Heritable Factors Influence Sexual Orientation in Women

J. Michael Bailey, PhD; Richard C. Pillard, MD; Michael C. Neale, PhD; Yvonne Agyei, MA

- Homosexual female probands with monozygotic cotwins, dizygotic cotwins, or adoptive sisters were recruited using homophile publications. Sexual orientation of relatives was assessed either by asking relatives directly, or, when this was impossible, by asking the probands. Of the relatives whose sexual orientation could be confidently rated, 34 (48%) of 71 monozygotic cotwins, six (16%) of 37 dizygotic cotwins, and two (6%) of 35 adoptive sisters were homosexual. Proband also reported 10 (14%) nontwin biologic sisters to be homosexual, although those sisters were not contacted to confirm their orientations. Heritabilities were significant using a wide range of assumptions about both the base rate of homosexuality in the population and ascertainment bias. The likelihood that a monozygotic cotwin would also be homosexual was unrelated to measured characteristics of the proband such as self-reported history of childhood gender nonconformity. Concordant monozygotic twins reported similar levels of childhood gender nonconformity. (Arch Gen Psychiatry. 1993;50:217-223)

More than 50 years ago, Hirschfeld1 noted that both male and female homosexuality appeared to be familial. Only recently, however, have researchers rigorously begun to test Hirschfeld's observations and systematically explored the nature of the familiality. Pillard and Weinrich2 found a significantly higher rate of homosexuality among brothers of homosexual men than among brothers of heterosexual men. Using a combination of twin and adoption methods, Bailey and Pillard3 found evidence that male sexual orientation is moderately heritable. Female homosexuality also appears to be familial. Pillard4 found 25% of sisters of homosexual female probands to be homosexual (including bisexuals), compared with 11% of sisters of heterosexual female probands. Bailey and Benishay5 found that, depending on the criterion, from 12% to 35% of sisters of homosexual probands were homosexual compared with 2% to 14% of sisters of heterosexual probands. Although familiarity may arise from shared environment as well as genetic factors, these findings support the desirability of testing genetic hypotheses directly.

Other than isolated case reports,6,7 the only twin study of female homosexuality of which we are aware consists of a series of four pairs of monozygotic (MZ) twins reared apart.7 None of the four pairs was concordant. Although the authors suggested that female homosexuality may be predominantly environmental, their sample was not sufficiently large to justify a strong conclusion.

The dearth of genetic data on females is unfortunate, as there is no strong reason to expect that genetic findings for males will be similar to those for females. The most influential biologic theories of sexual orientation posit that the development of attraction to females requires the masculinization of relevant (hypothalamic) brain structures, and that attraction to males results if relevant neural structures do not masculinize.8,9 Thus, different processes are hypothesized for male and female homosexuality, suggesting that if genetic factors contribute to female sexual orientation, they may differ from those for male sexual orientation.

The study reported herein has two broad goals: first, to determine if there is a genetic contribution to female sexual orientation, and second, to investigate the behavioral expression of this contribution. The study combines two methods from behavioral genetics: the twin method and the adoption method. Three groups of female probands were recruited: MZ twins, dizygotic (DZ) twins with female cotwins, and women with adoptive sisters (ie, sisters related to the proband by virtue of either the proband's or the sister's adoption). We predicted that the rate of homosexuality would be higher for MZ than for DZ cotwins, and would be lowest for adoptive sisters of homosexual probands. We considered the degree to which ascertainment bias may have affected results. We then examined if any of several characteristics might be an indicator of genetic loading for female homosexuality. Finally, we examined the degree to which MZ cotwins were similar for traits related to homosexuality.

SUBJECTS AND METHODS

Subject Recruitment

The method for this study was almost identical to that of Bailey and Pillard's genetic study of male sexual orientation.3 Probands were recruited through advertisements placed in lesbian-oriented publications in several cities across the United States: Chicago, Ill; Dallas, Houston, Austin, and San Antonio, Tex; Boston, Mass; and

Arch Gen Psychiatry—Vol 50, March 1993

Accepted for publication November 10, 1992.

From the Department of Psychology, Northwestern University, Evanston, Ill (Dr Bailey and Ms Agyei); the Family Studies Laboratory, Division of Psychiatry, Boston (Mass) University School of Medicine (Dr Pillard); and the Department of Human Genetics, Medical College of Virginia/Virginia Commonwealth University, Richmond (Dr Neale).

Reprint requests to Department of Psychology, Northwestern University, Evanston, IL 60208 (Dr Bailey).
Table 1.—Characteristics of Probands

<table>
<thead>
<tr>
<th></th>
<th>Twins (n=115)</th>
<th>Adoptive Sisters (n=32)</th>
<th>Combined (N=147)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (±SD) age, y</td>
<td>31.6±7.4</td>
<td>29.9±5.3</td>
<td>31.3±7.0</td>
</tr>
<tr>
<td>No. (%) of homosexual subjects</td>
<td>97 (84.3)</td>
<td>29 (90.6)</td>
<td>126 (85.7)</td>
</tr>
<tr>
<td>No. (%) of bisexual subjects</td>
<td>18 (15.7)</td>
<td>3 (9.4)</td>
<td>21 (14.3)</td>
</tr>
<tr>
<td>Mean (±SD) Kinsey rating</td>
<td>4.9±1.2</td>
<td>4.7±1.3</td>
<td>4.8±1.2</td>
</tr>
</tbody>
</table>

Table 2.—Relatives’ Sexual Orientations by Self- and Proband Reports

<table>
<thead>
<tr>
<th></th>
<th>Self-report by Relative</th>
<th>Proband’s Report of Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Heterosexual</td>
<td>Bisexual</td>
</tr>
<tr>
<td>Heterosexual</td>
<td>76</td>
<td>1</td>
</tr>
<tr>
<td>Bisexual</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Homosexual</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

*Values are numbers of relatives. Data represent relatives who reported their sexual orientation and whose sexual orientation probands were at least virtually certain.

Approximately .14. (If, on average, more than one individual read each copy of a publication, π would have been even lower.) Of the probands, 126 (85.7%) described themselves as “lesbian/homosexual” and 21 (14.3%) described themselves as “bisexual” (Kinsey ratings). Values were obtained for adult fantasy and behavior combined. These scores range from 0 (both fantasy and behavior completely heterosexual) to 6 (both fantasy and behavior completely homosexual). The mean (±SD) Kinsey rating, 4.8±1.2, indicated a fairly high level of homosexual orientation for the sample as a whole, but individual Kinsey ratings ranged from as low as 1 to as high as 6. Because the Kinsey ratings reflected overall adult behavior and fantasy, a woman might give herself a low rating because she had assumed a homosexual identity later in life. Indeed, this accounted for most of the probands with relatively low Kinsey scores (<3). However, three probands with low Kinsey scores admitted to relatively low levels of homosexual feelings, although they all considered themselves bisexual. It was decided to include these probands and to investigate empirically if the strength of homosexual feelings was an important variable, e.g., if it predicted concordance.

Probands had a total of 151 relatives of interest: 115 cotwins and 36 adoptive sisters. Three of these relatives were deceased. Permission was granted to contact 136 (91.9%) of the remaining relatives: 107 of the twins and 29 of the adoptive sisters. Questionnaires were returned by 122 relatives, representing 82.4% of the entire sample of living relatives and 89.7% of those whom probands consented to contact.

Assessment of Relatives’ Sexual Orientations

The sexual orientation of relatives was assessed in two ways. First, probands were asked if they believed their relatives’ sexual orientation to be heterosexual, homosexual, or bisexual. Additionally, probands were asked how certain they were about their assessment, using the following scale: “completely certain” indicated that the relative had told the proband her orientation outright; “virtually certain” meant that the proband felt quite sure but that this was based on behavior alone; “suspect but not sure” meant that the proband had some reason for making a guess, but felt appreciable uncertainty; “very uncertain” meant that the proband could do little more than guess. Additionally, those relatives who could be contacted were asked directly to rate themselves as “homosexual/lesbian,” “heterosexual,” or “bisexual.” Relatives also gave their combined Kinsey fantasy and behavior ratings, and, separately, their attraction to men and to women. Relatives’ self-ratings of sexual orientation were used when available. However, for 30 relatives these data were lacking. We had reason to believe, based on past studies, that probands would be generally quite accurate in assessing their relatives’ sexual orientations, provided that the proband expressed a high level of confidence. This was, in fact, confirmed for those relatives for whom both ratings were available, as is evident in Table 2. In those 113 confirmable cases when a proband was at least virtually certain about her relative’s orientation, a prediction of heterosexuality was 100% accurate and a prediction of nonheterosexuality (i.e., either homosexuality or bisexuality) was 97.4% accurate. For the entire table, combining homosexual and bisexual ratings, κ=.96 (P<.001 by Fisher’s Exact Test). Probands were

Recruitment Results

This procedure resulted in 147 proband interviews: 115 probands with female twins and 32 probands with adoptive sisters. Descriptive characteristics of the sample are included in Table 1. Probands ranged in age from 19 to 57 years, with a mean age of 31.3 years.

The probability of ascertaining an eligible proband, π, appears to have been somewhat low. A very rough estimate of π was obtained as follows: the circulation of the publications in which we advertised totaled approximately 60,000. Assuming that 1.4 individuals per 100 are twins from same-sex pairs, the readership included approximately 840 eligible twin subjects. Thus, π was

Los Angeles, Calif. The advertisements specified that desired subjects were lesbian or bisexual women at least 18 years old with either (1) female cotwins or (2) adoptive or genetically unrelated sisters. (The adoptive sister component of the study was added after approximately one third of the twin data were collected.) The advertisements also stated: “We hope you will call regardless of the sexual orientation of your twin or adoptive sister.” No mention was made of the possibility of the participation of cotwins or adoptive sisters. Potential subjects were instructed to call the laboratory, where they were asked clarifying questions. An additional criterion for the adoptive sister component was assessed at this point. Both probands and their adoptive sisters must have been younger than 3 years when they entered the common rearing environment.

Subjects who met the inclusion criteria were scheduled for a 1- to 2-hour interview. The Family Studies Laboratory in Chicago was the main interview site, but two of us (J.M.B. and Y.A.) traveled to several cities to interview subjects. Subjects who lived within a reasonable distance of interview sites were interviewed in person. In most cases (66%), however, a telephone interview was necessary. All interviews were conducted with informed consent. The session included questions determining the proband’s adolescent and adult sexual orientation, siblings’ sexual orientations (including twins and adoptive sisters), and childhood gender nonconformity (CGN). After completion of the interview, each proband was asked for permission to contact her twin or adoptive sister. Probands inspected the questionnaire to be sent and were assured that the method of their own recruitment and the information they supplied would not be divulged to the relatives.

The premise of the cover letter to relatives was that they were being asked to participate in a general behavioral genetics study of personality, attitudes, and behavior. Five questions regarding sexual orientation were embedded in more than 100 other items about social attitudes, personality, and childhood behavior (including CGN). Questionnaires were sent to consenting probands’ twins or adoptive sisters.

Questionnaires sent to cotwins contained items related to zygosity. A follow-up reminder letter was sent 1 week later. If relatives had not responded within approximately 1 month, attempts were made to contact them by telephone. Efforts were halted to gain cooperation only if at least two mailings of the questionnaire were unsuccessful and (1) no telephone number was available for the relative, (2) repeated telephone calls were unsuccessful, or (3) the relative was contacted and declined to participate.

Arch Gen Psychiatry—Vol 50, March 1993

Sexual Orientation in Women—Bailey et al
reasonably accurate in predicting if a nonheterosexual relative would label herself homosexual/lesbian or bisexual (κ = .65; \( P < .001 \) by Fisher's Exact Test); this contrasts with Bailey and Pillard's male study,\(^2\) in which this prediction was made at chance levels. The \( k \) for the combined table, keeping the three categories separate, was .89 (\( P < .001 \) by Fisher's Exact Test). However, for most of the following analyses, we employed a dichotomous measure of sexual orientation: heterosexual vs. homosexual, with bisexuals being included in the homosexual group.

Given the high degree of accuracy for confirmable cases when probands expressed a high degree of certainty, if a relative's self-rating was unavailable, the proband's assessment of her relative's sexual orientation was used, provided the proband was at least virtually certain. If a relative's self-rating was unavailable and the proband was less confident, that case was omitted from analyses of sexual orientation. Sexual orientation ratings were available for 145 of the 151 relevant relatives, including 110 cotwins and 35 adoptive siblings. In addition to twins and adoptive siblings, each twin proband was asked about the sexual orientation of non-twin biologic siblings in the manner described above. However, such siblings were not contacted directly. These data were not systematically collected for probands with adoptive siblings.

### Diagnosis of Twin Zygosity

Zygosity was determined using the questionnaire developed by Nichols and Bilbro,\(^9\) which contains items relating to physical similarity, past and present likelihood of twins being mistaken for each other, and twins' beliefs regarding their zygosity. Such questionnaires generally range in accuracy from 90% to 95%.\(^9,\)\(^10\)

Like most zygosity questionnaires, the Nichols-Bilbro questionnaire is intended to be answered by both twins of a pair, with both sets of responses entered into the diagnostic algorithm. Of the 97 pairs in which both twins completed the questionnaire (including one complete set of triplets), 65 were identified as MZ, 30 were identified as DZ, and zygosity could not be determined for two. Because of the desirability of using data from incomplete pairs, as well as the likelihood that acceptable accuracy would be obtained using only one twin's responses, zygosity of incomplete pairs was identified using only the proband's responses. An additional nine MZ pairs and nine DZ pairs were classified in this manner. Of the 92 complete pairs for which zygosity could be determined using the probands' responses alone, the same zygosity was determined for 87 probands (94.6%) using both full and partial data. Thus, diagnosis of zygosity using only the proband's information appeared to be nearly as accurate as diagnosis using information from both twins. The final twin relative subsample, including only those relatives whose sexual orientation and zygosity could be determined, consisted of 71 MZ twins and 37 DZ twins.

### Childhood Gender Nonconformity

All probands and cooperating relatives completed a 10-item scale of CGN. The items were taken from the Revised Childhood Gender Behaviors Questionnaire\(^8\) and ask about interest in stereotypically masculine and feminine activities during childhood as well as childhood gender identity (ie, comfort being a girl vs the desire to be a boy). The items were found to discriminate homosexual from heterosexual women. The internal consistency reliability of the scale, computed using only the probands' scores, was .82.

### RESULTS

#### Rates of Homosexuality in Relatives

The rates reported herein are probandwise concordance rates and not pairwise rates. Probandwise concordances refer to the proportion of relatives in each group who were homosexual. Probandwise rates are superior to pairwise rates, as they are independent of the probability of ascertainment\(^9\) and can be compared directly with population rates.\(^11\)

The rates of homosexuality (including bisexuality) among MZ cotwins, DZ cotwins, and adoptive sisters of probands are given in Table 3. Of the 71 MZ cotwins, 34 (48%) were either homosexual or bisexual, using the algorithm for the assessment of sexual orientation described earlier (method 1, in which the proband's report was used if the relative's self-report was unavailable), compared with six (16%) of the 37 DZ cotwins, and two (6%) of the 35 adoptive sisters. The rate for MZ cotwins was significantly greater than for both DZ cotwins (\( \chi^2 = 10.5; P < .001 \)) and adoptive sisters (\( \chi^2 = 18.6; P < .001 \)). Rates for DZ cotwins and adoptive sisters did not differ significantly (\( P > .26 \) by Fisher's Exact Test).

<table>
<thead>
<tr>
<th></th>
<th>Monozygotic Cotwins</th>
<th>Dizygotic Cotwins</th>
<th>Adoptive Sisters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homosexual, No.</td>
<td>(%/total)</td>
<td>(%/total)</td>
<td>(%/total)</td>
</tr>
<tr>
<td>34 (48/71)</td>
<td>6 (16/37)</td>
<td>2 (6/35)</td>
<td></td>
</tr>
<tr>
<td>Confirmed homosexual; No.</td>
<td>(%/total)</td>
<td>(%/total)</td>
<td>(%/total)</td>
</tr>
<tr>
<td>33 (51/65)</td>
<td>3 (10/29)</td>
<td>1 (4/25)</td>
<td></td>
</tr>
<tr>
<td>Mean ±SD age, ( \chi^2 )</td>
<td>31.9 ±7.5</td>
<td>31.7 ±7.5</td>
<td>30.3 ±9.5</td>
</tr>
</tbody>
</table>

*Data were computed using proband reports when relative self-reports were unavailable and when probands were at least virtually certain of relative's sexual orientation. Bisexual relatives were included in the homosexual group.

+ Data included only relatives who directly participated in the study.

### Focus on relatives for whom we had complete data (method 2, in which only relatives' self-reports were used), the picture was similar. The proportion of homosexuals among MZ cotwins exceeded that among DZ cotwins (33 [50.8%] of 65 vs three [10.3%] of 29; \( \chi^2 = 13.9; P < .001 \)). Similarly, the rate of homosexuality in MZ cotwins remained greater than that in adoptive sisters (one [4%] of 25; \( \chi^2 = 16.8; P < .001 \)). The difference in rates between DZ cotwins and adoptive sisters remained nonsignificant.

Because Bailey and Benishay\(^7\) found that rates of familial homosexuality varied widely depending on which of four different criteria was used to assess homosexuality in siblings, we computed concordance rates using all four of their criteria. Besides the two methods that we have already described, we also used an adult Kinsey score (combined feelings and behavior) of 2 or greater (method 3) and admission of any homosexual feelings (method 4). For method 3 the concordances of MZ cotwins, DZ cotwins, and adoptive sisters were 31 (48%) of 64, three (10%) of 29, and two (6%) of 24, respectively. For method 4, they were 31 (48%) of 64, six (20%) of 30, and three (14%) of 22. (Differences in sample size between methods 2 through 4 occurred because at least one subject did not respond to the relevant subsamples were homosexual relatives significantly older than heterosexual relatives. Indeed, across the entire sample, homosexual relatives were slightly, but not significantly, younger than heterosexual relatives (30.6 ± 6.5 vs 31.6 ± 8.6; \( t = -0.7; P = .44 \)). Thus, age did not appear to influence results importantly.

Because women who adopt homosexual identities frequently do so as adults\(^5,\)\(^6\) and because ages of study relatives range widely, the effect of age on concordance was deemed important to investigate. We did so by comparing the ages of homosexual and heterosexual relatives. In none of the three subsamples were homosexual relatives significantly older than heterosexual relatives. Indeed, across the entire sample, homosexual relatives were slightly, but not significantly, younger than heterosexual relatives (30.6 ± 6.5 vs 31.6 ± 8.6; \( t = -0.7; P = .44 \)). Thus, age did not appear to influence results importantly.

"Sexual Orientation in Women"—Bailey et al 219

Arch Gen Psychiatry—Vol 50, March 1993
Table 4.—Proband-Relative Correlations and Estimates of the Influence of Heritability and Environment Based on Several Sets of Assumptions*

<table>
<thead>
<tr>
<th>Assumptions</th>
<th>$B_i$</th>
<th>$P_i$</th>
<th>$r$ for Monozygotic Cotwins</th>
<th>$r$ for Dizygotic Cotwins</th>
<th>$r$ for Adoptive Sisters</th>
<th>Parameter Estimates</th>
<th>Significance Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.015</td>
<td>.85</td>
<td>.52</td>
<td>.25</td>
<td>.15</td>
<td>.62</td>
<td>26.6</td>
</tr>
<tr>
<td>1.0</td>
<td>.060</td>
<td>.77</td>
<td>.31</td>
<td>-.01</td>
<td>.31</td>
<td>.69</td>
<td>26.4</td>
</tr>
<tr>
<td>1.0</td>
<td></td>
<td>.100</td>
<td>.18</td>
<td>-.16</td>
<td>.31</td>
<td>.69</td>
<td>24.5</td>
</tr>
<tr>
<td>2.0</td>
<td>0.015</td>
<td>.72</td>
<td>.36</td>
<td>-.12</td>
<td>.28</td>
<td>.64</td>
<td>30.6</td>
</tr>
<tr>
<td>2.0</td>
<td>.060</td>
<td>.58</td>
<td>.11</td>
<td>-.16</td>
<td>.28</td>
<td>.55</td>
<td>27.6</td>
</tr>
<tr>
<td>3.0</td>
<td>0.015</td>
<td>.49</td>
<td>-.04</td>
<td>-.30</td>
<td>.36</td>
<td>.44</td>
<td>21.0</td>
</tr>
<tr>
<td>3.0</td>
<td>.060</td>
<td>.63</td>
<td>.37</td>
<td>-.23</td>
<td>.56</td>
<td>.62</td>
<td>32.7</td>
</tr>
<tr>
<td>3.0</td>
<td></td>
<td>.100</td>
<td>-.15</td>
<td>-.37</td>
<td>.73</td>
<td>.27</td>
<td>42.6</td>
</tr>
</tbody>
</table>

*See "Subjects and Methods" section for explanations of assumptions, correlations, parameters, and tests of significance.

Heritability of Sexual Orientation

The most informative index of the degree of genetic influence on a trait is heritability. Assuming a multifactorial model of transmission (i.e., that genetic influence is polygenic; environmental events are many and each has a small effect; and all influences contribute additively or that genotype-by-environment interaction in its statistical sense is absent), one can calculate heritabilities from rates of homosexuality in relatives. Computing heritability also requires an estimate of the base rate of homosexuality in the general population. Gebhard's reanalysis of Kinsey's data yielded an estimated 10% to 12% incidence of women with any homosexual experience, and an estimated 1% to 1.5% incidence of predominantly homosexual women. Unfortunately, neither criterion is identical to that used here, i.e., self-identification as either homosexual/lesbian or bisexual. It seems likely that the population base rate of women meeting our criterion for homosexuality is between Gebhard's estimates. Finally, the accuracy of a heritability estimate depends on assumptions about sampling. If, as is generally the case in volunteer twin samples, relatives who are most similar to each other are most likely to be ascertained, heritability estimates may be biased (although the direction of the bias may vary).

Heritability estimates were computed using the data from the top row of Table 3 (i.e., data about relatives for whom a sexual orientation assessment was available) as follows: first, tetrachoric correlations were computed for the three groups of relatives based on nine sets of assumptions, which depended on two parameters, $P_i$ and $B_i$. The parameter $P_i$ represents the base rate of female homosexuality in the general population and was assumed to be 1.5%, 5%, or 10%. The second parameter $B_i$, represents concordance-dependent ascertainment bias and is defined as the ratio of the likelihood that a proband will be ascertained if her relative is homosexual to the likelihood that she will be ascertained if her relative is heterosexual. Thus if $B_i=2$, a proband with a homosexual relative is twice as likely to be ascertained homosexual as a proband with a heterosexual relative. Parameter estimates were computed for three values of $B_i$: 1.0 (no differential ascertainment), 2.0, and 3.0.

For each set of assumptions, three parameters were estimated using the relatives' frequencies of homosexuality, tetrachoric correlations, and the model-fitting program MX (which specifically fits multifactorial threshold models): the heritability, $h^2$, or the proportion of phenotypic variance explained by additive genetic differences; $c^2$, the proportion of variance explained by those features of the environment shared by siblings; and $e^2$, the proportion of variance explained by the environment that siblings did not share. Genetic model-fitting capitalizes on the fact that phenotypic correlations between different types of relatives reflect different degrees of genetic and environmental similarity. For example, ignoring sampling error, nonadditive genetic variance, and ascertainment bias, the correlation between MZ cotwins is equal to the sum of the additive genetic and shared environmental parameters. The correlation between DZ cotwins is equal to the sum of the shared environmental parameter and half of the additive genetic parameter. The correlation between adoptive sisters is equal to the shared environmental parameter. The model-fitting program finds the parameters that yield an optimal fit to the data, i.e., that most closely generate the raw data. Further details regarding the model-fitting procedure can be obtained from the first author (see address on p. 217). (The theoretical rationale behind the computation of heritabilities for threshold characters can be found elsewhere.)

For a recent, detailed exposition of structural modeling applied to psychiatric genetics, see Kendler et al.

Results are presented in Table 4. Heritability estimates ($h^2$) ranged from .27 ($P_i=100, B_i=3.0$) to .76 ($P_i=.060, B_i=1.0$). The lowest heritability estimate resulted from assumptions that appeared implausible because they yielded negative correlations for both DZ twins and adoptive sisters, which would not be expected with most theoretical models. All other heritability estimates were at least .40. Thus, estimated heritability remained appreciable based on a wide variety of assumptions.

The estimated proportion of phenotypic variance attributed to shared environment ($c^2$) ranged from 0 (for seven models) to .23 ($P_i=.015, B_i=1.0$), and was in every case smaller than the estimated heritability. Estimated nonshared environmental variance ($e^2$) ranged from .15 ($P_i=.015, B_i=1.0$) to .73 ($P_i=100, B_i=3.0$). (Standard errors are not available for the parameter estimates because MX performs significance tests using the likelihood ratio test, which is preferable on statistical grounds to tests using standard errors. However, it should be remembered that all parameter estimates had a band of uncertainty.)

Both heritability and shared environmental parameters ($h^2$ and $c^2$, respectively) were tested for significance. To test the significance of $h^2$, that parameter was set to zero, and the chi-square from the resulting restricted model was compared with that from the unrestricted model in which $h^2$ was allowed to exceed zero. The difference of the two chi-square values yields a chi-square test with df=1, for the significance of $h^2$. The test for $c^2$ is analogous. No test was performed for the nonshared environmental parameter, $e^2$, because that parameter could not plausibly be zero given the number of heterosexual monozygotic cotwins who otherwise would all have been expected to be homosexual. Thus, the importance of $e^2$ was assumed and not in question in every model considered.
$h^2$ was highly significant ($P<.001$). In contrast, $c^2$ was significant in none.

The Rate of Homosexuality in Nontwin Siblings

Twin probands reported 73 nontwin biological sisters about whose sexual orientation they were at least virtually certain. Of these, 10 (14%) were thought to be homosexual or bisexual. This percentage is similar to both the 16% rate found for DZ cotwins and to the 12% rate found by Bailey and Benishay in their nontwin sibling study of female homosexuality, in which identical criteria were used. The probands also provided the sexual orientation of 104 nontwin brothers about whom they were adequately certain. Five nontwin brothers (5%) were thought to be homosexual. This percentage is significantly less than that of nontwin sisters judged to be homosexual ($x^2=4.4; P<.05$).

Possible Indicators of Genetic Loading

Although we found evidence that female sexual orientation is at least somewhat heritable, the question of what, precisely, is inherited remains. To elaborate on the nature of the genetic contribution, a potentially informative analysis was to compare MZ probands from concordant pairs with those from discordant pairs. If a trait was correlated with genes for homosexuality, that trait should have been more common among the former group. (This kind of analysis assumes that genes, rather than shared environmental factors, account for concordance of sexual orientation. Results presented in Table 4 support this assumption.) Two broad sets of variables have been proposed as potential markers of genetic diathesis in other domains, such as some forms of psychopathology: those related to early onset and those indicating phenotypic extremity (in psychopathology, severity). In the present study, measures indicating early onset of relevant behavior included CGN, the adolescent Kinsey score, and age of first homosexual feelings. Neither CGN nor age of first homosexual feelings differed significantly between MZ probands from concordant pairs and MZ probands from discordant pairs ($P>.4$ for all comparisons). Measures indicating relatively extreme homosexuality included the adult Kinsey score, the absence of positive sexual feelings for men, and the self-designation of lesbian/homosexual as opposed to bisexual. None of these variables was significant as a predictor of concordance ($P>.25$). Thus, neither early onset nor extremity appeared to be an indicator of genetic loading. However, these retrospective measures are probably of modest reliability, diminishing the statistical power of relevant analyses.

Other Similarities Between MZ Cotwins

Another general analysis that could, in principle, illuminate the pathways from genotype to phenotype concerns the similarity of MZ cotwins for characteristics related to sexual orientation. If cotwins from concordant pairs tend to be similar for other important variables, such as degree of CGN, these variables would appear to be influenced by genes and/or aspects of the environment shared by cotwins. Cotwins from concordant pairs had significant correlations in CGN ($r=33=.51; P<.005$; all $P$s from these analyses were one-tailed) and adolescent Kinsey score ($r=29=.35; P<.05$), but not for the degree of current heterosexual feelings ($r=32=.25; P>.08$), adult Kinsey scores ($r=32=.08; P>.30$), or the self-designation of bisexual vs lesbian/homosexual ($r=33=.10; P>.30$). Thus, cotwins from concordant pairs displayed a moderate degree of similarity regarding early precursors of adult homosexuality. In contrast, cotwins from discordant pairs were not similar on any of these measures ($r<.15$ and $P>.25$ for each analysis).

COMMENT

 Ascertainment Bias

The primary threat to the validity of the central finding, that genetic factors may play a role in the origin of female sexual orientation, is ascertainment bias. Because probands were not obtained through systematic sampling, and particularly given the evidently low probability of ascertainment, it is possible that patterns of volunteering yielded misleading results. However, not all kinds of ascertainment bias are of equal concern. For example, our heritability analyses examined the effects of concordance-dependent bias and found that heritabilities remained significant when a wide range of assumptions were made about that kind of bias. Moreover, concordance-dependent bias cannot lead to a false finding of nonzero heritability, although it does affect heritability estimates. Regarding the magnitude of concordance-dependent bias in the present study, it is noteworthy that the concordance rate for DZ cotwins (16%) was similar to that found for nontwin sisters of homosexual probands by Bailey and Benishay (12%) and to that found for nontwin sisters in the present study (14%). It is less plausible that the latter rates were seriously biased. Bailey and Benishay recruited subjects who were initially blind to the study’s focus on familiarity. Twin probands in the present study were aware of its focus on twins, and, hence, might be expected to consider their cotwins’ sexual orientations in deciding whether to participate. It seems less likely that they would weigh their other siblings’ orientations. Despite these differences, all three rates are fairly similar, suggesting that at least for DZ twins, concordance-dependent bias has not been a serious problem.

Serious errors could have resulted if concordance-dependent bias differed among the three groups. For instance, if DZ twins are less susceptible than MZ twins to concordance-dependent bias, a false finding of nonzero heritability could result. However, to account completely for the MZ-DZ difference found in the present study, by this type of bias, one would have to assume that if DZ twins were equally likely as MZ to be ascertained homosexual whether or not they were from concordant pairs (ie, $B=1$ for DZ twins), MZ twins from concordant pairs were 4.7 times more likely to be ascertained homosexual than twins from discordant pairs (ie, $B=4.7$ for MZ twins). If DZ twins were assumed to exhibit some concordance-dependent bias, the MZ bias would need to be increased accordingly. While this possibility cannot be ruled out, a more plausible interpretation is that the true MZ concordance exceeds the true DZ concordance.

One kind of ascertainment bias evident in our study was the overrepresentation of MZ probands, who constituted approximately two thirds of our sample. In comparison, approximately half of same-sex twin births are monozygotic. This overrepresentation is a common characteristic of volunteer-twin samples. The reason for the MZ bias is unclear, as are its implications. To the extent that it merely reflects the increased saliency of twinness in MZ than in DZ twins, and, hence, their increased willingness to volunteer, the bias has no implications. To take into account the extent to which overrepresentation reflects diminished probability that dissimilar pairs will cooperate compared with similar pairs, $B$ was incorporated in our model. A systematically ascertained twin and adoptive sample that avoids this ascertainment bias is highly desirable. Unfortunately, at present no promising way is known to obtain such a sample of homosexual women.

Another kind of ascertainment bias that occurred in the present study concerns the fact that probands were recruited via advertisements in homosexual publications. It is unknown if female homosexuals who read such publi-
cations differ in important respects from those who do not. Future studies of sexual orientation that avoid this bias, for example through the use of twin registries, are clearly desirable.

**Equal-Environments Assumption**

The validity of our method to study genetic variation depends on one key assumption. Because the method uses relatives who are reared together, it is necessary to assume that the trait-relevant environment is equally similar for MZ and DZ twins and adoptive sisters. Although this “equal-environments assumption” has been criticized in the context of twin research, available research supports its validity, at least for traits studied so far. For instance, although some MZ twins are treated quite similarly by their parents in predictable ways (e.g., being dressed alike), these MZ twins are no more similar in intelligence and personality than those treated dissimilarly by their parents. Twins whose zygosity is mislabeled by their parents are as phenotypically similar as twins whose zygosity is correctly assessed. Although the equal-environments assumption has generally been supported, future genetic studies of sexual orientation should examine the assumption directly.

**Implications for the Causes of Sexual Orientation**

Heritability remained significant and appreciable (>25%) for all sets of assumptions examined, although the assumptions explored herein do not exhaust the realm of possibilities. These findings should be considered in the context of prior research on this and related questions. Thus, for instance, the discordance of all female MZ pairs reported by Eckert et al. suggests the need for caution in drawing conclusions from our study, although their small sample size does not allow powerful tests of differences between the two studies. Similarly, King and McDonald’s recent twin study of male and female homosexuality combined found markedly lower concordance rates than either this study or our previous study of male sexual orientation. The other hand, another recent report found even higher concordance rates. Given the serious methodologic concerns, particularly that of ascertainment bias, the inconsistency of some past research, and the small number of related studies, we urge that our results be evaluated cautiously. Although our results are highly suggestive of nonzero heritability, they are not conclusive. This caveat applies even more strongly to the parameter estimates, which are strongly dependent on assumptions of unknown validity. Our results should be considered the first word on this subject, rather than the last. We hope this study will inspire further, more definitive studies in the area.

Assuming, however, that our finding of significant heritability is valid, an elucidation of the nature of the genetic variance could be an important step to unraveling the origins of female sexual orientation. Unfortunately, our attempts to do so were unsuccessful. Specifically, neither variables indicating an early onset of relevant behaviors, such as CGN, nor variables indicating extreme homosexuality, such as self-identification as lesbian rather than bisexual, appeared to be related to genetic loading for homosexuality. Nevertheless, the comparison of MZ probands from concordant and discordant pairs is a potentially fruitful method of identifying indicators of genetic loading and should be pursued with other variables and larger samples.

Monozygotic twins from concordant pairs were somewhat similar for the important developmentally-relevant variables of CGN and adolescent Kinsey score, suggesting that genes may influence developmental differences among homosexual women. To test for genetic influences, one could compare the correlations for CGN for concordant MZ pairs, concordant DZ pairs, and concordant adoptive sisters. Unfortunately, numbers of concordant DZ pairs and concordant adoptive sisters in this study were insufficient for a reasonably powerful test of correlation differences. In general, this kind of analysis is a useful way to discover genetic heterogeneity. If female homosexuality has different genetic routes and each of these routes has distinct phenotypic characteristics, MZ cotwins should be similar for the relevant characteristics. Finally, it is noteworthy that cotwins from discordant pairs were dissimilar in CGN. Thus, no evidence was present that relevant behaviors of the heterosexual twin in these pairs was even partially influenced by genetic processes affecting the homosexual twin.

In our previous article and in the comments just made, we have primarily focused on quantifying and elucidating the genetic influences on sexual orientation. These studies were designed to detect heritable variation, and, if it was present, to counter the prevalent belief that sexual orientation is largely the product of family interactions and the social environment. Although both male and female sexual orientation appear to be at least somewhat inheritable, environment also must be of considerable importance in their origins.

Reiss et al. have recently drawn attention to the power of genetic studies such as ours to demonstrate the importance and illuminate the nature of environmental influences. For example, MZ cotwins who differ in sexual orientation can do so only because relevant environmental factors differ. Thus, discordant MZ twin pairs are uniquely informative regarding potential environmental influences. Results of the present study allow one general conclusion regarding the nature of environmental influences on female sexual orientation: the effective environment appears to comprise factors not even shared by MZ cotwins (i.e., the nonshared environment), rather than shared factors (shared environment). This implies, for instance, that investigations of possible parental influence should focus less on stable attributes of parental personality than on idiosyncratic features of the relationship between parents and their homosexual children. Similarly, hypotheses concerning the prenatal environment should emphasize aspects of prenatal environment that might reasonably be expected to differ between MZ cotwins. The results of the present study do not illuminate the relative importance of psychosocial or biologic environmental factors. Either is consistent with the results reported herein. Methods of behavioral genetics, such as the study of discordant twins, may be helpful in identifying relevant environmental factors. To accomplish this, it is first necessary to specify candidates for study.

**Female vs Male Sexual Orientation**

Probands reported significantly more homosexual sisters than brothers, suggesting that male and female homosexuality are at least somewhat independent etiologically. Furthermore, because female homosexuality appears to be

Sexual Orientation in Women—Bailey et al.
substantially less common than male homosexuality, the relatively high frequency of homosexuality among sisters compared with brothers is especially striking. This finding supports the necessity of studying female homosexuality directly rather than assuming that findings for males can be extended to females. The optimal way of demonstrating different modes of transmission for the two sexes is via genetic modeling, in which the same model would be fit simultaneously to data for both males and females, and compared with the model in which relevant parameters were free to vary by sex. Data should ideally be collected in a manner that would minimize ascertainment bias and allow the estimation of the base rate of homosexuality for both sexes. Furthermore, opposite-sex DZ twins would be a highly desirable population to include in such an analysis.

How do the findings of the present study compare with those of Bailey and Pillard's genetic study of male sexual orientation, which employed a similar method? The most important similarity is that both male and female sexual orientation appeared to be influenced by genetic factors. However, in neither study was an indicator of genetic loading found. Both studies found a high concordance among MZ co-twins for degrees of recalled CGN. Consistent with previous findings, females were more like males to be bisexual. For instance, 1% of female probands considered themselves bisexual compared with 7% of male probands ($\chi^2=4.6; P=0.05$). Perhaps the most important difference between the present study of females and the previous study of males is that far more relevant research was available for males. Thus, the finding of significant heritability for male sexual orientation was not unexpected.

This research was supported in part by grant MH27227 from the National Institute of Mental Health and by a Northwestern University research grant.

We thank Natalie Beckerman, Deanna Benishay, Sue Bihore, Erin Dugan, Kristine Iniba, Tania Bishee, Gregory Kovacs, Amy Pyron, and Vicky Singh for their assistance in this research, and Sheri Berenbaum, PhD; William Byrne, MD, PhD; John Loeblin, PhD; Lee Willerman, PhD; and five anonymous reviewers for their helpful comments on the manuscript.

References