Received: 29 April 2022 | Revised: 1 November 2022 | Accepted: 3 November 2022

DOI: 10.1111/desc.13345

Developmental Science WILEY

SHORT REPORT

# Stereotypes as prototypes in children's gender concepts

Emily Foster-Hanson[1] | Marjorie Rhodes[2]

[1]Department of Psychology, Princeton University, Princeton, New Jersey, USA

[2]Department of Psychology, New York University, New York City, New York, USA

**Correspondence**
Emily Foster-Hanson, Department of Psychology, Princeton University, Princeton, NJ, USA.
Email: emily.fosterhanson@princeton.edu

## Abstract

How do gender stereotypes shape prototypes across development? In the current pre-registered study with children ages 3- to 10-years-old and adults ($N = 257$), participants made judgements about which members of gender categories (boys and girls) and animal categories (for comparison) were the most representative and informative about their kinds, using simplified scales of five category members varying on a stereotypical feature (e.g., girls wearing more or less pink). Young children chose boys and girls with extreme stereotypical features (e.g., the girl in head-to-toe pink) as both representative and informative of their categories and this tendency declined with age, similar to developmental patterns in prototypes of animal categories. Controlling for age, children whose parents reported more conservative social-political views also held more extreme gender (but not animal) prototypes. Thus, stereotypes play a central role in children's gender prototypes, especially young children and those living in socially-conservative households.

KEYWORDS
conceptual development, gender, prototypes, social cognition, stereotypes, variability

### Research Highlights

- Stereotypes play a central role in children's gender prototypes, especially young children and those in socially-conservative households.
- Children ages 3–10 and adults chose which girls, boys, and animals were most representative and informative.
- Younger children chose category members with more extreme stereotypical features (e.g., the girl in head-to-toe pink) than older children and adults.
- Children with more conservative parents also held more extreme gender prototypes.

## 1 | INTRODUCTION

Imagine you meet a friendly alien who is fascinated by earthlings and asks for your help to learn about them. The alien says it has heard of something called "girls" and asks to see an example. Who would you show the alien to represent what girls in general are like? Would you choose a girl with the most extreme stereotypical features associated with her gender (e.g., one in head-to-toe pink), or perhaps a girl with more average stereotypical features (e.g., one in an average amount of pink for girls)? For most adults, the answer is straightforward—category members with more average features are more prototypical because they better represent within-category variability (Rosch & Mervis, 1975). But what if the alien asked a 4-year-old? In the current work, we asked how gender prototypes (beliefs about which

  

category members are most representative and informative of their kinds) vary across childhood development. We predicted that young children would hold exaggerated prototypes focusing on extreme values of stereotyped features, and that prototypes would better reflect within-category variation with age.

Prototypes determine how people use categories to understand the world—prototypical category members (e.g., a robin for the category *birds*) are learned first (Bjorklund & Thompson, 1983), categorized more quickly (Rosch & Mervis, 1975), come to mind more easily (Anglin, 1986), and provide the basis for generalization (Osherson et al., 1990). Whereas stereotypes are beliefs about which specific *features* are associated with social categories (Ellemers, 2018; Hilton & von Hippel, 1996), prototypes are summary representations of what the categories themselves are like in general, including beliefs about all of the features associated with the category as well as how much they vary (Medin et al., 1987). For adults, stereotypes and prototypes are separable: Adults might endorse the stereotype that girls wear pink (and even think girls *should* wear pink) but understand that actual girls vary in how much pink they wear (Bodenhausen & Macrae, 1998; Kunda & Spencer, 2003). Adults' prototypes often reflect this variability by centering around averages rather than extreme exemplars (Kim & Murphy, 2011).

But there are several reasons to expect that stereotypes and prototypes are more closely related for young children. In early-developing representations of animal categories, children tend to have extreme prototypes that ignore within-category variation (Foster-Hanson & Rhodes, 2019a). For example, 5-year-old children think that the most representative and informative cheetah (i.e., the prototypical cheetah) is the one that runs the very fastest, with a gradual shift across childhood toward more average prototypes (e.g., the average-speeded cheetah; Foster-Hanson & Rhodes, 2019a). Children select extreme prototypes for animal categories even though they *recognize* that there is within-category variability (Rhodes & Liebenson, 2015). Across domains, young children often conflate ideas about how things are with how they should be (Kalish, 1998; Rakoczy & Schmidt, 2013; Shtulman & Phillips, 2018), suggesting perhaps that extreme prototypes stem from beliefs that category members *should* display those features. Consistent with this interpretation, children view average rather than extreme exemplars as prototypical for novel categories when they do not have any prior beliefs about which types of features members of a category are "supposed" to have (Foster-Hanson & Rhodes, 2019a, Study 3).

Gender is an early-emerging social category (Maccoby, 1988; Martin & Ruble, 2004), and children begin acquiring gender stereotypes by age 2 (Poulin-Dubois et al., 2002). Therefore, if early gender concepts are structured similarly to animal concepts (Rhodes & Gelman, 2009), children's early awareness of gender stereotypes might lead them to hold similarly narrow, idealized gender prototypes (e.g., thinking the most prototypical girl is the one wearing head-to-toe pink). Alternatively, children's greater familiarity and personal experience with gender categories (versus animals) might lead them to attend to within-category variability earlier in development, so this question is important to test directly.

Extreme stereotypical gender prototypes could have important consequences for children's lives. For example, idealized animal prototypes lead children to seek out narrow, idealized samples of evidence to learn from (e.g., examining just the two fastest cheetahs to learn about cheetahs in general), with a shift across childhood toward preferring diverse samples that account for within-category variability (e.g., faster and slower cheetahs, since actual cheetahs run different speeds; Foster-Hanson et al., 2020). In the context of gender, ignoring within-category variability could give rise to prejudice, including backlash against gender nonconformity (Okimoto & Brescoll, 2010) and the marginalization of some members of society (Purdie-Vaughns & Eibach, 2008). Relying on gender stereotypes to infer what people are like could also limit children's long-term opportunities (Fiske & Taylor, 2013), like stereotypes about intelligence, which emerge in childhood and impede equal representation in science, technology, engineering, mathematics (STEM fields; Bian et al., 2017). Understanding the role of stereotypes in children's gender concepts is crucial for mitigating these potentially harmful effects.

The current study measured gender prototypes and stereotypes in 3- to 10-year-old children and adults, using the methods developed by Foster-Hanson and Rhodes (2019a). Participants were introduced to an array of five exemplars that varied in their level of a stereotypical feature (e.g., girls wearing more or less pink; boys with shorter or longer hair). To provide a direct measure of prototypicality, some participants made explicit judgements about which exemplar was most representative of their gender categories. To test how explicit endorsement of gender stereotypes (e.g., thinking it's *better* for girls to wear pink) relates to gender prototypes across age, other participants instead chose which category member is "best," (following Foster-Hanson & Rhodes, 2019a). In addition, to provide converging evidence of prototypicality beliefs, participants also completed a secondary, less direct, measure of prototypicality, in which they selected which exemplars were most informative about the kind. For comparison, participants also made judgments about animal categories (as in Foster-Hanson & Rhodes, 2019a).

Although we expected general developmental changes in children's gender concepts, we also suspected that in the social domain, perhaps more so than for animals, the development of prototypes may vary based on social experiences. Children rely on the adults in their communities to learn about the social world (Harris et al., 2018), and adults' gender stereotyping is correlated with conservatism (Stern & Axt, 2021), so as a proxy for variation in children's social experiences we asked parents to report their social-political views at the end of the study to test how parental views related to individual variation in children's gender prototypes.

## 2 | METHOD

### 2.1 | Participants

We recruited 3- to 10-year-old children, grouped by age (3–4, 5–6, 7–8, 9–10) for even recruitment distribution. We stopped data collection

the week we obtained at least 48 children per age group, per our preregistered sampling plan (https://osf.io/krxzs); 223 children completed the study via an online platform for developmental research. Families received a $10 Amazon gift-card for participating. As preregistered, we excluded four children for failing all training or practice questions. Our final sample of 219 children ($M_{age}$ = 6.84; 99 male, 120 female) was 71% White, 3% Black, 6% Asian, and 19% multiracial; 11% Hispanic of any race. Participants were located across 170 different zip codes in the U.S., U.K., and Canada.

We also intended to include 48 adult participants for comparison; we recruited 56 adults through Amazon Mechanical Turk, expecting to exclude 15% based on previous studies in the lab. We excluded two adults for failing practice and attention check questions, and five for failing Winograd schema questions designed to detect bots (Levesque et al., 2012); one participant did not finish the study. This left 48 adults ($M_{age}$ = 39; 24 male, 24 female; 88% White, 4% Black, 4% multiracial; 12% Hispanic of any race).

## 2.2 | Procedure

First, participants completed a training session to practice with the 5-item scales. They saw five glasses with different amounts of water, and they clicked on the glass containing the least, most, and middle amounts. Then, half of participants were randomly assigned to make explicit judgments about which category member was most representative, by picking one to put in a book to teach an alien about the category (representative condition; Foster-Hanson & Rhodes, 2019a; Lei et al., 2021; Rhodes et al., 2008). The other half were randomly assigned to make explicit judgments about which category member should get a prize for being the "best" of its kind, as a measure of stereotype endorsement (best condition; Foster-Hanson & Rhodes, 2019a). Children were randomly assigned to make either explicit judgments about representativeness or about stereotype endorsement because we suspected that explicit stereotype endorsement responses might affect explicit prototype judgments if children had been asked to make both kinds of judgments in the same experimental session, particularly since both types of items used identical response formats.

In addition, to provide converging evidence of prototypicality beliefs and to test the possibility that extreme prototypes might lead to biased social learning because people choose to learn from prototypical examples (Osherson et al., 1990), participants were also asked to select which category member they saw as most informative. On this informativeness task, participants chose whether an unseen category member would share a novel property with the average or most extreme category member (Figure 1). This indirect measure of prototypes used a different structure than our explicit measures so we did not expect responses on this measure to be influenced by the explicit measure that preceded it, therefore we included it in both conditions to maximize data collection. Participants completed three blocks (girls, boys, animals) in counterbalanced order. Each block contained two trials showing different properties (e.g., clothing color, hair length); each trial included a representativeness or "best" choice (depending on condition) and an informativeness question (all conditions). Participants answered memory check and condition-specific practice questions to ensure they understood the procedure (for details see the SOM, https://osf.io/e7bvq/); each age group answered over 80% correctly, meaning the tasks were accessible even to the youngest participants.

The study was presented using Qualtrics as videos with verbally-narrated questions; participants clicked "buttons" to indicate their responses (parents clicked if children could not). Children participated on their home computers using unmoderated remote research (Rhodes et al., 2020); their computer's webcam recorded the study session and all videos were checked to ensure both parent and child were present and agreed to participate. During the consent procedure, parents were informed that the study was about the categories "boys" and "girls" and saw an example of the gender stimuli and questions; families were encouraged to participate in a different study if they felt the topic was sensitive for their children. As preregistered, a random 20% of videos were coded for interference using Datavyu (Datavyu Team, 2014); we identified only one instance of parental interference, so all data were retained for analysis (for coding details and explanation of this approach, see the SOM). At the end of the study, parents filled out a demographics questionnaire which included reporting their social-political views (7 = very liberal; 1 = very conservative; reverse-coded to measure social conservatism). Adults completed the same study as children; they also reported their social-political views, but a power analysis indicated that our adult sample provided insufficient power to test individual variation so we do not discuss adults' social-political views further.

We analyzed participants' responses on the 5-point scales using ordinal logistic regression (Christensen, 2019); means are average category members selected on the 1-5 scales with 95% CIs, higher numbers indicating more extreme features. We analyzed informativeness responses (binary choices of which feature to generalize) using binomial generalized linear mixed models (Kuznetsova et al., 2017); means are the probability of choosing the extreme (rather than average) category member with 95% confidence intervals. For all models, we analyzed effects of age categorically including adults, and continuously for children's data separately. We also tested main and interactive effects of category (girls, boys, animals), with random intercepts for participants, trials, and block order. In planned exploratory analyses, we tested children's responses for effects of participant gender (see the SOM) and parental conservatism. We report the results of Likelihood Ratio Tests for all measures; for more details see the SOM. Data, analysis code, and a sample video of a full study session are available at https://osf.io/e7bvq/.

# 3 | RESULTS

Category prototypes varied by age, with younger children choosing more extreme category members on both our explicit measure of representativeness ($X^2$ (4) = 22.71, $p < .001$) and our indirect measure of informativeness ($X^2$ (4) = 43.66, $p < .001$), similarly for


**FIGURE 1** Method. Participants completed three category blocks with two trials each (girls block: amount of pink, length of hair; boys block: amount of blue, length of hair; animals block: running speed for cheetahs; jumping height for kangaroos). We measured prototypes explicitly by asking participants to choose the most representative category member to put in a book to teach someone (half of participants). We measured prototypes indirectly by asking which category member is most informative, by asking participants to choose which an unseen category member would share a novel property with (the average, or the most extreme; all participants). We also measured explicit stereotype endorsement by asking which category member should get a prize for being the best (half of participants). The photo shows a 5-year-old participant in the representative condition (image reproduced here with the parent's consent); a video of this participant completing a full study trial is available on OSF, https://osf.io/e7bvq/; the complete video of this participant's session is available to authorized researchers on Databrary, https://nyu.databrary.org/volume/1549

animal and social categories (Figure 2). Analysis of children's data, with age included continuously, showed the same age-related change (representativeness, $X^2$ (1) = 19.73, $p$ < .001; informativeness, $X^2$ (1) = 20.10, $p$ < .001). There were no main or interactive effects of category.

Explicit judgments about which category members were "best" varied by an age by category interaction ($X^2$ (8) = 23.27, $p$ = .003; children only, $X^2$ (2) = 15.94, $p$ < .001): "Best" judgments were consistent across age about animals (simple slope = .01, $p$ = .81) and boys (simple slope = -0.04, $p$ = .34) but declined with age about girls (simple slope = -0.17, $p$ < .001). As shown in Figure 2, the different patterns across age reflect that the youngest children (ages 3–4) selected more extreme exemplars for girls than boys; we suspected this difference might reflect that a feature selected for the boy category—hair length—was not a salient stereotype about boys for the youngest children. Indeed, with boys' hair length trials removed, "best" judgments also declined with age for boys (simple slope = -0.11, $p$ = .02). Therefore, overall, explicit judgments that it's *better* for category members to display extreme features declined with age for gender but not animal categories. In our preregistered full model with only age and condition (representative, best) as predictors, there were main effects of age ($X^2$ (4) = 24.44, $p$ < .001) and condition ($X^2$ (1) = 21.10, $p$ < .001; reflecting higher ratings in the best condition) but no interaction ($X^2$ (4) = 5.86, $p$ = .21).

### 3.1 | Individual variation and parental conservatism

Children learn about the social world largely from adults, so we suspected that children's social prototypes might also vary in line with their parents' social-political views. To test this, we ran exploratory analyses of children's responses with their parents' self-reported social-political conservatism as a predictor variable. Controlling for age, children of more socially-conservative parents chose more extreme stereotypical category members as representative of gender categories ($X^2$ (1) = 5.20, $p$ = .02) but not animal categories ($X^2$ (1) = 0.02, $p$ = .89; Figure 3). "Best" and informativeness responses did not vary across parental conservatism.

## 4 | DISCUSSION

The current study revealed developmental change in gender prototypes. Younger children chose more extreme stereotypical "boys" and "girls" as representative and informative of their categories, with a gradual shift across childhood toward choosing more average category members. This pattern of developmental change in gender prototypes mirrors the pattern for animal prototypes documented in previous work (Foster-Hanson & Rhodes, 2019a), suggesting it reflects a

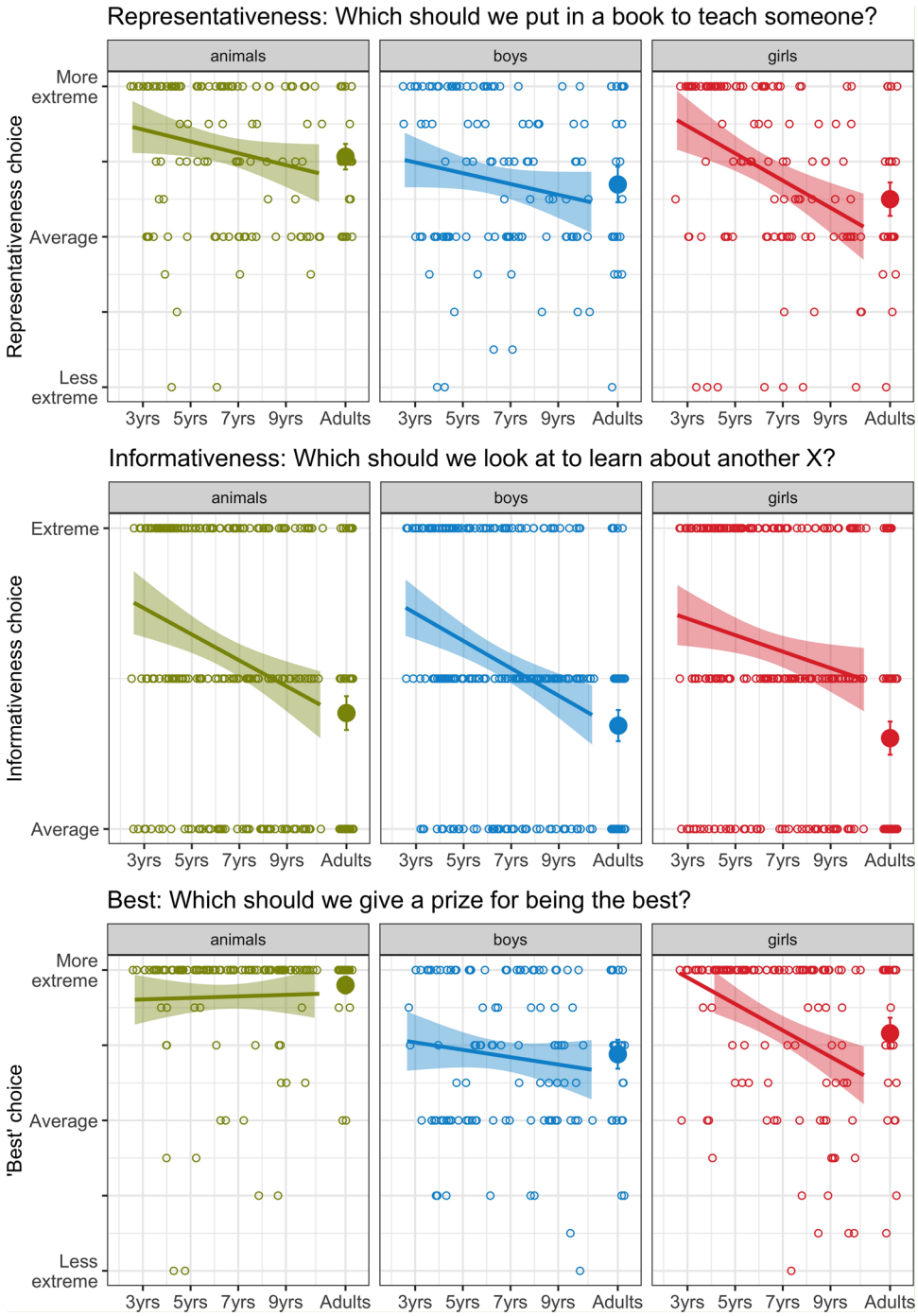


**FIGURE 2** Responses by age and category. Children's responses are included with age as a continuous variable, with lines and shaded bands showing group means with 95% Confidence Intervals. Adults' responses are included categorically, with large circles and error bars showing group means with 95% Confidence Intervals. Small circles show individual participants' responses averaged across the two trials per category block. For representativeness choices, participants chose which category member on a five-item scale they would put in a book to teach an alien about the category. For informativeness choices, participants chose whether an unseen category member would share a novel property with the average or extreme category member. For "best" choices, participants chose which category member on the five-item scale should get a prize for being the "best" of its kind.

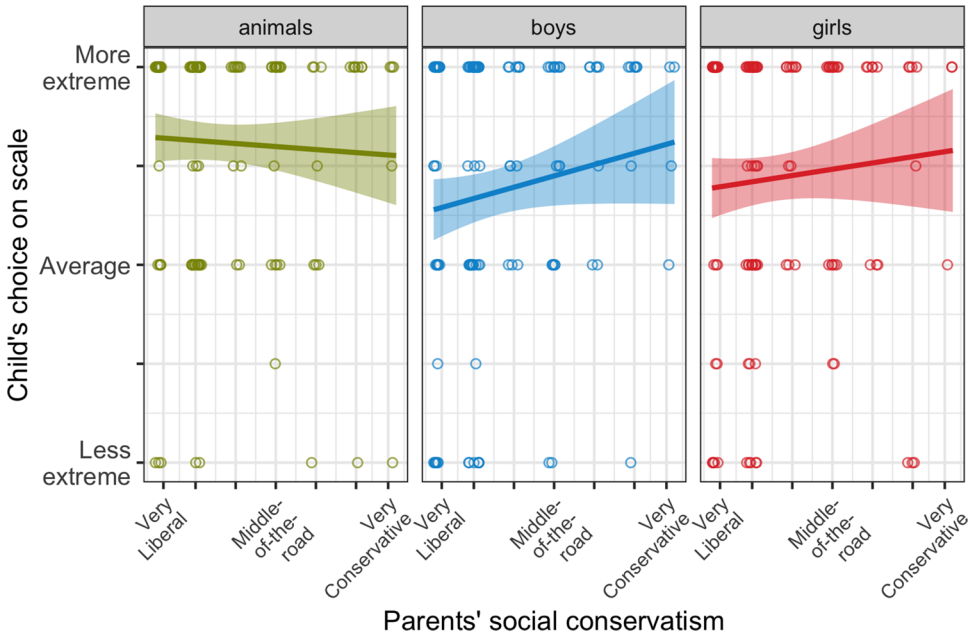


**FIGURE 3** Children's prototypes (i.e., representativeness choices) by their parents' social-political beliefs. Parents rated their beliefs on social issues on a seven-point scale, from very liberal to very conservative. Large lines and shaded bars show group means with 95% Confidence Intervals, and small circles show individual participants' average responses per category block

feature of children's general conceptual structure—with young children's prototypes characterized by a narrow focus on idealized category members (Foster-Hanson et al., 2020; Mervis & Pani, 1980), and a protracted developmental shift towards prototypes that capture within-category variability.

Why might children hold narrow, idealized prototypes of gender and animal categories? One possibility is that children's prototypes are shaped by their beliefs that category members *should* display stereotypical features, and they struggle to dissociate ideas about how things *are* with how they *should be* (Foster-Hanson & Rhodes, 2019b; Kalish, 1998; Rakoczy & Schmidt, 2013; Shtulman & Phillips, 2018). It is also possible that children *can* differentiate descriptive and prescriptive category information, but narrowly focusing on characteristic features is just less cognitively demanding than attending to within-category variability (Kruschke, 2001). If so, young children might indeed be choosing representative category members that they view as average, but they are mistakenly assuming that category members generally share stereotypical features (so the actual variation is clustered around extreme stereotypical features). In support of this interpretation, when idealized features are highly salient, adults' also view them as more typical (Foster-Hanson & Rhodes, 2019a, Study 3) and also more common (Foster-Hanson & Lombrozo, 2022). Future research should directly test these different interpretations.

Children might also focus on extreme, idealized category members as a consequence of essentialist beliefs (Medin & Ortony, 1989). Children develop essentialist beliefs about gender and animal categories by age 3 (Gelman, 2003; Rhodes & Gelman, 2009); these beliefs include expecting category members to share features due to a common causal "essence" (Gelman & Markman, 1986), often leading people to ignore category variability (Emmons & Kelemen, 2015). Relatedly, generic statements (e.g., "girls wear pink") are frequent in child-directed speech and correlate with essentialist beliefs (Gelman et al., 2004). These statements can be both descriptive and prescriptive (e.g., meaning both that girls do wear pink, and that they *should*; Foster-Hanson et al., 2019; Leslie, 2015), thus contributing to both essentialism and idealized prototypes.

In addition to general patterns of developmental change in gender prototypes, we also found individual variation, across age, related to variable social experiences as measured by parental conservatism. Gender essentialism is correlated with social conservatism (Skewes et al., 2018) as well as the production of generic language, so generics could be one mechanism through which parental views shape children's prototypes. However, other forms of cultural input could also give rise to idealized prototypes (e.g., praise for gender-typed behavior; Fagot, 1978). Future research should directly examine how cultural and social experience shapes prototypes.

For animal categories, children selected extreme examples (e.g., the fastest cheetah) as the "best" across age, whereas their prototypes shifted across age to more average exemplars (consistent with Foster-Hanson & Rhodes, 2019a). In contrast, judgments about who is the "best" boy or girl showed a more similar pattern to gender prototypes—both shifted towards more average values across age, suggesting declining endorsement of explicit gender stereotypes across childhood (Martin & Ruble, 2004). One interpretation of this pattern is that gender prototypes and stereotypes remain tightly linked across childhood—more than for representations of animal categories—both shifting to incorporate more variation with age (even into adolescence; Katz & Ksansnak, 1994). Alternately, we may have underestimated the strength of older children's and adults' gender stereotypes by using such an explicit task (Greenwald & Banaji, 1995), which could

have elicited social desirability concerns (Paulhus & Vazire, 2007). If so, gender stereotypes and prototypes may become more differentiated across age than suggested here. Additionally, the features used in the current study (pink and blue, hair length) were selected because they were salient for even 3-year-old children in a pretest, but stereotypes about other features might become more salient and explicit with age. Explicit stereotype endorsement might also be higher across age for features people view as functional, like the features of animals (Foster-Hanson & Lombrozo, 2022; Foster-Hanson & Rhodes, 2019a).

One limitation of the current work is that we chose to present gender categories as exemplified by White children, based on evidence that gender stereotypes are biased towards White boys and girls beginning in childhood (Lei et al., 2021). However, actual people children encounter in daily life hold membership in multiple social categories simultaneously, leaving open the question of how gender stereotypes shape children's intersectional social prototypes. This is an important open question for future research.

Prototypes shape how people reason about themselves and others, so idealized gender prototypes could have negative consequences, for instance by narrowing the choices children consider early in life (Bian et al., 2017; Fiske & Taylor, 2013). Idealized social prototypes could also lay the groundwork for prejudice, leading people to judge others based on narrow cultural stereotypes (Okimoto & Brescoll, 2010; Purdie-Vaughns & Eibach, 2008). Thus, future work should identify the cultural and cognitive factors that foster greater consideration of within-category variability across childhood in order to mitigate these harmful effects.

## ACKNOWLEDGMENTS

We thank Anjola Tega for assistance with video coding and Martin Wickstrom for illustrating our study stimuli. We are especially grateful to the families who participated in this research.

## CONFLICT OF INTEREST

None.

## DATA AVAILABILITY STATEMENT

All data, code, and materials are available on the Open Science Framework: https://osf.io/e7bvq/. Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the NIH under Award Number F31HD093431 (to E.F.H.) and Award Number R01HD087672 (to M.R.), and NSF grant BCS-1729540 (to M.R.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Science Foundation. The authors declare no conflict of interest. All procedures were approved by the IRB of New York University.

## ORCID

*Emily Foster-Hanson* https://orcid.org/0000-0002-0963-3934

## REFERENCES

Anglin, J. M. (1986). *Semantic and conceptual knowledge underlying the child's words*. Springer.

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. http://doi.org/10.1126/science.aah6524

Bjorklund, D. F., & Thompson, B. E. (1983). Category typicality effects in children's memory performance: Qualitative and quantitative differences in the processing of category information. *Journal of Experimental Child Psychology*, *35*(2), 329–344.

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. *Advances in Social Cognition*, *11*, 1–52.

Christensen, R. H. B. (2019). "ordinal—Regression Models for Ordinal Data.". R package version 2019.12-10. https://CRAN.R-project.org/package=ordinal

Datavyu Team (2014). *Datavyu: A video coding tool*. Databrary Project, New York University. *URL* http://datavyu.org

Ellemers, N. (2018). Gender stereotypes. *Annual Review of Psychology*, *69*, 275–298. http://doi.org/10.1146/annurev-psych-122216-011719

Emmons, N. A., & Kelemen, D. A. (2015). Young children's acceptance of within-species variation: Implications for essentialism and teaching evolution. *Journal of Experimental Child Psychology*, *139*, 148–160. https://doi.org/10.1016/j.jecp.2015.05.011

Fagot, B. I. (1978). The influence of sex of child on parental reactions to toddler children. *Child Development*, *49*(2), 459–465. https://doi.org/10.2307/1128711

Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture*. Sage.

Foster-Hanson, E., & Lombrozo, T. (2022a). What are men and mothers for? The causes and consequences of functional reasoning about social categories. *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

Foster-Hanson, E., & Lombrozo, T. (2022b). How "is" shapes "ought" for folk-biological concepts. *Cognitive Psychology*, *139*, 101507. https://doi.org/10.1016/j.cogpsych.2022.101507

Foster-Hanson, E., Leslie, S., & Rhodes, M. (2019, January 20). Speaking of kinds: How correcting generic statements can shape children's concepts. https://doi.org/10.31234/osf.io/28qf7

Foster-Hanson, E., Moty, K., Cardarelli, A., Ocampo, J. D., & Rhodes, M. (2020). Developmental changes in strategies for gathering evidence about biological kinds. *Cognitive Science*, *44*(5), e12837. http://doi.org/10.1111/cogs.12837

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

Foster-Hanson, E., & Rhodes, M. (2019b). Normative Social Role Concepts in Early Childhood. *Cognitive Science*, *43*(8). Portico. https://doi.org/10.1111/cogs.12782

Foster-Hanson, E., Roberts, S. O., Gelman, S. A., & Rhodes, M. (2021). Categories convey prescriptive information across domains and development. *Journal of Experimental Child Psychology*, *212*, 105231. https://doi.org/10.1016/j.jecp.2021.105231

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. 69(1), Oxford Cognitive Development.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*(3), 183–209. http://doi.org/10.1016/0010-0277(86)90034-X

Gelman, S. A., Taylor, M. G., Nguyen, S. P., Leaper, C., & Bigler, R. S. (2004). Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs. *Monographs of the Society for Research in Child Development*, i–142.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4. http://doi.org/10.1037/0033-295X.102.1.4

Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology*, *69*, 251–273. http://doi.org/10.1146/annurev-psych-122216-011710

Hilton, J. L., & Von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, *47*(1), 237–271. http://doi.org/10.1146/annurev.psych.47.1.237

Kalish, C. (1998). Reasons and causes: Children's understanding of conformity to social rules and physical laws. *Child Development*, *69*(3), 706–720. https://doi.org/10.1111/j.1467-8624.1998.tb06238.x

Katz, P. A., & Ksansnak, K. R. (1994). Developmental aspects of gender role flexibility and traditionality in middle childhood and adolescence. *Developmental Psychology*, *30*(2), 272. https://doi.org/10.1037/0012-1649.30.2.272

Kim, S., & Murphy, G. L. (2011). Ideals and category typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1092.

Kruschke, J. K. (2001). Toward a Unified Model of Attention in Associative Learning. *Journal of Mathematical Psychology*, *45*(6), 812–863. https://doi.org/10.1006/jmps.2000.1354

Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin*, *129*(4), 522. http://doi.org/10.1037/0033-2909.129.4.522

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). "lmerTest Package: Tests in linear mixed effects models." *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lei, R. F., Leshin, R. A., Moty, K., Foster-Hanson, E., & Rhodes, M. (2022). How race and gender shape the development of social prototypes in the United States. *Journal of Experimental Psychology: General*, *151*(8), 1956–1971.

Leslie, S. J. (2015). "Hillary Clinton is the only man in the Obama administration": Dual character concepts, generics, and gender. *Analytic Philosophy*, *56*(2), 111–141. https://doi.org/10.1111/phib.12063

Levesque, H., Davis, E., & Morgenstern, L. (2012). *The Winograd schema challenge*. AAAI Publications.

Maccoby, E. E. (1988). Gender as a social category. *Developmental Psychology*, *24*(6), 755. https://doi.org/10.1037/0012-1649.24.6.755

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. https://doi.org/10.1111/j.0963-7214.2004.00276.x

Medin, D. L., & Ortony, A. (1989). Psychological essentialism.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology*, *19*(2), 242–279. http://doi.org/10.1016/0010-0285(87)90012-0

Mervis, C. B., & Pani, J. R. (1980). Acquisition of basic object categories. *Cognitive Psychology*, *12*(4), 496–522. http://doi.org/10.1016/0010-0285(80)90018-3

Okimoto, T. G., & Brescoll, V. L. (2010). The price of power: Power seeking and backlash against female politicians. *Personality and Social Psychology Bulletin*, *36*(7), 923–936. http://doi.org/10.1177/0146167210371949

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*(2), 185. https://doi.org/10.1037/0033-295X.97.2.185

Paulhus, D. L., & Vazire, S. (2007). The self-report method. *Handbook of Research Methods in Personality Psychology*, *1*(2007), 224–239.

Poulin-Dubois, D., Serbin, L. A., Eichstedt, J. A., Sen, M. G., & Beissel, C. F. (2002). Men don't put on make-up: Toddlers' knowledge of the gender stereotyping of household activities. *Social Development*, *11*(2), 166–181. https://doi.org/10.1111/1467-9507.00193

Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional invisibility: The distinctive advantages and disadvantages of multiple subordinate-group identities. *Sex Roles*, *59*(5-6), 377–391. https://doi.org/10.1007/s11199-008-9424-4

Rakoczy, H., & Schmidt, M. F. (2013). The early ontogeny of social norms. *Child Development Perspectives*, *7*(1), 17–21. https://doi.org/10.1111/cdep.12010

Rhodes, M., Brickman, D., & Gelman, S. A. (2008). Sample diversity and premise typicality in inductive reasoning: Evidence for developmental change. *Cognition*, *108*(2), 543–556. http://doi.org/10.1016/j.cognition.2008.03.002

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. http://doi.org/10.1016/j.cogpsych.2009.05.001

Rhodes, M., & Liebenson, P. (2015). Continuity and change in the development of category-based induction: The test case of diversity-based reasoning. *Cognitive Psychology*, *82*, 74–95. http://doi.org/10.1016/j.cogpsych.2015.07.003

Rhodes, M., Rizzo, M. T., Foster-Hanson, E., Moty, K., Leshin, R. A., Wang, M., Benitez, J., & Ocampo, J. D. (2020). Advancing developmental science via unmoderated remote research with children. *Journal of Cognition and Development*, *21*(4), 477–493. http://doi.org/10.1080/15248372.2020.1797751

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

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. http://doi.org/10.1371/journal.pone.0200921

Shtulman, A., & Phillips, J. (2018). Differentiating "could" from "should": Developmental changes in modal cognition. *Journal of Experimental Child Psychology*, *165*, 161–182. http://doi.org/10.1016/j.jecp.2017.05.012

Stern, C., & Axt, J. (2021). Ideological differences in race and gender stereotyping. *Social Cognition*, *39*(2), 259–294. https://doi.org/10.1521/soco.2021.39.2.259

**How to cite this article:** Foster-Hanson, E., & Rhodes, M. (2022). Stereotypes as prototypes in children's gender concepts. *Developmental Science*, e13345. https://doi.org/10.1111/desc.13345