

Maternally Rated Childhood Gender Nonconformity in Homosexuals and Heterosexuals

J. Michael Bailey, Ph.D.,^{1,3} Joseph S. Miller, M.S.,¹ and Lee Willerman, Ph.D.²

Homosexual and heterosexual subjects provided self-ratings of childhood gender nonconformity. Additionally, their mother rated them on several adjectives describing childhood behavior, which included words related to gender nonconformity. Male homosexuals were remembered by their mothers as less masculine and more nonathletic. This finding did not appear to be due to a bias in mothers' memories. Though female homosexuals were recalled as more masculine than female heterosexuals, this appeared to reflect retrospective bias, as mothers who knew of their daughters' homosexuality were more likely to rate them as masculine. Both self-rated and maternally rated childhood gender nonconformity made independent contributions in predicting sexual orientation. Within the homosexual samples, maternal and self-ratings of subjects' childhood gender nonconformity failed to correlate significantly.

KEY WORDS: homosexuality; gender nonconformity; sexual orientation.

INTRODUCTION

Retrospective self-reports of childhood gender nonconformity have been associated repeatedly with adult homosexuality (Whitam, 1977; Bell *et al.*, 1981; Pillard, 1991). On average, homosexual males recall more effeminate behaviors during childhood; and homosexual females recall more masculine behaviors, relative to heterosexuals of the same sex. Ross (1980) suggested that retrospective distortion may contribute importantly to the

¹Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, Illinois 60208-2710.

²University of Texas, Austin, Texas.

³To whom correspondence should be addressed.

aforementioned association because of (i) unintentionally false reports due to poor memory, (ii) intentionally false reports because of implicit demand characteristics, and (iii) unintentionally distorted reports due to a theory-driven response style. Ross found evidence for a theory-driven response style using an American homosexual sample; however, a Swedish sample failed to exhibit the predicted effect. Although prospective studies by Green (1987) and Zuger (1984) clearly demonstrate that extensive effeminacy in childhood is strongly predictive of adult male homosexuality, the unusual degree of effeminacy of subjects in those studies necessitates caution in drawing inferences about the bulk of homosexual men. Thus despite the prospective studies, work that examines the validity of retrospective ratings of childhood gender nonconformity remains desirable.

Although homosexual and heterosexual samples differ appreciably in self-rated childhood gender nonconformity, many homosexuals deny such a history. Bell *et al.* (1981) suggested different etiologies for gender conforming versus gender nonconforming homosexuality. To investigate such hypotheses, it is necessary to have measures of childhood gender nonconformity that validly differentiate among homosexuals. Even if differences in self-reported childhood gender nonconformity between homosexuals and heterosexuals accurately reflect group differences, differences among homosexuals in childhood gender nonconformity may also exist.

We examine here the convergent validity of self-ratings of childhood gender nonconformity vis-à-vis their correspondence with maternal ratings. We focus primarily on two questions: First, are maternal ratings of childhood gender nonconformity associated with subjects' sexual orientation; and second, what are the relationships among self and maternal recollections of childhood gender nonconformity?

METHOD

Subjects

The subjects consisted of 58 heterosexual (Kinsey fantasy scores ≤ 1) males, 83 bisexual and homosexual (Kinsey fantasy scores ≥ 3) males, 51 heterosexual females, and 19 bisexual and homosexual females, as well as their mothers. Subjects include all those studied by Bailey *et al.* (1991) whose Kinsey fantasy scores did not equal 2, and whose mothers completed ratings of subjects' childhood personalities. (Four male and 3 female Kinsey 2s were excluded because of problems in classifying their orientations.) For a detailed explanation of subjects recruitment see Bailey *et al.* (1991). For brevity, we refer to both homosexual and bisexual subjects as "homosex-

Table I. Mean Subject Characteristics

	Male				Female			
	Heterosexual (<i>n</i> = 58)		Homosexual (<i>n</i> = 83)		Heterosexual (<i>n</i> = 51)		Homosexual (<i>n</i> = 19)	
	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD
Age	19.7	2.9	24.3	5.3	19.8	3.4	23.9	6.2
Kinsey fantasy	0.2	0.4	5.1	0.7	0.3	0.4	4.1	0.9
Kinsey behavior	1.1	0.3	5.3	1.1	1.0	0.6	4.2	2.2
Self-rated childhood gender nonconformity	4.6	1.0	8.4	2.5	6.1	2.2	8.6	2.9

ual." Table I gives characteristics of each subsample. Both male and female homosexuals were significantly older than their heterosexual counterparts (male, $t = 6.0$; female, $t = 3.6$; both $ps < 0.001$). This was not problematic, however, since after controlling for sexual orientation, age showed low and statistically nonsignificant correlations with all the variables investigated below.

Instruments

Each subject answered four items concerning childhood gender nonconformity (henceforth, CGN), adapted from Whitam (1977). The male items ask whether before age 13, the subject (i) was considered a sissy; (ii) wished he had been a girl; (iii) preferred playing with girls; and (iv) ever dressed in female clothes. The items were altered in an attempt to make them appropriate for female subjects. It should be noted, however, that some of the female versions of the items may not be equally good indicators of CGN. For example, the female versions of the "sissy" item substituted "tomboy." Tomboyism is considerably more common among girls than effeminacy among boys (Green, 1987). A total score for self-rated CGN was obtained by adding all four responses. Coefficient alpha for the self-rated scale was .78 for male and .71 for female. Table I contains the mean self-rated CGN scores for the four subgroups. As with the majority of past studies, there are large, significant differences between homosexuals and heterosexuals in self-rated CGN (male, $t = 12.5$; female, $t = 6.3$; both $ps < 0.001$).

Mothers were asked to rate subjects on 11 attributes of childhood personality, where childhood was defined as before age 13. For each item,

mothers were to rate their children on a scale from 1 to 5, with 5 indicating that the attribute was very accurate in describing their child's early behavior, and 1 indicating that the attribute's opposite was more applicable. Attributes included both relevant (e.g., "masculine" and "athletic") and irrelevant (e.g., "introverted") descriptors. The irrelevant descriptors were originally included in order to hide the nature of the study from mothers, since they were not told that the study concerned sexual orientation. The irrelevant items are pertinent here to the issue of discriminant validity, since they should show smaller differences as a function of sexual orientation than the theoretically relevant items.

RESULTS

Maternally Rated CGN and Sexual Orientation

Table II contains the mean maternal ratings for the 11 items, by sex and sexual orientation. Within each sex, the *t* values for the items provide an indication of the relative size of the differences between homosexual and heterosexual. The largest rating differences for males were for Masculine and Nonathletic, with homosexual subjects rated less Masculine and Nonathletic by their mothers. Smaller, though still significant differences were found for Healthy and Poorly adjusted, with homosexual subjects rated as less Healthy and more Poorly adjusted. The results for men thus support the validity of the association between sexual orientation and CGN, with the largest differences occurring precisely on those items that are theoretically most relevant.

Despite the smaller sample of female homosexuals (with the associated decrement in statistical power), there were even more significant differences. The largest differences were for Nonconforming, Masculine, and Submissive, with homosexuals rated more Nonconforming, more Masculine, and less Submissive. Smaller differences were found for Poorly adjusted, Healthy, and Passive, with homosexuals rated as more Poorly adjusted, less Healthy, and more Passive.

Differences between heterosexual males and females were also examined. Interestingly, only the adjective Masculine yielded a significant difference, and that difference was almost 3 SD ($t = 14.8$; $p < 0.001$).

For the analyses of maternally rated CGN which follow, it was desirable to use the most reliable and valid measures. For males, the adjectives Masculine and Nonathletic have both face validity and construct validity (i.e., differed significantly between heterosexuals and homosexual). Furthermore, the two items were moderately correlated ($r = .39$, $p <$

Table II. Maternal Ratings for Childhood Behavior of Homosexual and Heterosexual Males and Females^a

Item	Male					Female				
	Heterosexual		Homosexual		t	Heterosexual		Homosexual		t
	\bar{x}	SD	\bar{x}	SD		\bar{x}	SD	\bar{x}	SD	
Masculine	4.25	0.8	3.47	0.9	-5.20 ^c	1.96	0.8	2.68	1.1	2.99 ^c
Nonathletic	2.02	1.5	3.66	1.2	7.29 ^c	2.34	1.3	2.68	1.9	0.87
Submissive	2.40	1.0	2.73	1.2	1.80	2.55	0.8	1.79	1.1	-3.16 ^c
Nonconforming	2.35	1.2	2.46	1.2	0.52	2.47	0.8	3.16	1.1	2.83 ^c
Poorly adjusted	1.23	1.2	1.64	1.1	2.17 ^b	1.23	0.9	1.79	1.1	2.23 ^b
Low activity level	1.82	1.0	2.15	1.2	1.72	2.06	0.9	2.00	1.2	-0.22
Passive	2.61	1.0	2.86	1.2	1.31	2.75	0.9	2.16	1.4	-2.07 ^b
Calm	2.97	1.1	2.77	1.2	-1.01	3.08	1.0	2.74	1.2	-1.22
Rude	1.05	1.0	1.22	1.0	1.01	1.40	0.9	1.42	1.0	0.10
Healthy	4.18	0.9	3.73	1.2	-2.37 ^b	4.38	0.8	3.89	1.2	-1.98
Introverted	2.21	1.1	2.38	1.1	0.87	2.31	0.9	2.00	1.4	-1.10

^aHigher scores indicate that the adjective is more characteristic of the person.

^b $p < 0.05$.

^c $p < 0.01$.

0.001). The internal consistency reliability (coefficient alpha) for the resulting two-item scale, .56, was modest. For females, only Masculine had both face validity and construct validity as a measure of CGN per se. This adjective was the only one selected because none of the other maternally rated descriptors correlated significantly with Masculine for the females. Because female maternally rated CGN comprised only one item, its reliability was probably low.

Possible Rating Bias

It might be objected that mothers may also be susceptible to retrospective bias. If they hold a stereotype of homosexuals as gender atypical, then mothers may recall their children in accordance with this stereotype regardless of their children's actual behavior. All homosexual subjects were asked about the extent of their mother's knowledge of their sexual orientation. (Two male subjects and 1 female subject whom we classified as bisexual based on Kinsey scores classified themselves as heterosexual; they did not respond to the maternal knowledge item, and have been excluded from this analysis.) Five homosexual males said that their mothers did not suspect them of being homosexual; 23 said that their mothers might suspect; 2 said that their mothers definitely knew they were homosexual. The corresponding subsamples for females are 4, 4, 1, and 9.

If mothers rate their homosexual children as more gender nonconforming because of stereotypes, then mothers who are most certain of their children's sexual orientation should have the most biased recollections. For the males, the correlation between extent of mothers' knowledge and maternal ratings of their children's gender nonconformity was not significant ($r = .12$) suggesting the absence of a theory-driven response style on the maternal ratings.

For females, however, maternal knowledge did correlate with their ratings ($r = .71$; $p < 0.001$). Interpretation of this correlation alone, however, is ambiguous. Although it may result from mothers' biased memories, it could also occur if mothers of the most masculine women were most likely to learn of their daughters' homosexuality. To examine the latter possibility, we computed the partial correlation between maternal knowledge and maternally rated CGN, controlling for self-rated CGN. This correlation was .65 ($p < 0.01$), suggesting that the relationship for females between maternal knowledge and maternally rated CGN is not mediated by true childhood masculinity (insofar as the latter is repre-

sented by self-rated CGN), and more probably represents a stereotypic rating bias.

Self-Rated and Maternally Rated CGN and Sexual Orientation

The analysis of maternal ratings supported the validity of differences between homosexuals' and heterosexuals' self-rated CGN. We also investigated the extent to which maternal ratings correspond to self-ratings. Attending first to the difference in CGN between heterosexual and homosexual, do maternal ratings make a unique contribution in explaining sexual orientation? Or alternatively, do maternal ratings predict sexual orientation through their relationship with self-ratings? To investigate these possibilities, discriminant analyses (separately for each sex) were performed with sexual orientation (heterosexual vs. homosexual) as the dependent variable, and maternally rated and self-rated CGN as predictors.

For males the overall R^2 accounted for by the model was .44. The partial regression coefficients for both maternally rated ($\beta = .43$; $p < 0.001$) and self-rated CGN ($\beta = .36$, $p < 0.001$) were significant. Thus, both self-rated and maternally rated CGN were related to sexual orientation, even after controlling for each other. The proportion of variance in sexual orientation explained *jointly* by maternal and self-ratings was .18. (This represents the proportion of variance that could not be uniquely apportioned to either maternally or self-rated CGN.)

For females the overall R^2 accounted for by the model was .21. The partial regression coefficient for self-rated CGN was significant ($\beta = .34$; $p < 0.01$); however, the respective coefficient for maternally rated CGN was not significant ($\beta = .18$; $p > 0.10$). The proportion of the variance in sexual orientation explained *jointly* by maternal and self ratings was .09.

A related issue is whether maternal ratings are related to self-ratings, within the homosexual groups. Were homosexual men who reported being effeminate as children rated as having been less masculine by their mothers as well? As noted above, this is important in establishing the validity of classifying homosexuals as gender conforming versus gender nonconforming. The correlation between self-rated and maternally rated childhood gender nonconformity was .15 for homosexual males and .42 for homosexual females, neither correlation being significant. Thus, we found little evidence for reliable discrimination among male homosexual subjects according to their degree of CGN. Though statistically nonsignificant, the

corresponding analysis for females is numerically stronger and should be investigated using a substantially larger sample.

DISCUSSION

Results of the present study provide convergent validation for the hypothesis that a history of childhood gender nonconformity is common among homosexual men. Among homosexual men unselected of their gender-related behaviors, mothers recalled a higher level of childhood gender nonconformity than did the mothers of heterosexual men. Furthermore, this did not appear to be due to biased recollection.

Results for females were less straightforward. It is noteworthy that female homosexuals were rated as slightly more nonathletic, though the difference was not significant. This latter finding conflicts with those of Grellert *et al.* (1982), indicating that female homosexuals recall having played more "masculine" sports as children. Compared to males, results for females were suggestive of a more general nonconformity than mere gender nonconformity. The pattern of maternal ratings indicate that mothers perceived their homosexual daughters as having been willful and disobedient during childhood. In comparison to the heterosexuals, mothers of homosexual women rated them as somewhat more masculine during childhood only when the mothers knew of their daughters' homosexuality. Thus, rating bias may be responsible for this homosexual-heterosexual difference in the maternal ratings. However, the homosexual female sample was small, and only one item was used as the measure of maternally rated CGN.

A large number of studies have investigated the relationship between childhood gender nonconformity and male homosexuality, with uniformly positive results (Pillard, 1991). In contrast, characterizing dimensions of childhood gender nonconformity among homosexual men, *i.e.*, why some homosexual men were effeminate children, whereas others were conventionally masculine, has received much less attention. The study of this question may help to illuminate both the etiology of sexual orientation and the phenomenon of childhood gender nonconformity. This study did not provide encouraging results for the utility of Whitam's (1977) indicators to discriminate between gender conforming and gender nonconforming homosexual men, because these ratings did not correlate with maternal ratings of CGN in the present study. Of course, the sizes of those correlations were constrained by the low reliability of maternally rated CGN. Future studies need to be more ambitious with respect to the number of items for both self-rated and maternally rated scales.

Research on the relationship between homosexuality and childhood gender nonconformity is less advanced for females. Thus, longitudinal studies, which control for the possibility of retrospective bias, are still needed to demonstrate convincingly whether homosexual females are disproportionately more gender nonconforming during childhood.

REFERENCES

- Bailey, J. M., Willerman, L., and Parks, C. (1991). A test of maternal stress hypothesis of human male homosexuality. *Arch. Sex. Behav.* 20: 277-293.
- Bell, A. P., Weinberg, M. S., and Hammersmith, S. K. (1981). *Sexual Preference: Its Development in Men and Women*, Indiana University Press, Bloomington.
- Green, R. (1987). *The "Sissy Boy Syndrome" and the Development of Homosexuality*, Yale University Press, New Haven, CT.
- Grellert, E. A., Newcomb, M. D., and Bentler, P. M. (1982). Childhood play activities of male and female homosexuals and heterosexuals. *Arch. Sex. Behav.* 11: 451-478.
- Pillard, R. C. (1991). Masculinity and femininity in homosexuality: "Inversion" revisited. In Gonzorick, J. C., and Weinrich, J. D. (eds.), *Homosexuality: Research Findings for Public Policy*, Sage, Newbury Park, CA.
- Ross, M. (1980). Retrospective distortion in homosexual research. *Arch. Sex. Behav.* 9: 523-531.
- Whitam, F. L. (1977). Childhood indicators of male homosexuality. *Arch. Sex. Behav.* 6: 89-96.
- Zuger, B. (1984). Early effeminate behavior in boys: Outcome and significance for homosexuality. *J. Nerv. Ment. Dis.* 172: 90-97.