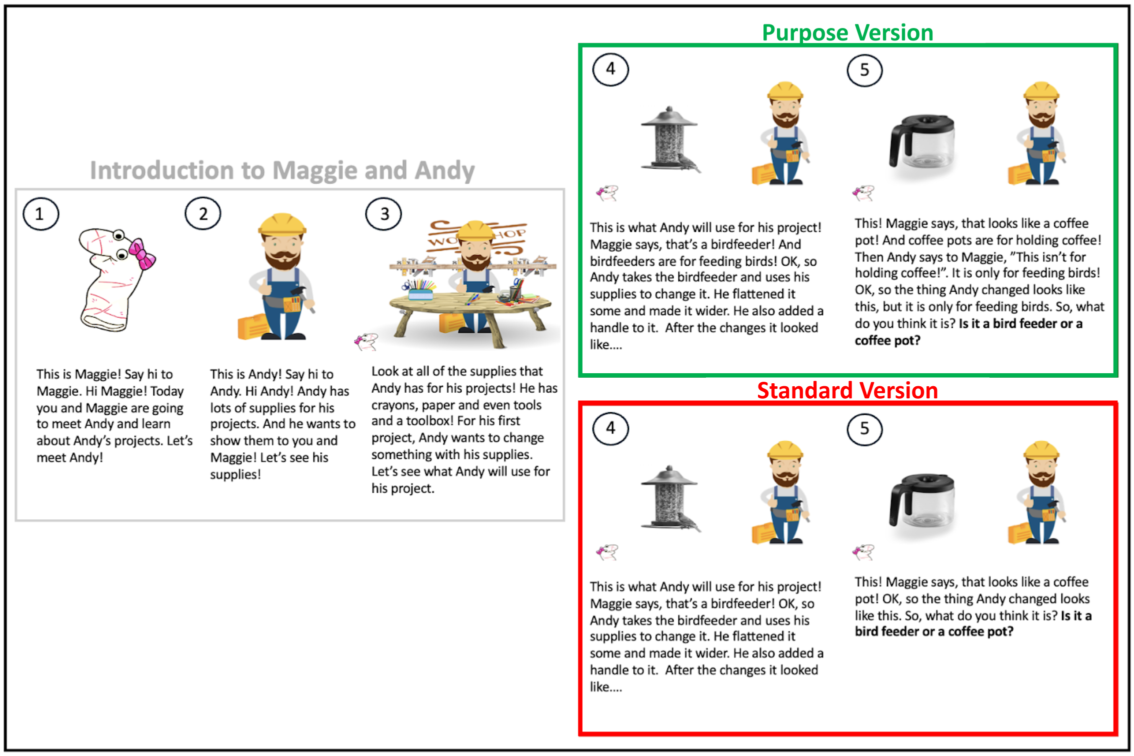

Figure 8. **Experiment 3 procedure**: Overview of transformation experiment procedure. Children began by (1) being introduced to Maggie, (2) introduced to Andy and (3) shown the supplies that Andy uses to change things (left; gray). Children in the purpose version (top; green) were then (4) introduced to the object Andy is going to change with Maggie identifying what it is and what it is for. Then (5) children are shown what the object looks like after the change, Maggie says what it looks like and what it is for, while Andy says that it is for what the original object was for. Children are then asked what the object is. The standard version (bottom; red) is the same except Maggie doesn't mention a purpose when she sees the initial object (4) or when she sees the changed object (5). Text was not shown on the slides to participants.

the object didn't change its category after the transformation. We found evidence of this (0.77 CrI[0.14, 1.48]). And for the standard version, we predicted, in line with Keil's original finding, that children would overall be more likely to think biological kinds didn't change their original category membership after the transformation. However, we didn't find evidence supporting this (odds ratio: 2.68 CrI[.25, 7.27]). We did, however, and as predicted, find that as children got older they were more likely to judge that the object didn't change its category after the transformation (1.14 CrI[.44, 2.07]). Lastly, we predicted that in the standard version, the difference between thinking biological kinds and artifacts retained their category membership would change with age such that children, as they got older, would be more inclined to think biological kinds retained their category membership. But we didn't find evidence of this (0.01 CrI[-0.54, 0.59]).

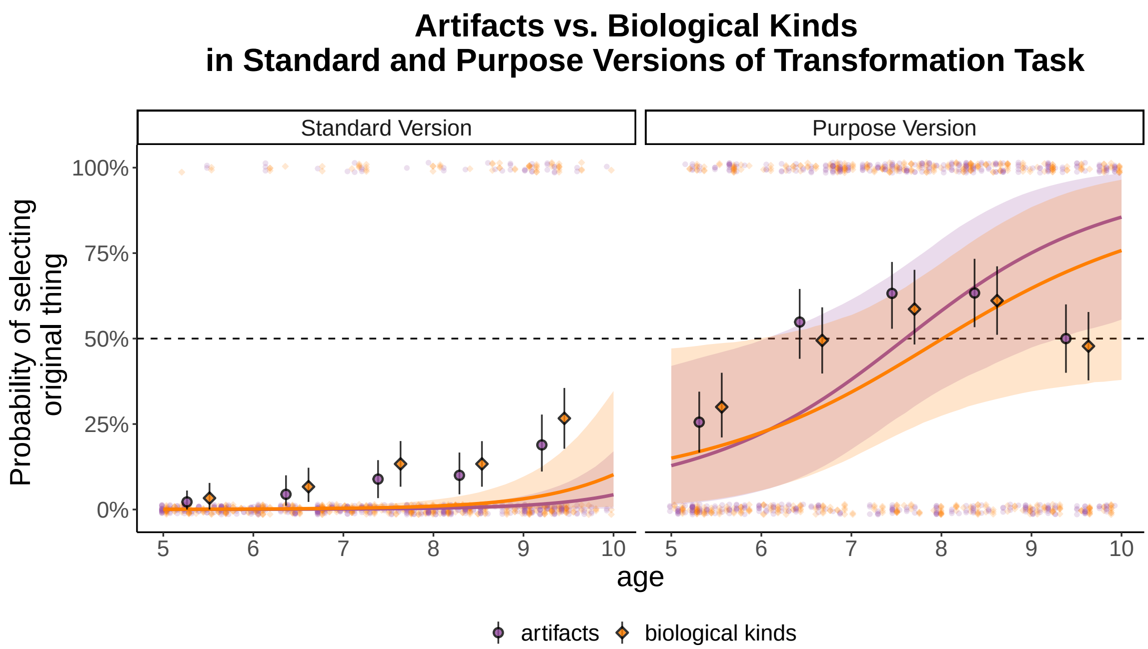

Figure 9. **Experiment 3 results**: Probability of judging that the changed object was still a member of its original category in the standard (left panel) and purpose (right panel) conditions. For each condition, regression lines show the fits of Bayesian logistic mixed effects models with 95% credible intervals with model predictions for artifacts (purple) and biological kinds (orange) over age. Large dots are means for each age group for the artifact (purple) and biological kind (orange) items with bootstrapped 95% confidence intervals. Small dots are individual responses for artifacts (purple) and biological kinds (orange).

Discussion

In contrast to the standard transformation task, whether a biological kind or artifact underwent radical transformation in its appearance, if told it preserved its original purpose, children were more inclined to think it retained its identity. In fact the developmental trajectories of children in the standard transformation task and those told the object preserved its purpose were almost non-overlapping.

General Discussion

We found that knowledge of what something is for plays a powerful role in categorization. At the typical age, between 3 and 4, where shape plays a dominant role in categorization (e.g., Landau et al., 1988), we found that, in contrast to providing a label or inductively irrelevant properties, 3 and 4 year old children can use information about an object's purpose to override a reliance on shape. And they use purposes for both biological kinds and artifacts, and at the basic (Experiment 1) and more challenging superordinate (Experiment 2) level, to override a reliance on shape. Purposes also help children override a reliance on perceptual features in radical transformation tasks (e.g., Keil, 1989). When an object underwent radical changes in its perceptual features, children were more inclined to think the object retained its identity when it preserved its purpose (Experiment 3). In fact, providing five year old children with purposes led them to be as likely as nine year old children, in the standard procedure with no purpose information, to think that an object that underwent radical perceptual transformation retained its identity.

That children use a thing's purpose to determine what it is for both biological kinds and artifacts in both categorization and transformation tasks, suggests that teleological information is readily accessible for determining category membership and identity. This suggests that children can elaborate placeholders for essences with teleological information.

We are not claiming that scientific essentialism has no place in children's reasoning about category membership and identity (Joo & Yousif, 2022; Neufeld, 2021; Toorman, 2023). But elaboration of scientific essences requires abstract knowledge of sophisticated concepts, such as DNA, that will have a protracted development, and placeholders alone lack content that can be leveraged in categorizing. While we didn't directly compare how well children can rely on scientific or teleological information in elaborating placeholders, determining how to specify scientific elaborations of placeholders is challenging. And providing scientific information about the insides of things or what they are composed of has been unsuccessful (Gelman & O'Reilly, 1988). In contrast, from an early age, teleological considerations can already serve to elaborate some placeholders.

Scientific essentialism treats placeholders as yet unspecified scientific mechanisms, that may or may not eventually be elaborated with scientific information. To swap one placeholder for another, people might also have teleological placeholders. People may not always know what the purpose of something is. But children's readiness to use teleological information in reasoning about category membership and identity suggests that it can provide an especially cognitively accessible and fruitful kind of content to fill out placeholders.

It is an interesting further question how the teleological notions that children rely on to elaborate placeholders develops. In our experiments, we used objects with familiar purposes. A question for future research is how beneficial it would be in categorization and transformation tasks to provide children with novel purposes. Would novel purposes readily fill out placeholders?

Another interesting, yet unresolved, question is what role different kinds of teleological notions play in the development and representation of placeholders. Purposes can be treated as either intrinsic or relational, e.g., concerning somethings ecological role (see e.g., Dink & Rips, 2017; Medin & Atran, 2004; Unsworth et al., 2012; Waxman, Medin, et al., 2013). Viewing purposes as either intrinsic or relational suggests a broader conception of essences. Scientific essentialism focuses only on intrinsic properties. But even in biology, essences can be construed as relational (Griffiths, 1999; Okasha, 2002).

Teleological thinking might promote a broader conception of essentialist thinking, and one that is continuous with some aspects of science. As one example, anatomy and ecology are scientific disciplines that understand and characterize kinds in terms of their functions, where these can be viewed as intrinsic or relational. Developing a conception of purposes on which they can be intrinsic or relational might promote some forms of thinking that underwrite scientific reasoning—including discipline specific forms of knowledge that even children recognize (Keil, Stein, Webb, Billings, & Rozenblit, 2008)—and even aid, as purposes sometimes do, in scientific discovery (Craver & Darden, 2013). Teleological essentialism suggests a reorientation of the landscape surrounding research on essentialist thinking and its development, opening new ways to think about how essences might be construed.

References

Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, *21*(4), 547–569.

Baldwin, D. A. (1989). Priorities in children's expectations about object label reference: Form over color. *Child Development*, 1291–1306.

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

Booth, A. E., Waxman, S. R., & Huang, Y. T. (2005). Conceptual information permeates word learning in infancy. *Developmental psychology*, *41*(3), 491–505.

Butler, L. P., & Markman, E. M. (2014). Preschoolers use pedagogical cues to guide radical reorganization of category knowledge. *Cognition*, *130*(1), 116–127.

Cimpian, A., & Markman, E. M. (2005). The absence of a shape bias in children's word learning. *Developmental Psychology*, *41*(6), 1003.

Craver, C. F., & Darden, L. (2013). *In search of mechanisms: Discoveries across the life sciences*. University of Chicago Press.

Deák, G. O., Ray, S. D., & Pick, A. D. (2002). Matching and naming objects by shape or function: age and context effects in preschool children. *Developmental psychology*, *38*(4), 503–518.

Diesendruck, G., & Bloom, P. (2003). How specific is the shape bias? *Child development*, *74*(1), 168–178.

Dink, J. W., & Rips, L. J. (2017). Folk teleology and its implications. In D. Rose (Ed.), *Experimental metaphysics* (pp. 207–235). Bloomsbury.

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

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

Gelman, S. A. (2013). Artifacts and essentialism. *Review of Philosophy and Psychology*, *4*, 449–463.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*(3), 183–209.

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

Gelman, S. A., & O'Reilly, A. W. (1988). Children's inductive inferences within superordinate categories: The role of language and category structure. *Child development*, 876–887.

Gelman, S. A., & Ware, E. A. (2012). Conceptual development: The case of essentialism. In *The oxford handbook of philosophy of cognitive science*. Oxford University Press.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the nonobvious. *Cognition*, *38*, 213–244.

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive development*, *14*(4), 487–513.

Gershkoff-Stowe, L., & Smith, L. B. (2004). Shape and the first hundred nouns. *Child development*, *75*(4), 1098–1114.

Graham, S. A., & Diesendruck, G. (2010). Fifteen-month-old infants attend to shape over other perceptual properties in an induction task. *Cognitive Development*, *25*(2),

111–123.
Griffiths, P. (1999). Squaring the circle: Natural kinds with historical essences. In R. A. Wilson (Ed.), *Species: New interdisciplinary studies* (pp. 209–228). MIT Press.
Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development*, *9*, 45–75.
Jara-Ettinger, J., Levy, R., Sakel, J., Huanca, T., & Gibson, E. (2022). The origins of the shape bias: Evidence from the tsimane'. *Journal of Experimental Psychology: General*, *151*(10), 2437.
Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development*, *8*(2), 113–139.
Joo, S., & Yousif, S. R. (2022). Are we teleologically essentialist? *Cognitive Science*, *46*(11), e13202.
Keil, F. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In *Mapping the mind: Domain specificity in cognition and culture* (pp. 234–254). Cambridge University Press.
Keil, F. C. (1989). *Concepts, kinds, and cognitive development.* MIT Press.
Keil, F. C., Stein, C., Webb, L., Billings, V. D., & Rozenblit, L. (2008). Discerning the division of cognitive labor: An emerging understanding of how knowledge is clustered in other minds. *Cognitive science*, *32*(2), 259–300.
Kelemen, D. (1999). The scope of teleological thinking in preschool children. *Cognition*, *70*(3), 241–272.
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.
Kripke, S. (1980). *Naming and necessity.* Harvard University Press.
Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive development*, *3*(3), 299–321.
Macario, J. F. (1991). Young children's use of color in classification: Foods and canonically colored objects. *Cognitive Development*, *6*(1), 17–46.
Markman, E. M. (1989). *Categorization and naming in children: problems of induction.* MIT Press.
Medin, D. L., & Atran, S. (2004). The native mind: biological categorization and reasoning in development and across cultures. *Psychological review*, *111*(4), 960.
Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In *Similarity and analogical reasoning* (pp. 179–195). Cambridge University Press.
Neufeld, E. (2021). Against teleological essentialism. *Cognitive Science*, *45*(4), e12961.
Okasha, S. (2002). Darwinian metaphysics: Species and the question of essentialism. *Synthese*, *131*, 191–213.
Peretz-Lange, R., & Kibbe, M. M. (2024). "shape bias" goes social: Children categorize people by weight rather than race. *Developmental Science*, *27*(2), e13454.
Putnam, H. (1975). The meaning of 'meaning'. In *Minnesota studies in the philosophy of science* (Vol. 7, pp. 215–271). University of Minnesota Press.
Rose, D., & Nichols, S. (2019). Teleological essentialism. *Cognitive science*, *43*(4), e12725.
Rose, D., & Nichols, S. (2020). Teleological essentialism: generalized. *Cognitive science*, *44*(3), e12818.
Samuelson, L. K., & Smith, L. B. (2005). They call it like they see it: Spontaneous naming

and attention to shape. *Developmental Science*, *8*(2), 182–198.

Scott, K., & Schulz, L. (2017). Lookit (part 1): A new online platform for developmental research. *Open Mind*, *1*(1), 4–14.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*(1), 13–19.

Toorman, J. (2023). Against arguments from diagnostic reasoning. *Cognitive Science*, *47*(11), e13376.

Unsworth, S. J., Levin, W., Bang, M., Washinawatok, K., Waxman, S. R., & Medin, D. L. (2012). Cultural differences in children's ecological reasoning and psychological closeness to nature: Evidence from menominee and european american children. *Journal of Cognition and Culture*, *12*(1-2), 17–29.

Waxman, S., Medin, D., & Ross, N. (2007). Folkbiological reasoning from a cross-cultural developmental perspective: early essentialist notions are shaped by cultural beliefs. *Developmental psychology*, *43*(2), 294.

Waxman, S. R. (1990). Linguistic biases and the establishment of conceptual hierarchies: Evidence from preschool children. *Cognitive Development*, *5*(2), 123–150.

Waxman, S. R., Medin, D. L., et al. (2013). Teleological reasoning about nature: intentional design or relational perspectives? *Trends in Cognitive Sciences*, *17*(4), 166–171.

Zhu, R., Goddu, M. K., Zhu, L. Z., & Gopnik, A. (2024). Preschoolers' comprehension of functional metaphors. *Open Mind*, *8*, 924–949.

Appendix A

Experiment 3 Norming Part 1: Saying what something's purpose is

The goal for the first part of our norming experiment was to determine whether there were purposes for some artifacts and animals that children tended to agree on.

Methods

Participants. We recruited 45 participants through Children Helping Science and Lookit who were between the ages of 4 and 6 and met our pre-registered inclusion criteria (*gender*: 25 female, 14 male, 3 no response/other; *language*: 30 English, 15 no response/other). Each age group included 15 participants. Participants were compensated $5 for their participation.

Materials. We created forty trials where 20 were artifacts and 20 were biological kinds. The full set of materials is shown in Figure A1

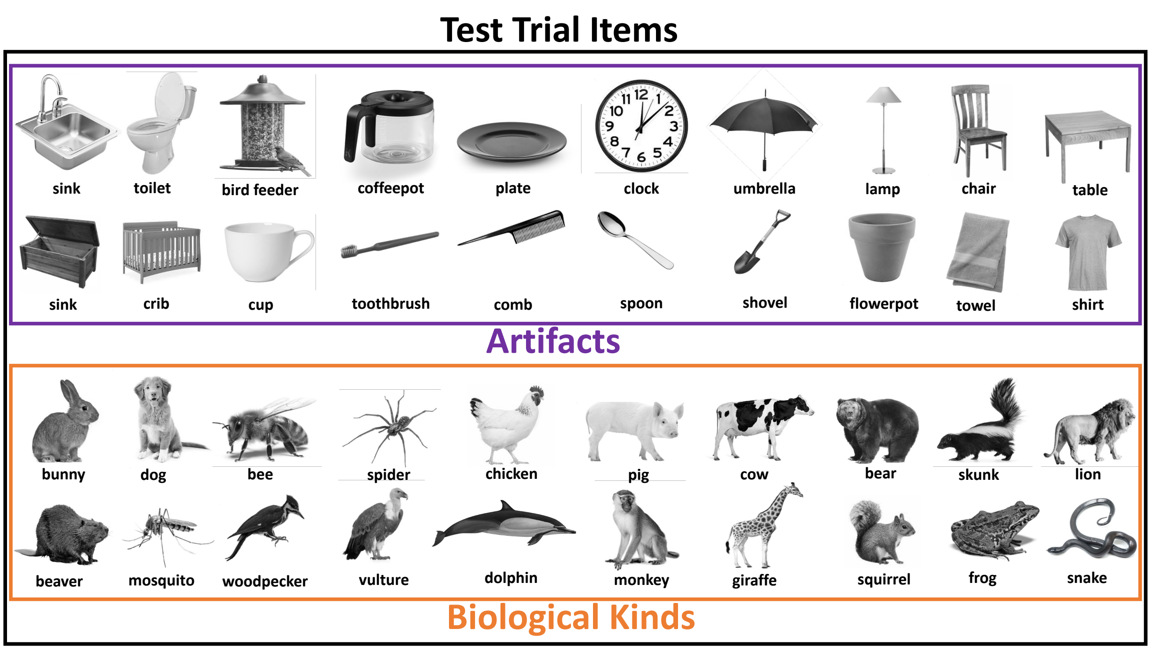

Figure A1. **Norming Part 1 items**: All forty test items that children were asked what they are for. The top twenty (purple) are artifacts and the bottom twenty (orange) are biological kinds.

Procedure. Participants were introduced to a puppet named Maggie, told that Maggie wants to learn what different things are for and that they would teach Maggie what different things are for. On each trial, participants were shown an object and told what it was (e.g., "See this? This is a sink."). Then they were asked whether they could tell Maggie what that object is for (e.g., "Can you tell Maggie what sinks are for?"). Children made their responses out loud.

Design. Participants were given all 40 items in a random order.

Results

Our pre-registration didn't include any hypotheses since we were only interested in identifying items that could be used in developing study materials. We computed the probability of each word produced for an item. Figure A2 shows the top items selected.

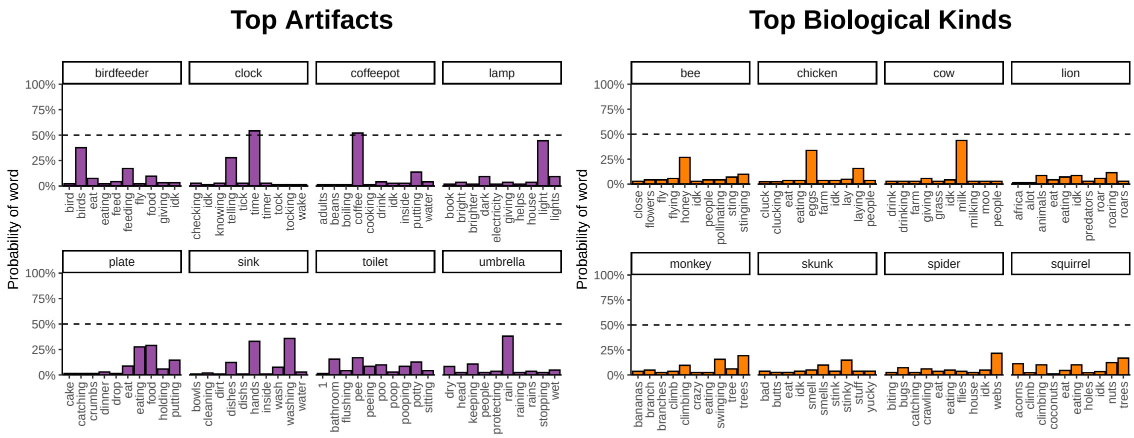

Figure A2. **Norming Part 1 top items**: Top items for which children tended to converge on a purpose. The top eight artifacts are on the left (e.g., clocks tell time) and the top eight biological kinds are on the right (e.g., chickens lay eggs).

Discussion

Having identified artifacts and biological kinds for which there is some convergence on what children say their purpose is, we now, as the second part of our norming study, use the purposes that children tended to mention in this study and ask whether children can identify which objects have these purposes.

Appendix B

Experiment 3 Norming Part 2: Identifying what things have some particular purpose

The goal for the second part of our norming experiment was to determine whether children could identify which objects had the purposes they attributed to them in the first part of our norming study.

Methods

Participants. We recruited 40 participants through Lookit who were between the ages of 4 and 5 and met our pre-registered inclusion criteria (*gender*: 21 female, 19 male; *language*: 39 English, 1 no response/other). Each age group included 20 participants. Participants were compensated $5 for their participation.

Materials. We created 28 trials that included 14 artifact pairs and 14 biological kind pairs. These pairs included all 16 items from part 1 of our norming as well as additional items used in that experiment. The full set of item pairs is shown in Figure B1

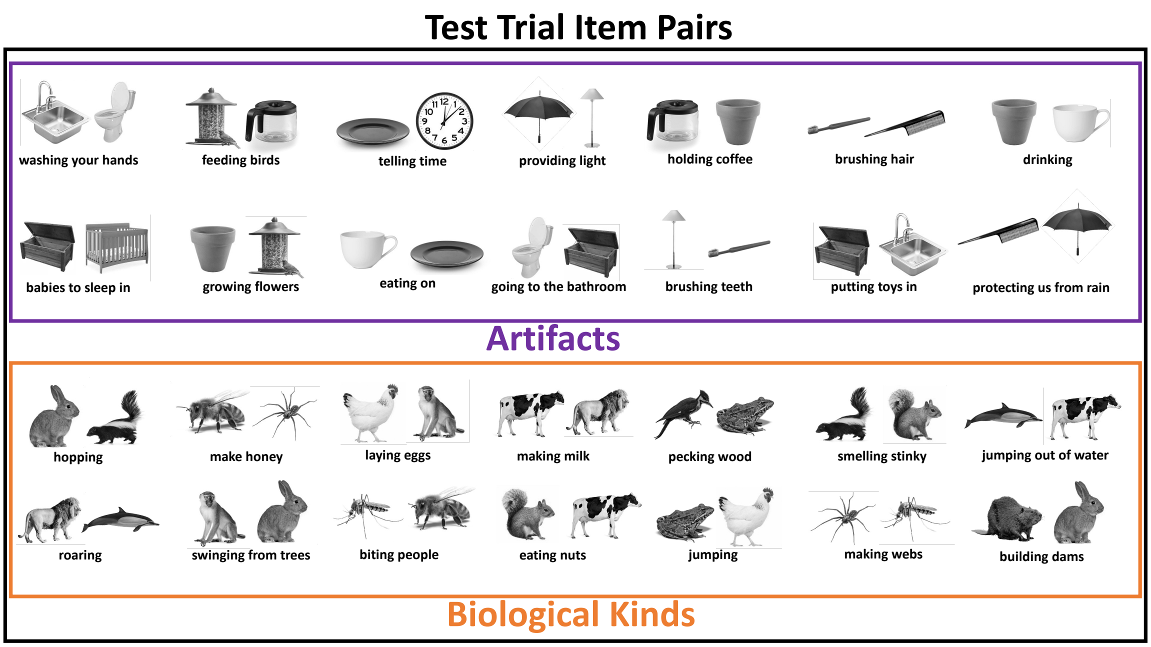

Figure B1. **Norming Part 2 items**: Test trial item pairs with fourteen artifact pairs (e.g., a sink and toilet where the target purpose was washing hands) shown on top (purple) and fourteen biological kind pairs (e.g., a rabbit and skunk where the target purpose was hopping) shown on bottom (orange).

Procedure. The procedure was similar to part 1 of our norming study. After being introduced to Maggie and told that they would teach her what different things are for, participants were then shown pairs of items. They were told what each item in the pair is (e.g., "See this? This is a sink. See this? This is a toilet"). Then they were asked which item in the pair is for some particular purpose (e.g., "Do you think sinks or toilets are for going to the bathroom?"). Responses were made out loud.

Design. The order of trials was randomized.

Results

Our pre-registration didn't include any hypotheses since we were only interested in identifying items that could be used in developing study materials. We computed the probability that the correct item was identified for a given purpose (see **??**).

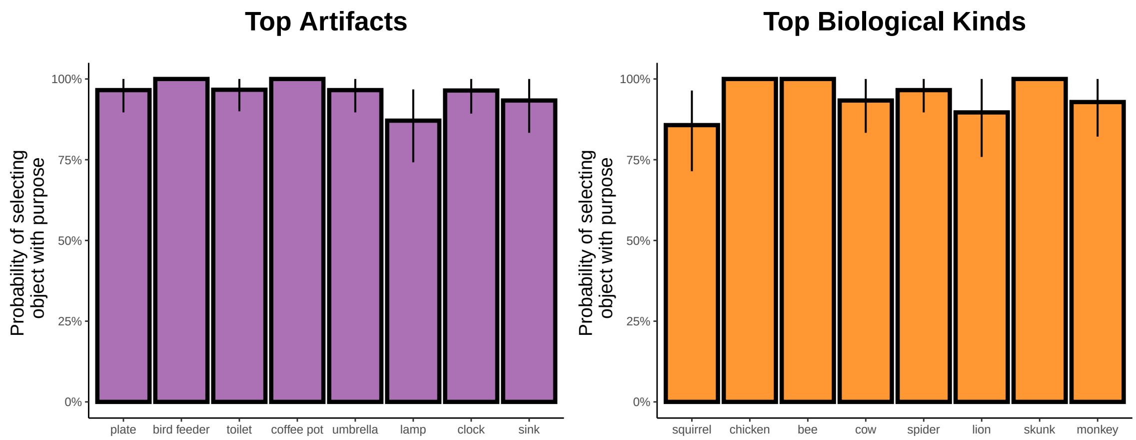

Figure B2. **Norming Part 2 results**: Correct item selected when given a purpose and shown a pair of objects. Artifacts (purple) are shown on the left and biological kinds (orange) are shown on the right.

Discussion

Children were overwhelmingly inclined to select the correct items when asked which of two had some particular purpose. Since they can identify which items have which purposes, we used these in our final experiment (Experiment 3) involving things transforming to look like other things.