UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Functional Explanations Link Gender Essentialism and Normativity

Permalink

https://escholarship.org/uc/item/3s78z492

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Foster-Hanson, Emily
Lombrozo, Tania

Publication Date

2024

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

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

Functional Explanations Link Gender Essentialism and Normativity

Emily Foster-Hanson (efoster3@swarthmore.edu)
Department of Psychology, 500 College Ave
Swarthmore, PA 19081 USA

Tania Lombrozo (lombrozo@princeton.edu)
Department of Psychology, Peretsman Scully Hall
Princeton, NJ 08540 USA

Abstract

Why do beliefs that gender differences are innate (i.e., gender essentialism) sometimes lead to normative judgments about how individual people ought to be? In the current study, we propose that a missing premise linking gender essentialism and normativity rests on the common folk-biological assumption that biological features serve a biological function. When participants ($N = 289$) learned that a novel feature of the gender category "mothers" was common and innate, they overwhelmingly assumed that it must have served some function across human history. When they learned that it served a historical function, they assumed that it must still be beneficial in today's environment. When participants learned that the feature was beneficial, they judged that contemporary mothers ought to have it, and they were more willing to intervene to ensure that they would by constraining the choices of individual mothers. Thus, we suggest that essentialist assumptions can shape normative social judgments via the explanations people tend to generate about why certain features of natural kind categories become common to begin with. This finding articulates one manifestation of the naturalistic fallacy, with implications for policy debates about bodily autonomy and choice.

Keywords: essentialism; gender; normativity; is-ought reasoning; folk-biology; intuitive theories

Introduction

Of all social categories, gender is the one most often conceptualized as a "natural kind" (Bach, 2012; Haslam et al., 2000; Kenrick, 1994; Maccoby, 1988; Martin & Ruble, 2004; McConnell-Ginet, 2013; Prentice & Miller, 2006; Taylor, 1996; Wood & Eagly, 2010). By age 3, participants across cultural contexts tend to view gender as marking essentially different kinds of people that exist in nature (Taylor, 1996; Skewes et al., 2018), and they expect gender differences in behavior to emerge naturally and inevitably due to innate predispositions (Brescoll & LaFrance, 2004; Meyer & Gelman, 2016; Prentice & Miller, 2007), regardless of a person's appearance (Gelman et al., 1986) or environment (Rhodes & Gelman, 2009; Taylor, 1996).

These essentialist beliefs (Medin & Ortony, 1989) about gender categories can have important downstream consequences for how people reason about themselves and others in daily life, including by giving rise to prescriptive judgments about what category members *should* be like (Allport, 1954; Haslanger, 2014; Keller, 2005; Martin & Parker, 1995; Skewes et al., 2018; Wilton et al., 2019). For example, priming essentialist explanations for gender differences increases endorsement of gender inequality (Morton et al., 2009) and gender stereotypes (Brescoll & LaFrance, 2004; Coleman & Hong, 2008), and impairs performance on counter-stereotypical tasks (Dar-Nimrod & Heine, 2006). People even judge immoral actions as less blameworthy when they are attributed to innate gender-based predispositions (e.g., male promiscuity; Ismail et al., 2012; see also Dar-Nimrod et al., 2014; Monterosso et al., 2005).

The assumption that people ought to do what is natural for their gender category thus offers a striking example of the common tendency to infer *ought* from *is* (Hume, 1740/2000; see also Black, 1964; Hudson, 1969; Knobe et al., 2013). Sometimes referred to as the naturalistic fallacy (Moore, 1903/2004), people are particularly inclined to assume that how things are *naturally* must be *good* (Friedrich et al., 1989; Kierniesky & Sobus, 1989). For example, many consumers are willing to pay more for foods labelled "natural" and oppose genetically-modified foods that they view as "unnatural" (Rutjens et al., 2018; Scott & Rozin, 2020).

On the surface, assuming that *what is* natural for gender categories reflects how individual people *ought to be* seems like an irrational bias (Daston, 2014; Haslanger, 2014). Essentialist beliefs comprise several inter-related descriptive expectations (e.g., that category membership is intrinsic and causally powerful, that category members homogenously share features, and that the boundaries between categories are discrete and objective; Gelman, 2003; Rhodes & Mandalaywala, 2017), which license a range of inferences and predictions about what members of gender categories are like in general. These descriptive inferences do *not* logically support prescriptive claims about what members of gender categories *should* be like—at least, not without the additional premise that what is natural is somehow valuable or good. For instance, the observation that most women *do* provide the majority of childcare does not necessarily mean that women *should* do so. However, this can follow if someone believes that (1) caring for children is more natural for mothers than fathers, and (2) what is natural for mothers reflects something that is valuable or good—for example, that the current gendered division of labor embodies the ideals of some creator, or that natural selection results in biological features that are valuable for the survival of the species overall.

In the current work, we propose that the missing premise linking gender essentialism and prescriptive judgments rests on people's intuitive theories about why gender differences

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

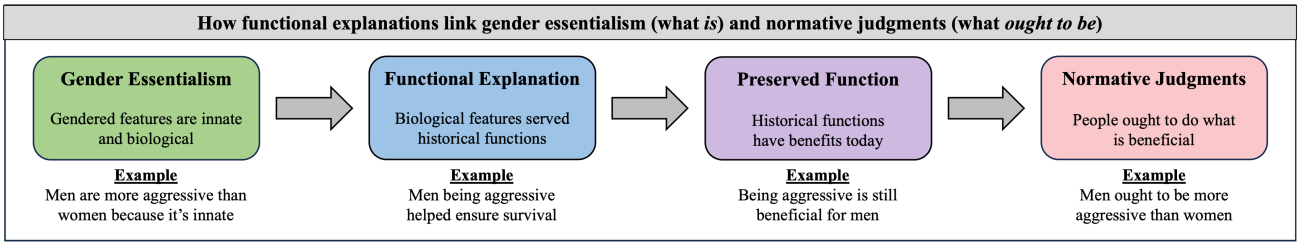

Figure 1: Causal model of the link between gender essentialism and prescriptive judgments about gender via rational inferences about the functions of biological features and their continued benefits in today's environment.

exist in the first place. Specifically, we propose that essentialist beliefs about gender are taken to license normative judgments about how things ought to be via the well-documented folk biological bias to assume that the way things are in nature must serve some function or purpose.

Beliefs about the functions of biological features are situated within people's broader background beliefs about the origins of different kinds of things. For example, humans create artifacts for specific purposes, so their features should fulfill the corresponding function (e.g., knives are sharp because people make them to cut things, so a good knife is a sharp one; Bloom; 1996). Beliefs about what is natural can similarly license normative judgments when the natural is explained teleologically (e.g., giraffes have long necks *so that* they can eat leaves on tall trees; Kelemen, 1999). These explanations could arise from belief in a divine creator that designed the natural world for a purpose (Evans, 2001), but they are also consistent with common naïve theories about nature. For instance, many people think of nature as agentive, so they view natural progress over time as akin to the goal-directed actions of an agent (i.e., Gaia beliefs; Blancke et al., 2014; Järnefelt et al., 2015; Kelemen, 2012, Moore et al., 2002). In this way, people might explain the features of natural kinds in terms of functions, and think that they should fulfill those functions because it is the intention of their creator (whether God or nature itself) that they should do so.

But common folk-biological beliefs may be sufficient to support a pervasive tendency to explain the biological world in terms of function, even without an explicit appeal to God or Gaia (Coley & Tanner, 2015; Gregory & Ellis, 2009; Kelemen & Rosset, 2009; Kelemen et al., 2013; Liquin & Lombrozo, 2018; Lombrozo et al., 2006; Mayr, 1982; Shtulman, 2006, 2017; Ware & Gelman, 2014). From childhood, people tend to assume that the features of biological kinds exist to serve a function (Kelemen, 1999; Lombrozo & Rehder, 2012), similar to human-made artifacts (Keil, 1994; Rips, 1989). These assumptions peak in childhood (Kelemen, 1999; 2004) but persist throughout the lifespan as an intuitive reasoning strategy (Kelemen & Rosset, 2009; Kelemen et al., 2013; Roberts et al., 2021). Further, while people tend to think of evolution as being more teleological than it is, many biologists and philosophers of biology in fact accept teleological claims regarding the products of natural selection (for relevant discussion, see Lombrozo & Carey, 2006). Crucially, functional explanations establish a normative standard against which category members can be judged as better or worse at fulfilling the function (Foster-Hanson & Lombrozo, 2022b; Lombrozo & Wilkenfeld, 2019).

Here we suggest that essentialist beliefs about gender—i.e., viewing gender categories as akin to natural kinds—can give rise to normative expectations about what people *ought* to do through the same mechanisms as other is-ought inferences about natural kinds (Foster-Hanson & Lombrozo, 2022b). That is, when people attribute gender differences to innate biological predispositions, they may also assume that the causal processes that led to these differences must be inherently beneficent—either because they serve the goals of an intelligent creator, or because of misconceptions that evolution means improvement—so what is common and natural must therefore also have been right and good throughout human history. Features that were historically beneficial are viewed as stable causes (Lombrozo & Carey, 2006; Vasilyeva et al., 2018), which people expect will continue to be beneficial unless the environmental context has radically changed (Foster-Hanson & Lombrozo, 2022b). Thus, beliefs in function and continued benefits license normative judgments that individual people *ought* to have these features. This proposal is summarized in Figure 1.

Method

To provide the strongest test of the chain of inferences described above and depicted in Figure 1, we taught participants about a novel, but plausibly biological, feature of the gendered category "mothers." We chose this category as a test case because we expected participants to generally hold high levels of essentialist beliefs about the category at baseline (Foster-Hanson & Lombrozo, 2022a; Park et al., 2015) and to willingly report their normative judgments about the category (Eagly & Karau, 2002; Foster-Hanson & Lombrozo, 2022a; Heilman & Okimoto, 2008; Thomas et al., 2016).

Participants

Out of 300 adult participants recruited through Prolific, we excluded 11 because they failed to correctly answer one or more manipulation check questions included throughout the study, leaving a final sample of $N = 289$, $M_{age} = 39.64$; 148 men and 137 women (1 each transgender), 3 nonbinary, and

1 unspecified. Participants were tested using Qualtrics and were paid a $12/hour prorated rate for participating.

Procedure

All participants were first told that 87% of biological mothers have a chemical called *virdexin* in their blood; fathers and women without children do not. We then first measured, and then manipulated, each of the predicted inferences in turn. After each new piece of information, we asked all participants about their baseline inferences, then we randomly assigned participants to learn additional information regardless of whether or not it aligned with their prior expectations, to test the downstream consequences of each inference within the chain. For example, after learning that 87% of mothers have *virdexin* in their blood, we asked all participants whether they thought *virdexin* was innate. Then, regardless of their answer, we randomly assigned some participants to learn that *virdexin* is innate (i.e., a byproduct of pregnancy, also found in the blood of previous generations of biological mothers) or not innate (i.e., caused by modern medical practices, and not in the blood of previous generations of mothers). This stepwise procedure was repeated for the inference from innateness to historical function, from historical function to contemporary benefit, and from benefit to normative judgments. Participants were required to correctly answer manipulation check questions throughout the study to proceed. We describe these measures in greater detail in the context of results, below.

We analyzed responses in R using the lme4 package (linear models for responses on 1-7 scales; binomial generalized linear models for binary responses); we report the results of likelihood ratio tests. Data and analysis code are available at https://osf.io/h6wrt/; we preregistered our hypotheses and procedure on OSF, https://osf.io/4ucqz.

Results

At baseline, roughly half (54%) of participants thought that most mothers had *virdexin* in their blood because it was biological or in their genes (comparison to chance was not significant, contrary to our preregistered prediction; $p = .22$). Participants were then randomly assigned to learn either that *virdexin* was innate (i.e., a byproduct of pregnancy, also in the blood of previous generations of biological mothers) or not innate (i.e., caused by modern medical practices, and not in the blood of previous generations of mothers). All other preregistered hypotheses were supported by the data.

Assumptions that innate features are functional

Participants were asked why they thought most mothers have *virdexin* in their blood: (a) it probably serves some function or purpose, (b) it's because of something people did, or (c) some other reason. As predicted, the overwhelming majority of participants who learned that *virdexin* was innate (91%) thought the chemical was functional, in contrast with participants who learned *virdexin* was not innate (14%; X^2 (1) = 177.96, $p < .001$; Figure 2).

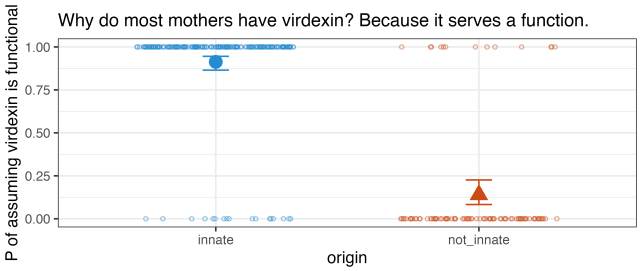

Figure 2: Proportion of participants responding that most mothers have *virdexin* in their blood because it served some function, by origin (innate, not innate). Large shapes are group means with 95% Cis; small circles are individual responses.

Participants were then told (by random assignment) either that *virdexin* has persisted throughout human history because it served an important function for human survival and reproduction, or that it has persisted throughout human history because of chance, and it did not serve any function for human survival or reproduction, but it did no harm either. Thus, participants were assigned to three different historical origin conditions: One third learned that the chemical was both innate and historically functional, one third learned that the chemical was innate but served no historical function, and one third learned that the chemical was not innate at all.

Assumptions that functional means beneficial today

All participants were then asked if they thought *virdexin* has benefits for mothers or infants in today's environment. In support of our preregistered hypothesis, most participants (91%) who learned that *virdexin* was innate and served a historical function thought it would continue to have benefits for both mothers and infants today (Figure 3). In contrast, few who learned it was innate but served no historical function thought it would have benefits for either, and participants who learned that *virdexin* was not innate responded at chance (main effect of condition, X^2 (1) = 99.95, $p < .001$; Figure 3). Thus, participants' assumption that innate features have served a historical function (which was at near-ceiling levels) supported subsequent inferences that the feature continues to offer benefits (again at near-ceiling levels), but just viewing the feature as innate was not sufficient to support these inferences.

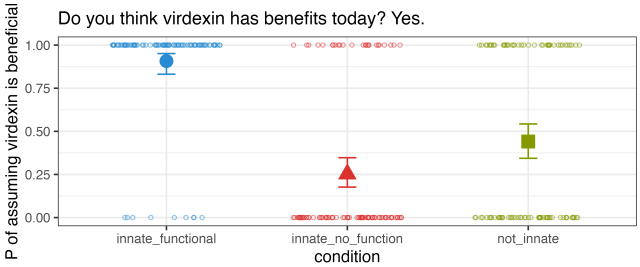

Figure 3: Proportion of participants assuming that *virdexin* has benefits for mothers or infants in today's environment, by historical origin condition. Large shapes are group means with 95% Cis; small circles are individual responses.

All participants then were told (by random assignment) either that *virdexin* has important benefits for infants, important benefits for mothers (in both cases, reducing pain after birth), or no benefits at all. To ensure that they understood the manipulation, participants in the innate conditions who learned that *virdexin* has no benefits today were first told that, due to changes in our lifestyle, traits that worked one way across human history might work differently today.

From *is* to *ought*

To measure how assumptions about historical functions and preserved benefits support normative judgments, we then measured both participants' agreement with a general "ought" claim and their willingness to intervene by restricting the choice of individual mothers to bring about the feature in question. Both showed a consistent pattern of results. To test the willingness to intervene, participants were told:

> Researchers recently concluded that a medication commonly used to treat nausea in pregnant mothers decreases *virdexin* levels. This medication is considered very safe and has been FDA approved for decades. Mothers are currently informed of the risks associated with the mediation. Would you support a national policy preventing doctors from prescribing this medication to pregnant mothers?

Participants responded on a scale from 1 (not at all) to 7 (very much). Support for this policy varied only by a main effect of benefit information (as predicted, $F(2) = 6.00$, $p = .003$). Participants agreed with the policy significantly more if there were benefits for mothers or infants than if there were no benefits, regardless of origin condition.

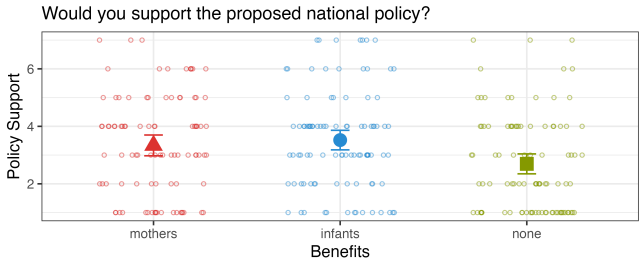

Figure 4: Agreement with the proposed policy, by benefits information. Large shapes are group means with 95% CIs; small circles are individual responses.

Participants then rated their agreement with the statement, "Mothers ought to have *virdexin* in their blood" (from 1 = Strongly disagree to 7 = Strongly agree). Agreement with the general "ought" claim similarly varied by benefits information, as predicted (main effect, $F(2) = 89.09$, $p < .001$). Agreement also varied by origin condition (which we did not predict; main effect, $F(2) = 16.82$, $p < .001$). Participants agreed with the statement the most when *virdexin* was described as innate, functional, and having benefits for mothers or infants in today's environment; they agreed the least when *virdexin* was described as not innate and having no benefits for mothers or infants in today's environment. The origin condition x benefits condition interaction was not significant ($p = .89$; Figure 5).

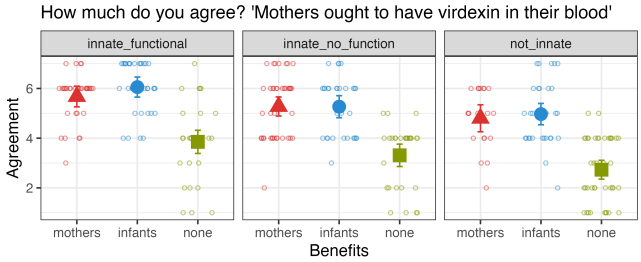

Figure 5: Agreement with the statement that mothers ought to have *virdexin* in their blood, by condition and information about benefits. Large shapes are group means with 95% CIs; small circles are individual responses.

Natural means normal

Finally, we asked participants to rate their agreement with the statement, "It's normal for mothers to have *virdexin* in their blood" (1 = Strongly disagree to 7 = Strongly agree). We predicted that essentialist beliefs would shape conceptions of what's normal or typical for a category, and that an innate feature would be viewed as more normal for the category than the same feature when described as not innate. As predicted, agreement varied only by origin condition, $F(2) = 115.68$, $p < .001$. Participants mostly agreed in both the innate functional condition ($M = 6.23$, 95% CI [6.01, 6.46]) and the innate no function condition ($M = 6.25$, 95% CI [6.02, 6.47]; pairwise contrast, $p = 1$), whereas participants in the not innate condition responded near the middle of the scale (i.e., 4, "neither agree nor disagree," $M = 4.11$, 95% CI [3.87, 4.35]; pairwise comparisons with the other two conditions, both $p < .001$).

Manipulation recall questions

At the end of the study, participants were asked several manipulation recall questions. These questions allowed us to ensure that our experimental manipulations had been interpreted as intended, and to test whether the direction of participants' *mis*understandings might shed additional light on our proposal.

First participants were asked if, in their opinion, they think *virdexin* has other benefits for mothers or infants (if they answered yes, they were asked what kinds of benefits they had in mind). Responses varied only by a main effect of benefit information (X^2 (2) = 185.90, $p < .001$): Participants who had been told that *virdexin* had benefits for mothers or infants were more likely to assume it had other benefits as well (e.g., benefits for fathers).

Participants were also asked if, in their opinion, they thought *virdexin* served some biological function in human evolution (again, if they said yes, they were asked what kind of biological function they had in mind). Responses to this question varied by origin condition (X^2 (2) = 74.35, $p < .001$):

Nearly all participants who had been told that *virdexin* was innate and functional believed that it served a historical function (in line with their condition manipulation; $M = 0.93$, 95% CI [0.85, 0.97]). However, about half of participants who had been told that the feature was innate but *not* functional did so as well ($M = 0.60$, 95% CI [0.50, 0.70]), and these responses were much less common among participants who had been told that it was not innate ($M = 0.34$, 95% CI [0.24, 0.46]; all pairwise contrasts between conditions, ps < .005). Thus, the strong tendency to infer that innate features must be functional caused some participants to assume so even though our condition manipulation explicitly told them otherwise (see Figure 2).

Controlling for origin condition, participants were also more likely to assume that *virdexin* must have served a historical function if it had benefits in today's environment (whether for mothers, $M = 0.75$, 95% CI [0.62, 0.84], or infants, $M = 0.73$, 95% CI [0.61, 0.82]) compared to when it had no benefits ($M = 0.56$, 95% CI [0.41, 0.70]; main effect of benefits information condition, $X^2(2) = 9.45$, $p = .009$; all pairwise comparisons, ps > .10). These results again lend support to our proposed chain of inferences: Some participants assumed that a feature with benefits in today's environment must have served a historical functional as well—despite having explicitly been told that it did not.

Finally, participants were asked if they thought most mothers have *virdexin* because it's biological or in their genes; responses to this question varied only by origin condition (X^2 (2) = 129.19, $p < .001$). Participants mostly agreed when they had been told the feature was innate (both when it was functional, $M = 0.86$, 95% CI [0.77, 0.91], and when it was not functional, $M = 0.72$, 95% CI [0.62, 0.80]; pairwise contrast, $p = .05$), but they mostly disagreed when told it was not innate ($M = 0.12$, 95% CI [0.07, 0.20]; pairwise contrasts with both innate conditions, ps < .001). Our manipulation check questions thus confirm that our experimental manipulations were mostly successful—and when they were not, participants' tendencies to misunderstand were in line with our proposed causal model.

Discussion

In the current study, we found empirical support for our proposal that a missing premise linking gender essentialism and normativity rests on the tendency to assume that common features of biological kinds must be functional. When participants learned that a novel feature of mothers was innate, they overwhelmingly assumed it must have served a historical function. When they learned that it served a historical function, they assumed it must still be beneficial in today's environment. When the feature was beneficial today, participants judged that mothers ought to have it, and they were more willing to intervene to ensure that they would. In this way, the naturalistic fallacy can shape normative social judgments via the explanations people tend to generate about why certain features of natural kind categories are common to begin with.

One implication of this proposal is thus that the naturalistic fallacy may not be as irrational as it first appears—rather than reflecting an *irrational inference*, it could be the product of a rational inference that happens to be operating over potentially mistaken assumptions about the world. These mistaken assumptions include thinking of nature as a goal-oriented agent, thinking of evolution as improvement, and thinking of all biological features as adaptations. Although the conclusions that people arrive at through these reasoning processes may sometimes be mistaken, they are the result of rational reasoning about causal forces; the mistake arises from inaccurate inputs to the rational system (e.g., assuming that a given feature is an adaptation, when it is just a spandrel or the product of culture, or assuming that evolution means improvement). In the case of gender, these potentially mistaken assumptions are further compounded by the essentialist tendency to assume that current gender differences must be due to innate, biological predispositions (Gelman & Taylor, 1996; Prentice & Miller, 2006; Skewes et al., 2018). For example, the fact that more mothers than fathers *are* primary caregivers in modern U.S. society (Craig & Mullen, 2011) might lead people to assume that mothers *ought* to be primary caregivers because (1) the current gender division of labor is "natural" (not necessarily true), (2) processes like natural selection have shaped human social structures over time (a reasonable premise), and (3) natural processes entail improvement over time (an inaccurate representation of nature). When combined, these mistaken assumptions can have dramatic downstream consequences for people's lives. For instance, reasoning that mothers ought to be primary caregivers could lead people to act in ways that could constrain the choices available to individual mothers, such as withholding support for paternal leave policies or public funding for childcare.

Beliefs that people ought to do what is natural for their gender category can perpetuate gender stereotypes in more subtle ways as well. As described by social role theory (Eagly, 1987; Eagly et al., 2000), social agents often internalize social stereotypes by incorporating them into their existing background knowledge and causal-explanatory frameworks, including their knowledge about the natural world. These beliefs about what is natural for different gender categories then shape how people explain, understand, and judge others as well as themselves. For example, beliefs about innate gender-based predispositions can determine which activities people engage in from early childhood (Bian et al., 2017), contributing to gender differences in acquired skills and experiences and perpetuating patterns of gender stratification (Bian et al., 2018; Chestnut et al., 2018; Meyer & Gelman, 2016). Though subtle, these beliefs could have especially pernicious consequences because their normative entailments are never directly stated, only implied, so people may not even be consciously aware of them (Becker & Wright, 2011; Benitez et al., 2024; Haslanger, 2014).

A related question for future research is whether the current effects require people to engage in explicit and intentional causal reasoning about natural processes. Do people infer

ought from *is* by consciously generating explicit functional explanations? Or might they also rely on intuitive assumptions that the way things are is "on purpose"—even without a clear conception of what that purpose is? Generating explicit functional explanations is costly, so when making judgments in daily life people might sometimes rely instead on more all-purpose conceptions of what's normal as a cognitive shortcut (an idea first posited by Aristotle, ca. B.C./1996; see also Bear et al., 2020; Bear & Knobe, 2017; Foster-Hanson & Lombrozo, 2022b; Foster-Hanson & Rhodes, 2019; 2022; Lane, 2020). From this perspective, people might implicitly link essentialism and normativity via a general assumption that what's natural must be good because natural processes entail improvement over time. In support of this notion, participants in the current study viewed innate features as more "normal" for the category, and systematic misunderstandings of our condition manipulations suggest that participants' intuitive assumptions sometimes prevented them from accepting the explicit information provided to them.

This question is also important to help interpret the current findings in light of evidence from developmental psychology. Young children often show more rigid essentialist beliefs about gender categories than the older children and adults in their own communities (Rhodes & Gelman, 2009). Similarly, young children are more likely to make teleological assumptions that the features of natural kinds exist to serve a function (Kelemen, 1999; 2003). But essentialist and teleological tendencies persist throughout the lifespan as intuitive reasoning strategies: When adults are under time pressure or have limited cognitive capacities, they are both more essentialist (Eidson & Coley, 2015) and more likely to endorse teleological explanations (Kelemen & Rossett, 2005; Kelemen et al., 2013; Lombrozo et al., 2007). Future research should directly test whether folk-biological assumptions about feature functions link essentialism and normativity through early-emerging intuitive theories that sometimes—but not always—give rise to explicit functional explanations about biological differences between social categories.

Future research should also test whether adults might sometimes flexibly and strategically endorse essentialist and functional explanations to support their pre-existing social attitudes and judgments. Appealing to nature can serve to justify a wide range of prescriptive social judgments, from views of same-sex marriage (O'Connor, 2017; Rozin, 2005) to gender stereotypes (Brescoll & LaFrance, 2004), and people sometimes endorse essentialist explanations for gender differences strategically to justify their own social motivations (Morton et al., 2009; see also Diesendruck, 2021). Here we suggest that both early-emerging essentialist *and* teleological tendencies could together comprise an intuitive explanatory stance that people might call on strategically to rationalize their social judgments (Cushman, 2020; Greene et al., 2005; Haidt, 2001). Similarly, people might expect members of society to fulfill specific functions as the result of collaborative social interaction, and then come to view those functions as natural and inevitable through processes of system justification (Jost & Banaji, 1994; Jost et al., 2004; Kay et al., 2009). Indeed, Daston (2014) describes the naturalistic fallacy itself as "a kind of covert smuggling operation in which cultural values are transferred to nature and nature's authority is then called upon to buttress those very same values" (Daston, 2014, p. 580). This proposal should be directly tested in future work.

One way in which future work might shed light on these various possible mechanisms (implicit or explicit assumptions about function; evidence-based or motivated reasoning) is by including a wider range of gendered categories. In the current study, we intentionally chose the category "mothers" as a test case both because we expected participants to report high levels of essentialist beliefs about the category at baseline (Park et al., 2015), and because we expected participants to openly and willingly report their normative judgments about the category (Eagly & Karau, 2002; Foster-Hanson & Lombrozo, 2022a; Heilman & Okimoto, 2008; Thomas et al., 2016). Although people engage in functional reasoning about a range of gendered categories, the extent to which they do so varies depending on their background beliefs about each category. For example, while most people agree about the functions of mothers, their beliefs about the functions of boys and girls, or men and women, vary depending on their background ideological commitments, including gender essentialism and benevolent sexism (Foster-Hanson & Lombrozo, 2022a). People might also be more or less willing to openly *report* their essentialist beliefs, functional assumptions, and normative judgments in experimental settings depending on their background beliefs about different categories and their social desirability motivations (Paulhus, 1984; Rudman & Kilianski, 2000). Future work might make use of these variations across categories to test the many open questions described above.

Evolutionary explanations for gender differences have been the topic of heated debate among researchers in recent decades (e.g., Archer, 1996; Buss, 1995; Tooby & Cosmides, 1992). At the heart of this debate lie fundamental disagreements about the extent to which gender differences are the product of adaptation via natural selection, or of culture and socialization, as well as disagreements about whether evolutionary explanations inevitably entail a normative justification of gender inequality (Eagly & Wood, 1999). Based on the current work, one possible source of these conflicts may be the fact that these very same explanations permeate intuitive assumptions about gender, and that they (and their normative entailments) stem from common—if often mistaken—intuitive theories about the natural world.

Acknowledgments

We are grateful to members of the Princeton University Concepts and Cognition Lab, especially Casey Lewry, and the University Center for Human Values for feedback on this project.

References

Allport GW. *The Nature of Prejudice*. Cambridge, MA: Perseus Publishing; 1954.

Archer, J. (1996). Sex differences in social behavior: Are the social role and evolutionary explanations compatible?. *American Psychologist*, *51*(9), 909.

Aristotle (1996). *Physics (*R. Waterfield & D. Bostock, Trans.). Oxford University Press. (Original work published ca. 350 B.C.).

Bach, T. (2012). Gender is a natural kind with a historical essence. *Ethics*, *122*(2), 231-272.

Bear, A., Bensinger, S., Jara-Ettinger, J., Knobe, J., & Cushman, F. (2020). What comes to mind?. *Cognition*, *194*, 104057. *https://doi.org/10.1016/j.cognition.2019.104057*

Bear, A., & Knobe, J. (2017). Normality: Part descriptive, part prescriptive. *Cognition*, *167*, 25-37. *https://doi.org/10.1016/j.cognition.2016.10.024*

Becker, J. C., & Wright, S. C. (2011). Yet another dark side of chivalry: Benevolent sexism undermines and hostile sexism motivates collective action for social change. *Journal of Personality and Social Psychology*, *101*(1), 62.

Benitez, J., Foster-Hanson, E., & Rhodes, M. (2024, April 24). Generic references to gender predict essentialism even when they express counter-stereotypic ideas. *https://doi.org/10.31234/osf.io/yprb5*

Bian, L., Leslie, S. J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*(6323), 389-391.

Bian, L., Leslie, S. J., Murphy, M. C., & Cimpian, A. (2018). Messages about brilliance undermine women's interest in educational and professional opportunities. *Journal of Experimental Social Psychology*, *76*, 404-420.

Black, M. (1964). The gap between "Is" and "Should." *Philosophical Review*, *73*, 165–181.

Blancke, S., Schellens, T., Soetaert, R., Van Keer, H., & Braeckman, J. (2014). From ends to causes (and back again) by metaphor: The paradox of natural selection. *Science & Education*, *23*(4), 793-808. *https://doi.org/10.1007/s11191-013-9648-8*

Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition*, *60*(1), 1-29.

Brescoll, V., & LaFrance, M. (2004). The correlates and consequences of newspaper reports of research on sex differences. *Psychological Science*, *15*, 515–520.

Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review*, *100*, 204–232.

Chestnut, E. K., Lei, R. F., Leslie, S. J., & Cimpian, A. (2018). The myth that only brilliant people are good at math and its implications for diversity. *Education Sciences*, *8*(2), 65.

Coleman, J. M., & Hong, Y.-Y. (2008). Beyond nature and nurture: The influence of lay gender theories on self-stereotyping. *Self and Identity, 7*(1), 34–53.

Coley, J. D., & Tanner, K. (2015). Relations between intuitive biological thinking and biological misconceptions in biology majors and nonmajors. *CBE Life Sciences Education, 14*(1), ar8. https://doi.org/10.1187/cbe.14-06-0094

Craig, L., & Mullan, K. (2011). How mothers and fathers share childcare: A cross-national time-use comparison. *American Sociological Review*, *76*(6), 834-861.

Cushman, F. (2020). Rationalization is rational. *Behavioral and Brain Sciences*, *43*, e28.

Dar-Nimrod, I., Cheung, B. Y., Ruby, M. B., & Heine, S. J. (2014). Can merely learning about obesity genes affect eating behavior?. *Appetite*, *81*, 269-276.

Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to Scientific Theories Affects Women's Math Performance. *Science, 314*(5798), 435.Keller, J. (2005). In Genes We Trust: The Biological Component of Psychological Essentialism and Its Relationship to Mechanisms of Motivated Social Cognition. *Journal of Personality and Social Psychology, 88*(4), 686–702.

Daston, L. (2014). The naturalistic fallacy is modern. *Isis*, *105*(3), 579-587.

Diesendruck, G. (2021). A Motivational Perspective on the Development of Social Essentialism. *Current Directions in Psychological Science*, *30*(1), 76-81.

Eagly, A. H. (1987). Sex differences in social behavior: A social-role analysis. Hillsdale, NJ: Erlbaum

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological review*, *109*(3), 573.

Eagly, A. H., & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist*, *54*(6), 408.

Eagly, A. H., Wood, W., & Diekman, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. *The Developmental Social Psychology of Gender*, *12*, 174.

Eidson, R. C., & Coley, J. D. (2014). Not so fast: Reassessing gender essentialism in young adults. *Journal of Cognition and Development*, *15*(2), 382-392.

Evans, E. M. (2001). Cognitive and contextual factors in the emergence of diverse belief systems: Creation versus evolution. *Cognitive Psychology, 42*(3), 217-266.

Foster-Hanson, E., & Lombrozo, T. (2022a). What are men and mothers for? The causes and consequences of functional reasoning about social categories. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 44, No. 44).

Foster-Hanson, E., & Lombrozo, T. (2022b). How "is" shapes "ought" for folk-biological concepts. *Cognitive Psychology*, *139*, 101507.

Foster-Hanson, E., & Rhodes, M. (2019). Is the most representative skunk the average or the stinkiest? Developmental changes in representations of biological categories. *Cognitive Psychology*, *110*, 1-15. https://doi.org/10.1016/j.cogpsych.2018.12.004

Foster-Hanson, E., & Rhodes, M. (2023). Stereotypes as prototypes in children's gender concepts. *Developmental science*, *26*(4), e13345.

Friedrich, J., Kierniesky, N., & Cardon, L. (1989). Drawing moral inferences from descriptive science: The impact of attitudes on naturalistic fallacy errors. *Personality and Social Psychology Bulletin*, *15*(3), 414-425.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford Cognitive Development.

Gelman, S. A., Collman, P., & Maccoby, E. E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development*, 396-404.

Gregory, T. R., & Ellis, C. A. (2009). Conceptions of evolution among science graduate students. *BioScience*, *59*(9), 792-799. *https://doi.org/10.1525/bio.2009.59.9.10*

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105-2108.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review*, *108*(4), 814.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, *39*(1), 113-127.

Haslanger, S. (2014). The normal, the natural and the good: Generics and ideology. *Politica & Società*, *3*(3), 365-392.

Heilman, M. E., & Okimoto, T. G. (2008). Motherhood: a potential source of bias in employment decisions. *Journal of Applied Psychology*, *93*(1), 189.

Hudson, W. D. (1969). *Is-Ought Question*. Springer.

Hume, D. (2000). *A treatise of human nature* (D. F. Norton & M. J. Norton, Eds.). New York, NY: Oxford University Press. (Original work published 1740).

Ismail, I., Martens, A., Landau, M. J., Greenberg, J., & Weise, D. R. (2012). Exploring the effects of the naturalistic fallacy: Evidence that genetic explanations increase the acceptability of killing and male promiscuity. *Journal of Applied Social Psychology*, 42(3), 735-750.

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

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, 33(1), 1-27.

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, 25(6), 881-919.

Kay, A. C., Gaucher, D., Peach, J. M., Laurin, K., Friesen, J., Zanna, M. P., & Spencer, S. J. (2009). Inequality, discrimination, and the power of the status quo: Direct evidence for a motivation to see the way things are as the way they should be. *Journal of Personality and Social Psychology*, 97(3), 421–434.

Keil, F. C. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In J. Tooby, A. M. Leslie, D. Sperber, A. Caramazza, A. E. Hillis, E. C. Leek, M. Miozzo, & L. Cosmides (Eds.), *Mapping the mind: Domain specificity in cognition and culture*, (pp. 234-254). Cambridge University Press.

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

Kelemen, D. (2012). Teleological minds: How natural intuitions about agency and purpose influence learning about evolution. In K. S. Rosengren, S. K. Brem, E. M. Evans & G. M. Sinatra (Eds.), *Evolution challenges: Integrating research and practice in teaching and learning about evolution,* (pp. 66-92). Oxford: Oxford University Press. https://doi. Org/10.1093/acprof:oso/9780199730421.003.0004

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

Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General*, *142*(4), 1074–1083. https://doi.org/10.1037/a0030399

Keller, J. (2005). In genes we trust: the biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology*, *88*(4), 686.

Kenrick, D. T. (1994). Evolutionary social psychology: From sexual selection to social cognition. In *Advances in Experimental Social Psychology* (Vol. 26, pp. 75-121). Academic Press.

Kierniesky, N. C., & Sobus, M. (1989). The naturalistic fallacy: Moral inferences drawn from research with children versus adults. *Psychological Reports*, *65*(2), 475-479. https://doi.org/10.2466/pr0.1989.65.2.475

Knobe, J., Prasada, S., & Newman, G. E. (2013). Dual character concepts and the normative dimension of conceptual representation. *Cognition*, *127*(2), 242-257. *https://doi.org/10.1016/j.cognition.2013.01.005*

Liquin, E. G., & Lombrozo, T. (2018). Structure-function fit underlies the evaluation of teleological explanations. *Cognitive Psychology*, *107*, 22-43.

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition*, 99(2), 167-204.

Lombrozo, T., Kelemen, D., & Zaitchik, D. (2007). Inferring design: Evidence of a preference for teleological explanations in patients with Alzheimer's disease. *Psychological Science*, *18*(11), 999–1006. https://doi.org/10.1111/j.1467-9280.2007.02015.x

Lombrozo, T., & Rehder, B. (2012). Functions in biological kind classification. *Cognitive Psychology*, *65*(4), 457-485.

Lombrozo, T., Shtulman, A., & Weisberg, M. (2006). The Intelligent Design Controversy: Lesson From Psychology and Education. *Trends in Cognitive Sciences*, 10(2), 56.

Lombrozo, T. & Wilkenfeld, D. (2019). Mechanistic versus functional understanding. In Stephen Grimm (Ed.0, *Varieties of Understanding: New Perspectives from Philosophy, Psychology, and Theology* (pp. 209-229). Oxford University Press.

Maccoby, E. E. (1988). Gender as a social category. *Developmental Psychology*, *24*(6), 755.

Martin, C. L., & Ruble, D. (2004). Children's search for gender cues: Cognitive perspectives on gender development. *Current Directions in Psychological Science*, *13*(2), 67-70.

Martin, C. L., & Parker, S. (1995). Folk theories about sex and race differences. *Personality and Social Psychology Bulletin, 21*(1), 45-57.

Mayr, E. (1982). The growth of biological thought: Diversity, evolution, and inheritance. Cambridge, MA: Harvard University Press.

McConnell-Ginet, S. (2013). Gender and its relation to sex: The myth of 'natural' gender. *The Expression of Gender*, 3-38.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. *Similarity and Analogical Reasoning*, *179*, 195.

Meyer, M., & Gelman, S. A. (2016). Gender essentialism in children and parents: Implications for the development of gender stereotyping and gender-typed preferences. *Sex Roles*, *75*, 409-421.

Monterosso, J., Royzman, E. B., & Schwartz, B. (2005). Explaining away responsibility: Effects of scientific explanation on perceived culpability. *Ethics and Behavior*, *15*(2), 139.

Moore, G. E. (2004). *Principia ethica*. Mineola, NY: Dover Publications. (Original work published 1903)

Moore, R., Mitchell, G., Bally, R., Inglis, M., Day, J., & Jacobs, D. (2002). Undergraduates' understanding of evolution: ascriptions of agency as a problem for student learning. *Journal of Biological Education*, *36*(2), 65-71. *https://doi.org/10.1080/00219266.2002.9655803*

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*(3), 653.

O'Connor, C. (2017). 'Appeals to nature' in marriage equality debates: A content analysis of newspaper and social media discourse. *British Journal of Social Psychology*, *56*(3), 493-514.

Park, B., Banchefsky, S., & Reynolds, E. B. (2015). Psychological essentialism, gender, and parenthood: Physical transformation leads to heightened essentialist conceptions. *Journal of Personality and Social Psychology*, *109*(6), 949.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *46*(3), 598.

Prentice, D. A., & Miller, D. T. (2006). Essentializing differences between women and men. *Psychological Science*, *17*(2), 129-135.

Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, *59*(3), 244-274.

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

Rips, L. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*. Cambridge: Cambridge University Press.

Roberts, A. J., Handley, S. J., & Polito, V. (2021). The design stance, intentional stance, and teleological beliefs about biological and non-biological natural entites. *Journal of Personality and Social Psychology*, *120*(6), 1720–1748.

Rozin, P. (2005). The Meaning of "Natural": Process more important than content. *Psychological Science*, 16, 652–658. *https://doi.org/10.1111/j.1467-9280.2005.01589.x*

Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin*, *26*(11), 1315-1328.

Rutjens, B. T., Sutton, R. M., & van der Lee, R. (2018). Not all skepticism is equal: Exploring the ideological antecedents of science acceptance and rejection. *Personality and Social Psychology Bulletin*, *44*(3), 384-405. *https://doi.org/10.1177/0146167217741314*

Scott, S. E., & Rozin, P. (2020). Actually, natural is neutral. *Nature Human Behaviour*, *4*(10), 989-990. *https://doi.org/10.1038/s41562-020-0891-0*

Shtulman, A. (2006). Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology*, 52(2), 170-194.

Shtulman, A. (2017). Scienceblind: why our intuitive theories about the world are so often wrong. New York: Basic Books.

Skewes, L., Fine, C., & Haslam, N. (2018). Beyond Mars and Venus: The role of gender essentialism in support for gender inequality and backlash. *PloS one*, *13*(7), e0200921.

Taylor, M. G. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development*, *67*(4), 1555-1571.

Thomas, A. J., Stanford, P. K., & Sarnecka, B. W. (2016). No child left alone: Moral judgments about parents affect estimates of risk to children. *Collabra*, *2*(1), 10.

Tooby J, Cosmides L. 1992. The psychological foundations of culture. In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, ed. JH Barkow, L Cosmides, J Tooby. New York: Oxford Univ. Press

Ware, E. A., & Gelman, S. A. (2014). You get what you need: An examination of purpose-based inheritance reasoning in

undergraduates, preschoolers, and biological experts. *Cognitive Science*, 38(2), 197-243.

Wilton, L. S., Bell, A. N., Carpinella, C. M., Young, D. M., Meyers, C., & Clapham, R. (2019). Lay theories of gender influence support for women and transgender people's legal rights. *Social Psychological and Personality Science, 10*(7), 883-894.

Wood, W., & Eagly, A.H. (2010). Gender. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (Vol. 1, 5th ed., pp. 629–667). Hoboken: Wiley.

Vasilyeva*, N., Blanchard*, T.J., & Lombrozo, T. (2018). Stable causal relations are better causal relations. *Cognitive Science, 42*(4), 1265-1296. *Equal contributions