Gender encoding in gender diverse and gender conforming children

Jessica J. Glazier

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2019

Committee:

Kristina Olson

Cheryl Kasier

Program Authorized to Offer Degree:

Psychology

©Copyright 2019

Jessica J. Glazier

University of Washington

Abstract

Gender encoding in gender diverse and gender conforming children

Jessica J. Glazier

Chair of the Supervisory Committee:

Kristina Olson

Department of Psychology

Previous research suggests that encoding people's gender may be universal, even in childhood. The present research provided a new test of this possibility by asking whether gender diverse children (i.e., children whose gender identity or expression differs from that expected based on their assigned sex) encode gender. We recruited gender diverse participants (N = 71), siblings of gender diverse children (N = 52), and gender conforming controls (N = 69) and found that the groups did not significantly differ in degree of gender encoding. These results converge with prior research to suggest that gender diverse children may process gender in ways that do not differ from gender conforming children, and provide further evidence that gender encoding may be a universal aspect of person perception.

Gender encoding in gender diverse and gender conforming children

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). To our knowledge, studies of gender encoding have focused only on participants who have a gender identity and expression aligning with both their sex assigned at birth and with how people treat them—people we henceforth term “gender conforming.” In the present 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.

Gender Cognition in Gender Diverse Children

Estimates of the number of gender diverse young people range, for example, from 2.7% of high school students (Rider, McMorris, Gower, Coleman, & Eisenberg, 2018) to 6% of children (Zucker & Lawrence, 2009). However, 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 from other gender diverse groups 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 endorse gender stereotypes less than gender conforming children (Ahlqvist et al., 2013; Martin & Dinella, 2012; Olson & Enright, 2018). Additionally, when asked about what gender they think a target was in the past, 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 (Olson et al., 2015). However, 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 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 are more likely to mistakenly recall it 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 related to 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).

More recently, a memory confusion protocol has been adapted for use with young children (e.g., Bennett & Sani, 2003; Bennett, Sani, Hopkins, Agostini, & Malucci, 2000). In Weisman et al. (2015)'s 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 binary gender with one another more often than they confused children of different genders, despite gender never being mentioned, suggesting that the task may assess automatic processing of gender information. 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 by examining whether gender diverse children encode binary gender. 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 or not children better encode the gender of individuals belonging to their gender in-group. Research has shown that children sometimes attend more to same-sex models (Bussey & Bandura, 1984; Slaby & Frey, 1975), though, neither Bennett and Sani (2003) nor Weisman et al. (2015) found evidence that (gender conforming) children were any more or less likely to encode the identities of people of their own gender than people of the other gender (i.e., they 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 (registration can be found at https://osf.io/v4xz9/?view_only=4d87713e420c42528691c071f2f02656). All 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 as detailed in the registration in the Supplemental 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 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 (e.g., as in Rae et al., 2019).

Participants completed the current task in addition to other tasks not relevant to the present questions during their visit; 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 below. 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 a wide variety of 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, they were moved to this group. 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, $t(137) = 21.61$, $p < .001$, $d = 3.67$; further, 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.

Procedure and design

Gender encoding. Participants completed the exact Gender Encoding Task from Weisman and colleagues (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 seconds. During the test phase, participants were prompted to recall which animal each child saw. The same two boys and two girls—all White smiling children with brown hair and a gray shirt who look gender-stereotypical—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 2 but less than 4 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 ANOVA and found no significant main effect of participant group, $F(2,189) = 2.41$, $p = .093$, $\eta_{p2} = .02$ (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 that they did encode gender.

Own-gender Encoding

We also registered to test whether children show an ingroup 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 “ingroup bias” score based on different methods of coding participants’ “own gender.” Each participant’s ingroup 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 ingroup bias score was calculated such that their “own gender” was the binary gender opposite their sex assigned at birth (e.g., assigned males were coded as being in-group members with girls). Using this scoring, there were no differences by group, $F(2,189) = 0.04$, $p = .960$, $\eta_{p2} < .001$, and no group showed a significant ingroup bias effect (gender diverse participants, M ingroup 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, $M = 0.30$, $t(68) = 1.51$, $p = .136$, $d = 0.18$).

After completing the study, we realized the first approach 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, $F(2,178) = 0.02$, $p = .983$, $\eta_{p2} < .001$, and no group showed a significant ingroup bias effect, (gender diverse participants, M ingroup 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, $M = 0.34$, $t(67) = 1.68$, $p = .098$, $d = 0.20$).

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 ingroup bias score was such that each participant's "own gender" was scored in line with their answers to another question they were 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 else" were excluded (n = 26). Again, there were no differences by group, $F(2,158) = 1.52$, $p = .223$, $\eta_{p2} = .02$, and siblings (M ingroup 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 ingroup bias effect. However, gender diverse participants *did* show a significant ingroup bias effect when this approach was utilized, $M = 0.64$, $t(51) = 2.40$, $p = .020$, $d = 0.33$.

Discussion

The current studies sought to investigate whether gender diverse children differ from gender conforming children in their 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 the gender conforming peers in these studies 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 sizes for the gender diverse children (d = 0.54) are 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 (analyses are reported in the Supplemental 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 could be universal. 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. This finding is in contrast to work suggesting that transgender children differ from cisgender children in their more deliberative, conscious beliefs about gender (Fast & Olson, 2018; Olson & Enright, 2018). This dissociation between more deliberative beliefs about gender and gender encoding may indicate that different processes could be driving responses to these different types of tasks. The encoding effect may reflect automatic processing of gender or may be caused by older, evolutionarily-driven processes thought to make gender especially critical (Cosmides, Tooby, & Kurzban, 2003; Kurzban et al., 2001), 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 (i.e., gender is mentioned). As such, assessments of gender beliefs may be more affected by socialization and 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. That gender diverse children, and children at gender neutral preschools, might show equivalent levels of gender encoding but lessened stereotypes compared to their gender conforming peers and children at gender typical schools (Olson & Enright, 2018; Shutts et al., 2017) could reflect these differential processes that contribute to mental representations of gender.

In addition to examining gender encoding in the current study, we explored whether children encoded their own gender better than the other gender (what we have called their ingroup bias score). We found little evidence that this was the case in any participant group or with any method of calculating the ingroup 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 this finding 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 ingroup encoding bias in this paper.

Limitations and Suggestions for Future Research

One limitation of the current work is that the encoding task only 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 nonbinary-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 nonbinary gender, making this a particularly important area for future research.

The gender encoding task used in the current study also had very limited racial representation, 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 decisively whether or not the tendency to encode gender generalizes to targets or participants of all races, again suggesting an important direction for future work.

Conclusions

The current studies provide evidence that, as young as the 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. http://dx.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

Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences*, *7*, 173-179. https://doi.org/10.1016/S1364-6613(03)00057-3

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

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*. 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*, *98*, 15387-15392. https://doi.org/10.1073/pnas.251541498

Maccoby, E. E. (1988). Gender as a social category. *Developmental Psychology*, *24*, 755. http://dx.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., & 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., 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., 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

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. http://dx.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., & 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

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

Figure 1.

The figure shows density of parent reports of child gender conformity using the Gender Identity Questionnaire (GIQC; Johnson et al., 2004). Scoring was done in accordance with Johnson et al. (2004) such that, when calculating the composite value, 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.

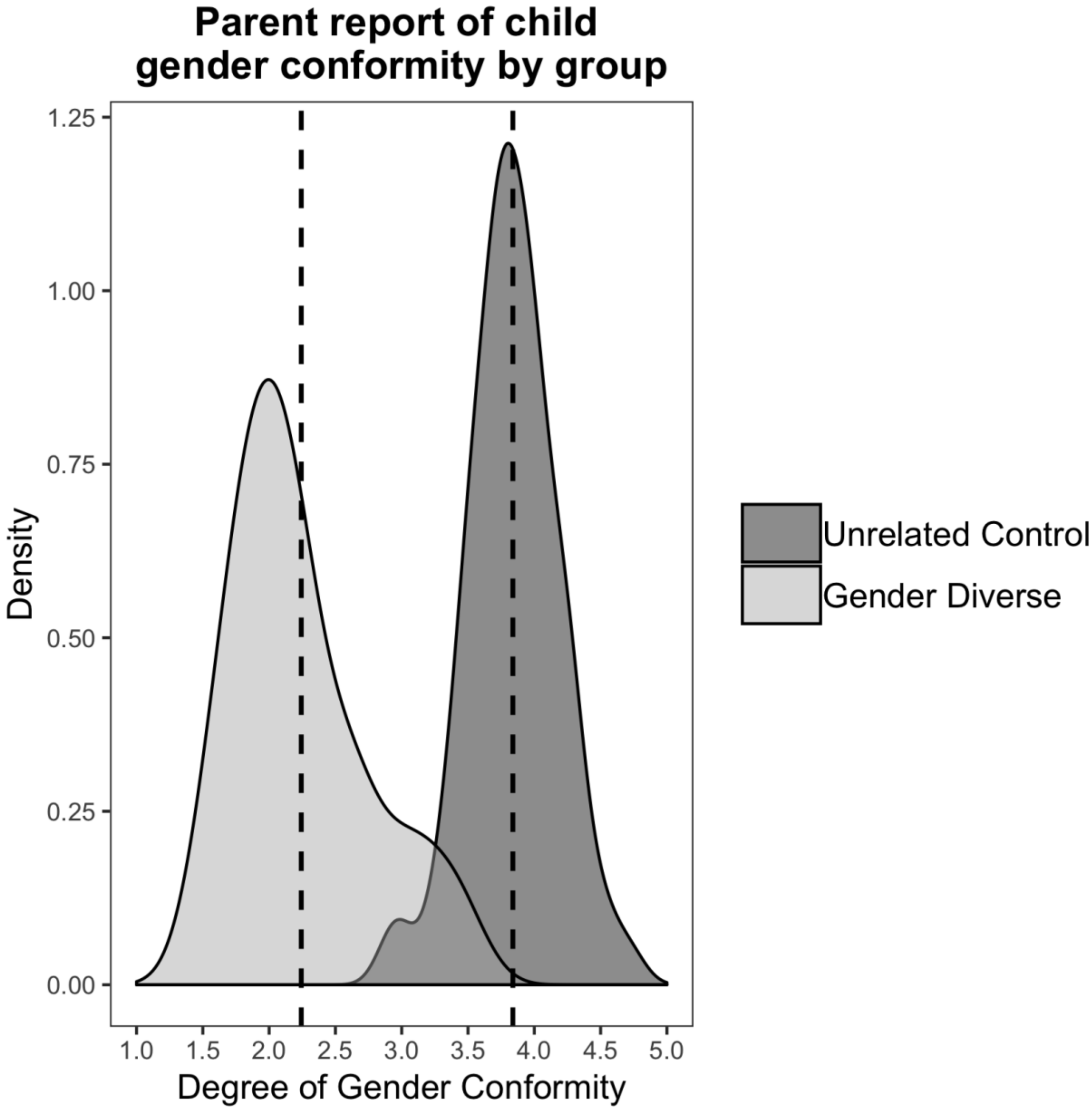
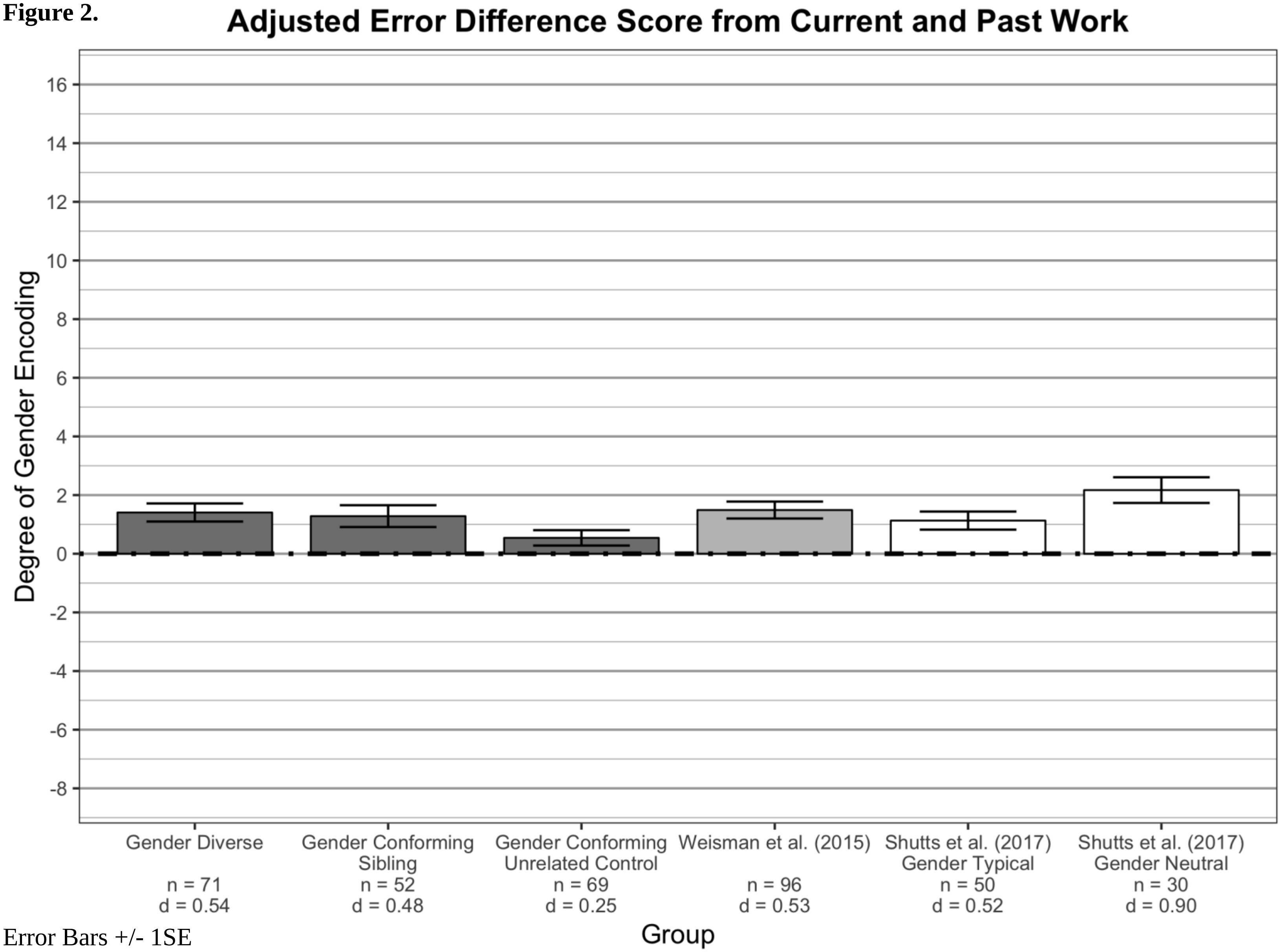


Figure 2.



Error Bars +/- 1SE

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			
Less than $25,000	2.9%	4.2%	3.8%
$25,001 to $50,000	1.4%	9.9%	9.6%
$50,001 to $75,000	14.5%	22.5%	23.1%
$75,001 to $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%

Supplemental Information

S1: Gender Encoding in Transgender and Gender Nonconforming participants separately

After data collection had already begun (164 out of 192 total participants from the main manuscript had been recruited), we registered that we would analyze data from socially transitioned transgender participants (as well as siblings of transgender children and unrelated cisgender controls). After registration of the analysis plan for transgender participants and their associated siblings and cisgender controls, we continued to collect data from gender nonconforming children and their associated siblings and controls with the intent to analyze and publish their data at a later time. However, after reviewers suggested increasing the sample size and diversity of gender identities within the transgender sample, we decided to include data from the gender nonconforming children as well as their associated siblings and unrelated cisgender controls to create a larger group of gender diverse children. Here, we report the registered analyses for the original transgender participants, siblings of transgender children, and their associated controls, separately from the gender nonconforming children, siblings of gender nonconforming children, and their associated controls.

Socially Transitioned Transgender Participants

Participants. We registered a stopping rule of 40 transgender participants who completed at least two of the four blocks of the gender encoding task used in the current study. We chose this number as we felt that it was reasonably achievable and comparable to other work including transgender preschoolers (Fast & Olson, 2018). One transgender girl (and her control) was excluded from analyses (and replaced) as she did not meet the inclusion criteria for number of trials completed and she was therefore replaced. Further, after data collection was completed, data from an additional participant (and her control) was discovered and we opted to include this participant (results are the same when only the first 40 usable participants and their controls are included in the analysis). Additionally, two control participants who were run in the larger study were dropped from this paper as they did not complete this task. Therefore, the full sample (*N* = 110; *M* age = 4 years, 11 months) included 41 transgender participants (10 transgender boys or assigned females), 30 cisgender siblings of transgender children (13 males), and 39 cisgender controls matched to the transgender participants by age (within 4 months) and gender (10 males). Children were considered transgender if they use the binary pronouns different from those they used at birth (e.g., a transgender girl or assigned male uses “she” and ”her”) in the household, school, and in meeting strangers. Both the siblings and cisgender controls used pronouns aligned with their sex assigned at birth. Participant demographics are presented in Table S1.

Results.

Gender encoding. As detailed in our registration, we conducted a single factor ANOVA and found a significant main effect of participant group, $F(2,107) = 3.97$, $p = .022$, $\eta_{p2} = .07$. Tukey comparisons showed that cisgender controls (*M* adjusted error difference score = 0.12, *SD* = 1.94) encoded gender to a lesser degree than transgender participants (*M* = 1.67, *SD* = 2.83, *p* = .017, *d* = 0.64). No other comparisons were significant. Additionally, 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). Transgender participants made significantly more same-gender errors than different-gender errors (*M* adjusted error difference score = 1.67, $t(40) = 3.77$, $p = .001$, $d = 0.59$) as did siblings (*M* = 1.07, $t(29) = 2.33$, $p = .027$, $d = 0.42$). However, cisgender controls did not make more same-gender errors than different-gender errors (*M* = 0.12, $t(38) = 0.40$, $p = .692$, $d = 0.06$).

Own-gender encoding. We examined whether children show an ingroup bias in memory, better encoding the gender of children of their own gender (as defined by the pronouns they used in everyday life) than children of the other binary gender [Importantly, for socially transitioned transgender children, like those in this study—the different approaches to "own gender" used in the main paper are not necessary as children's pronoun use matches the gender "opposite" their sex at birth. This is why only one approach to analyzing these data was necessary in the registration]. We registered and ran a single factor ANOVA and found no significant difference between groups, $F(2,107) = 0.80$, $p = .454$, $\eta_{p2} = .01$. As per our registration, we ran a one sample t-test for each group (comparing the ingroup bias score to 0, the expected value if children were responding randomly), and found that no group showed significantly better memory for own-gender than other-gender faces (transgender participants, *M* ingroup bias score = 0.54, $t(40) = 1.70$, $p = .098$, $d = 0.26$; siblings; $M = -0.07$, $t(29) = 0.16$, $p = .877$, $d = 0.03$; unrelated controls, $M = 0.33$, $t(38) = 1.25$, $p = .217$, $d = 0.20$).

Discussion. When analyzed separately, transgender children (and their siblings) encoded gender. Surprisingly, cisgender controls did not—a topic of investigation in the next study. Further, no groups showed an ingroup bias in memory.

A Follow-up Cisgender Sample

After we analyzed the data from the study described immediately above, we were concerned about the cisgender controls' results—namely that they did not appear to encode gender. This result was at odds with all past work. We therefore recruited a new, larger sample of cisgender children to investigate gender encoding in cisgender children.

Method.

Participants. We aimed to recruit 100 children for this study and our final sample included 100 cisgender children 3-5 years of age (*M* age = 5 years, 0 months; 28 males). Five additional participants were run but excluded from analyses due to experimenter error and one participant was excluded from analyses because their parent indicated that they were not cisgender. Participant demographics are presented in Table S2. All participants lived in the same metropolitan area in the Pacific Northwest as all unrelated controls in the main study and were either recruited from a database of families interested in participating and run in the lab or recruited from and run in local preschools.

Procedure, design and analytic approach. The procedure and design of the study were identical to that of the main study except that participants did not complete the Gender Encoding Task as part of a larger battery of measures; they only completed this measure. All analyses were computed in the same way except there was no omnibus test because only one group of participants was included in this sample.

Results. Participants encoded gender, meaning they made significantly more same-gender errors than different-gender errors (*M* adjusted error difference score = 1.01, $t(99) = 3.21$, $p = .002$, $d = 0.32$). We also used independent samples t-tests to compare the current sample to the socially transitioned transgender participants in the main study and to the unrelated cisgender controls associated with the transgender participants in the main study. We need to be cautious in interpreting these findings, as the current sample was run at a different time from the transgender participants and their associated controls in the main study, and the current sample did not complete other tasks in the same visit. We found that the current sample ($M = 1.01$) and the transgender participants from the main study ($M = 1.67$) did not significantly differ in the degree to which they encoded gender ($t(139) = 1.16$, $p = .247$, $d = 0.22$), while the current sample and the unrelated cisgender controls associated with the transgender participants from the main study (*M* = 0.12) did significantly differ in the degree to which they encoded gender ($t(110.8) = 2.00$, $p = .048$, $d = 0.34$).

Participants did not show significantly better memory for own-gender than other-gender faces ($M = 0.33$; $t(99) = 1.69$, $p = .095$, $d = 0.17$).

Discussion. Unlike the cisgender controls in the previous study (involving transgender participants), but similar to the transgender children and siblings of transgender children from the previous study as well as children in all past studies using this methodology, this larger sample of cisgender participants encoded the gender of the targets. These results suggest that it is possible that the lack of gender encoding in the unrelated cisgender controls in the previous study could have been a Type II error. Further evidence in favor of this conclusion comes in the following study.

Gender Nonconforming Children

Participants. The sample in the main paper included the transgender children (and siblings and controls) reported above as well as a group of 30 gender nonconforming children (10 assigned females), 22 gender conforming siblings of gender nonconforming children (9 males), and 30 gender conforming controls matched to the gender nonconforming participants by age (within 4 months) and with the "opposite" sex at birth (10 males; total $N = 82$; *M* age = 4 years, 10 months). While we had not registered this study, we had set a stopping goal of 30 participants and had initially planned to report them separately from the transgender group (as we do here in the supplement). Children were considered gender nonconforming if their parent identified them as gender nonconforming and they did not exclusively use the binary pronouns different from those they used at birth in all aspects of their life (e.g., they used a mix of pronouns, gender neutral pronouns, or exclusively used pronouns used at birth). Both the siblings and gender conforming controls used pronouns aligned with their sex assigned at birth. Participant demographics are presented in Table S3.

Results.

Gender encoding. A single factor ANOVA indicated that there was no significant main effect of participant group, $F(2,79) = 0.33$, $p = .721$, $\eta_{p2} = .01$. Additionally, one-sample t-tests (comparing the adjusted error difference score to 0, the expected value if children were responding randomly) showed that gender nonconforming participants made significantly more same-gender errors than different-gender errors (M adjusted error difference score = 1.05, $t(29) = 2.57$, $p = .016$, $d = 0.47$) as did siblings ($M = 1.57$, $t(21) = 2.52$, $p = .020$, $d = 0.54$), and gender conforming controls ($M = 1.08$, $t(29) = 2.50$, $p = .019$, $d = 0.46$).

Own-gender encoding. As in the main paper, own-gender encoding could be analyzed three different ways for gender nonconforming children. We report results to each approach here (noting that the first approach was our original plan and is in line with the registered approach used with transgender participants.

The first approach involved coding “own gender” as the binary gender “opposite” children’s sex assigned at birth. Utilizing this approach there were no differences by group, $F(2,79) = 2.40$, $p = .097$, $\eta_{p2} = .06$, and gender nonconforming participants (M ingroup bias score = -0.20, $t(29) = 0.62$, $p = .538$, $d = 0.11$) and gender conforming controls ($M = 0.27$, $t(29) = 0.85$, $p = .403$, $d = 0.15$) did not show a significant ingroup bias effect. However, siblings did show a significant ingroup bias effect ($M = 0.82$, $t(21) = 2.74$, $p = .012$, $d = 0.58$).

Our second approach involved coding “own gender” according to the pronouns participants use in everyday life (excluding participants who do not exclusively use pronouns associated with a single binary gender (*n* = 9)). According to this approach, there were no differences by group, $F(2,69) = 1.37$, $p = .260$, $\eta_{p2} = .04$, and gender nonconforming participants (*M* ingroup bias score = 0.03, $t(19) = 0.09$, $p = .927$, $d = 0.02$) and gender conforming controls ($M = 0.27$, $t(29) = 0.85$, $p = .403$, $d = 0.15$) did not show a significant ingroup bias effect. Again, siblings did show a significant ingroup bias effect ($M = 0.82$, $t(21) = 2.74$, $p = .012$, $d = 0.58$).

Our final approach to coding these items was to score "own gender" as participants' explicit self-identification on another measure at the same visit, excluding children who neither defined themselves as a boy or a girl (*n* = 13). Using this approach, there were no differences by group, $F(2,64) = 0.39$, $p = .677$, $\eta_{p2} = .01$, and gender nonconforming participants (*M* ingroup bias score = 0.40, $t(20) = 0.97$, $p = .345$, $d = 0.21$) and gender conforming controls ($M = 0.30$, $t(26) = 0.90$, $p = .375$, $d = 0.17$) did not show a significant ingroup bias effect, while siblings did show a significant ingroup bias effect ($M = 0.74$, $t(18) = 2.22$, $p = .040$, $d = 0.51$).

Discussion. Gender nonconforming children, their siblings, and gender conforming controls all significantly encoded gender. The latter group provided further evidence that the control group paired with the transgender sample may have been anomalous. We also found that the gender nonconforming and gender conforming controls did not show an ingroup memory bias, though surprisingly, and at odds with past work, the siblings did, and they did so across three different scoring techniques. We do not have a clear interpretation of this effect as it was unexpected and at odds with the other groups and with past findings, and we are cautious due to the unusually small sample size (21 or fewer participants depending on analysis approach). Until such an effect is replicated, we urge caution in interpreting these results. The findings of the gender nonconforming group and the control group are consistent with previous research (Bennett & Sani, 2003; Weisman, Johnson, & Shutts, 2015) suggesting that children do not differentially encode information about members of their gender in-group as compared to members of their gender out-group.

References

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

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

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

Table S1.

Participant Demographics for Socially Transitioned Transgender Participants and Associated Samples

	Participant Group		
	Cisgender Controls	Transgender	Cisgender Sibling
Child's Race			
White	69.2%	73.2%	63.3%
Multiracial	25.6%	22.0%	33.3%
Another Race	2.6%	4.9%	3.3%
Not Reported	2.6%	0.0%	0.0%
Household Annual Income			
Less than $25,000	2.6%	4.9%	3.3%
$25,001 to $50,000	0.0%	9.8%	13.3%
$50,001 to $75,000	12.8%	26.8%	23.3%
$75,001 to $125,000	33.3%	31.7%	33.3%
More than $125,000	46.2%	26.8%	26.7%
Not Reported	5.1%	0.0%	0.0%
Parent Education Level			
High school diploma	0.0%	0.0%	0.0%
Some college/Associate's degree	5.1%	12.2%	6.7%
College/Bachelor's degree	30.8%	26.8%	26.7%
Advanced degree (MA, MD, PhD, etc.)	35.9%	31.7%	36.7%
Not Reported	28.2%	29.3%	30.0%
Parent Political Ideology			
Liberal	71.8%	92.7%	83.3%
Moderate	23.1%	7.3%	16.7%
Conservative	2.6%	0.0%	0.0%
Not Reported	2.6%	0.0%	0.0%

Table S2.

Participant Demographics for Additional Cisgender Sample

	Additional Cisgender Sample
Child's Race	
White	66.0%
Multiracial	13.0%
Another Race	11.0%
Not Reported	10.0%
Household Annual Income	
Less than $25,000	1.0%
$25,001 to $50,000	2.0%
$50,001 to $75,000	9.0%
$75,001 to $125,000	25.0%
More than $125,000	52.0%
Not Reported	11.0%
Parent Education Level	
High school diploma	0.0%
Some college/Associate's degree	8.0%
College/Bachelor's degree	37.0%
Advanced degree (MA, MD, PhD, etc.)	45.0%
Not Reported	10.0%
Parent Political Ideology	
Liberal	55.0%
Moderate	31.0%
Conservative	1.0%
Not Reported	13.0%

Table S3.

Participant Demographics for Gender Nonconforming Participants and Associated Samples

	Participant Group		
	Gender Conforming Controls	Gender Nonconforming	Gender Conforming Sibling
Child's Race			
White	73.3%	70.0%	72.7%
Multiracial	20.0%	20.0%	13.6%
Another Race	6.7%	10.0%	4.5%
Not Reported	0.0%	0.0%	9.1%
Household Annual Income			
Less than $25,000	3.3%	3.3%	4.5%
$25,001 to $50,000	3.3%	10.0%	4.5%
$50,001 to $75,000	16.7%	16.7%	22.7%
$75,001 to $125,000	26.7%	40.0%	27.3%
More than $125,000	50.0%	30.0%	31.8%
Not Reported	0.0%	0.0%	9.1%
Parent Education Level			
High school diploma	0.0%	3.3%	4.5%
Some college/Associate's degree	13.3%	3.3%	18.2%
College/Bachelor's degree	30.0%	26.7%	9.1%
Advanced degree (MA, MD, PhD, etc.)	40.0%	50.0%	50.0%
Not Reported	16.7%	16.7%	18.2%
Parent Political Ideology			
Liberal	76.7%	80.0%	72.7%
Moderate	20.0%	20.0%	18.2%
Conservative	0.0%	0.0%	0.0%
Not Reported	3.3%	0.0%	9.1%