J.biosoc.Sci, (2006) **38**, 797–809, © 2005 Cambridge University Press
doi:10.1017/S002193200500101X First published online 18 April 2006

MASCULINITY–FEMININITY PREDICTS SEXUAL ORIENTATION IN MEN BUT NOT IN WOMEN

J. RICHARD UDRY AND KIM CHANTALA

University of North Carolina at Chapel Hill, NC, USA

Summary. Using the nationally representative sample of about 15,000 Add Health respondents in Wave III, the hypothesis is tested that masculinity–femininity in adolescence is correlated with sexual orientation 5 years later and 6 years later: that is, that for adolescent males in 1995 and again in 1996, more feminine males have a higher probability of self-identifying as homosexuals in 2001–02. It is predicted that for adolescent females in 1995 and 1996, more masculine females have a higher probability of self-identifying as homosexuals in 2001–02. Masculinity–femininity is measured by the classical method used by Terman & Miles. For both time periods, the hypothesis was strongly confirmed for males: the more feminine males had several times the probability of being attracted to same-sex partners, several times the probability of having same-sex partners, and several times the probability of self-identifying as homosexuals, compared with more masculine males. For females, no relationship was found at either time period between masculinity and sex of preference. The biological mechanism underlying homosexuality may be different for males and females.

Introduction

The origins of homosexuality are controversial. Debate continues as to the role of genes (DuPree *et al.*, 2004), the role of prenatal biological determinants (Cohen-Bendahan *et al.*, 2005), and the role of postnatal social environment (Gottschalk, 2003). Meanwhile, social–behavioural scientists still cling to sociological and psychological theories.

The prevalent biological theory is the prenatal hormone theory. The prenatal hormone theory holds that prenatal androgens play a crucial role in sexual differentiation of the central nervous system, and therefore sexually differentiated behaviour. This paper examines a prenatal hormone hypothesis. This hypothesis states that increasing prenatal androgens masculinize behaviour, and decreasing prenatal androgens feminize behaviour (Meyer-Bahlberg, 1984). The hormone

hypothesis opens the possibility that one sex could have a biology of homosexuality that is different or modified from the other sex. The primary source of prenatal testosterone for a female is her mother's blood, while for a male the primary source is his own testes. As gestation progresses, testosterone increases for males, but declines for females (Meulenberg & Hofman, 1991).

More than 50 years of research has documented that male homosexuals have more feminine traits than male heterosexuals (Terman & Miles, 1936). Feminine behaviour in pre-adolescent boys is associated with homosexual orientation in young adulthood (Green, 1987). It is also commonly believed that female homosexuals are more masculine than female heterosexuals. This paper explores the prediction of adult sexual orientation from adolescent masculinity–femininity.

Udry & Chantala (2002) reported that male adolescents who have sex partners of the same sex only are distinctly more feminine than those with partners of the opposite sex. On the other hand, they showed that adolescent females with partners of the same sex only do not differ in masculinity–femininity from females with partners of the opposite sex only. Lippa & Arad (1997), using college students, showed the same pattern as Udry & Chantala (2002): males showed a strong relationship between masculinity–femininity and sexual orientation, but females showed no significant correlations.

Research to test the hypothesis that adult homosexuals exhibited behaviour in childhood similar to the childhood behaviour of heterosexual adults of the opposite sex usually relies on the retrospective reports of homosexual vs heterosexual adults about their own childhood behaviour. Bailey & Zucker (1995) found all retrospective studies they examined to be confirmatory for both sexes: homosexual males were more feminine than heterosexual males, and homosexual females were more masculine that heterosexual females. Because there is a general belief in the population that the hypothesis is true, retrospective self-reports are not convincing evidence. If there is a general belief that homosexuals of each sex have behaviour that is more similar to the opposite sex, then this is the way adult homosexuals will 'remember' their own childhoods. Prospective studies routinely find confirmatory evidence for boys, according to Bailey & Zucker, but prospective evidence is lacking for girls. In the present study the hypothesis was tested for both sexes using a prospective method.

What causes people to differ in masculinity–femininity? Female fetuses with the genetic disorder congenital adrenal hyperplasia (CAH) experience an abnormally high level of androgenic hormones before birth, and during childhood and later show distinctively masculine behaviour preferences. When treated postnatally to counteract their excess androgens, they still maintain more masculine behaviour into adulthood (Dittmann *et al*., 1992), and show less commitment to heterosexuality (Hines *et al.*, 2004). Normal female fetuses also differ from one another in their exposure to androgens during the prenatal period, and in adulthood those women who were fetally exposed to higher androgens were more masculine than less exposed females (Udry, 2000). Male (XY) fetuses who are insensitive to their own androgens are during childhood and later as feminine in their behaviour as normal females. These examples show us that fetal androgens masculinize behaviour after birth, and absence of fetal androgens feminizes the offspring. Females, who do not normally make significant amounts of fetal testosterone, are (naturally) feminine. This is the

biological explanation of routine, naturally occurring sex differences in behaviour. This explanation does not negate the influence of gender socialization, but augments it.

In fetal males, prenatal testosterone is converted by the enzyme 5-*a*-reductase to dihydrotestosterone, which causes masculinization of the body. Prenatal testosterone is also converted in particular areas of the brain into oestrogen by the enzyme aromatase. This oestrogen masculinizes specific local areas in the brain, and therefore specific types of behaviour (Allen & Gorski, 1992). The specific local areas that are masculinized are the same across different individuals. Since normal female fetuses do not produce testosterone, they can only experience testosterone in any quantity through the testosterone in their mother's blood (Udry, 2000) or by some abnormal condition such as CAH. Maternal testosterone during mid-pregnancy is sufficient to cause later differential masculinization within the typical range in normal females.

The history of measurement of gender-typical behaviour (masculinity–femininity or MF) begins with Terman & Miles (1936). They selected questionnaire items demonstrated to have responses that were differentially selected by males or females. From these they constructed a scale that was bipolar, unidimensional and measured the extent to which respondents' answers were typical of their sex. This type of scale dominated measurement until the early 1970s. In the 1970s, academic political considerations led to a reconsideration of the nature of sex differences, and avoidance of the idea that the sexes were 'opposite'. Bem's BSRI (Bem Sex Role Inventory), a transformation of masculinity–femininity to orthogonal (uncorrelated) dimensions of Instrumentality and Expressivity, captured the day. With this measurement, there was no longer a method of comparing sex differences, because the two dimensions were uncorrelated. The BSRI dominated the gender field from the 1970s until the present.

Meanwhile, since the 1960s, biological researchers had begun to discover that hormones were related to sex differences in animals and humans in a way that made it useful to think of MF as a single continuum, and to think of humans being more or less masculine meaning the same as less or more feminine. This led them to measurement of MF by techniques similar to those pioneered by Terman & Miles (1936). Lippa (2001) and Cleveland *et al.* (2001) have helped the behavioural sciences return to the classic measurement. This return to a continuous measure is also useful when thinking of MF in connection with sex differences in the context of homosexuality. This invites us to remember that sex differences are not limited to the instrumental–expressive dimension, but rather may be thought of as encompassing any dimension of behaviour for which there are sex differences. While Lippa includes among his sex differences occupational preferences, hobbies, everyday activities and other dimensions, in this paper no behavioural domain restrictions are considered in the measurement of masculinity–femininity.

Methods

The data for analysis come from the National Longitudinal Study of Adolescent Health, hereafter Add Health, a panel study of a national representative sample of US adolescents initially in grades 7 to 12 in 1994. A stratified probability sample of 80 high schools (and where necessary, feeder schools to those high schools to include

grade levels 7 to 12) was selected from the Quality Education Database that lists all high schools in the 50 United States as the frame. Eighty per cent of the contacted schools agreed to participate by directing a self-administered questionnaire on adolescent health to all students present on a particular day. School refusals were replaced by schools from the same stratum. For each school, all students present on a particular day completed a one-period op-scan questionnaire. Information from the self-administered questionnaire was used to identify specific sub-groups for over-sampling in a second-stage sample for home interviews. About 90,000 students took the school questionnaire. Details of the sample design are provided at the Add Health website.

Respondents from the school rosters and school questionnaires of participating schools were selected for a stratified probability sample to be interviewed at home. Use of school rosters made it easy to include in-home interviews for the school absentees on the test day as well as dropouts. In the first wave of home interviews, 80% of the selected students (about 20,750) completed home interviews with permission from parents. Computer-assisted interviews were conducted by an interviewer, but sensitive questions were self-administered on the laptop computer by respondents with recorded questions heard through earphones. School questionnaires were administered in 1994–95. The first wave of home interviews was administered in 1995. The second wave of home interviews was administered in 1996, about a year later with the same respondents and a similar questionnaire. Omitted from the 1996 follow-up were those in grade 12 at the first wave. A third wave of home interviews was administered in 2001–02. In the third wave all respondents who participated in the first home interview were re-surveyed. About 15,000 interviews were completed in the third wave. This was about 77% of those interviewed at Wave I home interview. The sample is weighted to represent the adolescent population from which it was originally drawn.

To measure the effect of adolescent masculinity–femininity (MF) on young adult sexual preference, sexual orientation in adulthood was measured at Wave III when approximately 96% of the respondents were 19–24 years old. Masculinity–femininity was measured at first and second home interview, 1995 and 1996. Masculinity–femininity was measured by a non-obvious method, by selecting items from main questionnaires with responses that showed significant sex differences. These items were scattered about in the questionnaire. Table 1 gives the items included in the MF scale for Wave II. The dependent variables were measured 6 years later, so it doesn't matter whether respondents had a belief about the hypotheses. Masculinity–femininity norms were measured in the population studied in the year of the questionnaire and not some other year or some other group. This method avoids the problems considered by Bailey & Zucker (1995) about the validity of retrospective measures of MF.

Present measurement of MF

Add Health questionnaire items were selected the answers to which were biased by sex to construct a separate MF score for each wave of the study. Each wave of Add Health identified those questionnaire items with statistically significant sex differences. Items covered a broad range of topics. Logistic regression computed for each

Table 1. Difference of means for boys and girls expressed as fraction of boys' standard deviation (STD) for variables used in Wave II MF

Content of items	Regression coefficient	Direction of response pattern	(Boy–girl mean)/ (boy STD)	Boys' STD
Frequency of crying	−1·2525	0 to 4; 4=every day	−1·21	0·43
Frequency of moodiness	−0·2255	0 to 4; 4=every day	−0·47	0·87
Frequency of poor appetite	−0·2022	0 to 4; 4=every day	−0·38	0·74
How honestly answered questions	−0·2425	1 to 4; 4=completely honest	−0·21	0·86
Trouble paying attention	0·3109	0 to 4; 4=every day	0·11	1·06
Bothered by things	−0·1194	0 to 3; 3=most/all of the time	−0·34	0·63
How physically fit	−0·3676	1 to 5; 5=strongly disagree	−0·48	0·82
Past 12 months, frequency serious fighting	0·6654	0 to 3; 3=5or more times	0·27	0·60
Frequency of exercising	−0·1331	0 to 3; 3=5or more times	−0·06	1·06
Frequency of roller-blading/cycling	0·3056	0 to 3; 3=5or more times	0·29	0·99
How emotional you are	0·1217	1 to 5; 5=strongly disagree	0·44	0·98
Do you like yourself as you are?	−0·2042	1 to 5; 5=strongly disagree	−0·44	0·79
Live without thought for future	−0·2171	1 to 5; 5=strongly disagree	−0·26	1·11
How sensitive to others feelings	0·3175	1 to 5; 5=strongly disagree	0·32	0·80
Do you like to take risks?	−0·1762	1 to 5; 5=strongly disagree	−0·31	1·01
Upset by difficult problems	0·2025	1 to 5; 5=strongly disagree	0·36	1·06

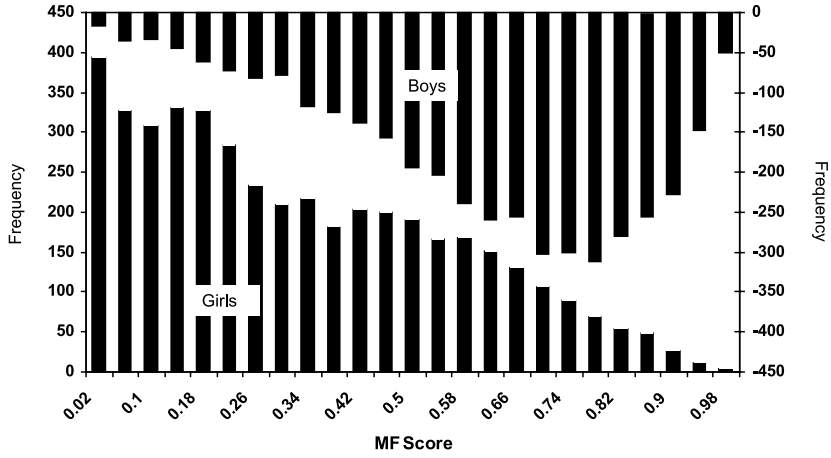

Fig. 1. Wave II MF distribution.

respondent the probability of being a boy. The mean MF declines with age for each sex. By Wave III, many respondents are married. Married males in Wave III have a lower (more feminine) MF than unmarried males. Figure 1 gives the distribution of MF for each sex for Wave II. Masculinity–femininity is the statistical equivalent of the probability of being a boy. For the present analysis the MF scores were selected from Wave II (1996) when everyone was in school in grades 8 to 12. Wave II MF scores were used to predict sex preferences and sexual behaviour in Wave III (2001–02), when respondents averaged 19–24 years old. The study was then replicated by repeating the analysis using Wave I values for MF.

Masculinity–femininity is considered as a biological consequence of prenatal testosterone on postnatal behaviour. The prenatal testosterone in males is produced by the fetal testis, and the prenatal testosterone in females is produced by the mother, whether or not through transformation to oestrogen via the aromatase mechanism (Lephart, 1996). The more prenatal testosterone, the more masculine the behaviour, irrespective of the sex of the fetus, and irrespective of where the testosterone comes from. The testosterone from the testicles of the fetus produces higher MF as experienced by males, and the testosterone from the mothers of the fetus produces lower MF as experienced by females. Yet MF values run from 0·00 to 0·99 for both males and females, although the mean for males is much higher, with males having a mean MF of about 0·66, while females have an MF mean of about 0·34.

Measurement of sexual orientation

In Wave III only, at about ages 19–24, when all respondents were no longer in high school, respondents were asked: 'Please choose the description that best fits how you think about yourself:

- 100% heterosexual (straight).
- Mostly heterosexual (straight) but somewhat attracted to people of your own sex.
- Bisexual – that is, attracted to men and women equally.

https://doi.org/10.1017/S002193200500101X Published online by Cambridge University Press

Table 2. Sexual orientation in Wave III

	Percentage of males*; $n=6759$†	Percentage of females*; $n=7563$†
100% heterosexual	94·03	85·10
Mostly heterosexual	3·18	10·65
Bisexual	0·57	2·55
Mostly homosexual	0·63	0·70
100% homosexual	1·18	0·47
No sex interest	0·41	0·52
Total	100·00	100·00

*Sampling weights were used to compute population percentages.
†There were 48 males and 85 females at Wave III who did not provide information on sexual orientation.

- Mostly homosexual (gay) but somewhat attracted to people of the opposite sex.
- 100% homosexual (gay).
- Not sexually attracted to either males or females.

Table 2 provides the distribution of sexual orientation at Wave III.

Measurement of sex of sexual attraction

Sexual attraction and behaviour were measured at Wave III, for the period since 1995. At Wave III, respondents were asked, 'Have you ever had a sexual attraction to a male?' and 'Have you ever had a sexual attraction to a female?'

Sex of sex partners

At Wave III, respondents listed their sexual partners since 1995, and the sex of each partner. From this list each respondent was classified as having had no sex partners, only same-sex partners, only opposite-sex partners, or sex partners of both sexes during this period.

The plan for analysis is to determine the degree of strength of the relationship between MF at adolescence (Wave II) and the different components of homosexuality at Wave III, and to determine how this differs for males and females. Logistic regression models were fitted to predict each sexual measure from Wave III using the MF score. Separate models were estimated for males and females. The strength of the relationship was determined between MF and the different components of each sexual measure from Wave III, and how this differs for males and females. For males, the degree of strength is measured by dividing the model-predicted risk of reporting a particular category of the sexual measure for the most feminine males (MF=0) by the risk of reporting that same category for the most masculine males (MF=1). Conversely, for females, the risk of reporting a particular category of the sexual measure for the most masculine females (MF=1) is divided by the risk of reporting

https://doi.org/10.1017/S002193200500101X Published online by Cambridge University Press

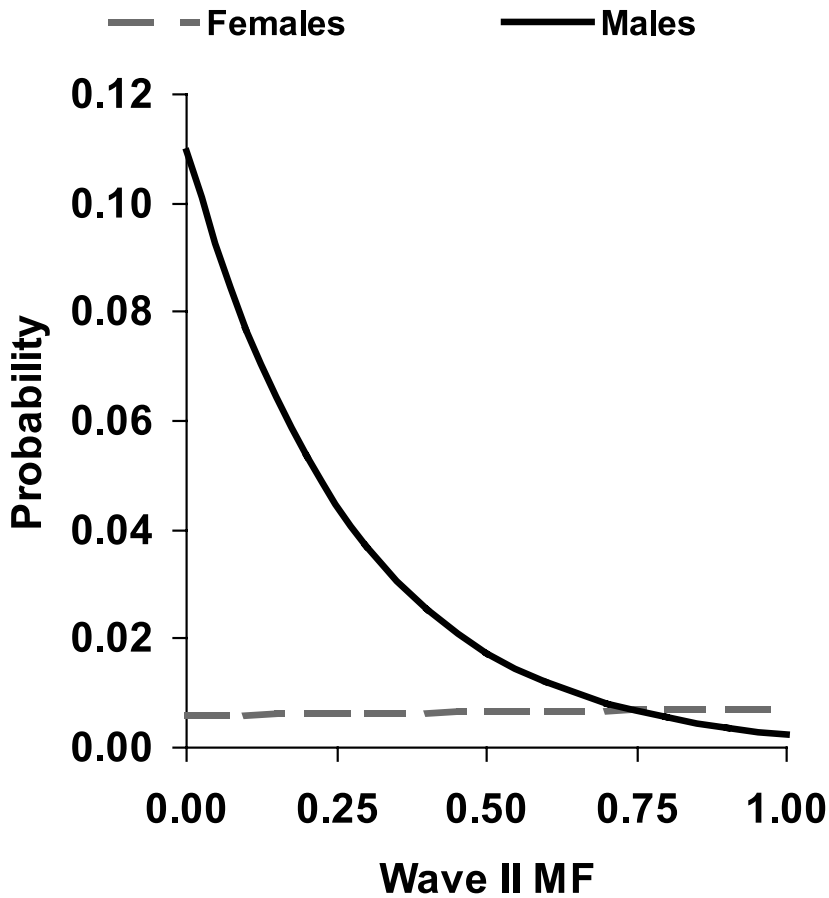

Fig. 2. Probability of reporting only same-sex partners versus only opposite-sex partners at Wave III predicted from Wave II MF score.

that same category for the most feminine females (MF=0). Thus the strength of the relationship between MF and the Wave III sexual measure is the ratio of the risk of reporting a particular category for the most traditional to the most non-traditional member (as measured by MF) of each sex.

Results

Figure 2 illustrates the basic type of finding. It shows the risk of reporting (at Wave III) only same-sex partners versus reporting only opposite-sex partners since 1995 as a function of the respondents' Wave II MF score. These risks are the probabilities predicted from separate logistic regression models for males and females. Because age at Wave III was not significant when included as a covariate, it was omitted from the models. The most feminine males (MF=0) have a risk of 0·11 for reporting only same-sex partners (Fig. 2) compared with having only opposite-sex partners. On the

https://doi.org/10.1017/S002193200500101X Published online by Cambridge University Press

Table 3. Risk ratio (or probability ratio) and 95% confidence intervals for report of Wave III sexual measure for females (most masculine compared with most feminine) and males (most feminine compared with most masculine)

Wave III comparison	Females: Masculine (MF=1) Feminine (MF=0)	Males: Feminine (MF=0) Masculine (MF=1)
Predicted from Wave I MF score		
Sexual orientation: homosexual vs heterosexual	2·0 (− 0·34, 4·39)	23·3 (− 2·52, 49·1)
Attraction: only same sex vs only opposite sex	0·77 (− 0·26, 1·79)	19·9 (− 11·1, 50·95)
Romantic partners: only same sex vs only opposite sex	2·2 (− 1·19, 5·59)	21·3 (− 5·32, 47·87)
Predicted from Wave II MF score		
Sexual orientation: homosexual vs heterosexual	0·51 (− 0·15, 1,18)	26·0 (− 0·17, 52·18)
Attraction: only same sex vs only opposite sex	0·19 (− 0·18, 0·55)	23·0 (− 7·9, 53·98)
Romantic partners: only same sex vs only opposite sex	1·23 (− 1·35, 3·82)	42·8 (− 17·6, 103·3)

other hand, the most feminine females (MF=0) and the most masculine females (MF=1) have nearly the same risk of reporting only same-sex partners (0·0060 and 0·0074 respectively). Therefore the effect of MF on the risk of same-sex partners is much stronger for males than for females.

Effect size of male MF for sex of romantic partner

Presented in Table 3 are the effects of Wave II MF on Wave III sex of sex partners, sex of attraction and sexual orientation. Next the findings were reproduced using Wave I MF effects on Wave III behaviour. These findings are also represented in Table 3. The findings are very similar to those using Wave II MF, indicating that for males, MF predicts sexual behaviour whether it is collected in Wave I or II, and even though it is predicting sexual behaviour 5 or 6 years later than MF was collected. For females, there is no relationship between Wave I or Wave II MF and Wave III sexual behaviour.

The effect size of MF measured at Wave III for the probability of having a same-sex partner vs an opposite-sex partner since 1995 is equal to 21·3 for males and 2·2 for females. This effect is nearly ten times as large for males as it is for females.

The effect size of MF on probability of a same-sex only attraction vs opposite-sex attraction at Wave III is 19·9 for males, and 0·8 for females, a ratio that is statistically not significantly different from 1. The effect of MF on the probability of a homosexual vs heterosexual orientation is 23·3 for males, and for females it is 2·0.

A review of these differences indicates that the effect of MF on sex of sex partners, sex of attraction and sexual orientation is several times the size for men that it is for women, for whom the effect is not significant.

For practical purposes, for females, there is no relationship between MF and same-sex behaviour or sexual orientation. In contrast, the effect of MF on sex of sexual orientation, sex of partners and sex of attraction is large and important for males.

Discussion

For males, the level of masculinity–femininity between 12 and 18 years of age predicts the degree of same-sex attraction, the number of same-sex partners and a homosexual orientation 6 years later. For females, there is no relationship between level of MF and homosexual orientation, same-sex attraction and number of same-sex partners since 1995. This is in spite of the fact that prenatal androgens are the primary source of variability in later MF for both males and females. This casts doubt on the presumption that female homosexuals are less feminine than females who are not homosexual. This lack of a relationship between masculinization among females and their preference for same-sex partners also suggests that the mechanism for creating homosexuality in females may be different from the mechanism for males.

Yet the level of androgen exposure experienced by females *in utero* is related to the degree of masculinization of their behaviour in their third decade of life (Udry, 2000). Thus the mechanism for masculinization is the same for females as for males. But for females, masculinization does not correlate with same-sex sexuality, even though it does correlate for males.

The fact that male MF is correlated with later same-sex behaviour but female MF is not may be related to the differences between the sexes in hormone release patterns in the prenatal period. Normal female fetuses do not produce testosterone in the prenatal period. Their only exposure to testosterone prenatally is from maternal blood.

It is easy to estimate the prenatal testosterone exposure difference between male and female fetuses. In mid-trimester, female fetuses average 29 ng/100 ml, while male fetuses average 249 ng/100 ml, or 10 times as much (Abramovich & Rowe, 1973). Assuming that the ratio of the concentration of testosterone in fetal males to females is related to the risk ratio of same-sex to opposite-sex partners for each sex, Fig. 1 estimates from MF the risk ratios. Females having MF scores from 0 to 0·75 have risk ratios from 0·005 to 0·01. On the other hand, males having MF scores from 0 to 0·75 have risk ratios from 0·01 to 0·075. The risk ratios for females are not statistically significant. On the other hand, the risk ratios for males cover a wide range of values (10 times the range for females). From mid to late pregnancy, the concentration of fetal testosterone decreases in female fetuses, but increases in male fetuses (Meulenberg & Hofman, 1991). From these estimates the amount of fetal testosterone in females is not enough to generate a significant relationship to sexual orientation later in life, even though it may generate a relationship to MF. For males, the amount of fetal testosterone is far greater than for females. This amount of testosterone is enough to generate both a relationship to MF in males, as well as a relationship to sexual orientation.

Let us assume for simplicity of theory that unusually high prenatal testosterone for females gives them an unusually high MF in childhood and adolescence, and gives

them a same-sex orientation in adulthood. Let us compare this theory with the Add Health data reported here. The mean Wave II MF for females who identified themselves as 100% homosexual at Wave III (about half of 1%) is 0·29, while for those 100% heterosexual it is 0·33. That is, those who were 100% heterosexual at Wave III were slightly *more* masculine at Wave II than those who were 100% homosexual at Wave III. If the 100% and the mostly homosexual are combined in Add Health, as was done by Brown *et al.* (2002), then the combination of 100% and mostly homosexual females has a mean Wave II MF of between 0·29 and 0·30, or slightly but not significantly more feminine than average.

Conclusion

Masculinity–femininity in adolescents predicts sexual preference and sexual orientation in male adults, but not female adults. The measurement of MF by the present technique may somehow eliminate aspects of sex differences that if included would show the MF relationship with sexual orientation among females. This seems unlikely. Lippa & Arad (1997) found no relationship between current MF and sexual orientation for females, using their measure GD (similar to MF).

The studies reviewed by Bailey & Zucker (1995) show a relationship in all studies between female childhood masculine sex attributes and adult homosexuality. The studies they examined were all retrospective. The measurements of MF for the retrospective studies were not the same across studies. Yet it is implausible that every single study showed homosexual females to (erroneously) report childhood behaviour more masculine than reported by heterosexual females.

Finally, the relationship between genes and homosexual behaviour remains unexplored in this paper. The problem is that researchers cannot consistently demonstrate genetic linkage between markers for homosexuality in males, and no genetic traces have shown up for behavioural genetic methods with females. The latest exploration has sought a connection between male homosexuality and the gene for aromatase, an enzyme that converts testosterone to oestrogen, and is thought to masculinize behaviour in non-human males through this conversion. However, male homosexual behaviour was not shown to be linked to the aromatase gene in humans (DuPree *et al.*, 2004). In fact it remains to be shown that human males require the aromatase mechanism, even though the mechanism seems to be widely observed in other species of males (Lephart, 1996).

Prenatal testosterone produces a higher (more masculine) MF score for males and females alike

Because, in female fetuses, prenatal testosterone comes from the maternal blood, which is less concentrated than the prenatal testosterone provided to the male fetus from his own testes, the additional prenatal testosterone in females required for same-sex orientation by the theory does not occur. But the prenatal testosterone shortfall required for same-sex orientation is consistent with the data in males. So masculinization in females does not correlate with same-sex sexuality, but in males, lower levels of masculinization may lead to same-sex orientation in adults.

What this means is that female homosexuals may be no more masculine than female heterosexuals. Yet male homosexuals are more feminine than male heterosexuals. The females who as fetuses got more testosterone became more masculine, but that did not affect their sexual orientation. So if variations in prenatal testosterone are related to same-sex sexuality in males, but not in females, what causes female homosexuality? It appears that female homosexuality may be a consequence of female socialization (Gottschalk, 2003) in addition to consequence of prenatal hormone experience.

Acknowledgments

This research is based on data from the Add Health project, a programme project designed by J. Richard Udry (PI), Kathleen Mullan Harris and Peter Bearman, and funded by grant P01-HD31921 from the National Institute of Child Health and Human Development to the Carolina Population Center, University of North Carolina at Chapel Hill, with cooperative funding participation by seventeen mostly federal agencies. Special appreciation goes to Ronald R. Rindfuss and Barbara Entwisle for assistance with the original design of the project. Persons interested in obtaining data files from The National Longitudinal Study of Adolescent Health should contact Add Health, Carolina Population Center, 123 West Franklin Street, Chapel Hill, NC 27516–2524 (email: addhealth@unc.edu).

References

Abramovich, D. R. & Rowe, P. (1973) Foetal plasma testosterone levels at mid-pregnancy and at term: Relationship to foetal sex. *Journal of Endocrinology* **56**, 621–622.

Allen, L. S. & Gorski, R. A. (1992) Sexual orientation and the size of the anterior commissure in the human brain. *Proceedings of the National Academy of Sciences of the USA* **89**, 7199–7202.

***Bailey, J. M., Miller, J. S. & Willerman, L.** (1993) Maternally rated childhood gender nonconformity in homosexuals and heterosexuals. *Archives of Sexual Behavior* **22**, 461–469.

***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.

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.

***Bell, A. P., Weinberg, M. S. & Hammersmith, S. K.** (1981) *Sexual Preference: Its Development in Men and Women*. Indiana University Press, Bloomington.

***Blanchard, R. & Freund, K.** (1983) Measuring masculine gender identity in females. *Journal of Consulting and Clinical Psychology* **51**, 205–214.

Brown, W. M., Finn, C. J., Cooke, B. M. & Breedlove, S. M. (2002) Differences in finger length ratios between self-identified 'butch' and 'femme' lesbians. *Archives of Sexual Behavior* **31**, 123–127.

Cleveland, H. H., Udry, J. R. & Chantala, K. (2001) Environmental and genetic influences on sex-typed behaviors and attitudes of male and female adolescents. *Personality and Social Psychology Bulletin* **27**, 1587–1598.

Cohen-Bendahan, C. C. C., van de Beek, C. & Berenbaum, S. A. (2005) Prenatal sex hormone effects on child and adult sex-typed behavior: Methods and findings. *Neuroscience and Biobehavioral Reviews* **29**, 353–384.

Dittmann, R. W., Kappes, M. E. & Kappes, M. H. (1992) Sexual behavior in adolescent and adult females with congenital adrenal hyperplasia. *Psychoneuroendocrinology* **17**, 153–170.

DuPree, M. G., Mustanski, B. S., Bocklandt, S., Nievergelt, C. & Hamer, D. H. (2004) A candidate gene study of CYP19 (aromatase) and male sexual orientation. *Behavior Genetics* **34**, 243–250.

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

***Grellert, E. A., Newcomb, M. D. & Bentler, P. M.** (1982) Childhood play activities of male and female homosexuals and heterosexuals. *Archives of Sexual Behavior* **11**, 451–478.

Green, R. (1987) *The 'Sissy Boy Syndrome' and the Development of Homosexuality*. Yale University Press, New Haven, CT.

***Gundlach, R. H. & Riess, B. F.** (1968) Self and sexual identity in the female: A study of female homosexuals. In Riess, B. F. (ed.) *New Directions in Mental Health*, vol. 1. Grune & Stratton, New York, pp. 205–231.

Hines, M., Brook, C. & Conway, G. S. (2004) Androgen and psychosexual development: Core gender identity, sexual orientation, and recalled childhood gender role behavior in women and men with congenital adrenal hyperplasia (CAH). *Journal of Sex Research* **41**, 75–81.

Kaye, H. E., Berl, S., Clare, J., Eleston, M. R., Gershwin, B. S., Gershwin, P. *et al. (1967) Homosexuality in women. *Archives of General Psychiatry* **17**, 626–634.

Lephart, E. D. (1996) A review of brain aromatase cytochrome P450. *Brain Research Reviews* **22**, 1–26.

Lippa, R. (2001) On deconstructing and reconstructing masculinity–femininity. *Journal of Research in Personality* **35**, 168–207.

Lippa, R. & Arad, S. (1997) The structure of sexual orientation and its relation to masculinity, femininity, and gender diagnosticity: Different for men and women. *Sex Roles* **37**, 187–208.

Meulenberg, P. M. M. & Hofman, J. A. (1991) Maternal testosterone and fetal sex. *Journal of Steroid Biochemistry and Molecular Biology* **39**, 51–54.

Meyer-Bahlburg, H. F. L. (1984) Psychoendocrine research on sexual orientation: Current status and future options. *Progress in Brain Research* **61**, 375–398.

***Saghir, M. T. & Robins, E.** (1973) *Male and Female Homosexuality: A Comprehensive Investigation*. Williams & Wilkins, Baltimore.

Terman, L. M. & Miles, C. C. (1936) *Sex and Personality*. McGraw-Hill Book Company, Inc., New York.

***Thompson, N. L. Jr, Schwartz, D. M., McCandless, B. R. & Edwards, D. A.** (1973) Parent–child relationships and sexual identity in male and female homosexuals and heterosexuals. *Journal of Consulting and Clinical Psychology* **41**, 120–127.

Udry, J. R. (2000) Biological limits of gender construction. *American Sociological Review* **65**, 443–457.

Udry, J. R. & Chantala, C. (2002) Risk assessment of adolescents with same-sex relationships. *Journal of Adolescent Health* **31**, 84–92.

***Whitam, F. L. & Mathy, R. M.** (1991) Childhood cross-gender behavior of homosexual females in Brazil, Peru, the Philippines, and the United States. *Archives of Sexual Behavior* **20**, 151–170.

*References marked with an asterisk indicate studies included in the meta-analysis by Bailey & Zucker (1995) that include females.