Gender Nonconformity of Identical Twins with Discordant Sexual Orientations: Evidence from Childhood Photographs

Tuesday M. Watts

Luke Holmes

Jamie Raines

Sheina Orbell

Gerulf Rieger

University of Essex

Author Note

Tuesday M. Watts, Luke Holmes, Jamie Raines, Sheina Orbell and Gerulf Rieger, Department of Psychology, University of Essex.

This research was supported by the American Institute of Bisexuality (RSC1466) and the University of Essex (DG00832).

We thank Joan A. W. Linsenmeier for her unparalleled advise and support.

Correspondence concerning this article should be addressed to Gerulf Rieger, Department of Psychology, University of Essex, Colchester C04 3SQ, UK. Email: gerulf@essex.ac.uk

Abstract

Childhood gender nonconformity (femininity in males, masculinity in females) predicts a non-straight (gay, lesbian, or bisexual) sexual orientation in adulthood. In previous work, non-straight twins reported more childhood gender nonconformity than their genetically identical, but straight, co-twins. However, self-reports could be biased. We therefore assessed gender nonconformity via ratings of photographs from childhood and adulthood. These ratings came from independent observers naïve to study hypotheses. Identical twins with discordant sexual orientations (24 male pairs, 32 female pairs) visibly differed in their gender nonconformity from mid-childhood, with higher levels of gender nonconformity for the non-straight twins. This difference was smaller than the analogous difference between identical twins who were concordant straight (4 male pairs, 11 female pairs) and identical twins unrelated to them who were concordant non-straight (19 male pairs, 8 female pairs). Further, twins in discordant pairs correlated in their observer-rated gender nonconformity. Non-genetic factors likely differentiated the discordant twins' gender-related characteristics in childhood, but shared influences made them similar in some respects. We further tested how recall of past rejection from others related to gender nonconformity. Rejection generally increased with gender nonconformity, but this effect varied by the twins' sexual orientation.

Keywords: sexual orientation; gender behavior; gender development; twins

Gender Nonconformity of Identical Twins with Discordant Sexual Orientations: Evidence from Childhood Photographs

Men who are more feminine, and less masculine, in their behaviors, appearances, and interests than most other men can be defined as gender nonconforming; similarly, gender nonconforming women are more masculine and less feminine than other women (Lippa, 2005). Gender nonconformity is more common in non-straight (gay, lesbian, and bisexual) than straight men and women, and this difference emerges in early childhood (Bailey & Zucker, 1995; Lippa, 2008; Rieger, Linsenmeier, Gygax, & Bailey, 2008). Another line of research suggests that across men and women, genetic variation explains approximately 30% of the differences in sexual orientation (Bailey et al., 2016). Furthermore, between 10% and 50% of the co-development of sexual orientation with gender nonconformity is explained by genetic variation (Alanko et al., 2010; Bailey, Dunne, & Martin, 2000; Burri, Cherkas, Spector, & Rahman, 2011).

Because sexual orientation and gender-related behavior are not fully determined by genetics, genetically identical twins can differ in both their sexual orientations and their level of gender nonconformity. By self-report, twins with discordant sexual orientations differ in their gender nonconformity to a degree similar to unrelated straight and non-straight individuals (Bailey & Pillard, 1991). However, self-reports may distort actual differences in gender nonconformity (Gottschalk, 2003). We therefore examined whether identical twins with discordant sexual orientations differed in their observable gender nonconformity by evaluating photographs taken in childhood and adulthood. We further compared their difference in gender nonconformity to the difference between pairs where both twins were straight and pairs unrelated to them where both twins were non-straight. Finally, we examined how social reactions during childhood related to gender nonconformity of these twins.

Sexual Orientation and Gender Nonconformity

Gay men, lesbians, and bisexual men and women report more gender-nonconforming behaviors and interests, on average, than straight adults of their sex (Lippa, 2008; Rieger, Linsenmeier, Gygax, Garcia, & Bailey, 2010; Swift-Gallant, Coome, Monks, & VanderLaan, 2017). In one meta-analysis, gay men reported more feminine and less masculine interests and self-concepts in adulthood than straight men; lesbians reported more masculine and less feminine interests and self-concepts than straight women (Lippa, 2005). These effects were large, $1.28 < d\text{'s} < 1.46$, $1.18 < 95\%\ CI < 1.56$, and $0.60 < d\text{'s} < 1.28$, $0.50 < 95\%\ CI < 1.38$, and, $.28 < d\text{'s} < 1.46$, $1.18 < 95\%\ CI < 1.56$, respectively. These differences emerge in childhood; in another meta-analysis, gay men recalled more gender nonconforming childhood behaviors and activities than straight men; the same was true for lesbians and straight women (Bailey & Zucker, 1995). These effects were also large, $d = 1.31$, $0.45 < 95\%\ CI < 3.08$, and $d = 0.96$, $0.26 < 95\%\ CI < 1.66$, respectively. Moreover, gender nonconformity in childhood predicts gender nonconformity in adulthood, suggesting some developmental stability of the trait (Rieger et al., 2008).

In adulthood, sexual orientation differences in gender nonconformity can be assessed by others based on motor behaviors, speech patterns, and physical appearances, even if displayed for only a few seconds (Johnson, Gill, Reichman, & Tassinary, 2007; Rieger et al., 2010). Observer ratings of adults confirm, in general, the link of sexual orientation with self-reported gender nonconformity. However, unlike the research on adulthood gender nonconformity, the majority of previous research on sexual orientation and childhood gender nonconformity has relied on retrospective self-reports, asking adults to recall their own childhood behavior. These retrospective reports may be subject to memory biases (Baumrind, 1995; Gottschalk, 2003), and

also be problematic simply due to their reliance on a memory system that was not fully developed at the time during which gender behavior emerged (Pillemer & White, 1989).

One way to address the limitations of retrospective work is to conduct prospective research. Studies have followed children who were referred to clinics due to extreme levels of gender nonconformity and concern over their gender identity. In adolescence and adulthood, these groups were substantially more likely to identify as bisexual, gay or lesbian, compared with individuals who were not gender nonconforming in childhood (Drummond, Bradley, Peterson-Badali, & Zucker, 2008; Green, 1985; Singh, 2012). One needs to be cautious regarding whether findings from clinical samples represent the relationship between these variables for most people. However, two longitudinal studies, using data from the general population, confirm that early gender-nonconforming behaviors (as early as 3 to 4 years old) predict a same-sex orientation in later life (Li, Kung, & Hines, 2017; Steensma, van der Ende, Verhulst, & Cohen-Kettenis, 2013). Such population-based, longitudinal work involves many logistical challenges. It takes years to conduct, needs hundreds of participants to capture enough with same-sex orientations, and can have substantial attrition rates. An alternative method is the assessment of gender nonconformity retroactively through the evaluation of information depicted in images from childhood. This method is free of the limitations of self-report (because it does not rely on subjective accounts), clinical samples (because it draws from a wider population), and longitudinal studies (because it does not take years to conduct). Research using this method also suggests that differences in observable gender nonconformity are predictive of an adulthood sexual orientation from ages 3 to 4 years onwards (Rieger et al., 2008).

Sexual Orientation of Identical Twins

Although monozygotic twins are genetically identical, their sexual orientations are not always concordant (Alanko et al., 2010; Långström, Rahman, Carlström, & Lichtenstein, 2010; Zietsch et al., 2012). Across several representative samples, and across men and women, the concordance rate was estimated as 24%; that is, about 24% of non-straight twins had a co-twin who was non-straight too (Bailey et al., 2016). In contrast, for the population as a whole, the occurrence of a non-straight sexual orientation may be as low as 3.5% across both sexes, although other estimates suggest 7% in men, and 13% in women (Gates, 2011; Savin-Williams & Vrangalova, 2013). Thus, the chance of a same-sex orientation for co-twins of non-straight twins is higher than that reported for the general population. Familial factors, including shared genes, likely influence the formation of a similar sexual orientation in these related individuals.

The aforementioned distributions of sexual orientations within identical twin pairs emphasize another pattern: A substantial proportion (76%) of identical twins who are non-straight have straight co-twins (Bailey et al., 2016). These twins are genetically identical and share many environmental influences. Thus, factors other than shared genes and shared environments, and which are unique to each twin, must account for their different sexual orientations. A study of such twin pairs could therefore point to possible effects of the unique environment on sexual orientation. Whatever these factors are, they could, in theory, also affect a correlate of sexual orientation, their degree of gender nonconformity.

Differences in Gender Nonconformity within Discordant Twin Pairs

Although identical twins with discordant sexual orientations have been part of several behavioral-genetic studies (e.g., Alanko et al., 2010), individual twins in these pairs were rarely systematically compared with respect to their degree of gender nonconformity. In one study of male pairs with discordant sexual orientations, non-straight twins reported more childhood

gender nonconformity than their straight co-twins. The average difference was similar to that for unrelated non-straight and straight men, suggesting substantial environmental (unique, non-genetic) effects on the link of sexual orientation with gender nonconformity (Bailey & Pillard, 1991). However, this finding is subject to the potential limitations of retrospective self-reports. Self-reported childhood gender nonconformity was also assessed in female identical twins with discordant sexual orientations, but no information was given on their differences or how they compared to unrelated straight and non-straight women (Bailey, Pillard, Neale, & Agyei, 1993).

To avoid the drawbacks of self-report, the present study investigated possible differences in the childhood gender nonconformity of discordant twins by examining visual images from their childhoods. For unrelated individuals, sexual orientation differences in observable gender nonconformity, based on evaluations of visual images, emerge by ages 3 to 4 and carry into adulthood (Rieger et al., 2008). If twins with discordant sexual orientations are like unrelated straight and non-straight individuals, then differences in their observable gender nonconformity could emerge at ages 3 to 4, and remain present in adulthood.

Correlation of Gender Nonconformity within Discordant Twin Pairs

Both straight and non-straight individuals vary in their degree of gender nonconformity; some score high relative to others of their sexual orientation, others score low (Rieger et al., 2010). These variations could be linked in related individuals. That is, familial influences (shared genes and/or the shared environment) could contribute to related levels of gender nonconformity. A non-straight twin who scores high on gender nonconformity, relative to other non-straight individuals, may have a straight co-twin who scores high on gender nonconformity, relative to other straight individuals. This would result in a correlation in the twins' gender nonconformity, even if they differed, on average, in their gender nonconformity.

In general, shared genes make people similar in their sexual orientation, gender behavior, and their co-expression, whereas the shared environment (any shared influences other than shared genes) does not substantially affect similarity in sexual orientation (Alanko et al., 2010; Burri et al., 2011; Kendler, Thornton, Gilman, & Kessler, 2000; Långström et al., 2010; Zietsch et al., 2012). Other work points to shared environmental influences on the expression of gender behavior (Iervolino, Hines, Golombok, Rust, & Plomin, 2005; Knafo, Iervolino, & Plomin, 2005). Thus, both shared genes and shared environments may affect twins' correlated traits.

Why, then, might twins have discordant sexual orientations, but have correlated levels of gender nonconformity? One possibility is that different timing of exposure to androgens during fetal development affect the formation of sexual orientation independently from the development of gender-related traits (Bailey & Zucker, 1995). Unique influences of androgens at one point in time could lead to twins' discordance for one trait (their sexual orientation), but shared androgen exposure at another time (and their shared genes) still result in related expression of another trait (their gender nonconformity). Thus, degree of gender nonconformity could co-develop in genetically identical individuals, even if they have discordant sexual orientations.

Although a correlation in self-reported gender nonconformity has been found for twin pairs with concordant non-straight orientations, it was not confirmed for discordant pairs (Bailey & Pillard, 1991; Bailey et al., 1993). Perhaps, discordant twins' gender-related traits are, in fact, unrelated to each other. Alternatively, failure to find a correlation in discordant pairs may have been due to the aforementioned limitations of self-report. We therefore examined these correlations both via self-report and the evaluation of gender nonconformity from photographs.

A Comparison of Discordant and Concordant Twin Pairs

A parsimonious model assumes that differences in gender nonconformity within twin pairs with discordant sexual orientations are as strong as those between unrelated individuals with different sexual orientations. Self-report supports this hypothesis (Bailey & Pillard, 1991). However, it is also possible that due to shared influences between twins of a pair, differences in gender nonconformity within discordant twin pairs may be smaller than those seen between unrelated straight and non-straight individuals. That is, either because of their shared genes or their shared environment (e.g., shared prenatal hormones or shared parental influences on their gender-typed dress, behaviors, and activities) twins within discordant pairs may be more similar to each other in their gender nonconformity than are unrelated straight and non-straight individuals. This could be the case even if non-genetic factors unique to each twin make twins within discordant pairs different to some extent in their gender nonconformity. We examined this possibility by testing within- and between-pair differences in gender nonconformity. That is, we first tested for the difference in gender nonconformity within twin pairs who were discordant for sexual orientation. We then compared this difference within discordant pairs to the difference in gender nonconformity between identical twins from pairs with concordant straight sexual orientations and identical twin pairs unrelated to them with concordant non-straight sexual orientations. Because being an identical twin was held constant across participants, any differences across twin types could be more easily interpreted with respect to the twins' similar or dissimilar sexual orientations.

We also examined whether, as previously seen in self-reports, the correlation in gender nonconformity is weaker in discordant twin pairs than concordant pairs (Bailey & Pillard, 1991; Bailey et al., 1993), or whether correlations are more similar across types of pairs if observer-ratings are used in order to examine such correlations.

Responses to Gender Nonconformity

Gender-nonconforming children can experience negative reactions and rejection from others (Maccoby, 1998; Smith & Leaper, 2006). Peers can react negatively to gender nonconformity, and boys are especially critical about gender nonconformity in other boys (Wallien, Veenstra, Kreukels, & Cohen-Kettenis, 2010; Young & Sweeting, 2004). Similarly, parents can be more detached and unsupportive if their children are gender nonconforming (Alanko et al., 2009; Landolt et al., 2004). There are exceptions to these reactions. In one study on very feminine boys, who likely became non-straight, the most feminine boys had mothers who responded less negatively, if not more positively, to their sons' gender nonconformity (Green, 1987). However, the overall conclusion across research is that more gender-nonconforming children experience more negative reactions; moreover, gender nonconformity and negative parental reactions may, in part, reinforce each other (Alanko et al., 2011).

Furthermore, because childhood gender nonconformity predicts adulthood gender nonconformity, adulthood gender nonconformity is also linked to past rejection from others (Rieger et al., 2008). We therefore expected that in twins, higher levels of gender nonconformity in childhood and adulthood would relate to recall of increased rejection.

The Present Study

We recruited twin pairs with either discordant or concordant sexual orientations to assess their gender nonconformity via ratings of photographs and self-report. We examined:

1) Whether non-straight twins in discordant pairs were more gender nonconforming than their straight co-twins.

2) Whether the difference in gender nonconformity within discordant pairs was smaller than the difference between concordant non-straight and unrelated concordant straight twins.

3) Whether, for discordant twin pairs, gender nonconformity of non-straight twins correlated with the gender nonconformity of their straight co-twins.

4) Whether the correlation of co-twins' levels of gender nonconformity was similar for discordant twin pairs and concordant twin pairs.

5) Whether, across discordant and concordant twins, higher levels of gender nonconformity related to recall of greater rejection by others.

Method

The University of Essex's Ethics Committee approved this study. The title of the approved project was "Gender nonconformity and sexual orientation: Genetic and environmental influences," with approval number "GR1303."

Recruitment and Participants

Twins. Advertisements for identical twins to participate in a study on sexual orientation were placed in the newsletter of the Department of Twin Research at Kings College London, on social media sites, and on three online news sites for gay men and lesbians (Gay Star News, Pink News, and Gay Times). We further recruited twins at three gay festivals. Each twin who contacted us was encouraged to recruit the co-twin. Twins self-identified as straight, bisexual, gay, or lesbian. They were asked twice during the study about their sexual identities, and all responses were consistent. The number of bisexual men and women (6 and 10 individuals) was low relative to the number of straight men and women (32 and 54) or gay men and lesbians (56 and 38). Furthermore, on 7-point Kinsey Scales (Kinsey, Pomeroy, & Martin, 1948), straight participants reported exclusive or almost exclusive preferences for the other sex, gay men and lesbians reported exclusive or almost exclusive preferences for the same sex, and bisexual

participants reported a stronger preference for the same sex than the other sex. For this reason, bisexual participants were grouped with gay men and lesbians into "non-straight".

For twins with discordant sexual orientations, 24 pairs were male and 32 pairs were female. Four male pairs were concordant straight, 19 male pairs were concordant non-straight, 11 female pairs were concordant straight, and 8 were concordant non-straight (Table 1). The number of male pairs who were concordant straight was low compared to the other categories. Analyses indicated that these men differed from concordant non-straight twins (and from discordant twins) in interpretable ways. In this sense, their smaller number did not appear to be problematic; however, in the Discussion, we speculate about why their numbers were so low.

Raters. Psychology students participated as raters of gender nonconformity for course credit. Twenty-five men self-identified as straight, 13 men as non-straight, 77 women as straight, and 17 women as non-straight. The higher proportion of women reflects that most psychology students at our institution are women. Ratings of gender behavior are minimally affected by the raters' sex or sexual orientation (Rieger et al., 2010). This was also the case for present ratings (see below), and different sizes of rater groups did not appear to affect our findings.

Self-Report Measures

Twin zygosity. In addition to asking twins whether they were identical, five standardized items about physical and visual similarity were administered to establish zygosity; such measures have at least 95% accuracy (Kasriel & Eaves, 1976; Martin & Martin, 1975). An example question is "During childhood, could you ever have fooled friends by pretending to be your twin?" To be considered monozygotic twins were expected to score below 2, on average.

Gender nonconformity. The Childhood Gender Nonconformity Scale measured self-reported gender nonconformity during childhood (Rieger et al., 2008). Both men and women

were given seven items. Example statements are "As a child I preferred playing with girls rather than boys" for males, and "As a child I often felt that I had more in common with boys than girls" for females. Gender nonconformity in adulthood was assessed with seven items for each sex with the Continuous Gender Identity Scale (Rieger et al., 2008). Example statements include "My mannerisms are more feminine than those for most men of my age" for men and "I assume most people see me as more masculine than other women" for women. For all measures, items were endorsed on 7-point scales, with higher scores representing higher levels of childhood or adulthood gender nonconformity. For both discordant and concordant twins of either sex, internal reliability (Cronbach's α) exceeded .84 for both childhood and adulthood measures.

Parental and peer rejections. Parental rejection was assessed using a 10-item version of the Recollection of Early Childrearing Scale (Ross, Campbell, & Clayer, 1982), measuring rejection versus acceptance by each parent during childhood. Example items include "My (mother/father) wished I had been like somebody else" and "I think my parent was mean and grudging toward me." Responses of peers during childhood were assessed using ten items from the Mother-Father-Peer Scale (Epstein, 1983). Sample statements include "When I was a child, other children were often unfair to me" and "When I was a child, other children often picked on me." Items were rated on 7-point scales, with higher scores representing higher levels of rejection. Cronbach's α for all three measures of rejection exceeded .72 for both discordant and concordant twins of either sex. Across twins, the three measures were correlated, p's < .0001, .40 < r's < .52, .30 < 95% CI's < .62. In addition to separate results for maternal, paternal, and peer rejection, we also report results for a composite of "overall rejection" across these measures.

Procedure

Collecting data and photographs. Approximately one fourth of individual twins (47/196) completed the following procedure in the lab; the remainder competed it remotely. An online link to the self-report measures was given or emailed to each twin. It was stressed that answers should not be discussed with co-twins or others. Furthermore, participants were asked to arbitrarily select and send, at a minimum, one to three photographs from any available age from childhood into adulthood. This included photographs that were recently taken. There was no upper limit on the number of photographs. The context of the photographs varied. Most common were close-ups of the individual, followed by photographs taken at school, birthday events, and holidays. Twins were asked to identify themselves (and their co-twin, if depicted) in each photograph. The majority of photographs were cropped before they were rated by observers, so that only the individual twin was shown (without the co-twin or other people).

Twins as children. During childhood (defined as 0-15 years) the number of photographs of individuals within pairs was almost the same. Using pairs as units, the mean (SD) number of photographs was 14.86 (5.30), 12.06 (3.47), and 10.08 (3.47) for discordant males, concordant straight males, and concordant non-straight males, respectively. The mean (SD) age in years in these photographs was 7.63 (5.02), 6.34 (4.93), and 6.60 (5.02), respectively. The mean (SD) number of photographs of female twins was 12.37 (3.54), 14.15 (5.20), and 9.14 (2.11) for discordant pairs, concordant straight pairs, and concordant non-straight pairs. Mean (SD) ages in these photographs were 7.73 (5.01), 8.57 (4.76), and 6.89 (5.07), respectively.

Twins as adults. Ages in photographs from adulthood ranged from 16 (the legal age of consent in the UK) to 27. The number of photographs was almost the same for twins within a pair. With pairs as units, the average (SD) number of photographs was 12.02 (4.95), 10.13 (4.22), and 9.85 (3.50) for discordant, concordant straight, and concordant non-straight males,

respectively. Mean (SD) ages in the photographs were 21.86 (2.67), 22.50 (1.07), and 21.95 (1.99), respectively. For females, the average (SD) number of photographs was 12.92 (3.84), 14.80 (5.57), and 10.00 (1.87) for discordant, concordant straight, and concordant non-straight twins. Mean (SD) ages were 22.15 (2.31), 21.33 (1.72), and 21.62 (2.60), respectively.

Ratings. Photographs were shown in the stimulus presentation software Inquisit. We sorted the photographs into 8 sets: 2 sets showing female children, 2 with male children, 2 with adult females, and 2 with adult males. For both discordant and concordant pairs, photographs of each twin were in separate sets than those of the co-twin.

In the lab, each rater viewed 4 sets of photographs (males as children, females as children, males as adults, females as adults) showing the same individuals in childhood and adulthood, but never their co-twins. That is, raters never viewed twin pairs together. Raters were neither told that those shown in the photographs were twins, nor that they varied in sexual orientation. The order in which sets of photographs were presented was counterbalanced across raters. Within each set, photographs were presented for 3 seconds each and in random order.

We used rating procedures that were similar to those that have previously resulted in reliable sexual orientation differences in observer-rated gender nonconformity (Rieger et al., 2008). Raters of photographs were instructed to indicate their impression of each individual's appearance and demeanor in comparison to their impression of most people of this age and sex. For example, after each photograph of a boy, they were told to "rate whether this boy appeared or behaved in a more feminine or masculine way". Ratings were completed on 7-point scales. For photographs of males, the score of 1 was "more masculine", 4 was "average," and 7 was "more feminine." A reversed scale was used for ratings of females. Thus, for both sexes, 7 represented maximum gender nonconformity. Raters were told that they might see the same

person more than once (because individuals provided several photos) and encouraged to judge photographs independently of each other. Ratings for each set took 10 to 20 minutes.

Straight male raters gave lower mean (SD) evaluations of gender nonconformity, 3.51 (0.98), than non-straight males, 3.65 (1.07); straight females, 3.60 (1.07), and non-straight females, 3.59 (1.09), were similar. Although a significant difference, its magnitude was minimal, $p = .004$, $R^2 = .002$, 95% CI [.000, .004]. Furthermore, average evaluations from the four groups of raters correlated strongly with each other, p's < .0001, .73 < r's < .89, .71 < 95% CI's < .90. Across all raters, inter-rater reliability (Cronbach's α) exceeded .93 for each combination of twin type (discordant or concordant), sex, and period (childhood or adulthood). Hence, for each photograph, an average rating of gender nonconformity was computed across all raters.

Results

Do Twins of Discordant Pairs Differ in Gender Nonconformity?

Within discordant pairs, non-straight twins could be viewed as more gender nonconforming than their straight co-twins. For each sex, a mixed-factorial regression analysis was conducted. The dependent variable was observer-rated gender nonconformity, across depicted ages (0 to 27 years). Independent variables were twins' sexual orientation and the age at which a photograph was taken. An interaction between sexual orientation and age tested whether differences in observer-rated gender nonconformity between straight and non-straight twins changed with age. Twin pairs and individuals were included as random effects to account for dependency of the data within pairs and repeated evaluations of each individual across ages.

For males, non-straight twins were rated as more gender nonconforming than their straight co-twins, $p < .0001$, β [95% CI] = .23 [.15, .31]. The interaction of sexual orientation with age indicated that their difference in gender nonconformity increased with age, $p < .0001$, β

= .16 [.08, .24] (Figure 1A). For females, non-straight twins were also rated as more gender nonconforming than their straight co-twins, $p < .0001$, $\beta = .23$ [.15, .30], and an interaction indicated that their difference increased with age, $p < .0001$, $\beta = .16$ [.09, .24] (Figure 1B).

Figure 1A further shows that for males, the confidence intervals for gender nonconformity of straight twins and their non-straight co-twins separated at 8 years. For females, the separation was at age 6 (Figure 1B). Hence, differences in observer-rated gender nonconformity between twins with discordant sexual orientations became significant these ages.

Analyses thus far focused on differences in observer-rated gender nonconformity as a function of age in the photographs. In the following, we examined main effects of sexual orientation on gender nonconformity for both observer ratings and self-report. These mixed-factorial analyses (including twins and individuals as random effects) were broken down by the two age periods (childhood and adulthood). In childhood, non-straight twins were rated as more gender nonconforming than their straight co-twins. This was found for males, $p = .006$, $\beta = .15$ [.04, .25], and females, $p = .0004$, $\beta = .16$ [.07, .24]. This difference was stronger for their self-reported childhood gender nonconformity, $p = .0002$, $\beta = .58$ [.31, .85], and $p = .0001$, $\beta = .46$ [.24, .68], respectively. In adulthood, sexual orientation effects on observer ratings for males and females, $p < .0001$, $\beta = .40$ [.30, .50], and $p < .0001$, $\beta = .46$ [.32, .61], were similar to effects on self-reported gender nonconformity, $p = .004$, $\beta = .43$ [.16, .71], and $p < .0001$, $\beta = .40$ [.22, .59]. Figure 2 shows a simple illustration of these effects. On a between-group level, Figure 2 makes it appear as if some sexual orientation differences in observer-rated gender nonconformity are not significant; however, within twin pairs all differences were significant, as described above.

Are Differences in Gender Nonconformity Smaller within Discordant Pairs than Between Concordant Pairs?

Shared factors could make sexual orientation differences in gender nonconformity within discordant pairs smaller than analogous differences in gender nonconformity between unrelated straight and non-straight individuals. To test this possibility, we compared the magnitude of within-pair differences in gender nonconformity to the magnitude of such differences between unrelated pairs. That is, these comparisons involved twins from pairs who were concordant straight and unrelated twins from pairs who were concordant non-straight.

We first focused on observer ratings of concordant pairs. We conducted mixed-factorial regression analyses for concordant twins similar to those previously described for discordant twins. For concordant male twin pairs, non-straight twins were rated as more gender nonconforming than unrelated straight twins, $p = .001$, $\beta = .58$ [.27, .89]. An interaction of sexual orientation with age suggested that this difference between straight twins and unrelated non-straight twins became stronger with age, $p = .0004$, $\beta = .14$ [.07, .23]. Figure 3A shows that straight twins becoming less gender nonconforming with age, whereas non-straight twins kept an intermediate level. This difference in gender behavior became significant between 1 and 2 years.

For concordant females, non-straight twins were rated as more gender nonconforming than unrelated straight twins, $p = .006$, $\beta = .39$ [.14, .65]. The interaction with age, $p = .06$, $\beta = .09$ [-.00, .18], indicated that differences between straight and unrelated non-straight twins in observer-rated gender nonconformity became stronger with age. This difference became significant at 3 years (Figure 3B).

In a subsequent analysis, we compared this difference in gender nonconformity in concordant twins to that in discordant twins. In a mixed-factorial regression analysis, an interaction between sexual orientation and pair type indicated that the effect of sexual orientation on observer-rated gender nonconformity varied between twins with discordant or concordant

sexual orientations, $p = .02$, $\beta = .13$ [.02, .24]. This interaction can be interpreted by comparing Figure 1 with Figure 3: Across males and females, the average difference in observer-rated gender nonconformity was weaker, and emerged at a later age, within discordant pairs as compared to this difference between concordant non-straight and straight pairs.

We then examined, similar to the previous comparisons within discordant pairs, main effects of sexual orientation on gender nonconformity for concordant pairs, looking at both observer ratings and self-report. For childhood, effects of sexual orientation on observer-rated gender nonconformity of males and females, $p = .004$, $\beta = .66$ [.26, 1.06], and $p = .03$, $\beta = .40$ [.05, .74], respectively, were similar to effects on self-reported gender nonconformity, $p = .006$, $\beta = .50$ [.16, .84], and $p = .0008$, $\beta = .61$ [.29, .92], respectively. For adulthood, the effects of sexual orientation on observer-rated gender nonconformity of males and females, $p = .005$, $\beta = .47$ [.17, .77], and $p = .005$, $\beta = .55$ [.20, .91], were somewhat stronger than those on self-reported gender nonconformity, $p = .04$, $\beta = .36$ [.02, .70], and $p = .02$, $\beta = .46$ [.09, .83]. Figure 4 is a simple illustration of these findings. This pattern varied from that of identical twins with discordant sexual orientations, who were, at least in their childhood, less different in their observer-rated than self-reported gender nonconformity. Moreover, in discordant pairs, the difference by measures appeared to be driven by straight twins under-reporting gender nonconformity, compared to how others rated them; for their non-straight co-twins a difference by measure was not as apparent (Figure 2). However, additional analyses (not shown here) suggested that the stronger discrepancy between measures of gender nonconformity in straight twins than non-straight twins was true for all straight twins and not specific to those who were discordant or concordant.

Are Twins of Discordant Pairs Correlated for Gender Nonconformity?

We next examined, for twins with discordant sexual orientations, whether the siblings were correlated in their degree of gender nonconformity. We conducted a mixed-factorial regression analysis, which tested for the correlation of gender nonconformity within pairs. Simultaneously, we tested for differences by sex, age period (childhood and adulthood), and measure (observer ratings or self-report) in this correlation. On average, the correlation of gender nonconformity differed significantly by measure, $p < .0001$, $\beta = .15$ [.10, .21]. When broken down by measure, the correlations were stronger for observer ratings than for self-report, $p < .0001$, $\beta = .47$ [.36, .58], and $p = .06$, $\beta = .19$ [-.01, .38], respectively. However, this difference by measure was further moderated by sex and age period. Table 2 shows that in male discordant pairs, and across age periods, observer ratings of their gender nonconformity were more strongly correlated than their self-reports. In females, stronger correlations for observer ratings than self-reports were found in childhood, but the effect was reversed in adulthood.

Are Correlations of Gender Nonconformity Similar for Discordant and Concordant Pairs?

In our next set of analyses we compared the correlation of gender nonconformity within discordant twin pairs to this correlation within concordant pairs. The focus, then, was the degree to which the siblings' correlation in gender nonconformity varied by twin type.

In the previous section, we reported correlations for discordant pairs. Next, we computed correlations of gender nonconformity for concordant pairs. We combined, for each sex, straight and non-straight pairs, because some pair numbers were low (Table 1). To ensure that correlations in gender nonconformity were not enhanced by grouping concordant straight and non-straight twins, we partialled out the effect of their sexual orientation. The computed correlations suggested that concordant pairs reported related levels of gender nonconformity, and that observers' ratings of them were similarly related (Table 3). In contrast, for discordant pairs,

correlations in gender nonconformity were generally stronger for observer ratings than for self-report (Table 2). A mixed-factorial regression analysis suggested that this difference in siblings' correlations, depending on twin type and measure, was significant, $p = .004$, $\beta = .07$ [.02, .11].

Does Gender Nonconformity Relate to Rejection?

Our final prediction was that across discordant and concordant twins, higher levels of gender nonconformity related to recall of increased rejection by others.

We first examined whether our measures of gender nonconformity could be collapsed into a composite score. Across individual twins of either type or sex, the four measures of gender nonconformity (observer ratings and self-report from childhood and adulthood) were significantly related, even after the effect of sexual orientation was partialled out, $p's \leq .0002$, $.13 < \beta's < .54$, $.06 < 95\%$ CI's $< .64$. Furthermore, across measures of gender nonconformity (observer ratings and self-report from childhood and adulthood), relationships with recalled rejection were similar in effect. For the sake of simplicity, we used a composite score of each twin's gender nonconformity across measures in the following analyses.

Across all twins, four mixed-factorial regression analyses were conducted. Each of the measures of rejection (from mothers, fathers, peers, and overall) was predicted by the twins' gender nonconformity, sexual orientation, twin type, and their interactions. We further tested whether these effects differed by sex. Twin pairs were a random effect.

Table 4 shows that, in general, greater gender nonconformity was linked to increased rejection. However, the interaction of gender nonconformity with sexual orientation was significant (or close to significant) for maternal, peer, and overall rejection: the relationship of gender nonconformity with rejection was stronger in magnitude for straight twins than non-straight twins. For example, for overall rejection, the separate effects for straight and non-straight

twins were $p = .0003$, $\beta = .57$ [.27, .87], and $p = .02$, $\beta = .20$ [.02, .38], respectively. Further, there was a tendency for an interaction among gender nonconformity, sexual orientation, and twin type: the link of gender nonconformity to rejection was strongest for straight twins of concordant pairs. For example, for overall rejection the effect for concordant straight twins was stronger than for the combination of all other twins, $p = .003$, $\beta = .73$ [.28, 1.19], and $p = .004$, $\beta = .26$ [.08, .44], respectively. There were no significant sex differences in these effects.

Discussion

Based on evaluations of photographs, non-straight twins were rated as more gender nonconforming than their identical, but straight, co-twins. This difference emerged in childhood, but later than the difference between concordant non-straight twins and unrelated concordant straight twins, and it was also smaller in magnitude. Twin siblings' levels of observer-rated gender nonconformity were correlated, and the correlation was similar for pairs with discordant and concordant sexual orientations. These patterns were somewhat different for their self-reports, partly because straight twins reported less gender nonconformity than what was observed by raters. Finally, in general, gender nonconformity related to recall of past rejection from others.

Possible Reasons for Differences between Twins

Using a within-pair design, present findings suggest that identical twins with discordant sexual orientations visibly differed in their gender nonconformity from mid-childhood on (Figure 1). Given that certain types of information are not available in photographs, and that identical twins look the same in many ways, the detection of a difference in gender nonconformity between straight twins and their identical, but non-straight, co-twins seems especially striking. Factors other than genetics must account for this difference. These factors could be prenatal. About 30% of identical twin pairs develop with different placentas, which could lead to varied

hormonal exposure and different epigenetic effects (Patterson, 2007). Both of these factors could be relevant for their development of different degrees of gender nonconformity and discordant sexual orientations, as it is possible that epigenetic mechanisms mediate long-term effects of hormonal expose on the formation of sexual orientation (Ngun & Vilain, 2014).

Different social influences could also lead to discordant sexual orientations and gender behaviors in identical twins. However, evidence for socialization effects on the origin of sexual orientation is weak (Bailey et al., 2016; Rahman, 2005). Our data suggested that, across all twins, gender nonconformity related to self-reports of past rejection from others (Table 4). In this sense, it is unlikely that the social environment encouraged one twin over the other to become gender nonconforming (or, subsequently, non-straight). However, rejection from others was more strongly related to gender nonconformity of those who became straight than those who became non-straight. It is difficult to understand how this finding might relate to the development of discordant twins, since it seemed to be somewhat driven by the effect found in concordant straight twins. Yet, some speculation might be insightful. One study suggested that boys who were particularly gender nonconforming had mothers who were more tolerant of these behaviors (Green, 1987). Perhaps, in some cases, reactions to gender nonconformity are less negative, for example, if parents conclude that they cannot change these behaviors because they notice that the child will become non-straight. In contrast, children who are expected to become straight and to follow traditional gender norms may experience harsher reactions if they exhibit gender nonconformity. Such differential reactions could foster gender nonconformity in children who become non-straight, whereas they might suppress it in children who become straight.

Possible Reasons for Similarity of Twins

Identical twins with discordant sexual orientations were in two respects similar in their observer-rated gender nonconformity. First, their differences were generally smaller, compared to the difference between concordant straight and unrelated concordant non-straight twins (Figures 1 & 3). Secondly, discordant twins correlated, in general, in their level of observer-rated gender nonconformity (Table 2). That is, those non-straight twins who were more gender nonconforming than other non-straight twins tended to have straight co-twins who were more gender nonconforming that other straight co-twins.

The present research does not provide information on the exact influences that made twins similar to each other. Given previous research, approximately 30% of the variation in gender nonconformity is explained by genetic variation (Burri et al., 2011), suggesting that shared genes can lead to a shared expression of gender nonconformity. Moreover, variations in genes contribute between 10% and 50% to the covariance of sexual orientation with gender nonconformity (e.g., Alanko et al., 2010; Burri et al., 2011). Thus, the twins' shared genes could make them similar in their gender nonconformity, but this must not result in identical sexual orientations. As aforementioned, it is also possible that different exposures to androgens during early development resulted in the twins' discordant sexual orientations, whereas similar exposure to androgen during later development, in addition to shared genes, yielded a similarity in their gender nonconformity (Bailey & Zucker, 1995).

Socialization could also contribute to similarity in gender behavior. Identical twins possibly share environmental influences that are unique to their twinship (Iervolino et al., 2005). For example, parents may emphasize their similarity by dressing or posing them in a coordinated manner. Twins themselves may coordinate their appearances (e.g., by wearing similar clothing) and behaviors (Kendler & Gardner, 1998; Rose, Kaprio, Williams, Viken, & Obremski, 1990).

Perhaps, then, one twin who was gender nonconforming in childhood, and became non-straight, influenced the straight co-twin to be more gender nonconforming; and vice versa, the straight twin affected the non-straight co-twin to express less gender nonconformity. These influences could result in a reduced distinction between the twins and a stronger coordination of their gender-related traits.

Discordant twins differentiated significantly in their gender nonconformity between ages 6 and 8, and later than straight versus non-straight concordant pairs, who differentiated between ages 1 and 3 (Figures 1 & 3). This is consistent with the hypothesis that discordant twins are more similar to each other than unrelated individuals. Yet, one must be careful in interpreting exact ages for these differentiations. First, in discordant pairs, the confidence intervals for these differentiations included the ages of 3 to 4, the suggested age for such differentiation between unrelated straight and non-straight individuals (Li et al., 2017; Rieger et al., 2008). Second, for male concordant pairs, this differentiation in gender nonconformity was earlier than the suggested age of 3, and whether this unexpectedly early differentiation is robust is unclear.

Observer Ratings and Self-Report

The discussion thus far has focused on observer ratings of gender nonconformity. Findings were somewhat different for self-reported gender nonconformity, which showed larger differences between discordant twins, and were less correlated within discordant pairs, compared with observer ratings. These findings for self-report were comparable to previous results based on self-reports from discordant twins (Bailey & Pillard, 1991; Bailey et al., 1993). Within discordant pairs, average discrepancies between observer ratings and self-report were more pronounced in straight than non-straight twins (Figure 2). Because these straight twins have closely related siblings (i.e., twins) who are non-straight, it is tempting to suggest that, perhaps

unaware to them, they have a predisposition to a same-sex orientation that has not fully developed, but is expressed in its correlate - their visible gender nonconformity. Yet, this discrepancy between observer ratings and self-report was also found among concordant straight twins. For straight twins in concordant pairs, this difference by measure cannot be explained based on an unaware predisposition to a co-twins' non-straight sexual orientation, since their co-twins identify as straight too. A more reasonable interpretation is that straight participants, independent of whether they had non-straight relatives, were less likely than non-straight participants to realize or acknowledge their level of gender nonconformity (Gottschalk, 2003).

The Utility and Limitations of Using Photographs

The present study used an uncommon method to assess gender nonconformity – via the evaluations of photographs. Of course, photographs do not encompass all facets of gender nonconformity. For one, some aspects of self-perceived gender nonconformity, including feelings and self-concepts, may not be captured at all in photographs, whereas they may be effectively covered with self-report. Furthermore, other-perceived gender nonconformity includes not only appearance but also gesture and posture, voice patterns, expressed interests, and activities (Rieger et al., 2008; Rieger et al., 2010). The images used did not provide sound, and gesture, posture, or activities, if evident, were static. For this reason, we had also asked twins to provide videos. However, only one pair provided one video from adulthood, whereas all twins readily provided photographs taken at several ages. We compensated for this loss of information in photographs by asking participants for several photographs across the ages. In fact, twins provided, on average, almost five times more photographs from their childhood (12.00 per participant, on average) than participants provided childhood videos in a previous study (2.50 per participant, on average, Rieger et al., 2008). Likewise, in adulthood, the average number of

photographs (11.67) greatly exceeded the one video taken from each participant in the previous study. Hence, the several photographs of the same twin potentially increased the chance to detect reliable signs of his or her gender nonconformity.

Furthermore, the several facets of gender nonconformity (appearance, movements, voice patterns, interests, and activities) are modestly related (Rieger et al., 2010). Perhaps for this reason (in addition to the several photographs per participant) present effects of sexual orientation on observer-rated gender nonconformity, using photographs, were similar to the previously reported effects using ratings of videos (which included sound and movements, Rieger et al., 2008). For example, for unrelated straight and non-straight individuals, the main effect of sexual orientation on observer-rated gender nonconformity from ages 0 to 15 was comparable, $p = .001$, $\beta = .37$ [.17, .59], and $p < .0001$, $\beta = .38$ [.23, .53], respectively.

Still, we do not know what exact features raters focused on. Informal discussions with raters indicated that they evaluated appearance (e.g., clothes or make up), behaviors (e.g., posture, facial expressions, and body language), and activities (e.g., handled toys). They relied more strongly on appearances, because behaviors and activities was less easily detectible or not present. Several components of appearance may be influential. A gender-nonconforming self-presentation may, for example, include a short haircut on non-straight females or stylish dress for non-straight males (e.g., Krakauer & Rose, 2002). Non-straight individuals can also be sex-reversed in sex-dimorphic features such as the structure of the nose or chin (Skorska, Geniole, Vrysen, McCormick, & Bogaert, 2015) or in their height, bone structure, and body shape (Martin & Nguyen, 2004). It is therefore possible that in the present study, style, morphology, and anatomy all contributed to evaluations of gender nonconformity from photographs.

Another limitation is relevant particularly for our analyses of correlations. Identical twins are very similar in their overall appearance, especially when they are young. Perhaps correlations between siblings in their observer-rated gender nonconformity were inflated because the ratings reflect assessments of not only gender nonconformity but also other aspects of appearance. We aimed to avoid such a confound. Raters were unaware that they were evaluating twins, rating each twin of a pair independently, without seeing both members of the pair, and with explicit instructions to focus on masculinity and femininity. Because observer ratings related to the twins' self-reported gender nonconformity, these observer ratings seemed valid on a general level. However, this relationship with self-report was not perfect. Hence, we cannot rule out that observer ratings were, to some level, affected by unwanted effects, and that unknown aspects of appearance, which made co-twins similar, inflated correlations of observer-ratings within pairs.

On a different note, the Internet made it easy to share photographs, without most twins being required to visit the lab or use physical mail. This was a true benefit since this research focused on a rare group, with several twins living far from the lab, and we wanted to minimize attrition. In sum, we believe that evaluations of photographs offer an effective mechanism for studying the link of sexual orientation with observable gender behavior.

Further Limitations

Another limitation of the present study was the small number of male twins who were concordant straight (4 pairs). Statistically, several interpretable effects involving these twins were detected. Moreover, comparisons of such twins with discordant twins have not been made before. In this respect, the inclusion of these twins was informative, despite their low numbers. However, one should consider why their numbers were low. In our experience, straight males are less interested in participating in research on sexual orientation, compared with other groups.

Furthermore, with our recruitment methods (e.g., advertising for participation on gay news sites) straight twins are less easily reached than non-straight twins. We aimed to compensate for this by further recruitment via a twin registry and social media sites, but here, too, the advertisement pointed to a study on sexual orientation. In discordant pairs, non-straight twins may have motivated the straight co-twins to participate. In concordant straight pairs, encouragement by a co-twin seems less likely. This limitation could be overcome by advertising broadly for research on twins without disclosing the nature of the study. Yet, this would likely yield small numbers of the rarer groups, concordant-non-straight and discordant twins. Ethically, this approach is also problematic, especially when participants share personal information, including photographs.

Another limitation is a potential participant bias in selecting photographs. For example, if participants felt the need to confirm stereotypes associated with their sexual orientation group, straight twins may have predominantly chosen photographs that made them appear gender conforming, and non-straight twins may have predominantly chosen images that made them appear gender nonconforming. In order to minimize the risk that such bias affected findings, we asked participants to send any photographs, regardless of quality and content, and we did not explicitly tell participants that this study was on gender nonconformity and sexual orientation.

A further limitation was the measure of zygosity. We used questions which have previously shown at least 95% accuracy in assessing zygosity when compared with genotyping (Kasriel & Eaves, 1976; Martin & Martin, 1975). For the majority of individual twins, their average scores were below 2, suggesting monozygosity. However, individual twins from three pairs who identified as “identical” scored 2 or higher, suggesting dissimilarity with their co-twin. For these pairs, zygosity was re-assessed by re-contacting both twins, by observing them when

visiting our lab, and by examining their photographs, to confirm that they were likely identical. A superior method to assess zygosity would be DNA analyses, for example, from saliva samples.

Another limitation concerns statistical power. No significant differences in observer-rated gender nonconformity were found in discordant twins for ages 3 to 4, the predicted age range for such differentiation. For discordant twins and this age range, we had collected 22 photographs from straight males, 23 from non-straight males, 34 from straight females, and 35 from non-straight females. Because non-straight twins appeared slightly more gender nonconforming than their straight co-twins at that age, although not significantly so, perhaps a larger sample of photographs would elicit significant differences. Notably, for concordant twins at this age, we had 13 photographs from straight males, 27 from non-straight males, 18 from straight females, and 13 from non-straight females. Hence, concordant twins provided fewer photographs at that age than discordant twins, on average; still, significant differences emerged by that age for them.

Conclusion

Identical twins with discordant sexual orientations visibly differed in their gender nonconformity, starting in childhood. Moreover, the method used - the evaluation of gender-related characteristics seen in photographs - pointed to subtle similarities of these twins, possibly because of shared influences. With the increased collection of data through twin registries, it may become possible to examine these patterns longitudinally in more detail and to identify factors that affected the twins' similar and dissimilar development. Thus, further studies of twins with discordant sexual orientations, which go beyond self-report, will provide a unique window into the developmental of sexual orientation.

Table 1.

Distribution of Numbers, Age, and Ethnicities by Twin Type (Twins with Discordant or Concordant Sexual Orientations) and Sex.

Males	Discordant	Concordant Straight	Concordant Non-Straight
Number of Pairs	24	4	19
Average Age	31.08 [26.21, 35.96]	23.00 [20.40, 25.60]	31.42 [26.71, 36.13]
Percentage Caucasian	85 [67.52, 94.08]	50 [15.00, 84.99]	82 [58.97, 93.81]
Females	Discordant	Concordant Straight	Concordant Non-Straight
Number of Pairs	32	11	8
Average Age	29.38 [26.10, 32.65]	28.45 [23.62, 33.29]	27.75 [20.60, 34.90]
Percentage Caucasian	91 [77.04, 96.95]	100 [75.75, 100.00]	70 [39.68, 89.22]

Note. Units are pairs. Numbers in brackets are 95% confidence intervals.

Table 2.

Correlations within Discordant Twin Pairs for Observer-Rated and Self-Reported Gender Nonconformity in Childhood and Adulthood.

Measure	Discordant Males	Discordant Females
Observer-rated Childhood	.70*** [.52, .87]	.54*** [.37, .71]
Observer-rated Adulthood	.63* [.22, 1.03]	.17 [-.19, .54]
Self-Reported Childhood	-.18 [-.62, .26]	.12 [-.24, .49]
Self-Reported Adulthood	.00 [-.45, .46]	.42* [.08, .76]

Note. Numbers are standardized regression coefficients, *β's*, with 95% confidence intervals in brackets. For observer ratings, twins were included as a random effect to account for repeated measures within pairs. Higher scores indicate stronger correlations (technically, regression coefficients) of gender nonconformity between twins. *$p < .05$. ***$p < .0001$.

Table 3.

Correlations within Concordant Twin Pairs for Observer-Rated and Self-Reported Gender Nonconformity in Childhood and Adulthood.

Measure	Concordant Males	Concordant Females
Observer-rated Childhood	.57*** [.38, .76]	.67*** [.50, .84]
Observer-rated Adulthood	.55** [.24, .89]	.71* [.21, 1.20]
Self-Reported Childhood	.56* [.15, .98]	.58** [.16, .99]
Self-Reported Adulthood	.41† [-.09, .91]	.46* [.01, .92]

Note. Numbers are standardized regression coefficients, *β's*, with 95% confidence intervals in brackets. Concordant straight and concordant non-straight pairs are combined and the effect of sexual orientation is partialled out. For observer ratings, twins were included as a random effect to account for repeated measures within pairs. Higher scores indicate stronger correlations (technically, regression coefficients) of gender nonconformity between twins. $^{\dagger}p < .10$. $*p < .05$. $**p < .001$. $***p < .0001$.

Table 4.

Multiple Regression Analyses for Gender Nonconformity, Sexual Orientation, and Twin Type predicting Degree of Rejection across 196 Individuals.

Measure	Maternal Rejection[1]	Paternal Rejection[1]	Peer Rejection[1]	Overall Rejection[1]
Gender Nonconformity (GN)[2]	.33 [.11, .55]*	.21 [.00, .43]*	.38 [.18, .59]***	.40 [.21, .60]***
Sexual Orientation (SO)[3]	-.08 [-.29, .13]	.04 [-.17, .26]	-.14 [-.34, .06]	-.08 [-.27, .12]
GN X SO	-.19 [-.42, .03]†	-.17 [-.39, .05]	-.29 [-.49, -.08]*	-.26 [-.45, -.06]*
Twin Type (TT)[4]	-.11 [-.33, .12]	.04 [-.2, .29]	.12 [-.1, .34]	.04 [-.18, .26]
GN X TT	.04 [-.18, .27]	.09 [-.14, .32]	.17 [-.05, .38]	.11 [-.09, .32]
SO X TT	-.10 [-.32, .12]	-.10 [-.33, .13]	-.10 [-.31, .11]	-.11 [-.31, .10]
GN X SO X TT	-.09 [-.33, .15]	-.14 [-.37, .09]	-.23 [-.45, -.01]*	-.20 [-.41, .01]†

Note. R^2*'s* for the four models are .12, .10, .18, and .16, respectively. Numbers are standardized regression coefficients, *β's*, with 95% confidence intervals in brackets. [1]Higher scores indicate more rejection. [2]Higher scores indicate more gender nonconformity, derived from a composite across all measures. [3]A score of 0 indicates "straight," 1 indicates "non-straight". [4]A score of 0 indicates "discordant," 1 indicates "concordant." Moderations by sex were not significant and are not shown. Twin pairs were a random effect. †$p < .10$. *$p < .05$. ***$p < .0001$.

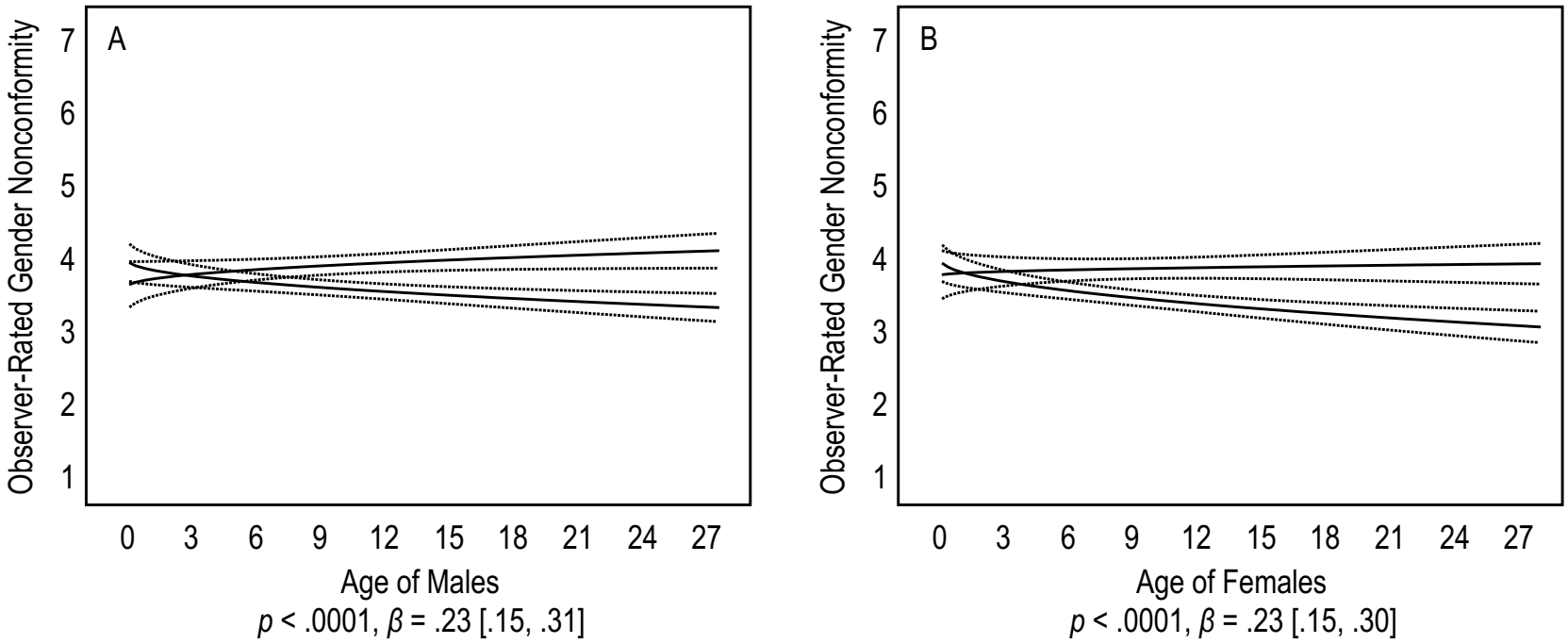

Figure 1. Observer-rated gender nonconformity from photographs of A) 24 male twin pairs and B) 32 female twin pairs with discordant sexual orientations. For each photo, ratings of gender nonconformity were averaged across raters. Upper and lower triple-lines represent regression coefficients with 95% confidence intervals for non-straight twins and straight co-twins, respectively. The x-axis represents the twins' age. On the y-axis, 1 is the least, and 7 the most gender-nonconforming score. Estimates are restricted to the age of 0 or older. Statistics represent the main effect of sexual orientation on gender nonconformity.

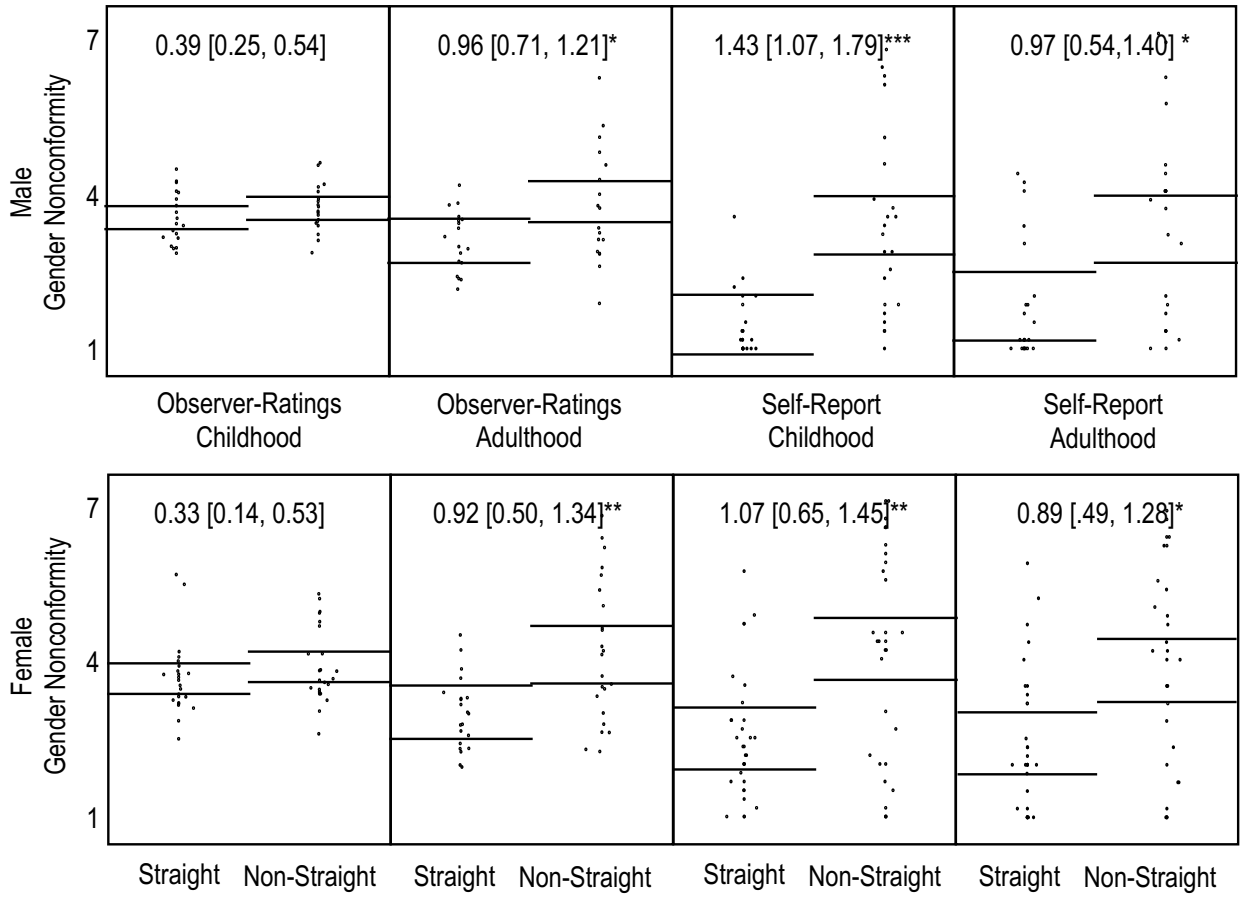

Figure 2. Observer-rated and self-reported gender nonconformity of 24 male and 32 female twin pairs with discordant sexual orientations. Dots are gender nonconformity scores of individuals, averaged across all ratings or their self-report. Lines are the means' 95% confidence intervals. On the y-axis, 1 is the least, and 7 the most gender-nonconforming score. Numbers are Cohen's *ds* with 95% confidence intervals for between-group comparisons. All effects are significant once data dependencies within pairs are accounted for. * $p < .05$. ** $p < .001$. *** $p < .0001$.

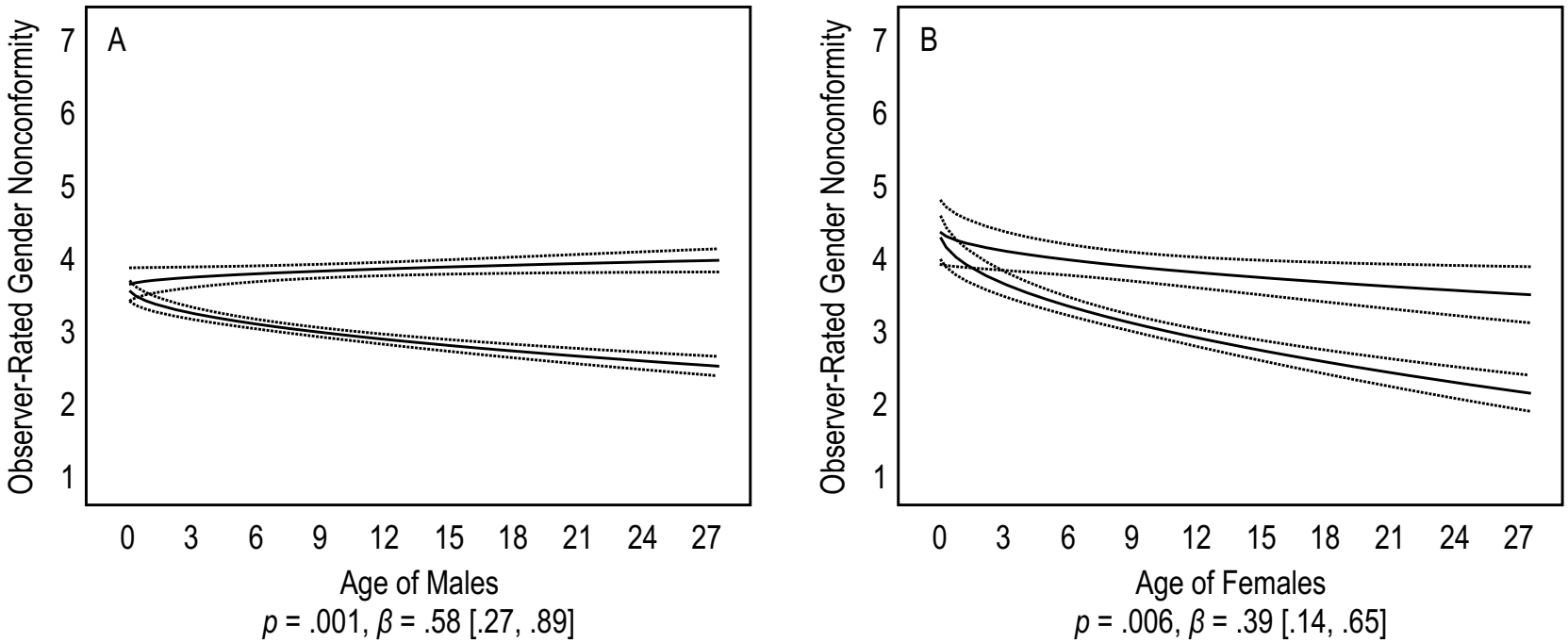

Figure 3. Observer-rated gender nonconformity from photographs of A) 23 male twin pairs and B) 19 female twin pairs with concordant straight or concordant non-straight sexual orientations. For each photo, ratings of gender nonconformity were averaged across raters. Upper and lower triple-lines represent regression coefficients with 95% confidence intervals for non-straight twins and unrelated straight twins, respectively. The x-axis represents the twins' age. On the y-axis, 1 is the least, and 7 the most gender-nonconforming score. Estimates are restricted to the age of 0 or older. Statistics represent the main effect of sexual orientation on gender nonconformity.

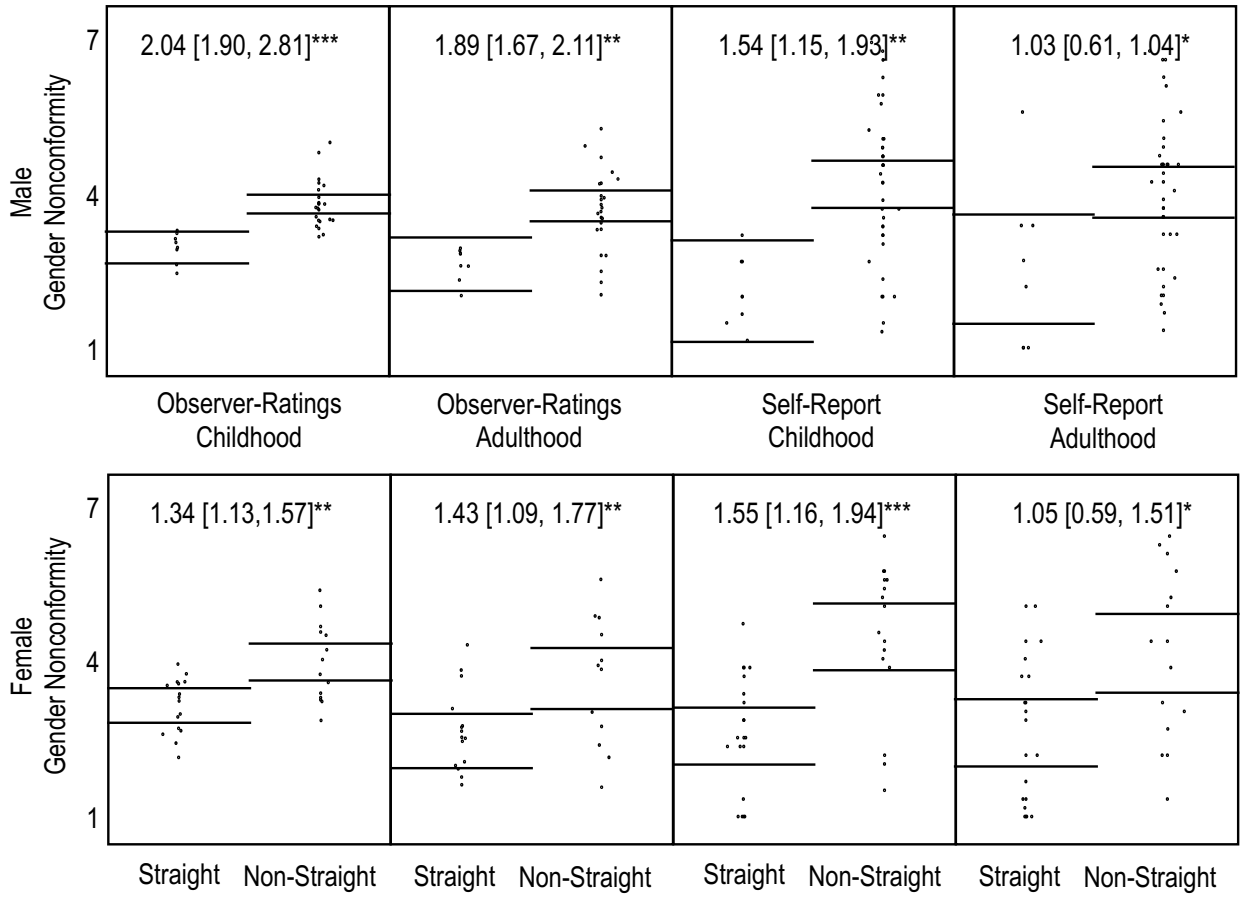

Figure 4. Observer-rated and self-reported gender nonconformity of 23 male and 19 female twin pairs with concordant straight or concordant non-straight sexual orientations. Dots are gender nonconformity scores of individuals, averaged across all ratings or their self-report. Lines are the means' 95% confidence intervals. On the y-axis, 1 is the least, and 7 the most gender-nonconforming score. Numbers are Cohen's *ds* with their 95% confidence intervals for between-group comparisons. * $p < .05$. ** $p < .001$. *** $p < .0001$.

References

Alanko, K., Santtila, P., Harlaar, N., Witting, K., Varjonen, M., Jern, P., et al. (2010). Common genetic effects of gender atypical behavior in childhood and sexual orientation in adulthood: A study of Finnish twins. *Archives of Sexual Behavior, 39*, 81-92. doi: 10.1007/s10508-008-9457-3

Alanko, K., Santtila, P., Salo, B., Jern, P., Johansson, A., & Sandnabba, N. K. (2011). Testing causal models of the relationship between childhood gender atypical behaviour and parent–child relationship. *British Journal of Developmental Psychology, 29*, 214-233. doi: 10.1348/2044-835X.002004

Alanko, K., Santtila, P., Witting, K., Varjonen, M., Jern, P., Johansson, A., et al. (2009). Psychiatric symptoms and same-sex sexual attraction and behavior in light of childhood gender atypical behavior and parental relationships. *Journal of Sex Research, 46*, 494-504. doi: 10.1080/00224490902846487

Bailey, J. M., Dunne, M. P., & Martin, N. G. (2000). Genetic and environmental influences on sexual orientation and its correlates in an Australian twin sample. *Journal of Personality and Social Psychology, 78*, 524-536. doi: 10.1037/0022-3514.78.3.524

Bailey, J. M., & Pillard, R. C. (1991). A genetic study of male sexual orientation. *Archives of General Psychiatry, 48*, 1089-1096. doi: 10.1001/archpsyc.1991.01810360053008

Bailey, J. M., Pillard, R. C., Neale, M. C., & Agyei, Y. (1993). Heritable factors influence sexual orientation in women. *Archives of General Psychiatry, 50*, 217-223. doi: 10.1001/archpsyc.1993.01820150067007

Bailey, J. M., Vasey, P. L., Diamond, L. M., Breedlove, S. M., Vilain, E., & Epprecht, M. (2016). Sexual orientation, controversy, and science. *Psychological Science in the Public Interest, 17*, 45-101. doi: 10.1177/1529100616637616

Bailey, J. M., & Zucker, K. J. (1995). Childhood sex-typed behavior and sexual orientation: A conceptual analysis and quantitative review. *Developmental Psychology, 31*, 43-55. doi: 10.1037/0012-1649.31.1.43

Baumrind, D. (1995). Commentary on sexual orientation: Research and social policy implications. *Developmental Psychology, 31*, 130-136. doi: 10.1037/0012-1649.31.1.130

Burri, A., Cherkas, L., Spector, T., & Rahman, Q. (2011). Genetic and environmental influences on female sexual orientation, childhood gender typicality and adult gender identity. *PLoS ONE, 6*, e21982. doi: 10.1371/journal.pone.0021982

Drummond, K., Bradley, S. J., Peterson-Badali, M., & Zucker, K. J. (2008). A follow-up study of girls with gender-identity disorder. *Developmental Psychology, 44*, 34-45. doi: 10.1037/0012-1649.44.1.34

Epstein, S. (1983). *The mother-father-peer scale.* Amherst, MA.

Gates, G. J. (2011). How many people are lesbian, gay, bisexual and transgender? *The Williams Institute*, Retrieved from: https://escholarship.org/uc/item/09h684x682.

Gottschalk, L. (2003). Same-sex sexuality and childhood gender non-conformity: A spurious connection. *Journal of Gender Studies, 12*, 35-50. doi: 10.1080/0958923032000067808

Green, R. (1985). Gender identity in childhood and later sexual orientation: Follow-up of 78 males. *American Journal of Psychiatry, 142*, 339-341. doi: 10.1176/ajp.142.3.339

Green, R. (1987). *The "sissy boy syndrome" and the development of homosexuality*. New Haven, CT: Yale University Press.

Iervolino, A. C., Hines, M., Golombok, S. E., Rust, J., & Plomin, R. (2005). Genetic and environmental influences on sex-typed behavior during the preschool years. *Child Development, 76*, 826-840. doi: 10.1111/j.1467-8624.2005.00880.x

Johnson, K. L., Gill, S., Reichman, V., & Tassinary, L. G. (2007). Swagger, sway, and sexuality: Judging sexual orientation from body motion and morphology. *Journal of Personality and Social Psychology, 93*, 321-334. doi: 10.1037/0022-3514.93.3.321

Kasriel, J., & Eaves, L. (1976). The zygosity of twins: Further evidence on the agreement between diagnosis by blood groups and written questionnaires. *Journal of Biosocial Science, 8*, 263-266. doi: 10.1017/S0021932000010737

Kendler, K. S., & Gardner, C. O. (1998). Twin studies of adult psychiatric and substance dependence disorders: are they biased by differences in the environmental experiences of monozygotic and dizygotic twins in childhood and adolescence? *Psychological Medicine, 28*, 625-633.

Kendler, K. S., Thornton, L. M., Gilman, S. E., & Kessler, R. C. (2000). Sexual orientation in a U.S. national sample of twin and nontwin sibling pairs. *American Journal of Psychiatry, 157*, 1843-1846. doi: 10.1176/appi.ajp.157.11.1843

Kinsey, A. C., Pomeroy, W. B., & Martin, C. E. (1948). *Sexual behavior in the human male*. Philadelphia, PA: Saunders.

Knafo, A., Iervolino, A. C., & Plomin, R. (2005). Masculine girls and feminine boys: Genetic and environmental contributions to atypical gender development in early childhood. *Journal of Personality and Social Psychology, 88*, 400-412. doi: 10.1037/0022-3514.88.2.400

Krakauer, I. D., & Rose, S. M. (2002). The impact of group membership on lesbians' physical appearance. *Journal of Lesbian Studies, 6*, 31-43. doi: 10.1300/J155v06n01_04

Landolt, M. A., Bartholomew, K., Saffrey, C., Oram, D., Perlman, D., & Bartholomew, K. (2004). Gender nonconformity, childhood rejection, and adult attachment: A study of gay men. *Archives of Sexual Behavior, 33*, 117-128. doi: 10.1023/B:ASEB.0000014326.64934.50

Långström, N., Rahman, Q., Carlström, E., & Lichtenstein, P. (2010). Genetic and Environmental Effects on Same-sex Sexual Behavior: A Population Study of Twins in Sweden. *Archives of Sexual Behavior, 39*, 75-80. doi: 10.1007/s10508-008-9386-1

Li, G., Kung, K. T. F., & Hines, M. (2017). Childhood gender-typed behavior and adolescent sexual orientation: A longitudinal population-based study. *Developmental Psychology*. doi: 10.1037/dev0000281

Lippa, R. A. (2005). Sexual orientation and personality. *Annual Review of Sex Research, 16*, 119-153. doi: 10.1080/10532528.2005.10559831

Lippa, R. A. (2008). Sex differences and sexual orientation differences in personality: Findings from the BBC Internet survey. *Archives of Sexual Behavior, 37*, 173-187. doi: 10.1007/s10508-007-9267-z

Maccoby, E. E. (1998). *The two sexes: Growing up apart, coming together*. Cambridge, MA: Belknap Press.

Martin, J. T., & Nguyen, D. H. (2004). Anthropometric analysis of homosexuals and heterosexuals: Implications for early hormone exposure. *Hormones and Behavior, 45*, 31-39. doi: 10.1016/j.yhbeh.2003.07.003

Martin, N. G., & Martin, P. G. (1975). The inheritance of scholastic abilities in a sample of twins I. Ascertainment of the sample and diagnosis of zygosity. *Annals of Human Genetics, 39*, 213-218. doi: 10.1111/j.1469-1809.1975.tb00124.x

Ngun, T. C., & Vilain, E. (2014). The biological basis of human sexual orientation: Is there a role for epigenetics? In D. Yamamoto (Ed.), *Epigenetic shaping of sociosexual interactions: From plants to humans* (pp. 167-184). Waltham, MA: Academic Press.

Patterson, P. H. (2007). Maternal effects on schizophrenia risk. *Science, 318*, 576. doi: 10.1126/science.1150196

Pillemer, D. B., & White, S. H. (1989). Childhood events recalled by children and adults. *Advances in Child Development and Behavior, 21*, 297-340. doi: 10.1016/S0065-2407(08)60291-8

Rahman, Q. (2005). The neurodevelopment of human sexual orientation. *Neuroscience & Biobehavioral Reviews, 29*, 1057-1066. doi: 10.1016/j.neubiorev.2005.03.002

Rieger, G., Linsenmeier, J. A. W., Gygax, L., & Bailey, J. M. (2008). Sexual orientation and childhood gender nonconformity: Evidence from home videos. *Developmental Psychology, 44*, 46-58. doi: 10.1037/0012-1649.44.1.46

Rieger, G., Linsenmeier, J. A. W., Gygax, L., Garcia, S. C., & Bailey, J. M. (2010). Dissecting "gaydar": Accuracy and the role of masculinity-femininity. *Archives of Sexual Behavior, 39*, 124-140. doi: 10.1007/s10508-008-9405-2

Rose, R. J., Kaprio, J., Williams, C. J., Viken, R., & Obremski, K. (1990). Social contact and sibling similarity: Facts, issues, and red herrings. *Behavior Genetics, 20*, 763-778. doi: 10.1007/BF01065919

Ross, M. W., Campbell, R. L., & Clayer, J. R. (1982). New inventory for measurement of parental rearing patterns: An English form of the EMBU. *Acta Psychiatrica Scandinavica, 66*, 499-507. doi: 10.1111/j.1600-0447.1982.tb04508.x

Savin-Williams, R. C., & Vrangalova, Z. (2013). Mostly heterosexual as a distinct sexual orientation group: A systematic review of the empirical evidence. *Developmental Review, 33*, 58-88. doi: 10.1016/j.dr.2013.01.001

Singh, D. (2012). *A follow-up study of boys with gender idenity disorder.* (PhD), University of Toronto, Toronto.

Skorska, M. N., Geniole, S. N., Vrysen, B. M., McCormick, C. M., & Bogaert, A. F. (2015). Facial structure predicts sexual orientation in both men and women. *Archives of Sexual Behavior, 44*, 1377-1394. doi: 10.1007/s10508-014-0454-4

Smith, T. E., & Leaper, C. (2006). Self-perceived gender typicality and the peer context during adolescence. *Journal of Research on Adolescence, 16*, 91-103. doi: 10.1111/j.1532-7795.2006.00123.x

Steensma, T. D., van der Ende, J., Verhulst, F. C., & Cohen-Kettenis, P. T. (2013). Gender variance in childhood and sexual orientation in adulthood: A prospective study. *The Journal of Sexual Medicine, 10*, 2723–2733. doi: 10.1111/j.1743-6109.2012.02701.x

Swift-Gallant, A., Coome, L. A., Monks, D. A., & VanderLaan, D. P. (2017). Handedness is a biomarker of variation in anal sex role behavior and recalled childhood gender nonconformity among gay men. *PLoS ONE, 12*, e0170241. doi: 10.1371/journal.pone.0170241

Wallien, M. S. C., Veenstra, R., Kreukels, B. P. C., & Cohen-Kettenis, P. T. (2010). Peer group status of gender dysphoric children: A sociometric study. *Archives of Sexual Behavior, 39*, 553-560. doi: 10.1007/s10508-009-9517-3

Young, R., & Sweeting, H. (2004). Adolescent bullying, relationships, psychological well-being, and gender-atypical behavior: A gender diagnosticity approach. *Sex Roles, 50*, 525-537. doi: 10.1023/B:SERS.0000023072.53886.86

Zietsch, B. P., Verweij, K. J. H., Heath, A. C., Madden, P. A. F., Martin, N. G., Nelson, E. C., et al. (2012). Do shared etiological factors contribute to the relationship between sexual orientation and depression? *Psychological Medicine, 42*, 521-532. doi: 10.1017/S0033291711001577