CHILD DEVELOPMENT

Child Development, November/December 2020, Volume 91, Number 6, Pages 1877–1885

# Gender Encoding in Gender Diverse and Gender Conforming Children

Jessica J. Glazier
University of Washington

Selin Gülgöz
Fordham University

Kristina R. Olson
University of Washington

Previous research suggests that people encode gender starting in childhood. The present research asked whether gender diverse children (i.e., children whose gender identity or expression differs from that expected based on assigned sex) encode gender. Results showed that 3- to 5-year-old gender diverse participants ($N = 71$), siblings of gender diverse children ($N = 52$), and gender conforming controls ($N = 69$) did not significantly differ in degree of gender encoding. These results converge with prior research to suggest that gender diverse children process gender in ways that do not differ from gender conforming children, and provide further evidence that gender encoding may be a common aspect of person perception in societies that support a binary view of gender.

Gender is one of our most basic social categories, and children use gender as a basis for many decisions, such as who to interact and make friends with (Maccoby, 1988; Shutts, Roben, & Spelke, 2013). Noticing and encoding the gender of a person one meets (i.e., categorizing others based on gender) is thought to occur automatically (e.g., Kurzban, Tooby, & Cosmides, 2001) and the tendency to encode gender begins early in development (e.g., Weisman, Johnson, & Shutts, 2015) within cultures that view gender as binary. To our knowledge, studies of gender encoding have focused on participants whose gender identity and expression align with both their sex assigned at birth and with how people treat them (henceforth, gender conforming people). In this study, we asked whether "gender diverse" children (whose gender identities and expressions diverge from their assigned sex) encode gender to the same degree as gender conforming children.

The authors would like to thank Kara Weisman and Kristin Shutts for feedback on the manuscript as well as Riley Lowe, Daniel Alonso, Lily Durwood, Elizabeth Enright, Annie Fast, Madeleine DeMeules, Deja Edwards, and Gabrielle Lindquist for their assistance with data collection. This research was supported by the National Science Foundation (grant numbers BCS-1715068, BCS-1523632, and SMA-1837857) and the National Institute of Child Health and Human Development (grant number HD092347).

Correspondence concerning this article should be addressed to Jessica J. Glazier, Department of Psychology, University of Washington, 119A Guthrie Hall (GTH), Seattle, WA, 98195-1525. Electronic mail may be sent to jglazier@uw.edu.

### *Gender Cognition in Gender Diverse Children*

Estimates of the number of gender diverse youth range from 2.7% of high school students (Rider, McMorris, Gower, Coleman, & Eisenberg, 2018) to 6% of children (Zucker & Lawrence, 2009), though estimates vary depending on definition. Nonetheless, research on gender diverse children's gender cognition has been limited. The few studies examining gender cognition of gender diverse children have focused on deliberative processing of gender, finding that gender diverse children are more likely to prefer objects, clothes, and peers associated with the other binary sex than their gender conforming peers (Ahlqvist, Halim, Greulich, Lurye, & Ruble, 2013; Bailey, Bechtold, & Berenbaum, 2002; Fast & Olson, 2018; Olson, Key, & Eaton, 2015; Zucker, Doering, Bradley, & Finegan, 1982; tomboys differ in that they *also* tend to like objects, clothes, peers, associated with their assigned sex). Additionally, researchers have found differences between gender conforming and gender diverse children in how they *reason about* gender. Several studies suggest that gender diverse children are more accepting of gender nonconforming behaviors and may endorse gender stereotypes less than gender conforming children (Ahlqvist et al., 2013; Martin & Dinella, 2012; Olson & Enright, 2018). Additionally, when

 0009-3920/2020/9106-0004
DOI: 10.1111/cdev.13399

asked about what gender they think a target was in the past, is currently, or would be in the future, gender diverse children are more likely to give an answer inconsistent with the target's current gender than gender conforming controls, suggesting that gender diverse children may believe that gender is less stable across time compared to gender conforming peers (Fast & Olson, 2018; Zucker et al., 1999). Importantly, the tasks used in the above studies are directly about gender and children have time to deliberatively answer; therefore, these tasks could be considered more explicit, direct, or controllable measures.

Little work has examined more implicit, indirect, or automatic processing of gender in gender diverse children. The only implicit assessment of gender in gender diverse children that we are aware of is the Implicit Association Test, which has been used to assess children's gender preferences and identity (Gülgöz et al., 2019; Olson et al., 2015). In the present work, we were interested in low-level processing of others' gender. Specifically, we investigated whether gender diverse children automatically encode gender and do so at rates different from gender conforming children. Gender diverse children are likely to be treated differently and to have different discussions and experiences of gender than their gender conforming peers (e.g., Carter & McCloskey, 1984; Rahilly, 2015). As one example, many gender diverse children discuss with their parents that gender can change over time (Olson et al., 2019), a topic few gender conforming children are likely to discuss with their parents. These unique experiences with gender could impact not only deliberative, conscious beliefs about gender (e.g., Fast & Olson, 2018) but might impact incidental and automatic processing of gender. To investigate this question, we assessed gender encoding, or spontaneously noticing and remembering a person's gender.

## *Gender Encoding*

Perhaps the best-known test of gender encoding is the "Who Said What?" task (Taylor, Fiske, Etcoff, & Ruderman, 1978), a memory confusion protocol in which participants witness a conversation between men and women. Participants are later asked to recall which statements were made by which targets. The primary finding using this method is that people often mistake who said what in a systematic way: they confuse targets of the same gender with one another more often than they confuse targets of different genders (e.g., if a phrase was uttered by a woman, people more often misremember it as being said by another woman than by a man). This confusion is thought to occur because participants automatically encode gender when observing the initial conversation—that is, though there is no indication the task is about gender, participants notice and remember the gender of the conversation partners. The gender effect observed in this task is particularly immune to intervening influences (Kurzban et al., 2001). Other measures have found converging evidence of gender encoding as well (e.g., Tomelleri & Castelli, 2011).

Recently, a memory confusion protocol has been adapted for use with children (e.g., Bennett & Sani, 2003; Bennett, Sani, Hopkins, Agostini, & Malucchi, 2000). In Weisman et al.'s (2015) task, participants first learn about a series of gender-stereotypically presented children who visit different animals at a zoo, and then are asked to recall which animal was seen by which child. Using this task, Weisman and colleagues found that 4- to 6-year-old children confused children of the same gender with one another more often than they confused children of different genders, despite gender never being mentioned, suggesting the task may assess automatic processing of gender. Most relevant to the present work, Shutts, Kenward, Falk, Ivegran, and Fawcett (2017) showed that 3- to 6-year-old Swedish children attending a gender-neutral preschool (where teachers avoided gendered language and countered gender stereotypes in activities) showed the same levels of gender encoding as their peers in traditional preschools, suggesting that gender encoding is not always affected by differences in gender experience.

## *Current Work*

The present work investigated whether gender diverse children's low-level processing of gender differs from that of gender conforming children. For exploratory purposes, we also included a group of gender conforming siblings of gender diverse children. Past work has suggested that siblings of gender diverse children often, but not always, respond similarly to gender diverse children on measures of gender cognition (Fast & Olson, 2018; Olson & Enright, 2018; Olson et al., 2015).

Additionally, we also took this opportunity to examine whether children better encode the gender of individuals in their gender in-group. Research has shown that children sometimes attend more to same-sex models (Bussey & Bandura, 1984; Slaby & Frey, 1975), though, both Bennett and Sani (2003) and Weisman et al. (2015) showed that (gender conforming) children were equally likely to make errors on own- versus other-gender trials (we use binary language as the encoding task only represents two genders). The children in the present work were 3–5 years of age, because this is the age with whom most work on gender encoding in childhood has focused (Bednarek & Shutts, 2017; Shutts et al., 2017; Weisman et al., 2015), making comparison to past work possible.

## Method

### *Participants*

After data collection had begun (164 of 192 participants) but before analyzing the data, we registered our methods, research questions, and analysis plan (https://osf.io/fd2g4/?view_only=dca9cf9e65bc4464a0f772366c9986a4). Data were collected between October 2015 and June 2019. The most important difference between the official registration and this manuscript is that we initially proposed to analyze the gender diverse sample as two different groups: socially transitioned transgender children and gender nonconforming children. We do so in Supporting Information; however, based on peer review and to maximize power, we combined the samples in the main text.

Participants were children 3–5 years of age ($N = 192$; $M_{\text{age}} = 4$ years, 11 months). The final sample included 71 gender diverse participants (20 assigned female at birth), 52 gender conforming siblings of gender diverse children (22 males), and 69 gender conforming controls (20 males; see Table 1 for demographics) matched to the gender diverse participants by age (within 4 months) but with the "opposite" assigned sex at birth (as in Rae et al., 2019).

During their visit, participants completed the current task and other tasks not relevant to the present questions; the other tasks are published elsewhere or data collection is ongoing. To be included in the present analyses, as per our registration, participants must have completed at least two out of four blocks of the gender encoding task described later. One gender diverse girl and her control, as well as two additional controls, were excluded from analyses for this reason. The sample sizes of the current work are comparable to those of past work using this task; see Figure 2.

The gender diverse children in the present work were recruited through community groups, conferences, media coverage, and word-of-mouth. Our criteria to count as gender diverse was that parents referred their children for our study on transgender children or gender diversity, or in discussions with a control, if a parent indicated their child was gender diverse. Confirming these categorizations, we compared the scores of our gender diverse sample ($M = 2.24$) and our comparison gender conforming group ($M = 3.84$) on a parent-report measure of gender conformity (Johnson et al., 2004) and the two groups strongly differed, $t(137) = 21.61$, $p < .001$, $d = 3.67$; furthermore, the vast majority of controls scored higher than the vast majority of gender diverse participants; see Figure 1.

Siblings of gender diverse children were recruited whenever possible. Gender conforming control participants were recruited from a greater metropolitan area in the Pacific Northwest and were run in a developmental psychology laboratory.

Table 1
*Participant Demographics*

| | Participant group | | |
|---|---|---|---|
| | Gender conforming controls, % | Gender diverse, % | Gender conforming sibling, % |
| Child's race | | | |
| White | 71.0 | 71.8 | 67.3 |
| Multiracial | 23.2 | 21.1 | 25.0 |
| Another race | 4.3 | 7.0 | 3.8 |
| Not reported | 1.4 | 0.0 | 3.8 |
| Household annual income | | | |
| <$25,000 | 2.9 | 4.2 | 3.8 |
| $25,001–$50,000 | 1.4 | 9.9 | 9.6 |
| $50,001–$75,000 | 14.5 | 22.5 | 23.1 |
| $75,001–$125,000 | 30.4 | 35.2 | 30.8 |
| More than $125,000 | 47.8 | 28.2 | 28.8 |
| Not reported | 2.9 | 0.0 | 3.8 |
| Parent education level | | | |
| High school diploma | 0.0 | 1.4 | 1.9 |
| Some college/ Associate's degree | 8.7 | 8.5 | 11.5 |
| College/Bachelor's degree | 30.4 | 26.8 | 19.2 |
| Advanced degree (MA, MD, PhD, etc.) | 37.7 | 39.4 | 42.3 |
| Not reported | 23.2 | 23.9 | 25.0 |
| Parent political ideology | | | |
| Liberal | 73.9 | 87.3 | 78.8 |
| Moderate | 21.7 | 12.7 | 17.3 |
| Conservative | 1.4 | 0.0 | 0.0 |
| Not reported | 2.9 | 0.0 | 3.8 |

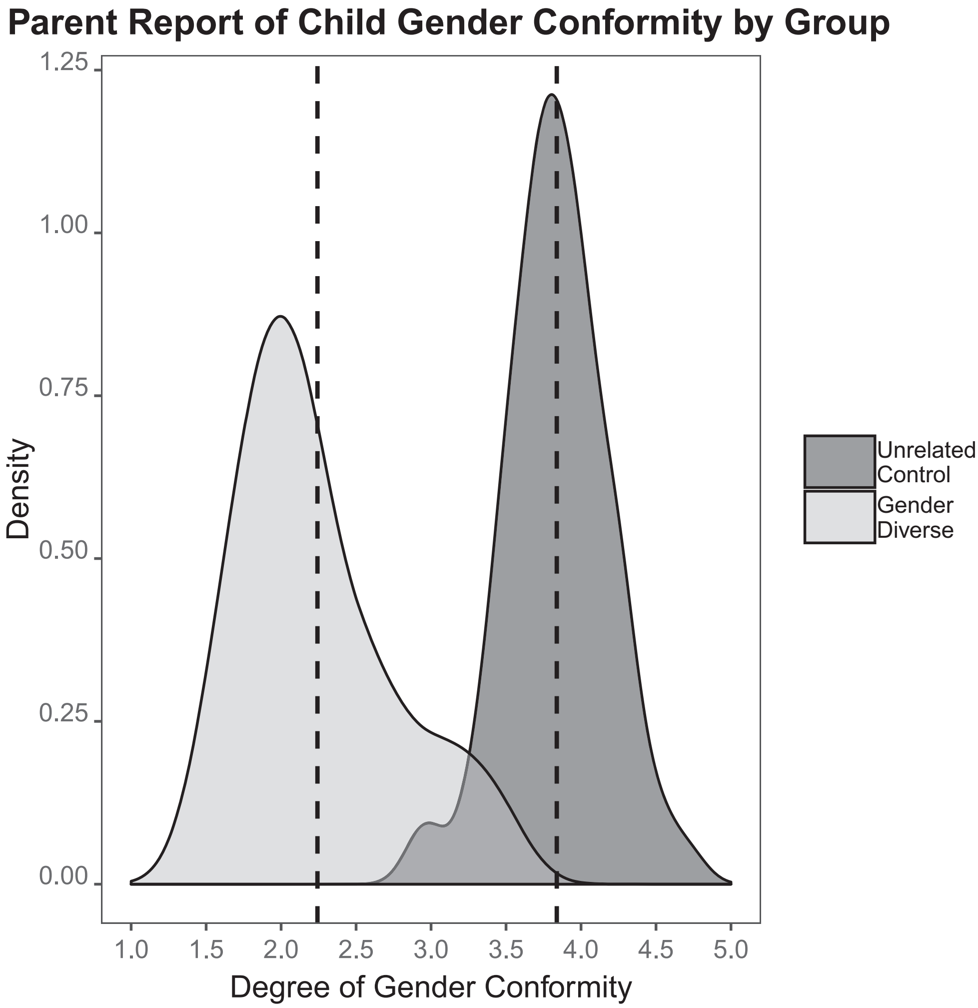


*Figure 1.* The figure shows density of parent reports of child gender conformity using the Gender Identity Questionnaire (Johnson et al., 2004). Scoring was in accordance with Johnson et al. (2004) such that the same items were reverse scored and the same two items were dropped. Responses were coded from 1 to 5, where 5 indicates responses most aligned with one's sex at birth. Means for each group are represented by the black, dashed lines.

### *Procedure and Design*

#### *Gender encoding*

Participants completed the exact Gender Encoding Task from Weisman et al. (2015), which consisted of four blocks, each including a familiarization and a test phase. Participants were told they would meet four children who went to the zoo several days in a row and would be asked to remember which child saw which animal. During the familiarization phase, participants saw each child paired with an animal for 8 s. During the test phase, participants were prompted to recall which animal each child saw. The same two boys and two girls—all White, gender-stereotypical, smiling children with brown hair and a gray shirt—appeared in each block. Children appeared in different orders in each familiarization phase but a fixed order during the test phase.

#### *Scoring*

To ask whether children mistake people of the same binary gender with one another more often than they mistake people of different genders with each other, in line with Weisman et al. (2015), and as specified in our registration, an adjusted error difference score was calculated ([Total same-gender errors] – [Total different-gender errors/2]). A "same-gender error" was an error in which the participant thought an animal was seen by the target who shared a gender with the target who had actually seen that animal. A "different-gender error" was an error in which the participant thought an animal was seen by a target of a different gender than the target who had actually seen that animal. We divided the total number of different-gender errors by two as there were twice as many errors possible for a different gender than of the same gender. If a participant completed at least two but less than four blocks, we re-computed their score to represent a value out of 16 trials. The final score could range from −8 (different-gender errors made on all trials) to 16 (same-gender errors made on all trials), with an expected value of 0 if children were responding randomly.

## Results

### Gender Encoding

We conducted a single factor analysis of variance (ANOVA) and found no significant main effect of participant group, $F(2, 189) = 2.41$, $p = .093$, $\eta_p^2 = .02$ (see Figure 2). Although the ANOVA indicated that there were no significant differences between groups, we registered testing the gender encoding effect in each individual group, using a one-sample *t*-test (comparing the adjusted error difference score to 0, the expected value if children were responding randomly). Gender diverse participants ($M$ adjusted error difference score = 1.41, $t(70) = 4.56$, $p < .001$, $d = 0.54$), siblings ($M = 1.28$, $t(51) = 3.45$, $p = .001$, $d = 0.48$), and controls ($M = 0.54$, $t(68) = 2.06$, $p = .043$, $d = 0.25$), all made significantly more same-gender errors than different-gender errors, indicating they did encode gender.

### Own-Gender Encoding

We also registered to test whether children show an in-group bias in memory, better encoding the gender of children of their own gender than of the other gender. However, when working with participants who may not identify as a binary gender, examining this question is challenging due to the binary nature of the task. To attempt to address this question, we analyzed these data three different ways by calculating an "in-group bias" score based on different methods of coding participants' "own gender." Each participant's in-group bias score was weighted such that it represented a value out of 16 trials, with a possible range of −8 (participants made same-gender errors on all trials involving other-gender targets and none on trials involving own-gender targets) to 8 (participants made same-gender errors on all trials involving own-gender targets and none on trials involving other-gender targets), with an expected value of 0 if participants were responding randomly. Importantly, the results were generally similar no matter how they were analyzed, with one exception.

First, we coded "own gender" using the principle by which we had matched gender diverse and control participants (the approach originally planned via our registration). For all gender diverse children, their in-group bias score was calculated such that their "own gender" was the binary gender opposite their assigned sex (e.g., assigned males were coded as being in-group members with girls). Using this scoring, there were no differences by group, $F(2,189) = 0.04$, $p = .960$, $\eta_p^2 < .001$, and no group showed a significant in-group bias effect (gender diverse participants, $M$ in-group bias score = 0.23, $t(70) = 0.98$, $p = .331$, $d = 0.12$; siblings; $M = 0.31$, $t(51) = 1.09$, $p = .281$, $d = 0.15$; controls, $M = 0.30$, $t(68) = 1.51$, $p = .136$, $d = 0.18$).

### Non-Registered, Exploratory Analyses

After completing the study, we realized the "own gender" analyses could be problematic as some gender diverse children may think of themselves as the gender associated with their assigned sex (but defy our cultural expectations concerning gender expression). To accommodate this possibility, we next coded "own gender" as the binary gender associated with the pronouns children use in everyday life, excluding participants who do not exclusively use pronouns associated with a single binary gender ($n = 10$). Again, using this coding, there were no differences by group, $F(2,179) = 0.03$, $p = .975$, $\eta_p^2 < .001$, and no group showed a significant in-group bias effect (gender diverse participants, $M$ in-group bias score = 0.37, $t(60) = 1.53$, $p = .132$, $d = 0.20$; siblings; $M = 0.31$, $t(51) = 1.09$, $p = .281$, $d = 0.15$; controls, $M = 0.30$, $t(68) = 1.51$, $p = .136$, $d = 0.18$).

One additional concern was that children's pronouns might not align with how they think about their identity. Therefore, the third way in which we calculated the in-group bias score was such that each participant's "own gender" was scored in line with their answers to another question asked on the same day—whether they were a boy, a girl, or something else. All children who selected either "boy" or "girl" were used in this analysis and scored according to their self-identified binary category, while children who responded "something

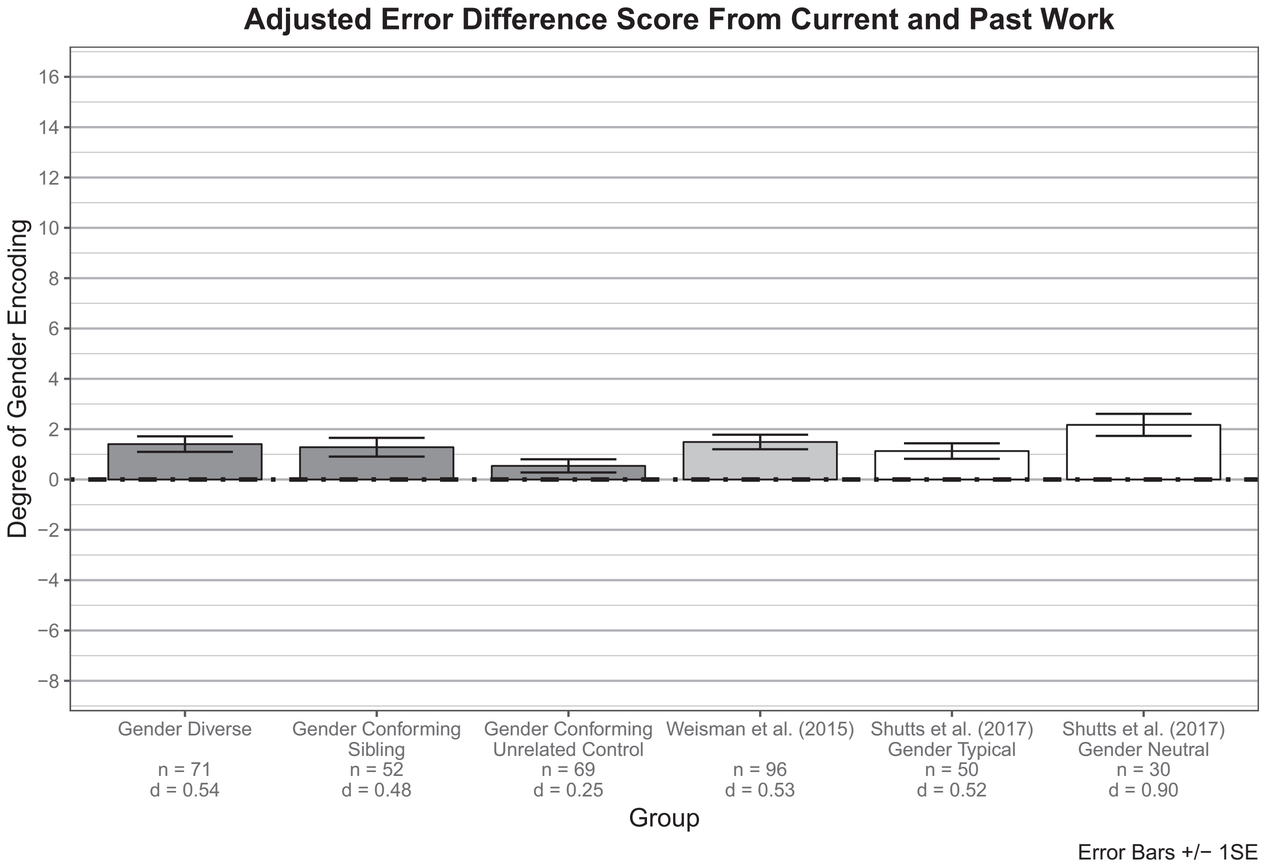


*Figure 2.* The figure shows the adjusted error difference score for each of the three samples in the current work as well as for all samples from past work using the gender encoding task. In accordance with Weisman et al. (2015), scores could range from −8 (different-gender errors made on all trials) to 16 (same-gender errors made on all trials), with an expected value of 0 if children were responding randomly. Sample size and effect size (Cohen's *d*) are indicated beneath each group.

else" or who did not respond were excluded ($n = 31$). Again, there were no differences by group, $F(2,158) = 1.52$, $p = .223$, $\eta_p^2 = .02$, and siblings ($M$ in-group bias score = 0, $t(46) = 0$, $p = 1$, $d = 0$) and controls ($M = 0.29$, $t(61) = 1.41$, $p = .164$, $d = 0.18$) did not show a significant in-group bias effect. However, gender diverse participants *did* show a significant in-group bias effect when this approach was utilized, $M = 0.64$, $t(51) = 2.40$, $p = .020$, $d = 0.33$.

## Discussion

The current studies investigated whether gender diverse children differ from gender conforming children in automatic encoding of gender. The groups did not differ from one another and, when examined separately, gender diverse children, siblings, and controls were all more likely to confuse two children of the same binary gender with one another than they were to confuse two children who differed in gender, suggesting that they encode gender. Importantly, gender diverse children neither appeared to be especially hyper- or hypo-attentive to gender compared to gender conforming peers in this study or in past research (e.g., Weisman et al., 2015). This is especially apparent when the effect sizes of the current study are compared to past research, as the effect size for the gender diverse children ($d = 0.54$) is remarkably similar to that of Weisman et al. (2015; $d = 0.53$) and the children enrolled in a gender typical preschool in Shutts et al. (2017; $d = 0.52$; see Figure 2).

The sample of gender diverse children in the current study was heterogeneous, including socially transitioned transgender children who may identify in a relatively binary way and children who might be described as "gender nonconforming"—those whose gender identities may be less binary and who have not socially transitioned. It is reasonable to wonder if there are differences in degree of gender encoding between socially transitioned and gender nonconforming children, as there could ostensibly be a relation between the degree to which a person's gender identity is binary and degree to which they encode binary gender. Analyzed separately, we found that both socially transitioned children and gender nonconforming children encoded gender (reported in Supporting Information), suggesting this effect is not driven by the slightly larger sample of socially transitioned children. These results provide further evidence that the tendency to encode gender may arise regardless of individual differences in gender experience. Still, as with all findings, replication of this effect, ideally with other measures assessing automatic gender encoding, would be important.

That gender diverse children did not differ from gender conforming children in terms of their gender encoding suggests that gender diverse children may automatically process gender information similarly to others. One possible explanation for these findings is that children, regardless of gender identity, may come to encode binary gender at an early age simply because gender is a prominent social category. All current and past tests of this effect in children have recruited participants from cultures that support a binary view of gender and therefore commonalities in gender encoding on this task may reflect shared cultural learning experiences, rather than being impacted by one's personal gender identity (similar arguments have been made in research on racial attitudes, e.g., Spencer, 1982, 1984). Replications of this effect in cultures that do not support a binary view of gender would be necessary to help determine whether gender encoding is learned based on one's cultural input or is an innate, universal aspect of person perception.

The finding that gender diverse children process gender similarly to others, however, is in contrast to work suggesting that transgender children sometimes differ from cisgender children in their more deliberative, conscious beliefs about gender (Fast & Olson, 2018; Olson & Enright, 2018). This dissociation between more deliberative gender beliefs and gender encoding may indicate that different processes could be driving responses to these different types of tasks. As alluded to above, the encoding effect may reflect automatic processing of gender, whereas assessments of beliefs about gender stability may reflect more thoughtful, reasoned logic, as they require more linguistic input and output and are more explicitly about gender. As such, assessments of gender beliefs may be affected by conscious considerations about gender. We know that young infants can perceptually divide targets into categories of male and female within a few months of birth (Patterson & Werker, 2002; Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002), whereas knowledge of gender stereotypes and beliefs about the stability of gender over time, appear to emerge later (Kohlberg, 1966; Martin, Wood, & Little, 1990; Slaby & Frey, 1975; Thompson, 1975). The different developmental trajectories of these capacities further suggest they may be distinct, dual processes, representing both high- and low-level processing of gender.

In the current study, we also explored whether children encoded their own gender better than the other gender (what we have called their in-group bias score). We found little evidence that this was the case in any participant group or with any method of calculating the in-group bias score, with the exception of the gender diverse participants when "own gender" was coded using explicit self-identification. We have some reason to be cautious in overinterpreting the latter effect, however. This was a post-hoc analysis, involved dropping many participants, and did not converge with the results from other scoring approaches or the other subject groups (or past work, e.g., Bennett & Sani, 2003; Weisman et al., 2015), leaving us with less confidence in the results. At most, these results suggest a provocative next direction. Our own take-away is that we have little to no evidence of in-group encoding bias in this paper.

### *Limitations and Suggestions for Future Research*

One limitation of this work is that the encoding task examined binary gender encoding and the stimuli themselves represent stereotypical presentations of girls and boys. Moreover, the task is based on the assumption that gender can be reliably visually assessed. Thus, we do not yet know if children encode the gender of people who are less binary or gender-stereotypical in appearance or identity. Furthermore, one could argue that we did not test whether non-binary-identifying children better encode their *own* gender, as we did not include non-binary targets in the encoding task. We are not aware of any research that has examined gender encoding of non-binary gender, making this a particularly important area for future research.

The gender encoding task used in the current study also had no racial diversity, as all four children within the task were White. To our knowledge, this has been true of all studies assessing gender encoding in children (all of which also include primarily White participants). As a result, we cannot say whether the tendency to encode gender generalizes to targets or participants of all races, suggesting another important direction for future work.

## Conclusions

The current studies provide evidence that, as young as preschool years, gender diverse children —like their gender conforming peers—encode binary gender, even without being instructed to do so. These data provide the clearest evidence to date that gender diverse and gender conforming children process gender similarly, at least when gender categorization is assessed incidentally.

## References


Ahlqvist, S., Halim, M. L., Greulich, F. K., Lurye, L. E., & Ruble, D. (2013). The potential benefits and risks of identifying as a tomboy: A social identity perspective. *Self and Identity*, *12*, 563–581. https://doi.org/10.1080/15298868.2012.717709

Bailey, J. M., Bechtold, K. T., & Berenbaum, S. A. (2002). Who are tomboys and why should we study them? *Archives of Sexual Behavior*, *31*, 333–341. https://doi.org/10.1023/a:1016272209463

Bednarek, G. T., & Shutts, K. (2017, November 21). *Automatic encoding of gender by three-year-old children*. https://doi.org/10.31234/osf.io/we8s5

Bennett, M., & Sani, F. (2003). The role of target gender and race in children's encoding of category-neutral person information. *British Journal of Developmental Psychology*, *21*, 99–112. https://doi.org/10.1348/026151003321164645

Bennett, M., Sani, F., Hopkins, N., Agostini, L., & Malucchi, L. (2000). Children's gender categorization: An investigation of automatic processing. *British Journal of Developmental Psychology*, *18*, 97–102. https://doi.org/10.1348/026151000165599

Bussey, K., & Bandura, A. (1984). Influence of gender constancy and social power on sex-linked modeling. *Journal of Personality and Social Psychology*, *47*, 1292. https://doi.org/10.1037/0022-3514.47.6.1292

Carter, D. B., & McCloskey, L. A. (1984). Peers and the maintenance of sex-typed behavior: The development of children's conceptions of cross-gender behavior in their peers. *Social Cognition*, *2*, 294–314. https://doi.org/10.1521/soco.1984.2.4.294

Fast, A. A., & Olson, K. R. (2018). Gender development in transgender preschool children. *Child Development*, *89*, 620–637. https://doi.org/10.1111/cdev.12758

Gülgöz, S., Glazier, J. J., Enright, E. A., Alonso, D. J., Durwood, L. J., Fast, A. A., ⋯ Olson, K. R. (2019). Similarity in transgender and cisgender children's gender development. *Proceedings of the National Academy of Sciences of the United States of America*, *166*, 24480–24485. https://doi.org/10.1073/pnas.1909367116

Johnson, L. L., Bradley, S. J., Birkenfeld-Adams, A. S., Kuksis, M. A. R., Maing, D. M., Mitchell, J. N., & Zucker, K. J. (2004). A parent-report gender identity questionnaire for children. *Archives of Sexual Behavior*, *33*, 105–116. https://doi.org/10.1023/b:aseb.0000014325.68094.f3

Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In E. E. Maccoby (Ed.), *The development of sex differences* (pp. 82–173). Stanford, CA: Stanford University Press.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 15387–15392. https://doi.org/10.1073/pnas.251541498

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

Martin, C. L., & Dinella, L. M. (2012). Congruence between gender stereotypes and activity preference in self-identified tomboys and non-tomboys. *Archives of Sexual Behavior*, *41*, 599–610. https://doi.org/10.1007/s10508-011-9786-5

Martin, C. L., Wood, C. H., & Little, J. K. (1990). The development of gender stereotype components. *Child Development*, *61*, 1891–1904. https://doi.org/10.1111/j.1467-8624.1990.tb03573.x

Olson, K. R., Blotner, C., Alonso, D., Lewis, K., Edwards, D., & Durwood, L. (2019). Family discussions of early childhood social transitions. *Clinical Practice in Pediatric Psychology*, *7*, 229. https://doi.org/10.1037/cpp0000289

Olson, K. R., & Enright, E. A. (2018). Do transgender children (gender) stereotype less than their peers and siblings? *Developmental Science*, *21*, e12606. https://doi.org/10.1111/desc.12606

Olson, K. R., Key, A. C., & Eaton, N. R. (2015). Gender cognition in transgender children. *Psychological Science*, *26*, 467–474. https://doi.org/10.1177/0956797614568156

Patterson, M. L., & Werker, J. F. (2002). Infants' ability to match dynamic phonetic and gender information in the face and voice. *Journal of Experimental Child Psychology*, *81*, 93–115. https://doi.org/10.1006/jecp.2001.2644

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, *31*, 1109–1121. https://doi.org/10.1068/p3331

Rae, J. R., Gülgöz, S., Durwood, L., DeMeules, M., Lowe, R., Lindquist, G., & Olson, K. R. (2019). Predicting early-childhood gender transitions. *Psychological Science*, *30*, 669–681. https://doi.org/10.1177/0956797619830649

Rahilly, E. P. (2015). The gender binary meets the gender-variant child: Parents' negotiations with childhood gender variance. *Gender & Society*, *29*, 338–361. https://doi.org/10.1177/0891243214563069

Rider, G. N., McMorris, B. J., Gower, A. L., Coleman, E., & Eisenberg, M. E. (2018). Health and care utilization of transgender and gender nonconforming youth: A population-based study. *Pediatrics*, *141*, e20171683. https://doi.org/10.1542/peds.2017-1683

Shutts, K., Kenward, B., Falk, H., Ivegran, A., & Fawcett, C. (2017). Early preschool environments and gender: Effects of gender pedagogy in Sweden. *Journal of Experimental Child Psychology*, *162*, 1–17. https://doi.org/10.1016/j.jecp.2017.04.014

Shutts, K., Roben, C. K. P., & Spelke, E. S. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development*, *14*, 35–62. https://doi.org/10.1080/15248372.2011.638686

Slaby, R. G., & Frey, K. S. (1975). Development of gender constancy and selective attention to same-sex models. *Child Development*, *46*, 849–856. https://doi.org/10.2307/1128389

Spencer, M. B. (1982). Preschool children's social cognition and cultural cognition. A cognitive developmental interpretation of race dissonance findings. *The Journal of Psychology*, *112*, 275–286. https://doi.org/10.1080/00223980.1982.9915385

Spencer, M. B. (1984). Black children's race awareness, racial attitudes and self-concept: A reinterpretation. *Journal of Child Psychology and Psychiatry*, *25*, 433–441. https://doi.org/10.1111/j.1469-7610.1984.tb00162.x

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, *36*, 778. https://doi.org/10.1037/0022-3514.36.7.778

Thompson, S. K. (1975). Gender labels and early sex role development. *Child Development*, *46*, 339–347. https://doi.org/10.2307/1128126

Tomelleri, S., & Castelli, L. (2011). On the nature of gender categorization. *Social Psychology*, *43*, 14–27. https://doi.org/10.1027/1864-9335/a000076

Weisman, K., Johnson, M. V., & Shutts, K. (2015). Young children's automatic encoding of social categories. *Developmental Science*, *18*, 1036–1043. https://doi.org/10.1111/desc.12269

Zucker, K. J., Bradley, S. J., Kuksis, M., Pecore, K., Birkenfeld-Adams, A., Doering, R. W., ⋯ Wild, J. (1999). Gender constancy judgments in children with gender identity disorder: Evidence for a developmental lag. *Archives of Sexual Behavior*, *28*, 475–502. https://doi.org/10.1023/a:1018713115866

Zucker, K. J., Doering, R. W., Bradley, S. J., & Finegan, J. A. K. (1982). Sex-typed play in gender-disturbed children: A comparison to sibling and psychiatric controls. *Archives of Sexual Behavior*, *11*, 309–321. https://doi.org/10.1007/bf01541592

Zucker, K. J., & Lawrence, A. A. (2009). Epidemiology of gender identity disorder: Recommendations for the standards of care of The World Professional Association for Transgender Health. *International Journal of Transgenderism*, *11*, 8–18. https://doi.org/10.1080/15532730902799946


## Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

**Appendix S1.** Gender Encoding in Transgender and Gender Nonconforming Participants Analyzed Separately