Arch Sex Behav (2010) 39:1207–1208
DOI 10.1007/s10508-010-9634-z

LETTER TO THE EDITOR

The Importance of Measuring Internalized Homophobia/ Homonegativity

Michael W. Ross · B. R. Simon Rosser · Derek Smolenski

Published online: 13 May 2010

Newcomb and Mustanski (2009) are to be congratulated for attempting a meta-analysis of internalized homophobia/ homonegativity (IH) and for recognizing the potential importance of this construct for HIV-related research in men who have sex with men (MSM). However, we disagree with their key conclusion that "researchers would be well-served to consider abandoning the investigation of IH as a predictor of risky sexual behavior in order to focus resources on more promising lines of research." There are seven concerns.

First, their approach to meta-analysis used correlation coefficients; however, correlation coefficients in meta-analyses *are not* and *cannot* be effect sizes. Such standardized effect sizes "cannot reliably meet the goals of meta-analysis" (Cummings, 2004, p. 597) and Greenland, Maclure, Schlesselman, Poole, and Morgenstern (1991) note that the arguments "against traditionally standardized coefficients apply with even greater force against the use of correlation coefficients as measures of effect" (p. 392). They are inappropriate to use because (1) they are bounded by the distributional properties of each individual study; (2) following from that, a one SD change in one sample is, by definition, not equivalent to one SD in another study; (3) even if a correlation is high, the effect size could be quite modest. In contrast, a lower or no correlation could indicate a curvilinear relationship or variation around a stronger fixed effect; and (4) analysis of correlations negates the ability to interpret the *magnitude* of any observed association, leaving only the *direction*.

Second, the nine different IH measures used are simply not comparable, absent a study in which the different measures are given to the same population and the degree of association measured. The outcome measures also vary across studies, and there are insufficient data to determine how the outcomes were actually measured and what their distributions were. The fact that some IH scales factor into several subscales (Ross & Rosser, 1996), some but not all of which may be related to risk behaviors, further complicates a comparison of IH measures (as Newcomb and Mustanski note).

Third, one study used compulsive sexual behavior (CSB) as the outcome measure. That is inappropriate since it superimposes a deterministic relationship between CSB and risky sex, not the appropriate probabilistic one. Unfortunately, the strong correlation there probably contributed to the observed effect.

Fourth, the recommendation not to study IH further is based upon their observing no direct relationship between IH and HIV risk. However, Newcomb and Mustanski readily acknowledge and reference studies showing clear indirect relationships that, in turn, may be helpful to the design of innovative interventions. The existence of several moot correlations could indicate conflicting pathways that haven't been explored. Should we determine some causal association of IH with other proximal determinants of risk behavior, then we will have a good idea of how to appropriately intervene on those factors. Ignoring IH in further risk behavior studies would prevent development of culturally specific and relevant interventions for MSM.

Fifth, the rather narrow choice of outcomes assumes that IH has a direct effect only on sexual behavior. Some data already indicate that IH or at least one component of it is associated with drug and alcohol abuse in MSM, and we know that drug and alcohol abuse are related to risky sexual behavior. It would

M. W. Ross (✉)
Center for Health Promotion and Prevention Research, University of Texas School of Public Health, 7000 Fannin, Suite 2622, Houston, TX 77030, USA
e-mail: Michael.W.Ross@uth.tmc.edu

B. R. S. Rosser · D. Smolenski
HIV/STI Intervention and Prevention Studies Program, Division of Epidemiology and Community Health, University of Minnesota School of Public Health, Minneapolis, MN, USA

make clinical as well as research sense to include other risk outcomes, such as drug and alcohol abuse, and measures of mood and affect as effects of IH which may themselves translate into sexual risk behavior. Further, the impact of IH on drug and alcohol abuse and other negative mental health outcomes (Rosser, Bockting, Ross, Miner, & Coleman, 2008) are in themselves clinically important for MSM, whether or not they translate into risky sexual behavior, and we concur with Newcomb and Mustanski on this.

Sixth, we already have data that suggest that IH operates through intervening variables. We demonstrated (Ross, Rosser, Neumaier, & the Positive Connections Team, 2008) that IH is significantly associated with HIV/STI risk through at least two intervening pathways (but has no *direct* effect). Serodiscordant unprotected anal intercourse is a function of men being less disclosing of their HIV serostatus, itself significantly associated with IH; and lower condom use through lower condom self-efficacy is itself also significantly associated with IH. There are almost certainly other important indirect pathways through which IH operates on risk behaviors waiting to be elucidated.

Finally, the most troubling aspect of their recommendation to abandon IH and HIV risk research is that the studies reviewed were overwhelmingly North American. Studies not in English were excluded. Worldwide, homonegativity and IH remain huge threats to the well-being of MSM. As we write, there is a wave of homonegativity sweeping Africa. Uganda has introduced legislation which provides for the death penalty for male homosexual acts. In Senegal, it is reported that the bodies of gay men have been dug up and desecrated. Outside North America and selected West European and Australasian countries, homonegativity appears an overwhelming and ever-present physical and psychological threat to MSM. To recommend not studying IH and its still poorly understood associations with risk behavior, simply because studies in the U.S. failed to observe direct relationships between IH and sexual risk behavior, makes no sense. Our study on IH in Ugandan MSM (Ross et al., 2010) will, we hope, stimulate other researchers to explore the construct and its direct and indirect health impacts in non-Western contexts and in non-Western MSM subcultures.

In summary, while Newcomb and Mustanski have produced a most thought-provoking article summarizing some studies on direct effects of IH on HIV risk behaviors, we encourage researchers to explore IH further, particularly through indirect effects and in contexts where homonegativity appears a major barrier to HIV prevention for MSM. We do not believe, as they suggest, that the "current utility of this construct for understanding sexual risk taking of MSM is called into question" for use by HIV/STI, sexuality and mental health researchers. Outside of the West, homophobia and internalized homophobia do still matter.

References

Cummings, P. (2004). Meta-analysis based on standardized effects is unreliable. *Archives of Pediatric and Adolescent Medicine, 156*, 595–597.

Greenland, S., Maclure, M., Schlesselman, J. J., Poole, C., & Morgenstern, H. (1991). Standardized regression coefficients: A further critique and review of some alternatives. *Epidemiology, 2*, 387–392.

Newcomb, M. E., & Mustanski, B. (2009). Moderators of the relationship between internalized homophobia and risky sexual behavior in men who have sex with men: A meta-analysis. *Archives of Sexual Behavior*. doi:10.1007/s10508-009-9573-8.

Ross, M. W., & Rosser, B. R. S. (1996). Measurement and correlates of internalized homophobia: A factor analytic study. *Journal of Clinical Psychology, 52*, 15–21.

Ross, M. W., Rosser, B. R. S., Neumaier, E. R., & the Positive Connections Team. (2008). The relationship of internalized homonegativity to unsafe sexual behavior in HIV seropositive men who have sex with men. *AIDS Education and Prevention, 20*, 547–557.

Ross, M. W., Smolenski, D. J., Kajubi, P., Mandel, J. S., McFarland, W., & Raymond, H. F. (2010). Measurement of internalized homonegativity in gay and bisexual men in Uganda: Cross-cultural properties of the Internalized Homonegativity Scale. *Psychology, Health and Medicine, 15*, 159–165.

Rosser, B. R. S., Bockting, W. O., Ross, M. W., Miner, M. H., & Coleman, E. (2008). The relationship between homosexuality, internalized homo-negativity, and mental health in men who have sex with men. *Journal of Homosexuality, 55*, 185–203.

Arch Sex Behav (2010) 39:1209–1211
DOI 10.1007/s10508-010-9655-7

LETTER TO THE EDITOR

The Importance of Measuring Internalized Homophobia/ Homonegativity: Reply to Ross, Rosser, and Smolenski (2010a)

Michael E. Newcomb · Brian Mustanski

Published online: 16 July 2010

Ross, Rosser, and Smolenski (2010a) provide a commentary regarding our recently published meta-analysis of the relationship between internalized homophobia (IH) and risky sexual behavior in men who have sex with men (MSM) (Newcomb & Mustanski, 2009). Ross et al. raised a number of methodological and conceptual concerns about the meta-analysis, and therefore disagree with our conclusion that researchers would be well-served to focus their resources on other predictors of risky sexual behavior that are more promising in terms of their predictive ability. We would like to address a number of these concerns in order to further support our conclusion that the literature does not support a current significant bivariate association between IH and risky sexual behavior. We further believe that few viable models have been proposed to examine the potential indirect effects of IH on risk.

First, Ross et al. raise concerns about the use of correlation coefficients (ES_r) in meta-analysis and state that "correlation coefficients in meta-analysis are not and cannot be effect sizes." In fact, correlation coefficients are commonly used as effect size in meta-analysis and are adequate measures of effect size for meta-analysis given that extra steps are taken in order to overcome certain problematic properties of this statistic. In their widely-used book on meta-analysis, Lipsey and Wilson (2001) stated:

> …when the research findings to be meta-analyzed involve bivariate relationships in which both the variables are continuous, the product-moment correlation coefficient is the straightforwardly appropriate effect size statistic. Indeed, virtually all such research findings will be reported in terms of the correlation coefficient in original studies, so it will rarely be sensible to consider any other effect size statistic to represent them" (p. 63).

Lipsey and Wilson further discuss the need to standardize ES_r using Fisher's r to z transformation in order to overcome a problematic standard error formulation which is used in weighting effect sizes based on sample size. By using standardized and weighted ES_r statistics, the analyst is able to meaningfully combine results across studies and make conclusions about combined overall effect size.

Ross et al. also noted that the IH measures used in the meta-analysis were not comparable. While it is true that there has been some disagreement in the literature about how best to measure IH, one of the purposes of meta-analysis is to combine results from multiple studies on the same relationship in order to account for methodological and measurement differences between these studies. Combining studies using multiple measures of IH that may assess slightly different components of the construct allow the analyst to make more general conclusions about IH than would be possible by examining each study individually. As such, we are better able to make general conclusions about the current state of the literature on the association between IH and risky sexual behavior because we did not rely on a single measure of IH. In fact, it is common practice in meta-analysis to use different measures of the same construct. For example, a widely-cited meta-analysis of the effects of school-based intervention programs on aggressive behavior included multiple different measures of aggressive behavior, such as physical aggression, externalizing problems, fighting, bullying, acting out, disruptiveness, and/or conduct and discipline problems (Wilson, Lipsey, & Derzon, 2003).

M. E. Newcomb (✉)
Department of Psychology (MC 285), University of Illinois at Chicago, 1007 W. Harrison Street, Chicago, IL 60607, USA
e-mail: mnewcomb@psych.uic.edu

B. Mustanski
Department of Psychiatry, University of Illinois at Chicago, Chicago, IL, USA

When multiple measures are used, it is also possible to test for differences in effects between measures. To consider this, we included the most commonly used measure of IH, the Nungesser Homosexual Attitudes Instrument (NHAI), as a moderator of the relationship in a follow-up analysis. This variable was not a significant moderator, indicating that the NHAI was not differentially associated with risky sexual behavior as compared to other measures.

Ross et al. also expressed concern about varying measures of risky sexual behavior, specifically the inclusion of a study assessing compulsive sexual behavior as the outcome variable. The majority of the studies included in the meta-analysis assessed unprotected anal sex as the risk variable, but several studies included measures of unprotected vaginal and oral sex and number of sexual partners as components of sexual risk. Again, meta-analysis should help to account for measurement differences across multiple measures of the same or similar constructs. While it is true that compulsive sexual behavior is less frequently used as a sexual risk variable, removing this study from the meta-analysis did not alter the pattern or significance of any of the findings, including the moderating effect of year of data collection on the relationship between IH and risky sexual behavior.

Ross et al. suggested that, even if there is no direct relationship between IH and risky sexual behavior, the indirect relationships reported in some studies repudiate our conclusions about the value of further research on this association. In our article, we cited studies that have found evidence for indirect effects of IH on risky sexual behavior in MSM. However, relatively few studies have provided evidence for and replicated these indirect effects, and none have provided a strong rationale for why an indirect effect may exist in the absence of a direct bivariate relationship. Such "inconsistent mediation" occurs when the indirect and direct effect are in opposite directions and therefore cancel each other out in the estimation of the direct relationship (i.e., a "suppression" effect) (MacKinnon, Krull, & Lockwood, 2000). In these more rarely reported cases of suppression, it is particularly important for researchers to explicate a reasonable model for understanding the reason and implications for suppression. We found no such explanations or even patterns of correlations consistent with suppression in the published papers we reviewed reporting indirect effects. In reading the article cited by Ross et al. as evidence for indirect effects, we found no report, description, or explanation of an indirect effect between IH and sexual risk taking (Ross et al., 2010b). We believe much stronger empirical evidence and theoretical justification for the existence of an important indirect relationship between IH and risky sexual behavior is needed before public health decisions should be made based on the existence of an indirect relationship.

Furthermore, if an indirect relationship was firmly established, this makes the case for focusing research and interventions on those more proximal processes that mediate the relationship between IH and risky sexual behavior. Ross et al. state, "Should we determine some causal association of IH with other proximal determinants of risk behavior, then we will have a good idea of how to appropriately intervene on those factors." Here, we seem to agree that such proximal factors should become the focus of scarce public health resources over more distal factors as IH would then be. Where we do disagree with Ross et al. is in their statement that "Ignoring IH in further risk behavior studies would prevent development of culturally specific and relevant interventions for MSM." First, we believe there is a danger in developing intervention content based on the assumption that the internalization of anti-gay and heterosexist stigma is central to the understanding of MSM risk behaviors, when the evidence does not bear this out. This may lead to a misdirection of scarce intervention resources. For example, it is possible that the external manifestations of this stigma, such as legislation that prevents two men from being married, may represent critical structural factors that increase risk for HIV among MSM. However, this policy may not need to be psychologically internalized by the individual in order for it to affect risk taking. In our opinion, this deserves further investigation. Second, we do not believe that a cultural factor needs to be correlated with a risk outcome in order to be considered for inclusion in making an intervention culturally specific and relevant. For example, in the CDC's ADAPT framework for adapting an existing HIV intervention to a target group (McKleroy et al., 2006), changes were deemed necessary in order to make the intervention culturally appropriate and to address unique risk factors and behavioral determinants. If IH is a core part of the culture, then it may very well need to be addressed in an HIV prevention program even if it is not a risk determinant. Individuals may disagree about the extent to which IH is currently a core part of gay culture.

Ross et al. also noted that our meta-analysis only examined one potential outcome associated with IH, and that an indirect relationship between IH and risky sexual behavior may exist via the association between IH and other outcomes, such as mental health and substance use. As is noted in our discussion in the meta-analysis of the relationship between IH and risky sexual behavior, we also conducted a meta-analysis of the relationship between IH and internalizing mental health problems, which found a moderate effect size for this relationship (Newcomb & Mustanski, 2010). We agree with Ross et al. that the effects of IH on mental health are important to consider in both research and clinical work. We also considered conducting meta-analyses of the relationships between IH and alcohol and drug use, but there were not enough studies on these relationships to warrant meta-analysis. While we agree that there may be a relationship between IH and substance use, we caution readers from drawing the conclusion that this is evidence for an indirect link between IH and risky sexual behavior. The direct association between substance use and sexual risk has not been consistent in the literature, particularly for alcohol use (Cooper, 2006), and conflicting findings have been reported. Similarly, the association between sexual risk

and mental health or affective states has been inconsistent (Crepaz & Marks, 2001; Mustanski, 2007).

Finally, we agree with Ross et al. that IH may be differently experienced in non-Western cultures and that it plausibly may predict risky sexual behavior in other cultures. We only included studies in our meta-analysis from Western cultures that were conducted in English as both language and culture may affect the experience of IH and thus render it impossible to make meaningful comparisons across studies. Our conclusion that "researchers would be well-served to consider abandoning the investigation of IH as a predictor of risky sexual behavior" applies to the countries or cultures in which the studies included in the meta-analysis were conducted. However, we encourage researchers to continue to be mindful of the difference between IH, which is an internal psychological process, and external anti-gay behaviors and policies. Victimization resulting from homophobia does not always lead to IH, and determining which individuals under which conditions internalize heterosexism would be helpful for clinicians working with this community. We do very much agree with Ross et al. about the importance of studying the health and development of LGBT individuals and MSM in non-Western contexts. It is not our intent to argue that IH is not worth studying outside the West in research on risky sexual behavior or that IH doesn't matter as a construct that is relevant to mental health and social justice. However, it is our intent to encourage researchers to be mindful of the science and rationale influencing their decisions to do so.

In conclusion, we stand by the conclusions of our meta-analysis about an overall lack of direct association between IH and risky sexual behavior in MSM in studies published in English with Western samples. We also believe there is not well developed evidence or explanations for indirect effects in the absence of direct effects. While the experience of IH may have important implications for other health-related outcomes (e.g., mental health), we maintain that is it critical for future research to focus on investigating predictors of sexual risk that are more promising in terms of empirical or theoretical support given scarce public resources available to study these processes.

References

Cooper, M. L. (2006). Does drinking promote risky sexual behavior? *Current Directions in Psychological Science, 15*, 19–23.

Crepaz, N., & Marks, G. (2001). Are negative affective states associated with HIV sexual risk behaviors? A meta-analytic review. *Health Psychology, 20*, 291–299.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*, 173–181.

McKleroy, V. S., Galbraith, J. S., Cummings, B., Jones, P., Harshbarger, C., Collins, C., et al. (2006). Adapting evidence-based behavioral interventions for new settings and target populations. *AIDS Education and Prevention, 18*, 59–73.

Mustanski, B. (2007). The influence of state and trait affect on HIV risk behaviors: A daily diary study of MSM. *Health Psychology, 26*, 618–626.

Newcomb, M. E., & Mustanski, B. (2009). Moderators of the relationship between internalized homophobia and risky sexual behavior in men who have sex with men: A meta-analysis. *Archives of Sexual Behavior,* doi:10.1007/s10508-009-9573-8.

Newcomb, M. E., & Mustanski, B. (2010). *The relationship between internalized homophobia and internalizing mental health problems: A meta-analytic review* (under review).

Ross, M. W., Rosser, B. R. S., & Smolenski, D. (2010a). The importance of measuring internalized homophobia/homonegativity [Letter to the Editor]. *Archives of Sexual Behavior,* doi:10.1007/s10508-010-9634-z.

Ross, M. W., Smolenski, D. J., Kajubi, P., Mandel, J. S., McFarland, W., & Raymond, F. H. (2010b). Measurement of internalized homonegativity in gay and bisexual men in Uganda: Cross-cultural properties of the Internalized Homonegativity Scale. *Psychology, Health, and Medicine, 15*, 159–165.

Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology, 71*, 136–149.

Arch Sex Behav (2010) 39:1213–1215
DOI 10.1007/s10508-010-9658-4

LETTER TO THE EDITOR

Depressive Symptoms Among Same-Sex Oriented Young Men: Importance of Reference Group

Ritch C. Savin-Williams · Kenneth M. Cohen · Kara Joyner · Gerulf Rieger

Published online: 21 July 2010

A recent meta-analysis indicated strong support for the commonly observed finding that, in contrast to heterosexual men, same-sex oriented men are at greater risk for mental health problems, especially affective disorders (King et al., 2008). The most common explanations for these findings are the increased prevalence of societal prejudice, stigma, and violence that allegedly generate "minority-stress" effects (King et al., 2008; Meyer, 2003; Saewyc, 2007). Hatzenbuehler (2009, p. 707) noted that stigma-related stress "gets under the skin" by inducing elevations in emotional dysregulation, which, in turn, confers risk for increased psychopathology among nonheterosexual populations. Although this explanation is reasonable, few other hypotheses have been proposed for the apparent elevation of depressive symptoms among same-sex attracted individuals.

The alternative considered here is that mental health discrepancies among male sexual orientation groups are illusory because they are derived from inappropriate (i.e., same biological sex) group comparisons. We adopt in a sex-inversion perspective—that cross-sex comparisons are more suitable for situations in which nonheterosexual men are assessed on sexually dimorphic variables.

An extensive literature documenting cross-sex behavior among numerous domains has consistently reported that, as a group, nonheterosexual men score significantly more similar to the opposite sex than heterosexual men on various measures of gender nonconformity (Bailey & Zucker, 1995; Cohen, 2002; Lippa, 2005; Rieger & Savin-Williams, 2010; Udry & Chantala, 2005). Indeed, the effect size for male sexual orientation difference appears to be among the largest sex-dimorphic discrepancies in the developmental literature (Bailey & Zucker, 1995). In a recent review, Bailey (2009) noted that same-sex oriented men are, on average, more sex atypical in their self-concepts, motor behavior, body movements, occupational careers, and recreational interests than heterosexual men. According to a recent Finnish twin study, substantial genetic effects account for variation in both sex-atypical behavior and adult sexual orientation, suggesting that at least part of this co-variation is due to biological influences (Alanko et al., 2010).

Given the robust finding that in some respects same-sex oriented men are more similar to heterosexual women than heterosexual men, we speculate that the basic principle of a correctly chosen reference group is critical in ascertaining the existence of male sexual orientation differences in mental health. To our knowledge, sex-atypical endorsement of sexually dimorphic mental health patterns has not been systematically addressed. We hypothesize that same-sex oriented men more closely resemble mental health patterns observed among heterosexual women than men. This would particularly be the case for mental health disorders that are especially sex dimorphic, such as depression. Indeed, a meta-analysis of sex differences in depression revealed that, beginning in early adolescence and continuing through adulthood, the prevalence of depression is more than twice as high among women as men (Twenge & Nolen-Hoeksema, 2002). This sex difference was replicated in King et al.'s (2008) meta-analysis for depression, anxiety, and suicidality.

To address this possibility, we analyzed data drawn from Wave 3 of Add Health, a comprehensive, representative school-based study of U.S. youth (Harris et al., 2003; Udry & Bearman, 1998). For Wave 3 (2001–2002), the project re-interviewed 15,197 of the original Wave 1 in-home participants, now aged

R. C. Savin-Williams (✉) · K. M. Cohen · G. Rieger
Human Development, Cornell University, MVR Hall,
Ithaca, NY 14853, USA
e-mail: savin-williams@cornell.edu

K. Joyner
Department of Sociology, Bowling Green State University,
Bowling Green, OH, USA

18–28 years. The response rates for Waves 1 and 3 were 79% and 77%, respectively. All included participants had to report the same sex at both waves to qualify for the sample. After excluding participants with missing data on main variables (sexual orientation and depression) and no longitudinal grand sample weight, the final sample was 13,992, of which 6,615 were heterosexual and nonheterosexual men and 6,351 were heterosexual women (the remainder were nonheterosexual women, not included in this study).

Interviews were conducted by professional staff who entered subjects' answers on a laptop computer. For sensitive questions, such as sexual orientation, interviewers handed participants the laptop on which they completed a computer-assisted self-interview; this enabled the young adults to enter responses to questions that appeared on screen and were heard on tape with earphones. Participants were asked to choose the description that best fits how they think about themselves. Options were 100% heterosexual (straight); mostly heterosexual (straight), but somewhat attracted to people of your own sex; bisexual—that is, attracted to men and women equally; mostly homosexual (gay), but somewhat attracted to people of the opposite sex; 100% homosexual (gay); or not sexually attracted to either males or females. The small number of participants who failed to report attractions were dropped ($n = 49$). Due to the small sample size of mostly homosexuals, they were combined with 100% homosexuals. Depressive symptoms were assessed with a 10-item derivative CES-D scale (Meadows, Brown, & Elder, 2006). Scale alpha was 0.81. Included were questions about the frequency of events during the past week such as "You felt depressed." Items were rated on a 4-point scale, ranging from 0 (never/rarely) to 3 (most of the time/all of the time).

Analyses incorporated survey design procedures to take into account Add Health's weighting (using the grand sample longitudinal weight for participants in Waves 1 and 3) and cluster design. Means and their 95% confidence intervals (CIs) were calculated to examine how depressive symptoms differed by sex and sexual orientation. Considered as a group and then as three sexual orientation subgroups (mostly heterosexual, bisexual, and homosexual), nonheterosexual men were compared to heterosexual men. These comparisons were then repeated with heterosexual women as the reference group.

As expected, when compared to heterosexual men, nonheterosexual men and heterosexual women reported significantly higher values for depressive symptoms (Table 1). By nonheterosexual subgroup, mostly heterosexual and homosexual men endorsed higher values of depressive symptoms than heterosexual men. Despite reporting the highest depressive value, bisexual men did not differ significantly from heterosexual males because of their small sample size. Nonheterosexual men as a group and all nonheterosexual subgroups did not differ significantly from heterosexual women in depressive symptoms.

Given the widely documented sex atypicality of same-sex attracted men across multiple domains, it may not be surprising that the elevation of affective symptoms among nonheterosexual men closely resembled that of heterosexual women. Although this finding has not been previously reported, it can be detected by re-analyzing King et al.'s

Table 1 Sexual orientation and depressive symptoms by male and female reference group

	Males ($n = 6615$)			
	Percentage/*M*	*SD*	95% CI	*t*
Sexual orientation (males)				
100% heterosexual ($n = 6238$)	0.944		(.936–.952)	–
All nonheterosexuals ($n = 377$)	0.056		(.048–.064)	–
Mostly heterosexual ($n = 204$)	0.032		(.026–.038)	–
Bisexual ($n = 41$)	0.006		(.003–.008)	–
Homosexual ($n = 132$)	0.018		(.013–.023)	–
Depressive symptoms[a]				
100% heterosexual males	4.61	0.07	(4.47–4.77)	–
All nonheterosexual males	6.05	0.31	(5.42–6.68)	4.65***
Mostly heterosexual males	6.11	0.48	(5.16–7.08)	3.22**
Bisexual males	6.32	0.97	(4.34–8.31)	1.81
Homosexual males	5.84	0.45	(4.94–6.75)	2.69**
100% heterosexual females ($n = 6351$)	5.90	0.09	(5.71–6.10)	–
All nonheterosexual males	6.05	0.31	(5.42–6.68)	0.45
100% heterosexual males	4.61	0.07	(4.47–4.77)	−12.15***
Mostly heterosexual males	6.11	0.48	(5.16–7.08)	0.44
Bisexual males	6.32	0.97	(4.34–8.31)	0.44
Homosexual males	5.84	0.45	(4.94–6.75)	−0.15

** $p < .01$; *** $p < .001$ (two-tailed significance test of difference in depressive symptoms

[a] Absolute range, 0–30. Symptoms between this group and 100% heterosexual males or females

(2008) meta-analysis: the percent of depression among nonheterosexual men (13.7%) was considerably closer to heterosexual women (10%) than to heterosexual men (5.9%). Thus, heterosexual women, rather than heterosexual men, may constitute a more appropriate reference group when exploring female-typical mental health concerns.

Whereas a minority-stress model would explain these findings by highlighting a shared pattern of devaluation, prejudice, and discrimination among nonheterosexual men and heterosexual women, we believe that a sex-inversion perspective is a more parsimonious interpretation. The underlying mechanisms, as well as reasons cross-sex shifts do not extend to all sex-dimorphic behaviors (e.g., desire for novel sex partners), remain uncertain. Wilson and Rahman (2005) suggested that neuropsychological differences between the sexes—which involve the hypothalamus, limbic system, and neurotransmitters, as well as higher neurological centers—may explain similarities between homosexual men and heterosexual women. We additionally suspect elevated female-typical sex-dimorphic traits that lead to increased endorsement of affective symptoms, such as interpersonal sensitivity, emotion-focused ruminative coping modality, and reporting style.

Once re-contextualized, the depressive symptoms reported by nonheterosexual men are not remarkable or even unexpected. By questioning the appropriate reference group for contrasting levels of mental health, our findings contribute to a growing call for depathologizing individuals who are not heterosexual (Savin-Williams, 2005; Savin-Williams & Joyner, 2010).

Acknowledgments This research used data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis. Use of this acknowledgment requires no further permission from the persons named.

References

Alanko, K., Santtila, P., Witting, K., Varjonen, M., Jern, P., Johansson, A., et al. (2010). Common genetic effects on childhood gender atypical behavior and adult sexual orientation. *Archives of Sexual Behavior, 39*, 81–92.

Bailey, J. M. (2009). What is sexual orientation and do women have one? In D. A. Hope (Ed.), *Nebraska symposium on motivation: Vol. 54: Contemporary perspectives on lesbian, gay, and bisexual identities* (pp. 43–63). New York: Springer.

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.

Cohen, K. M. (2002). Relationships among childhood sex-atypical behavior, spatial ability, handedness, and sexual orientation in men. *Archives of Sexual Behavior, 31*, 129–143.

Harris, K. M., Florey, F., Tabor, J., Bearman, P. S., Jones, J., & Udry, J. R. (2003). *The national longitudinal study of adolescent health: Research design*. Chapel Hill, NC: Carolina Population Center. Available at: http://www.cpc.unc.edu/projects/addhealth/design.

Hatzenbuehler, M. L. (2009). How does sexual minority stigma "Get under the skin"? A psychological mediation framework. *Psychological Bulletin, 135*, 707–730.

King, M., Semlyen, J., Tai, S. S., Killaspy, H., Osborn, D., Popelyuk, D., et al. (2008). A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. *BMC Psychiatry, 8*. doi:10.1186/1471-244X-8-70.

Lippa, R. A. (2005). Sexual orientation and personality. *Annual Review of Sex Research, 16*, 119–153.

Meadows, S. O., Brown, J. S., & Elder, G. H. (2006). Depressive symptoms, stress, and support: Gendered trajectories from adolescence to young adulthood. *Journal of Youth and Adolescence, 35*, 89–99.

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin, 129*, 674–697.

Rieger, G., & Savin-Williams, R. C. (2010). *Sexual orientation, gender nonconformity, and well- being*. Manuscript submitted for publication.

Saewyc, E. M. (2007). Contested conclusions: Claims that can (and cannot) be made from the current research on gay, lesbian, and bisexual teen suicide attempts. *Journal of LGBT Health Research, 3*, 79–87.

Savin-Williams, R. C. (2005). *The new gay teenager*. Cambridge, MA: Harvard University Press.

Savin-Williams, R. C., & Joyner, K. (2010). *Diversity among sexual orientation groups in depressive symptom levels: Examining the mediating role of risky behavior*. Unpublished manuscript, Cornell University, Ithaca, NY.

Twenge, J. M., & Nolen-Hoeksema, S. (2002). Age, gender, race, socioeconomic status, and birth-cohort differences on the Children's Depression Inventory: A meta-analysis. *Journal of Abnormal Psychology, 111*, 578–588.

Udry, J. R., & Bearman, P. S. (1998). *The national longitudinal study of adolescent health*. Retrieved April 30, 2003, from http://www.cpc.unc.edu/projects/addhealth/.

Udry, J. R., & Chantala, K. (2005). Risk factors differ according to same-sex and opposite-sex interest. *Journal of Biosocial Science, 37*, 481–497.

Wilson, G. D., & Rahman, Q. (2005). *Born gay: The psychobiology of sex orientation*. London: Peter Owen.

Arch Sex Behav (2010) 39:1217–1219
DOI 10.1007/s10508-010-9670-8

LETTER TO THE EDITOR

The Right Comparisons in Testing the Minority Stress Hypothesis: Comment on Savin-Williams, Cohen, Joyner, and Rieger (2010)

Ilan H. Meyer

Published online: 31 August 2010

Savin-Williams, Cohen, Joyner, and Rieger (2010) suggest that purported "mental health discrepancies among male sexual orientation groups are illusory." They suggest that their findings regarding gay men's depressive symptoms rebut minority stress explanation of sexual orientation mental health disparities. But as it pertains to minority stress theory, their rationale is flawed.

Minority stress is a frequently used framework for understanding observed health disparities between sexual minority and heterosexual populations (Herek & Garnets, 2007). Minority stress suggests that (1) lesbians, gay men, and bisexuals (LGB) comprise a disadvantaged social group that is subject to stigma and prejudice; (2) stigma and prejudice related to sexual orientation predispose LGB to excess stress; and (3) in turn, this excess stress may lead to adverse health outcomes and, thus, observed health disparities (Meyer, 2003). Studies using the minority stress perspective have consistently documented mental health disparities between LGB and heterosexual populations. This is evidenced, for example, in studies that used meta-analysis to summarize this literature (King et al., 2008; Meyer, 2003).

Savin-Williams et al. argued that the observed disparity is faulty because researchers have used an inappropriate reference group in studying mental health outcomes in gay and bisexual men. Typically, researchers compare gay and bisexual men with heterosexual men when they document mental health disparities; Savin-Williams et al. suggest that gay and bisexual men should be compared with heterosexual *women*, not men. The reason for this is that "cross-sex comparisons are more suitable for situations in which nonheterosexual men are assessed on sexually dimorphic variables." This claim is based on vast research that showed that gay and bisexual men are more like heterosexual women than heterosexual men on various measures (e.g., Bailey, 2009; Bailey & Zucker, 1995). For example, Savin-Williams et al. cite Bailey (2009), who has showed that "same-sex oriented men are, on average, more sex atypical in their self-concepts, motor behavior, body movements, occupational careers, and recreational interests than heterosexual men." Following this rationale, Savin-Williams et al. go on to show that gay and bisexual men's level of depressive symptoms is not different from that of heterosexual women. Thus, they conclude "once recontextualized, the depressive symptoms reported by nonheterosexual men are not remarkable or even unexpected."

Even if one would concede the premise of the argument—that "same-sex oriented men are more similar to heterosexual women than heterosexual men"—there is no merit to the claim that these findings undermine the minority stress perspective. The main reason for this is that Savin-Williams et al. make their argument based on one sexual minority subgroup—gay and bisexual men, excluding lesbians and bisexual women—and one outcome—depressive symptoms, excluding anxiety and substance use disorders. This cherry-picked comparison is not sufficient to test minority stress theory. Minority stress rests on sociological theory that links social structure with health outcomes (through the impact of stress). Therefore, it makes predictions about differentially situated groups (disadvantaged versus advantaged groups) and predicts similar patterns across various mental disorders (Schwartz & Meyer, 2010). Savin-Williams et al. err on both counts.

First, minority stress suggests that sexual minorities are socially disadvantaged in our society due to homophobia and heterosexism as a *group*—that is, across all subgroups, such as those defined by gender, race/ethnicity, etc. For the study

I. H. Meyer (✉)
Mailman School of Public Health, Columbia University,
Allan Rosenfield Building, 722 West 168th Street,
Room 533, New York, NY 10032, USA
e-mail: im15@columbia.edu

of minority stress, therefore, the groups compared ought to be all sexual minorities—men *and* women—versus all heterosexuals. Similarly, if gender inequality was studied, the reference group for women would be men; if race/ethnic inequality was studied, the reference group for race/ethnic minorities would be whites. This is because in each comparison we are interested in the *average* effect on the disadvantaged versus advantaged, that is, across diverse subgroups within (Schwartz & Meyer, 2010). Savin-Williams et al.'s hypothesis is refuted when one examines LGB as a group versus heterosexuals as a group. Consistent with the minority stress hypothesis, in such comparisons, LGB populations have higher rates of mental disorders than heterosexuals (King et al., 2008; Meyer, 2003).

Second, minority stress (and social stress theory more generally) is a sociological theory that predicts that disadvantaged social status affects the aggregate of mental disorders, rather than any particular disorder—we are interested in whether a disadvantaged group member has *any* disorder that is caused by minority stress. This is because social disadvantage and resultant stress are thought to be *generic* pathogens. Minority stress does not predict a *specific* impact on, say, depression versus anxiety and substance use disorders. Although this is not an infallible rule, researchers should provide a good reason to exclude one disorder or another. A reason to exclude a disorder would be, for example, that stress has no part in causing the disorder. In such a hypothetical situation, if stress does not play a causal role, the researcher could not reasonably hypothesize that minority stress would lead to excess disorder in the disadvantaged group.

This problem in Savin-Williams et al.'s proposed reference group hypothesis is reminiscent of a 1970s debate in the sociology of mental health about gender role stress and disorder. Gove and Tudor (1976) claimed that a higher level of depression in women compared with men is evidence of women's gender role stress. Critics argued, as we argue here, that selecting depression alone of all mental disorders is wrong because it provides opportunistic support for a favorite theory. When all mental disorders were compared, it became evident that women and men have similar levels of disorders, refuting the proposed gender role hypothesis (Dohrenwend & Dohrenwend, 1976). An alternative hypothesis was offered for the differing patterns of disorders, suggesting that men and women's stress response differed: Women internalize stress, resulting in higher levels of mood disorders, and men externalize stress, resulting in higher levels of substance use disorders and antisocial behaviors (Rosenfield, 1999).

If we relied on these gendered patterns of difference between internalizing and externalizing disorders that are characteristic of men and women in the general population, we would expect, based on Savin-Williams et al.'s proposed hypothesis, that gay and bisexual men would be similar to heterosexual women in patterns of substance used disorders. In fact, this is not the case: gay and bisexual men have a much higher prevalence of substance use problems, including substance use disorders, than heterosexual women (King et al., 2008; Meyer, 2003). As a result, when *all* mental disorders are considered, gay and bisexual men have higher levels of disorder than heterosexual women. This finding is consistent with the minority stress hypothesis but not with Savin-Williams et al.'s reference group hypothesis.

In summary, examining the evidence on sexual orientation, stress, and disorder, we find that sexual minorities have greater exposure to stress (Meyer, Schwartz, & Frost, 2008) and they have the expected resultant higher rates of disorder when compared with heterosexuals (King et al., 2008; Meyer, 2003). When we examine Savin-Williams et al.'s reference group hypothesis carefully in light of this evidence, we must refute it.

Finally, Savin-Williams et al. contend that their arguments "contribute to a growing call for depathologizing individuals who are not heterosexual." I reject the implication that minority stress theory pathologizes LGB individuals. Minority stress theory positions the source of stress, and therefore mental health problems, as stemming from prevailing societal-level sexual stigma, prejudice, and discrimination and not a reflection of individual traits. Although some politically-motivated persons may use evidence that LGB individuals have a higher prevalence of disorder than heterosexuals to pathologize, stigmatize, and discriminate against LGB persons, such arguments are misguided as they defy logic. During the debates that led to the removal of homosexuality as a mental disorder from DSM-II in 1973, Marmor (1980) noted how illogical it is to associate findings about prevalence of pathology in the group with pathologizing the group itself:

> …the basic issue…is not whether some or many homosexuals can be found to be neurotically disturbed. In a society like ours where homosexuals are uniformly treated with disparagement or contempt—to say nothing about outright hostility—it would be surprising indeed if substantial numbers of them did not suffer from an impaired self- image and some degree of unhappiness with their stigmatized status….It is manifestly unwarranted and inaccurate, however, to attribute such neuroticism, when it exists, to intrinsic aspects of homosexuality itself. (p. 400)

Minority stress theory points to pathogenic social conditions that stigmatize LGB people and treat them as inferior to heterosexuals. Even at the risk that research findings can be misused by some, studies on the psychiatric epidemiology of LGB individuals are important to help guide funding by governmental and other agencies and to direct research and prevention efforts.

References

Bailey, J. M. (2009). What is sexual orientation and do women have one? In D. A. Hope (Ed.), *Nebraska symposium on motivation: Vol. 54: Contemporary perspectives on lesbian, gay, and bisexual identities* (pp. 43–63). New York: Springer.

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.

Dohrenwend, B. P., & Dohrenwend, B. S. (1976). Sex differences and psychiatric disorders. *American Journal of Sociology, 81*, 1447–1454.

Gove, W. R., & Tudor, J. F. (1976). Adult sex roles and mental illness. *American Journal of Sociology, 78*, 812–835.

Herek, G. M., & Garnets, L. D. (2007). Sexual orientation and mental health. *Annual Review of Clinical Psychology, 3*, 353–375.

King, M., Semlyen, J., Tai, S. S., Killaspy, H., Osborn, D., Popelyuk, D., et al. (2008). A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. *BMC Psychiatry, 8*, doi:10.1186/1471-244X-8-70.

Marmor, J. (1980). Epilogue: Homosexuality and the issue of mental illness. In J. Marmor (Ed.), *Homosexual behavior: A modern reappraisal* (pp. 391–401). New York: Basic Books.

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin, 129*, 674–697.

Meyer, I. H., Schwartz, S., & Frost, D. (2008). Social patterning of stress exposures. *Social Science and Medicine, 67*, 368–379.

Rosenfield, S. (1999). Gender and mental health, do women have more psychopathology, men more, or both the same (and why)? In A. V. Horwitz & T. L. Scheid (Eds.), *A handbook for the study of mental health, social contexts, theories, and systems* (pp. 348–360). Cambridge: Cambridge University Press.

Savin-Williams, R. C., Cohen, K. M., Joyner, K., & Rieger, G. (2010). Depressive symptoms among same-sex oriented young men: Importance of reference group [Letter to the Editor]. *Archives of Sexual Behavior*. doi:10.1007/s10508-010-9658-4.

Schwartz, S., & Meyer, I. H. (2010). Mental health disparities research: The impact of within and between group analyses on tests of social stress hypotheses. *Social Science and Medicine, 70*, 1111–1118.

Arch Sex Behav (2010) 39:1221–1224
DOI 10.1007/s10508-010-9640-1

LETTER TO THE EDITOR

Gender Outcome in 46,XY Complete Androgen Insensitivity Syndrome: Comment on T'Sjoen et al. (2010)

Heino F. L. Meyer-Bahlburg

Published online: 15 June 2010

The case report by T'Sjoen et al. (2010) constitutes the third published case of an originally female-assigned individual with 46,XY complete androgen insensitivity syndrome (CAIS) who appeared to have developed a male gender identity. Of the three individuals, this third one underwent the most sophisticated genetic work-up, and the social history makes a contribution of social coercion to the identity development least likely. Therefore, this report raises challenging diagnostic, phenomenologic, and etiologic questions.

T'Sjoen et al. "believe this patient qualifies as female-to-male transsexual." Their diagnosis was predated by the psychotherapist's diagnosis of "childhood-onset gender identity disorder" made in the course of 5 years of psychotherapy that had started at age 23 years, i.e., 2 years before the initiation of androgen treatment and 4 years before mastectomy. The psychiatric nomenclature used in this case from Belgium was presumably that of the International Classification of Diseases, 10th revision (ICD-10; World Health Organization, 1992). In this context, the use of the diagnosis of F64.0 Transsexualism for a case of CAIS is puzzling, because the respective guidelines text clearly states that the "transsexual identity…must not be…associated with any intersex, genetic, or sex chromosome abnormality" (World Health Organization, 1992, p. 215; website p. 168). Perhaps the choice of diagnosis was influenced by the fact that this exclusion criterium has not been listed specifically for F64.2 Gender identity disorder of [prepubertal] childhood (World Health Organization, 1992, pp. 215–217; website pp. 168–170). In the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) of the American Psychiatric Association (2000), a "concurrent…physical intersex condition" is an exclusion criterium for Gender Identity Disorder regardless of pubertal stage. I have presented detailed arguments why the DSM-IV-TR diagnostic category of Gender Identity Disorder, as defined for persons without physical intersexuality (recently re-named "disorders of sex development" [DSD; Hughes et al., 2006]), should not be applied to persons with gender-identity problems in the context of a DSD (Meyer-Bahlburg, 1994, 2009). As T'Sjoen et al. did not discuss this issue in their report, it is unclear whether the transsexualism diagnosis was just an oversight on their part, or whether they assumed that, in this particular case, the patient-initiated gender change was distinctly different from that found in other cases of intersexuality and really more comparable to that of non-DSD transsexualism.

The case history, as presented retrospectively, is much in line with what one sees clinically in many cases of DSD with patient-initiated gender change. Although assigned female at birth and undiagnosed until age 17 years, this child showed substantial cross-gender behavior since the age of 3 years, with verbal expressions of the wish to be of the other gender. However, this patient must have agreed to a short course of estrogen treatment at age 18, which was soon stopped again because of the patient's aversive reaction to body feminization. Yet, subsequently, it took another 5 years and a revisit of her medical records for the patient to realize the potential implications of her 46,XY karyotype, which "supported the desire to make her body congruent with a long-standing male gender identity," but even then there was substantial "hesitation" about her presentation, and another 2 years of psychotherapy before beginning androgen treatment, etc. Transient periods of gender uncertainty are experienced by many patients with DSD (Meyer-Bahlburg et al., 2004b), and the slow gradual procession to gender change in later adolescence and adulthood also appears more common among

H. F. L. Meyer-Bahlburg (✉)
New York State Psychiatric Institute and Department of Psychiatry, Columbia University, 1051 Riverside Drive, Unit 15, New York, NY 10032, USA
e-mail: meyerb@childpsych.columbia.edu

patients with DSD than among persons with non-DSD GID, although we are lacking systematic comparison studies of this issue (Meyer-Bahlburg, 1994, 2009).

Prior to the report of a case of 46,XY CAIS with male gender identity by Kulshreshtha et al. (2009), the absence of documented similar cases in the world literature (see the review by Mazur, 2005) had led to a virtual textbook dogma that CAIS was associated with female psychological development. If this dogma were valid, one would expect that psychological studies of gender-related behavior in girls and women with 46,XY CAIS show means and variances comparable to those of non-DSD girls or women, if not even lower, given the normal variability in androgen exposure of girls and women and the effects of such exposure on gender-behavior variability (Auyeung et al., 2009; Udry, 2000). However, a careful examination of the limited psychological studies of CAIS that provide details on gender-related behavior and identity reveals a more variable gender outcome.

Early case reports (see citations in Negulici, Christodorescu, & Alexandru, 1968) described masculine traits in some patients with what was then called "testicular feminization," but it is not always clear whether they would be diagnosed as CAIS by today's standards. Already the first systematic psychological study of a group of patients with CAIS ($n = 10$) listed the "sex role preference" as "ambivalent" for one of these patients, while the other nine were coded as "content or prefers to be female"; two patients had "little or no interest" in clothing, make-up, and jewelry, and one preferred for clothing "slacks and shirts with dresses excluded"; sexual orientation was heterosexual in those nine with any information on this domain (Masica, Money, & Ehrhardt, 1971). A preceding publication on the same sample showed more variability in aspects of maternalism, with some patients acknowledging an influence of the awareness of their medical condition (Money, Ehrhardt, & Masica, 1968). Neither report provided comparison data from a non-DSD control group. Wisniewski et al. (2000) published data on 14 adult women with CAIS from the same medical institution (which, therefore, may have included some of the participants in the earlier study). All reported satisfaction with being a woman; one was homosexual in adulthood. There was variability in self-rated femininity and masculinity with at least one scoring in the masculine range for childhood (retrospectively), but the data cannot be evaluated for lack of a non-DSD control group.

A British study (Hines, Ahmed, & Hughes, 2003) compared 22 women over the age of 15 years to 22 non-DSD control women pairwise matched on sex-of-rearing, age, and race. Although there were no significant group differences in the means of the 13 self-report measures of gender behavior and identity, the CAIS women had larger variances on 12 of these measures, particularly strongly so for current gender identity ($p < .01$) and lifetime gender identity ($p < .05$, F-tested by myself). Unfortunately, 12 of the CAIS women were recruited through support groups and their CAIS classification was based on self-report. Moreover, the other 10 participating women with CAIS who were recruited from a clinical data base included only one-third of those eligible, and Hines et al. reported some degree of selection bias against those with "poor adjustment to the diagnosis." Thus, their inclusion might have further contributed to the variability in gender-related behavior.

A study of adults with DSD in Germany (Richter-Appelt, Discher, & Gedrose, 2005) included five women with CAIS. Their means on a female and a male gender identity/behavior composite scale were shifted in the masculine direction, with markedly enlarged variance, compared to non-DSD controls. Similarly, the women with CAIS retrospectively reported somewhat increased masculine and decreased feminine play activities during childhood. Brinkmann, Schweizer, and Richter-Appelt (2007) added that two of the women with CAIS reported "extremely low" femininity scores and three very low scores on a 4-item scale measuring certainty about belonging to their assigned gender. All five women were heterosexual (Schützmann, Brinkmann, Schacht, & Richter-Appelt, 2009).

The only study with results in the opposite direction was the one of children with XY karyotype and DSD, also from Germany, by Jürgensen, Hiort, Holterhus, and Thyen (2007), who found that their subgroup of six children with complete hypoandrogenization ranging in age from 3 to 12 years (five with CAIS and one with a complete steroid biosynthesis defect) had scores on several scales of gender behavior that were very similar in means and somewhat lower in variability compared to non-DSD control girls.

Although these diverse studies suffered from small sample sizes and possible selection biases, in combination they suggest more heterogeneity in the development of gender-related behavior and identity than one would expect on the basis of the stereotypic notion of CAIS femininity. From that perspective, the occasional development of a male gender identity and behavior in 46,XY CAIS appears less surprising. The causes of such heterogeneity in gender development of individuals with 46,XY CAIS are not yet understood. Regarding identity, clinical experience, pertinent case reports, and material from support group websites (e.g., Garrett & Kirkman, 2009) lets one expect that the combination of the awareness of the presence of a 46,XY karyotype, normal testes (prior to gonadectomy), and only a rudimentary vagina, and of the absence of ovaries and a uterus, and thereby the lack of child-bearing capacity, would easily raise self-doubts about their female identity for many women. The development of gender identity ought to be reconceptualized in any case as a more complex process than is commonly done (Egan & Perry, 2001), and recent findings from neuroscience (Keenan, Wheeler, & Ewers, 2003) need to be integrated with the psychosocial perspective. On the other hand, gender-related behavior tends to be more closely correlated with prenatal hormone history than gender identity is (Meyer-

Bahlburg et al., 2004a), and gender-atypical behavior in CAIS is harder to explain, especially during childhood, when most girls with CAIS (and in many cases their parents) do not yet know their medical status.

I agree with T'Sjoen et al. that their findings challenge the role of a functioning androgen-receptor pathway in the development of male gender identity and especially behavior. This challenge is even more acute, if one does not rely on rodent models alone as T'Sjoen et al. do, but takes also into consideration that in non-human primates androgens have an even more prominent role; in rhesus monkeys, androgens appear to be solely responsible for both masculinizaton and defeminization of sex-related behavior (Thornton, Zehr, & Loose, 2009; Wallen, 2009; Wallen & Hassett, 2009). On the other hand, our knowledge of the sexual differentiation of the brain has increased considerably: regarding the many genes involved in brain differentiation (Arnold, 2009), as well as the potential diverse mechanisms of steroid action, including genomic and non-genomic mechanisms (McEwen, 2009), tissue-specific steroid-hormone co-regulators (Thakur & Paramanik, 2009), neurotransmitters, and cell-to-cell signaling (McCarthy, Wright, & Schwarz, 2009). In addition, in the case of CAIS, one has to consider the possibility of genetic mosaicism due to post-zygotic mutation of the androgen receptor gene as suggested by T'Sjoen et al. Thus, there is much room for speculation and a clear need for additional empirical research to explain variability in gender outcome of individuals with CAIS.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Arnold, A. P. (2009). The organizational-activational hypothesis as the foundation for a unified theory of sexual differentiation of all mammalian tissues. *Hormones and Behavior, 55*, 570–578.

Auyeung, B., Baron-Cohen, S., Ashwin, E., Knickmeyer, R., Taylor, K., Hackett, G., et al. (2009). Fetal testosterone predicts sexually differentiated childhood behavior in girls and in boys. *Psychological Science, 20*, 144–148.

Brinkmann, L., Schweizer, K., & Richter-Appelt, H. (2007). Geschlechtsidentität und psychische Belastungen von erwachsenen Personen mit Intersexualität. *Zeitschrift für Sexualforschung, 20*, 129–144.

Egan, S. K., & Perry, D. G. (2001). Gender identity: A multidimensional analysis with implications for psychosocial adjustment. *Developmental Psychology, 37*, 451–463.

Garrett, C. C., & Kirkman, M. (2009). Being an XY female: An analysis of accounts from the website of the Androgen Insensitivity Syndrome Support Group. *Health Care for Women International, 30*, 428–446.

Hines, M., Ahmed, S. F., & Hughes, I. A. (2003). Psychological outcomes and gender-related development in complete androgen insensitivity syndrome. *Archives of Sexual Behavior, 32*, 93–101.

Hughes, I. A., Houk, C., Ahmed, S. F., Lee, P. A., LWPES Consensus Group, & ESPE Consensus Group. (2006). Consensus statement on management of intersex disorders. *Archives of Disease in Childhood, 91*, 554–563.

Jürgensen, M., Hiort, O., Holterhus, P.-M., & Thyen, U. (2007). Gender role behavior in children with XY karyotype and disorders of sex development. *Hormones and Behavior, 51*, 443–453.

Keenan, J. P., Wheeler, M. A., & Ewers, M. (2003). The neural correlates of self-awareness and self-recognition. In T. Kircher & A. David (Eds.), *The self in neuroscience and psychiatry* (pp. 166–179). Cambridge: Cambridge University Press.

Kulshreshtha, B., Philibert, P., Eunice, M., Khandelwal, S. K., Mehta, M., Audran, F., et al. (2009). Apparent male gender identity in a patient with complete androgen insensitivity syndrome [Letter-to-the-Editor]. *Archives of Sexual Behavior, 38*, 873–875.

Masica, D. N., Money, J., & Ehrhardt, A. A. (1971). Fetal feminization and female gender identity in the testicular feminizing syndrome of androgen insensitivity. *Archives of Sexual Behavior, 1*, 131–142.

Mazur, T. (2005). Gender dysphoria and gender change in androgen insensitivity or micropenis. *Archives of Sexual Behavior, 34*, 411–421.

McCarthy, M. M., Wright, C. L., & Schwarz, J. M. (2009). New tricks by an old dogma: Mechanisms of the organizational/activational hypothesis of steroid-mediated sexual differentiation of brain and behavior. *Hormones and Behavior, 55*, 655–665.

McEwen, B. S. (2009). Introduction: The end of sex as we once knew it. *Physiology & Behavior, 97*, 143–145.

Meyer-Bahlburg, H. F. L. (1994). Intersexuality and the diagnosis of gender identity disorder. *Archives of Sexual Behavior, 23*, 21–40.

Meyer-Bahlburg, H. F. L. (2009). Variants of gender differentiation in somatic disorders of sex development: Recommendations for Version 7 of the World Professional Association for Transgender Health's *Standards of Care*. *International Journal of Transgenderism, 11*, 226–237.

Meyer-Bahlburg, H. F. L., Dolezal, C., Baker, S. W., Carlson, A. D., Obeid, J. S., & New, M. I. (2004a). Prenatal androgenization affects gender-related behavior but not gender identity in 5–12 year old girls with congenital adrenal hyperplasia. *Archives of Sexual Behavior, 33*, 97–104.

Meyer-Bahlburg, H. F. L., Migeon, C. J., Berkovitz, G. D., Gearhart, J. P., Dolezal, C., & Wisniewski, A. B. (2004b). Attitudes of adult 46, XY intersex persons to clinical management policies. *Journal of Urology, 171*, 1615–1619.

Money, J., Ehrhardt, A. A., & Masica, D. N. (1968). Fetal feminization induced by androgen insensitivity in the testicular feminizing syndrome: Effect on marriage and maternalism. *The Johns Hopkins Medical Journal, 123*, 105–114.

Negulici, E., Christodorescu, D., & Alexandru, S. (1968). Psychological aspects of the testicular feminization syndrome. *Psychosomatic Medicine, 30*, 45–50.

Richter-Appelt, H., Discher, C., & Gedrose, B. (2005). Gender identity and recalled gender related childhood play-behaviour in adult individuals with different forms of intersexuality. *Anthropologischer Anzeiger, 63*, 241–256.

Schützmann, K., Brinkmann, L., Schacht, M., & Richter-Appelt, H. (2009). Psychological distress, self-harming behavior, and suicidal tendencies in adults with disorders of sex development. *Archives of Sexual Behavior, 38*, 16–33.

T'Sjoen, G., De Cuypere, G., Monstrey, S., Hoebeke, P., Freedman, F. K., Appari, M., et al. (2010). Male gender identity in complete androgen insensitivity syndrome. *Archives of Sexual Behavior*. doi:10.1007/s10508-010-9624-1.

Thakur, M. K., & Paramanik, V. (2009). Role of steroid hormone coregulators in health and disease. *Hormone Research, 71*, 194–200.

Thornton, J., Zehr, J. L., & Loose, M. D. (2009). Effect of prenatal androgens on rhesus monkeys: A model system to explore the

organizational hypothesis in primates. *Hormones and Behavior, 55*, 633–644.

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

Wallen, K. (2009). The organizational hypothesis: Reflections on the 50th anniversary of the publication of Phoenix, Goy, Gerall, and Young (1959). *Hormones and Behavior, 55*, 561–565.

Wallen, K., & Hassett, J. M. (2009). Sexual differentiation of behavior in monkeys: Role of prenatal hormones. *Journal of Neuroendocrinology, 21*, 421–426.

Wisniewski, A. B., Migeon, C. J., Meyer-Bahlburg, H. F. L., Gearhart, J. P., Berkovitz, G. D., Brown, T. R., et al. (2000). Complete androgen insensitivity syndrome: Long-term medical, surgical, and psychosexual outcome. *Journal of Clinical Endocrinology and Metabolism, 85*, 2664–2669.

World Health Organization. (1992). *Multiaxial version of ICD 10: Clinical descriptions and diagnostic guidelines.* Geneva: Author. (Website edition: www.who.int/classifications/icd/en/bluebook.pdf).

Arch Sex Behav (2010) 39:1225–1227
DOI 10.1007/s10508-010-9661-9

LETTER TO THE EDITOR

Problems with Ascertainment

Charles Moser

Published online: 23 July 2010

The Paraphilias subworkgroup (PSWG; see Zucker, 2010) has proposed a number of revisions for the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM), to be published by the American Psychiatric Association (APA). One supposedly new proposal is the recognition that the paraphilias are not *ipso facto* psychiatric disorders, but the existence of nonpathological unusual sexual interests was explicit and identical in both DSM-IV and DSM-IV-TR:

> A Paraphilia must be distinguished from the nonpathological use of sexual fantasies, behaviors, or objects as a stimulus for sexual excitement in individuals without a Paraphilia. Fantasies, behaviors, or objects are paraphilic only when they lead to clinically significant distress or impairment (e.g., are obligatory, result in sexual dysfunction, require participation of nonconsenting individuals, lead to legal complications, interfere with social relationships). (APA, 1994, p. 525; APA, 2000, p. 568; boldfacing in both originals removed)

In practice, the distinction between pathological and nonpathological unusual sexual interests has been ignored.[1] The "pathological" criteria often describe individuals with normophilic sexual interests and do not distinguish normophilic individuals from individuals with a paraphilia (Moser & Kleinplatz, 2005a, 2005b).

The above differential diagnosis language was adopted after a similar attempt to distinguish a nonstandard sexual interest from a mental disorder characterized by distress related to that interest was abandoned. The diagnosis of "Ego-dystonic Homosexuality" (APA, 1980) referred to individuals who were distressed by their homosexual arousal pattern, but a homosexual arousal pattern was not (and is not) considered a mental disorder. This diagnosis, which did not meet "the definition of a disorder" (Drescher, 2010, p. 435) was removed from the next edition of the DSM (see APA, 1987, p. 426), and thus "the APA implicitly accepted a normal variant view of homosexuality" (Drescher, 2010, p. 435).

One of the arguments supporting the removal of Ego-dystonic Homosexuality from the DSM was that its existence implied that other types of ego-dystonic characteristics were also mental disorders. Distress or impairment related to one's race, height, or normophilic sexual interests (e.g., ego-dystonic masturbation) could lead inappropriately to the creation of new psychiatric disorders (see Drescher, 2010). It is not clear why the DSM editors have not recognized that this argument also applies to the Paraphilias.

Despite these problems, the PSWG seems intent on trying to perpetuate the distinction between the interest versus the disorder which results from the interest yet again. According to this proposal, individuals are to be ascertained with a paraphilia, which is not a mental disorder. If the ascertained paraphilia "causes" distress or impairment, then the individual would be diagnosed with a paraphilic disorder (see Blanchard, 2010), which would be classified as a mental disorder.

The specific types of distress and impairment associated with paraphilic disorders have not been delineated (Moser, 2009). Apparently, the PSWG believes the distress must result from the sexual interest itself. Classifying a disorder according to the precipitant of the distress, rather than the characteristics of the

C. Moser (✉)
Department of Sexual Medicine,
Institute for Advanced Study of Human Sexuality,
45 Castro Street, #125, San Francisco, CA 94114, USA
e-mail: Docx2@ix.netcom.com

[1] This Letter focuses on the non-criminal paraphilias (fetishism, partialism, sexual masochism, consensual sexual sadism, and transvestism). The reader should not construe the present Letter to support the decriminalization of any sex crime.

distress produced, confuses the meaning of a mental disorder; it focuses the treatment away from the distress to whatever precipitated the distress. Becoming unemployed may result in distress for many individuals, but "unemployment disorder" is not a mental disorder.

Assuming the distress or impairment resolves, would the individual's status revert to paraphilic ascertainment, a paraphilic disorder in remission, or continue with the paraphilic disorder diagnosis? An update on the DSM-5 website suggests that the PSWG is leaning towards the "In Remission" option (APA, 2010), which confuses the goals of treatment. Is it the remission of the ascertained paraphilia or remission of the distress or impairment associated with the paraphilic disorder? Usually, the "In Remission" specifier is applied to mental disorders with characteristic relapsing courses, but which aspect of diagnosis is expected to relapse (the ascertained interest or the diagnosed distress or impairment)?

The creation of the ascertainment category does not imply that the ascertained individual is healthy—only that the individual does not meet the definition of a mental disorder *yet*. Whether the "patient" indicates the sexual interest causes the distress or impairment is not germane to the diagnosis. It is the mental health professional who "ascertains" the diagnosis. Most people seek mental health treatment because they are experiencing distress or impairment in some form. It is difficult to imagine that unusual sexual interests, denigrated by society, would not add to the distress or impairment resulting from an unrelated problem. Thus, the distinction between ascertaining a paraphilic interest and diagnosing a disorder could be meaningless in practice.

Blanchard (2010) suggests that once an individual is ascertained with a paraphilia, then that ascertainment can continue even if the intense and recurrent sexual arousal is no longer present. If one can have a paraphilia without the requisite intense and recurrent sexual arousal, then a paraphilic disorder without *current* distress or impairment also may exist. The duration of the distress or impairment is not indicated and it is common for individuals with paraphilic (or homosexual) interests to experience some distress as part of the "coming out" process. A moment of angst could label someone as mentally disordered forever.

Societal discrimination against lesbians, gays, and bisexuals is known to increase the incidence of non-sexual psychiatric disorders in these individuals (Hatzenbuehler, McLaughlin, Keyes, & Hasin, 2010). Assuming societal discrimination has a similar effect on individuals ascertained with a paraphilia, one should not diagnose a paraphilic disorder in addition to non-sexual psychiatric diagnoses. If the revelation of someone's homosexual (or heterosexual) interests precipitates the dissolution of his or her primary relationship resulting in a clinical depression, most mental health professionals would diagnose just the clinical depression. If the dissolution was precipitated by the revelation of his or her paraphilic interests, it seems inappropriate to diagnose both clinical depression and a paraphilic disorder.

Those with paraphilic interests could also be disadvantaged even if they do not seek mental health treatment. For example, in a child custody proceeding, an estranged spouse could expose the other spouse's ascertained paraphilia to those evaluating parental fitness. The distress from the exposure of one's sexual interests, the impending divorce, or the fear of losing custody of the child may now satisfy the distress or impairment criterion for a paraphilic disorder. An individual with a diagnosed paraphilic disorder, or even an ascertained paraphilia, can be seen as a less suitable parent and lose custody (Klein & Moser, 2006).

It is reasonable to assume that those individuals whose distress or impairment is associated with their paraphilic interests have a mental disorder or might benefit from treatment, but this reintroduces the problems associated with the Ego-dystonic Homosexuality diagnosis. There is no proposal to ascertain homosexual (or heterosexual) interests. Individuals who have distress or impairment related to normophilic interests can be diagnosed with the less stigmatizing Adjustment Disorder, Identity Problem, or Sexual Disorder NOS, without specifying the problematic interest. No empirical evidence or rationale is given to support the different treatment of distress or impairment for normophilic versus paraphilic interests.

It can be argued that the option of ascertainment represents a political compromise between the status quo and the formal removal of the paraphilias from DSM-5, analogous to the compromise which replaced Homosexuality with Ego-dystonic Homosexuality (Drescher, 2010). Nevertheless, political compromises should not override the current state of the science. The DSM editors and PSWG should indicate clearly the data on which they base their decisions. If the data exist, they should be presented. If the data do not exist, that should be stated unambiguously. The lack of empirical data supporting the proposed changes or the continued pathologizing of the paraphilias in the DSM is a glaring omission in a supposedly scientific document.

I object strongly to PSWG member Långström's (2010) statement, "I am not convinced that psychiatric nosology should change primarily because of the potential or actual misuse of diagnoses in the judicial system" (p. 323). It clearly is psychiatry's responsibility to minimize the misuse of its diagnoses; a basic dictum of all medicine is "First, do no harm." The usefulness of the paraphilic disorder diagnoses (or ascertainments) must be weighed against the damage from their misuse to justify their continued listing in the DSM-5. Why should psychiatry be exempt from these risk–benefit analyses?

It is clear that the paraphilic disorder diagnoses have been misused (Klein & Moser, 2006; Kolmes, Stock, & Moser, 2006; Wright, 2006). The usefulness of these diagnoses is not apparent and the arguments defending their usefulness are egregiously absent. From a historical perspective, it is doubtful that the creation of the ascertainment category will prevent further misuse of these diagnoses. It appears the concept of ascertainment will

confuse rather than clarify the issues with these diagnoses. If psychiatry cannot create a clear nosology to minimize the misuse of these diagnoses and appropriately exclude misdiagnoses, then these diagnoses (and ascertainments) should be removed until they are able to accomplish that task (Moser & Kleinplatz, 2005a).

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., revised). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

American Psychiatric Association. (2010). *DSM-5 development, sexual and gender identity disorders*. Retrieved June 13, 2010, from http://www.dsm5.org/ProposedRevisions/Pages/SexualandGenderIdentityDisorders.aspx

Blanchard, R. (2010). The DSM diagnostic criteria for transvestic fetishism. *Archives of Sexual Behavior, 39*, 363–372.

Drescher, J. (2010). Queer diagnoses: Parallels and contrasts in the history of homosexuality, gender variance, and the *Diagnostic and Statistical Manual*. *Archives of Sexual Behavior, 39*, 427–460.

Hatzenbuehler, M. L., McLaughlin, K. A., Keyes, K. M., & Hasin, D. S. (2010). The impact of institutional discrimination on psychiatric disorders in lesbian, gay, and bisexual populations: A prospective study. *American Journal of Public Health, 100*, 452–459.

Klein, M., & Moser, C. (2006). SM (sadomasochistic) interests as an issue in child custody proceedings. *Journal of Homosexuality, 50*(2/3), 233–242.

Kolmes, K., Stock, W., & Moser, C. (2006). Investigating bias in psychotherapy with BDSM clients. *Journal of Homosexuality, 50*(2/3), 301–324.

Långström, N. (2010). The DSM diagnostic criteria for exhibitionism, voyeurism, and frotteurism. *Archives of Sexual Behavior, 39*, 317–324.

Moser, C. (2009). When is an unusual sexual interest a mental disorder? [Letter to the Editor]. *Archives of Sexual Behavior, 38*, 323–325.

Moser, C., & Kleinplatz, P. J. (2005a). DSM-IV-TR and the paraphilias: An argument for removal. *Journal of Psychology and Human Sexuality, 17*(3/4), 91–109.

Moser, C., & Kleinplatz, P. J. (2005b). Does heterosexuality belong in the DSM? *Lesbian & Gay Psychology Review, 6*, 261–267.

Wright, S. (2006). Discrimination of SM-identified individuals. *Journal of Homosexuality, 50*(2/3), 217–231.

Zucker, K. (2010). Reports from the DSM-V Work Group on Sexual and Gender Identity Disorders [Editorial]. *Archives of Sexual Behavior, 39*, 217–220.

Arch Sex Behav (2010) 39:1229–1230
DOI 10.1007/s10508-010-9651-y

LETTER TO THE EDITOR

Depathologizing Consensual Sexual Sadism, Sexual Masochism, Transvestic Fetishism, and Fetishism

Susan Wright

Published online: 15 July 2010

The DSM-V Paraphilias Subworkgroup's suggested revision to differentiate between paraphilias and Paraphilic Disorders appears to be a step forward in depathologizing unusual sexual interests. Paraphilia diagnoses are regularly misused in criminal and civil proceedings as an indication that individuals cannot control their behavior; these individuals turn for assistance to the National Coalition for Sexual Freedom (NCSF), a national advocacy organization that advances the rights of, and advocates for, consenting adults in the BDSM-leather-fetish, swing, and polyamory communities.

One recent child custody case referred to NCSF illustrates the common misunderstanding that legal and social service professionals have with the DSM-IV-TR, and is the first documented reaction to the proposal to differentiate between paraphilias and Paraphilic Disorders. The children were removed in July 2009 while psychological evaluations were performed on the mother and the children, which concluded there was no mental illness.[1] However, the case worker with the Department of Social Services (DSS) Children's Division in the midwestern state where this case occurred sent the following January 21, 2010 letter to the mother's court appointed psychologist:

> With regards to [mother's] alternative lifestyle […] can she separate this from her parenting? There has [sic] been some questions arise [sic] from other team members regarding her sexual sadism. These are as follows: We were made aware at the last FST meeting that while all parties involved have seen the information provided regarding [mother's] blog and website, no action has been taken to determine how it affects the children or is factored into the stated case goal of reunification with [mother]. The following information is relevant: A. Sexual Sadism is considered a form of paraphilia in accordance with the DSM-IV-TR. B. [Mother] admitted in court on March 9, 2009 that she was a "domme"—slang for a female sexual sadist. C. Sexual Sadism involves inflicting pain and suffering on another individual in order to achieve sexual arousal…. Sexual sadism on the web has the following information: The essential feature of sexual sadism is a feeling of sexual excitement resulting from administering pain, suffering or humiliation on another person. In extreme cases, sexual sadism can lead to serious injury or death for the other person. According to the DSM[,] these catastrophic results are more likely when the paraphilia is diagnosed as severe, and when it is associated with antisocial personality disorder.[2] They may experience distressed or impaired functioning because of the sadistic behaviors or fantasies. This distress and impairment may be due to the fact that the partner is not consenting. The diagnosis of sexual sadism is complicated by several factors, beginning with the fact that most persons with the disorder do not voluntarily enter therapy. [Mother] indicated she gave up this lifestyle in March. However[,] the blog and stories that were found were posted to her website in May. There are concerns that she is still a moderator of the [BDSM] yahoo group. I have attached pages from her website in hopes that you can explore with [mother] her current involvement with this alternative lifestyle.

S. Wright (✉)
National Coalition for Sexual Freedom, 410 Guilford Ave., #127, Baltimore, MD 21202, USA
e-mail: susan@ncsfreedom.org

[1] All names and locations have been removed to protect the identity of those involved.

[2] Antisocial personality disorder was bolded in the original letter despite there being no evidence the mother has antisocial personality disorder.

The DSS letter concludes with the recommendation that: "Even though [mother] is complying with attendance in therapy, we feel the above issues need to be explored and addressed."

At the final permanency hearing in February 2010, the mother's lawyer submitted to the judge the proposed revisions for the DSM-V to separate the paraphilias from Paraphilic Disorders, resulting in a court determination to re-evaluate her entire case. The judge specifically chastised the Department of Social Services for not being aware of the proposed changes for the DSM-V. Based on the proposed revisions, in March the mother was awarded custody of three of the children, with the father retaining custody of one child in order to take advantage of his health care coverage.

As this example shows, when individuals who practice BDSM are brought to the attention of authorities, they are regularly misdiagnosed with a mental disorder. In 2009, NCSF was asked for help by 132 people regarding child custody/divorce issues directly involving their alternative sexual practices (NCSF, 2009). The year before, a total of 157 individuals contacted NCSF for help with child custody/divorce issues (NCSF, 2008). In total, almost 500 people each year request help from NCSF because of discrimination or persecution due to their alternative sexual practices.

Therefore, the implications of "ascertaining a paraphilia" and "diagnosing a paraphilic disorder" are critical to depathologizing consensual paraphilias. I am concerned that if sexual sadism receives its own diagnosis code separate from Sexual Sadism Disorder, social services and legal professions will continue to think that anyone who practices consensual sexual sadism (or sexual masochism, fetishism, and transvestic fetishism) therefore has a mental disorder.

The consensual paraphilias should be mentioned as the healthy comparison to a Paraphilic Disorder, much like various sexual behaviors are referred to in the proposed Hypersexual Disorder. Cybersex and masturbation don't have separate diagnostic codes in the DSM, and it is equally erroneous to assign separate codes for the paraphilias when they are not mental disorders or of clinical concern. For the same reason, the consensual paraphilias shouldn't be listed among the V-Codes.

Separating sexual behaviors (paraphilias) from the mental disorders (Paraphilic Disorders) is the first step in depathologizing consensual alternative sex. The second step is defining what exactly constitutes clinically significant distress. NCSF often consults with individuals who suffer distress and impairment in their social and occupational lives (i.e., interpersonal difficulties) because their desires conflict with current societal standards. These standards stem in a large part from the DSM itself: pathologizing unusual sexual interests has led to increased discrimination and discouraged individuals from seeking treatment for physical and mental health problems (Wright, 2008).

A distinction must be made in the DSM-V between distress imposed by societal stigma, and distress that is generated internally. As seen in the above referenced child custody case, mental health professionals are not the only ones who consult the DSM. When attorneys, judges, and social workers read the diagnoses in the DSM, they see "distressed or impaired" as the determiner of mental illness. Without a comprehensible definition, they look at the individual who is on trial or in a child custody battle, and that individual certainly appears distressed. They even speculate that if the individual gave up their BDSM practices, then their life wouldn't be in disarray, so clearly they must be suffering a mental disorder because their sexual behaviors are obligatory or "obsessive."

Therefore, the current list by which distress and impairment are diagnosed must be rejected: (1) are obligatory, (2) result in sexual dysfunction, (3) require participation of non-consenting individuals, (4) lead to legal complications, or (5) interfere in social relationships. Legal complications and interpersonal difficulties are common consequences of the stigma and discrimination against BDSM practices. In the Second National Survey of Violence & Discrimination Against Sexual Minorities, a total of 1,146 (37.5%) of the participants indicated that they had either been discriminated against or had experienced some form of harassment or violence (Wright, 2008). As a result, 60% of the 3,000 respondents are not "out" about their BDSM interests; the stress of being closeted and/or coming out promotes distress and impairment in these individuals, similar to that experienced by homosexuals.

In addition, once a Paraphilic Disorder is diagnosed, can it ever be in remission? If so, what are the mechanisms for determining that? If the distress and impairment are resolved, does the individual go back to the ascertainment category? As of now, once a mental disorder is diagnosed, it appears to apply for the lifetime of the individual.

Finally, it must be made clear that Paraphilic Disorders are extremely rare. In particular, the descriptive text for Sexual Sadism Disorder needs to clearly state that it is limited to forensic populations, and, as Krueger (2010) stated, "virtually all of the published papers using DSM criteria for Sexual Sadism have been done on studies of forensic populations." This will help prevent the conflation of those who practice consensual paraphilias with those who have a Paraphilic Disorder.

References

Krueger, R. B. (2010). The DSM diagnostic criteria for sexual sadism. *Archives of Sexual Behavior, 39*, 325–345.

National Coalition for Sexual Freedom. (2008). *2008 Incident Response Report*. Available at https://www.ncsfreedom.org/index.php?option=com_keyword&id=308.

National Coalition for Sexual Freedom. (2009). *2009 Incident Response Report*. Available at https://www.ncsfreedom.org/index.php?option=com_keyword&id=318.

Wright, S. (2008). *Second National Survey of Violence & Discrimination Against Sexual Minorities*. Available at http://www.ncsfreedom.org/images/pdfs/BDSM_Survey/2008_bdsm_survey_analysis_final.pdf.

Arch Sex Behav (2010) 39:1231–1233
DOI 10.1007/s10508-010-9650-z

LETTER TO THE EDITOR

What's Wrong With Sex?

Patrick Singy

Published online: 13 July 2010

Several psychiatrists and psychologists have recently made important criticisms of specific diagnostic innovations proposed by the *DSM-5*'s Sexual and Gender Identity Disorders Work Group (DeClue, 2009; Frances, 2010; Green, 2010; O'Donohue, 2010). As a historian and philosopher of science, my goal here is more general: I want to show that what this Work Group is trying to accomplish undermines the definitions of "paraphilia" and "mental disorder" that have been operative since the *DSM-III*. If the revisions proposed by the Work Group are implemented, the *DSM-5* will be closer to the *DSM-I* and *DSM-II* than to their successors. In order to understand why this is so, we need first to take a short historical detour and to look at how "mental disorder" and "paraphilias" have traditionally been defined. Only then will we be in a position to grasp the magnitude of what the *DSM-5* is trying to accomplish.

As is well known, the *DSM-I* (American Psychiatric Association, 1952) and *DSM-II* (American Psychiatric Association, 1968) did not define "mental disorder." In the *DSM-II*, the paraphilias (called "sexual deviations") were simply described as sexual interests directed "primarily toward objects other than people of the opposite sex, toward sexual acts not usually associated with coitus, or toward coitus performed under bizarre circumstance" (American Psychiatric Association, 1968, p. 44). Needless to say, it is rather bizarre to use "bizarre" as a nosological criterion, and probably not the best way to make psychiatry look like a serious science. In 1973, however, psychiatrists voted to remove homosexuality from the *DSM*, and at that point the need arose to spell out what a mental disorder is.

P. Singy (✉)
The Bioethics Program, Union Graduate College, 80 Nott Terrace, Schenectady, NY 12308, USA
e-mail: pbsingy@gmail.com

According to the *DSM-III* (American Psychiatric Association, 1980), a mental disorder must meet two criteria: (1) it must cause harm, either to oneself (in the form of distress or disability) or to others; (2) it must be a dysfunction. The fact that both criteria have to be met is a guarantee that doctors cannot transform every problem into a disease. Childbirth, for instance, is not a disease even when it causes harm, because it is not dysfunctional. Nor is homosexuality a disease since, even if it were dysfunctional (as some have claimed), it is not harmful.

The weakness of this understanding of mental disorder hinges mostly on the idea of dysfunction, which, as long as it is not better defined, can easily be interpreted in sociocultural terms. The following example clearly illustrates how this can have troublesome consequences. In 1851, Dr. Samuel A. Cartwright described a condition that he called "drapetomania," a disease that made slaves try to flee captivity (Cartwright, 1851). Is "drapetomania" a real disease? If we apply the *DSM-III* definition of mental disorder and use a sociocultural understanding of function, the answer must be "yes": slaves who try to escape from their masters indeed do not "function" in a racist society (and they harm their masters by depriving them of their property).

"Function" must therefore be clearly distinguished from sociocultural values, lest psychiatry be exactly what anti-psychiatrists have for decades accused it of being: a mere instrument of social control. This is why some of the psychiatrists who in the past have been in charge of the *DSM*, such as Robert L. Spitzer and Michael B. First, have recently pushed for a narrower understanding of "function" (Spitzer, 1999; Wakefield & First, 2003). They want to make psychiatry into a science whose job would not be to mirror and to fuel social prejudices, but to cure genuine diseases. They have usually relied on the influential work of Wakefield (1992).

According to Wakefield, a function should be understood as an internal mechanism selected by evolution. The theory of

evolution grounds the concept of "mental disorder" in biological and psychological science, and thus makes it less directly dependent on sociocultural values. Now we can say, for instance, that Cartwright's "drapetomania" is not only revolting by modern cultural standards (we hope) but also scientifically wrong. From an evolutionary point of view, trying to run away from an oppressor is certainly not dysfunctional.

So how will the *DSM-5* be different from its predecessors? The recent papers published by the Sexual and Gender Identity Disorders Work Group betray an ambition to weaken drastically, if not to dismantle altogether, both the "dysfunction" and the "harm" criteria.

Let's look at "dysfunction" first. The Work Group is perfectly content to turn its back on efforts to make the criterion of "dysfunction" independent of cultural values. For instance, Thornton, an official advisor to the *DSM-5*, claims that a paraphilia is "an abnormal sexual interest," adding matter-of-factly that "what counts as 'abnormal' is culturally relative" (Thornton, 2010, p. 411).

Blanchard (2010), the Chair of the Paraphilias Subwork Group, defines a paraphilia as "any powerful and persistent sexual interest other than sexual interest in copulatory or precopulatory behavior with phenotypically normal, consenting adult human partners" (p. 367). Although it is difficult to make sense of this hodgepodge definition, one thing seems clear: some kind of cultural norm must be the glue that holds together the odd conjunction of copulation (a behavior), consent (a legal concept), and phenotypical normality (whatever that means).

The official *DSM-5* website (www.dsm5.org) confirms the reduction of perversions to cultural abnormality. No definition of "paraphilia" is actually given in the paraphilia section of the website, where one would naturally expect to find it. But under the "Rationale" heading explaining the addition of the disease of "hypersexuality," this new (non-paraphilic) disease is contrasted with paraphilias, which are said to be "characterized by persistent, socially anomalous or deviant sexual arousal."

All these statements point in the same disturbing direction: the *DSM-5* is ready, and perhaps even eager, to crack a moral whip and to make psychiatrists into the guardians of cultural values. The pathological would simply be whatever society deems bizarre and morally unacceptable. In this respect, the *DSM-5* would more closely resemble the *DSM-I* and *DSM-II* than later editions.

But it gets worse. As if reducing the "dysfunction" criterion to cultural disapproval were not problematic enough (both scientifically and ethically), the Work Group is also trying to undermine the critical importance of the "harm" criterion. If it succeeds, a "harmless paraphilia" would indeed no longer be an oxymoron: people who enjoy sex in a way that harms no one, yet happens to be outside the cultural norm, would now have a paraphilia and would find their place in the *DSM-5* (Blanchard, 2010, p. 367; Zucker, 2010; www.dsm5.org). By contrast, from the *DSM-III* to the current *DSM-IV-TR*, such people have not been considered paraphiliacs and have been left alone.

Strictly speaking, the *DSM-5* claims that paraphilias, unlike paraphilic disorders, are not actual disorders, precisely because they are not harmful. But, obviously, neither are paraphilias completely unrelated to paraphilic disorders; otherwise, they would not even be mentioned in the *DSM*. It is, in fact, easy to predict that the distinction between paraphilias and paraphilic disorders will quickly crumble under relentless pressure from sex therapists and the pharmaceutical industry. Sure, the Work Group claims that "a paraphilia by itself would not automatically justify or require psychiatric intervention" (www.dsm5.org). But as a preventive measure, wouldn't it be safer to deal with a paraphilia now, when it is still benign, rather than later, when it will have grown into a full-blown paraphilic disorder? After all, if your doctor were to discover an unsuspected cancerous tumor in your body, would you wait until it causes painful symptoms to treat it? There is no reason to think that psychiatrists would not use the same kind of reasoning with paraphilias.

The *DSM* is taking a turn for the worse. If the Sexual and Gender Identity Disorders Work Group gets its way, it won't be long before those who don't copulate the straight and narrow way are bombarded with advertisements for pills and encouraged to start therapy before it's too late, weeping on talk shows as they share with us their painful struggles with abnormality. We are heading toward the creation of an epidemic of perverts—not because more people will engage in bizarre sexual behaviors as a result of the publication of the *DSM-5*, of course, but because a much wider range of sexual behaviors will be deemed pathological.

Obviously, sexual minorities should be especially upset about the direction taken by the *DSM-5*. What is needed is a new wave of powerful activism that would finish the job started in 1973 and force psychiatrists to remove not just one perversion from the *DSM*, but all of them. The law is sufficient to deal with those who engage in illegal sexual activities, and religious and cultural norms do not need the help of psychiatry to normalize harmless erotic quirks.

But psychiatrists themselves should also be concerned about the innovations of some of their colleagues. Historically, the scientific status of psychiatry has always been fragile. Tainted by accusations of racism, misogyny, and homophobia, it has been at pains to establish itself as a science worthy of the name. If the *DSM-5* does not quickly go through a radical revision, its publication will be yet another blow to the scientific reputation of psychiatry.

References

American Psychiatric Association. (1952). *Diagnostic and statistical manua: Mental disorders*. Washington, DC: Author.

American Psychiatric Association. (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington, DC: Author.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

Blanchard, R. (2010). The DSM diagnostic criteria for Transvestic Fetishism. *Archives of Sexual Behavior, 39*, 363–372.

Cartwright, S. (1851). Report on the diseases and physical peculiarities of the Negro race. *The New-Orleans Medical and Surgical Journal, 7*, 691–715.

DeClue, G. (2009). Should hebephilia be a mental disorder? A reply to Blanchard et al. (2008) [Letter to the Editor]. *Archives of Sexual Behavior, 38*, 317–318.

Frances, A. (2010, March 17). DSM5 and sexual disorders—Just say no. *Psychiatric Times.*

Green, R. (2010). Sexual preference for 14-year-olds as a mental disorder: You can't be serious!! [Letter to the Editor]. *Archives of Sexual Behavior, 39*, 585–586.

O'Donohue, W. (2010). A critique of the proposed DSM-V diagnosis of pedophilia [Letter to the Editor]. *Archives of Sexual Behavior, 39*, 587–590.

Spitzer, R. L. (1999). Harmful dysfunction and the DSM definition of mental disorder. *Journal of Abnormal Psychology, 108*, 430–432.

Thornton, D. (2010). Evidence regarding the need for a diagnostic category for a coercive paraphilia. *Archives of Sexual Behavior, 39*, 411–418.

Wakefield, J. C. (1992). Disorder as harmful dysfunction: A conceptual critique of DSM-III- R's definition of mental disorder. *Psychological Review, 99*, 232–247.

Wakefield, J. C., & First, M. B. (2003). Clarifying the distinction between disorder and nondisorder. In K. A. Phillips, M. B. First, & H. A. Pincus (Eds.), *Advancing DSM: Dilemmas in psychiatric diagnosis* (pp. 23–55). Arlington, VA: American Psychiatric Association.

Zucker, K. J. (2010, February 22). DSM5: Pandora replies to Dr Frances. *Psychiatric Times.*

Arch Sex Behav (2010) 39:1235–1237
DOI 10.1007/s10508-010-9654-8

LETTER TO THE EDITOR

Disregarding Science, Clinical Utility, and the DSM's Definition of Mental Disorder: The Case of Exhibitionism, Voyeurism, and Frotteurism

Andrew C. Hinderliter

Published online: 12 August 2010

In Långström's (2010) methodology section, he says that he "carefully read prior versions of the DSM diagnostic criteria for these paraphilias from the DSM-III to DSM-IV-TR" (p. 317), but his review makes some serious errors. Långström writes:

> In the DSM-III-R (American Psychiatric Association, 1987), the core criterion (A) for Exhibitionism (Over a period of at least 6 months, recurrent, intense sexual urges and sexually arousing fantasies involving intense sexual arousal from exposing one's genitals to an unsuspecting stranger) remained the same as in DSM-III (American Psychiatric Association, 1980). However, a qualifying diagnostic B criterion was added, as were [sic] true for all paraphilia diagnoses. This criterion was added to emphasize that psychiatric disorders or diagnoses had to include acting out against others or substantial distress. (p. 317)

Here is what DSM-III's (American Psychiatric Association, 1980) diagnostic criteria for Exhibitionism said: "Repetitive acts of exposing the genitals to an unsuspecting stranger for the purpose of achieving sexual excitement, with no attempt at further sexual activity with the stranger" (p. 272). Unlike DSM-III-R, DSM-III made no reference to sexual fantasies or sexual urges—the focus was on sexual behaviors. His analysis of Voyeurism in DSM-III is even more problematic.

> In the DSM-III-R, the core criterion for Voyeurism (Over a period of at least 6 months, recurrent, intense sexual urges and sexually arousing fantasies, involving the observation of an unsuspecting person who is naked, disrobing, or engaging in sexual activity) remained the same as in DSM-III. (Långström, 2010, p. 320)

What DSM-III says:

> A. The individual repeatedly observes unsuspecting people who are naked, in the act of disrobing, or engaging in sexual activity and no sexual activity with the observed people is sought.
> B. The observing is the repeatedly preferred or exclusive method of achieving sexual excitement. (p. 273)

The focus is on sexual *behaviors*, and—unlike DSM-III-R—DSM-III required these to be preferred to other sexual interests. Also, DSM-III-R did not require that "no sexual activity with the observed person is sought." These errors are more than inattention to detail. They suggest that the entire review process is seriously flawed and that previous criticisms of the Paraphilias have been ignored—it is often critics of the Paraphilias who give the closest analysis of their text.

These factual errors aside, Långström (2010) gives compelling reason for removing exhibitionism, voyeurism, and frotteurism from the DSM, but fails to draw this conclusion. In an article of eight pages, he reviews *three* diagnoses; given that his methodology indicates he did a thorough literature search, this must mean that there is only scant empirical work to support these diagnoses—implying that only the thinnest of scientific foundation exists. Regarding frotteurism, Långström acknowledges that it "has not been a subject of much clinical or scientific interest" (p. 320). Essentially, he admits that this diagnosis has minimal scientific foundation, minimal clinical value, and has resulted in minimal research.

Långström cites a number of studies indicating that voyeuristic behavior (and interest) is extremely common, but notes that, overall, there is not much research on prevalence. He suggests this is because "voyeurism is relatively easy to relate to for many

A. C. Hinderliter (✉)
Department of Linguistics, University of Illinois, 4080 Foreign Languages Building, Urbana, IL 61801, USA
e-mail: ahinder2@illinois.edu

individuals and, therefore, does not elicit the same strong negative emotions as do some of the other paraphilias" (p. 319). Here, he implies that the basis for classifying the paraphilias as disorders is prejudice–and then acknowledges that there is not even much prejudice against this paraphilia!

Not only does Långström (2010) tacitly admit that these three diagnoses have minimal scientific support, he recognizes that they can harm patients: "Seen from an overall perspective of DSM psychiatric nosology, [the issue of false-negatives] may be a particular problem for the potentially criminal paraphilias since shame, social stigma, and *negative legal consequences* prevent people from being open about their symptoms" (p. 232, emphasis added). If diagnosis of a Paraphilia can harm patients, the imperative to Do No Harm demands the deletion of these diagnoses unless it can be shown that their benefit considerably outweighs this harm. Furthermore, if these diagnoses discourage patients from being honest with clinicians, their clinical value is seriously drawn into question.

Långström's criticism of Sweden's decision to remove several of the Disorders of Sexual Preference from the ICD is especially telling in this regard: "[Now] clinicians in Sweden need to find other ways to diagnose *and occasionally treat* individuals with several of the deleted, but arguably still existing, disorders" (p. 322, emphasis added). The use of *occasionally* to modify *treat* but not *diagnose* suggests that the primary purpose of these diagnoses is not therapeutic. Is it punitive?

For frotteurism, Långström (2010) speculates that it was "its relevance to decision-making in criminal justice rather than in clinical settings that motivated its emergence as a specified paraphilia" (p. 321). Basically, he acknowledges that it is in the DSM simply because the associated act is illegal. With this type of reasoning, perhaps "Embezzlement Disorder" and "Insider Trading Disorder" should be added to the DSM? First and Frances (2008) emphasize the importance of not "blur[ing] the distinction between mental disorder and ordinary criminality" (p. 1240), which has been especially problematic in the diagnosis of the paraphilias.

From DSM-III to the present, all versions of the DSM have explicitly stated in their definitions of mental disorder that social deviance, whether commendable or not, is not itself a mental disorder. According to Spitzer and Williams (1982), this statement "was added to express indignation at the abuse of psychiatry as when, in the Soviet Union, political dissidents without sign of mental illness [were] labeled as having mental disorders and under that guise incarcerated in mental hospitals" (p. 21). In the U.S., 20 states and the Federal Government have Sexually Violent Predator laws that permit the (often lifelong) civil commitment of sex offenders *after* completing their sentences, including offenders with "no real mental disorder" (Frances, 2009), a situation which has been supported by a misreading of DSM-IV-TR's (American Psychiatric Association, 2000) phrase "sexually arousing fantasies, sexual urges, or behaviors" in the definition of Paraphilia. This has been used to diagnosis people with a Paraphilia on the basis of behavior alone. Långström rejects First and Frances' (2008) plea to remove "or behaviors" from the Paraphilias, which they say was a mistake to begin with. Practically, this means he endorses regarding *deviant sexual behavior alone*—without corresponding sexual fantasies or urges—as mental disorder. (That is what *or* means.) Combined with Långström's tacit acceptance of his speculation about why frotteurism was added to the DSM, this amounts to a rejection of the DSM's claim that social deviance itself is not a mental disorder.

Långström justifies rejecting First and Frances' plea saying, "I am not convinced that psychiatric nosology should change primarily because of the potential or actual misuse of diagnoses in the judicial system" (p. 323), which raises the controversial issue of what role legal consequences should play in nosological debates. Some authors (e.g., Stern, 2010) argue that ideological and political issues should have no role in what (supposedly) should be an entirely scientific issue. Besides suggesting a serious case of what Spitzer, tongue in cheek, called Politics Science-Dichotomy Syndrome (Spitzer, 1985), this assumes that whether something is a disorder is entirely an objective, scientific issue, a position many disagree with (e.g., Sedgwick, 1982; Spitzer, 1981; Wakefield, 1992). Here are many values involved in psychiatric nosology, some scientific and some not (Sadler, 2005). Indeed, the primary purpose of the helping professions is to help people, not to be "scientifically objective" in a way that disregards the consequences diagnostic decisions will have on real people, as though that were somehow a virtue. The primary purpose of the DSM is to help them do this. As Pincus and McQueen (2002) state: "The goals of the DSM-IV can be divided into four main categories: clinical, research, educational, and information management. The clinical goals of the DSM-IV were paramount because the book is used primarily by clinicians" (p. 9).

Brotto's (2010) literature review for Sexual Aversion Disorder (SAD) notes that it is a diagnosis that was added in DSM-III-R with little empirical justification, likely because of one person's historical influence. Given the minimal clinical and research interest it has produced—and because what treatment has developed treats it as a specific phobia—Brotto regards maintaining the status quo for SAD as the least desirable of the options discusses. Brotto noted that "With DSM-V and the emphasis placed on any changes being based on empirical science, SAD clearly would not have made its way into the DSM" (p. 276). The same holds for the three diagnoses that Långström reviews (and likely a number of the other paraphilias as well). However, rather than suggest that the scientific and clinical merit of these diagnoses be subject to critical review, he supports maintaining the status quo with a single modification to reflect a proposal being made for all the potentially criminal paraphilias. How long must these diagnoses remain in the DSM without empirical support before an accounting of them is demanded?

Frances (2010), chair of the DSM-IV Task Force, titled an editorial that bluntly states his assessment of the work of the Sexual and Gender Identity Disorder Workgroup: "DSM-5 and Sexual Disorders—Just Say No." He writes:

> Each of the Work Group's suggestions is based on the thinnest of research support—usually a handful of studies often done by members of the committee making the suggestion. None has been subjected to, or could possibly survive, anything resembling a serious risk/benefit or forensic analysis.

Though he makes comments about the Gender Identity Disorder Subworkgroup, most of his criticism is reserved for the proposals of the Paraphilias Subworkgroup.[1] I would add that not only should all proposed changes to the DSM be subject to serious risk/benefit analysis, so should proposals to maintain the status quo. It is doubtful that Långström's proposal to maintain the status quo could survive a serious risk/benefit analysis. The DSM should be scientifically rooted, clinically useful, and socially responsible. Långström's literature reviews inspire little confidence that the Paraphilias Subworkgroup is taking any of these values seriously.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., revised). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Brotto, L. A. (2010). The DSM diagnostic criteria for Sexual Aversion Disorder. *Archives of Sexual Behavior, 39*, 271–277.

First, M. B., & Frances, A. (2008). Issues for DSM-V: Unintended consequences of small changes: The case of paraphilias [Editorial]. *American Journal of Psychiatry, 165*, 1240–1241.

Frances, A. (2009, June 26). A warning sign on the road to DSM-V: Beware of unintended consequences. *Psychiatric Times*. http://www.psychiatrictimes.com/dsm-5/content/article/10168/1425378.

Frances, A. (2010, March 17). DSM5 and sexual disorders—just say no. *Psychiatric Times*. http://www.psychiatrictimes.com/dsm-5/content/article/10168/1539921.

Långström, N. (2010). The DSM diagnostic criteria for exhibitionism, voyeurism, and frotteurism. *Archives of Sexual Behavior, 39*, 349–356.

Pincus, H. A., & McQueen, L. (2002). The limits of an evidence-based classification of mental disorders. In J. Z. Sadler (Ed.), *Descriptions and prescriptions: Values, mental disorders, and the DSMs* (pp. 9–24). Baltimore: The John Hopkins University Press.

Sadler, J. Z. (2005). *Values and psychiatric diagnosis*. Oxford: Oxford University Press.

Sedgwick, P. (1982). *Psycho politics*. New York: Harper & Row.

Spitzer, R. L. (1981). The diagnostic status of homosexuality in DSM-III: A reformulation of the issues. *American Journal of Psychiatry, 138*, 210–215.

Spitzer, R. L. (1985). DSM-III and the politics-science dichotomy syndrome: A response to Thomas E. Schacht's "DSM-III and the politics of truth". *American Psychologist, 40*, 522–526.

Spitzer, R. L., & Williams, J. B. W. (1982). The definition and diagnosis of mental disorder. In W. R. Gove (Ed.), *Deviance and mental illness* (pp. 15–31). Beverly Hills, CA: Sage.

Stern, P. (2010). Paraphilic coercive disorder in the DSM: The right diagnosis for the right reasons. *Archives of Sexual Behavior*. doi: 10.1007/s10508-010-9645-9.

Wakefield, J. C. (1992). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist, 47*, 373–388.

[1] Unfortunately, Francis fails to mention that none of his criticisms are about the Sexual Dysfunctions Subworkgroup's proposals.

Arch Sex Behav (2010) 39:1239–1244
DOI 10.1007/s10508-010-9657-5

LETTER TO THE EDITOR

DSM-5 Proposals for Paraphilias: Suggestions for Reducing False Positives Related to Use of Behavioral Manifestations

Michael B. First

Published online: 10 August 2010

During each of the revisions of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders*, Workgroups comprised of experts in their respective fields are convened with the goal of considering changes to the diagnostic criteria. Based on my experiences working on the DSM-III-R and DSM-IV revisions, there is a natural tendency for Workgroups to focus on making changes aimed at broadening the diagnostic umbrella of their assigned categories with the goal of increasing diagnostic coverage, i.e., reducing what they consider to be false negatives, an effort which inevitably comes at the cost of increasing false positives (Pincus, Frances, Davis, First, & Widiger, 1992).

In the overall scheme of the DSM, false positives (i.e., erroneously giving a diagnostic label to an individual for whom it is not justified) are problematic because of the unnecessary stigma and inappropriate treatment that may result. However, because of their central role as a proxy for the "mental abnormality" requirement in sexually violent predator commitment statutes (First & Halon, 2008), a false positive diagnosis of a paraphilia has a uniquely negative outcome, namely inappropriate and potentially indefinite civil commitment to a secure forensic psychiatric facility. For this reason, the potential false positive implications of changes to the paraphilia criteria demand special consideration. Most of the proposed changes, as embodied in the draft DSM-5 proposals in the www.dsm5.org website, have a significant potential to lead to false positive diagnoses and should consequently be reexamined and reconsidered before the DSM-5 drafts are finalized in mid-2012. Proposed changes that have significant potential for false positives include (1) expanding the "official" roster of paraphilias to include sexual arousal to pubescent children (hebephilia) and sexual arousal to coercive sex (paraphilic coercive disorder); (2) continuing the inadvertent and erroneous DSM-IV-TR inclusion of behavior in the core definition of paraphilias, and (3) allowing a paraphilia diagnosis simply by exceeding an arbitrary number of sexual offenses. Because of space limitations, this Letter will focus only on considerations of the latter two problems; concerns about the inclusion of hebephilia and paraphilic coercive disorder have been raised elsewhere (Franklin, 2009; Green, 2010; Knight, 2010; Plaud, 2009; Quinsey, 2010; Tromovitch, 2009; Zander, 2009).

The definitions of the various paraphilias in DSM-IV-TR reflect the same overall diagnostic construct. The first component of the definition lays out the core psychopathology of a paraphilia, namely the fact that the person is intensely aroused by deviant sexual stimuli. The second part requires that the deviant pattern carries negative consequences for the individual or society: for those paraphilias which involve the participation of an unwilling victim (i.e., exhibitionism, voyeurism, frotteurism, pedophilia, sexual sadism), the diagnosis is made if the person has acted on his urges or else if the urges or fantasies cause marked distress or interpersonal difficulty; for the remaining paraphilias (e.g., fetishism, sexual masochism, and transvestic fetishism), the diagnosis is made if the urges, fantasies, or behaviors cause clinically significant distress or impairment in functioning.

As described in First and Frances (2008) and First and Halon (2008), however, a wording change introduced to the diagnostic criteria for paraphilias during the last stages of the DSM-IV production process inadvertently opened the door for false positive diagnoses of paraphilias, a fact that only came to light once the diagnostic criteria for paraphilias came under intense scrutiny in the context of civil commitment litigation of sexually violent predators. It is first important to understand that this

M. B. First (✉)
Division of Clinical Phenomenology, Department of Psychiatry, New York State Psychiatric Institute and Columbia University, 1051 Riverside Drive, Unit 60, New York, NY 10032, USA
e-mail: mbf2@columbia.edu

decision to adjust the wording was made under the mistaken belief that the new wording and the old wording were diagnostically equivalent, i.e., that it would identify exactly the same individuals as having a paraphilia as did DSM-III-R. In reviewing the limited empirical data base on paraphilias, the DSM-IV Sexual Disorders Workgroup recommended making no changes in the diagnostic criteria for any of the paraphilias, retaining the DSM-III-R wording of criterion A ("over a period of at least 6 months, recurrent intense sexual urges and sexually arousing fantasies involving…"). The only proposed change to the Paraphilias section of DSM-IV was the addition of a "with gender dysphoria" specifier to the criteria for Transvestic Fetishism. However, as part of a system-wide effort in the final stages of the preparation of DSM-IV to incorporate the clinical significance criterion (CSC) (Frances, 1998) into the criteria sets for most DSM-IV disorders, criterion B for all of the paraphilias (i.e., "the person has acted on these sexual urges, or the sexual urges or fantasies cause marked distress or interpersonal difficulty") was replaced by the standard CSC wording (i.e., "the fantasies, sexual urges, or behaviors cause clinically significant impairment in social, occupational, or other important areas of functioning"). Moreover, to make up for removal from criterion B of the behavioral indicator "acted on these urges," the phrase "or behaviors" was amended to criterion A in acknowledgement of the fact that it is typically behaviors which prompt a person to receive treatment for his or her paraphilia. Thus, the final wording of criterion A in DSM-IV and DSM-IV-TR was as follows: "over a period of at least 6 months, recurrent, intense sexually arousing fantasies, sexual urges, or behaviors involving [paraphilic focus]." This seemingly minor change in criterion A went virtually unnoticed until enterprising forensic evaluators in sexually violent predator commitment cases started using this change in wording to justify making a diagnosis of a paraphilia in cases in which the only evidence available to support the diagnosis was the nature of the sexual offenses themselves because the person either did not cooperate with the forensic evaluation or else denied having a deviant sexual arousal pattern.

There are both conceptual and practical problems related to including behaviors as part of the core definitional component of a paraphilia. As noted by Blanchard (2010) in his literature review of pedophilia, paraphilias "are erotic preferences or orientations that inhere in the individual and that have some existence independent of specific observable actions" (p. 310). A paraphilia is thus fundamentally a disturbed internal mental process (i.e., a deviant focus of sexual arousal) which is conceptually distinguishable from its various clinical manifestations, such as the person having sexual fantasies centered around the paraphilic focus, seeking and using pornography thematically linked to the paraphilic focus, having sexual urges triggered by stimuli related to paraphilic focus, or engaging in sexual behavior revolving around the paraphilic focus. Including behaviors as part of the core definition of a paraphilia conflate the underlying phenomenology of a paraphilia with its clinical manifestations.

From a practical perspective, offering "behaviors" as one of the defining elements of a paraphilia is potentially a significant source of false positive diagnoses, given that, of all of the clinical manifestations of a paraphilia, the person's behavior is the least specific and most fallible as an indicator of an underlying paraphilic focus of sexual arousal. Given a particular behavior, there are many potential mental states that could be driving that behavior. Take, for example, the behavior of exposing one's genitals in public. In addition to being a manifestation of an exhibitionistic paraphilic sexual arousal pattern, this behavior could be a manifestation of disinhibition or poor impulse control related to substance intoxication, a manic episode, or personality change due to a dementing illness. Thus, the particular mental state underlying this inappropriate behavior must be elucidated before one can attribute the behavior to a particular motivation, such as acting on a paraphilic urge. This "disconnect" between sexual behavior and underlying mental state is further illustrated by studies suggesting that paraphilic sexual interest may be the underlying explanation in only a minority of cases of sexual offenses. For example, Marshall and Fernandez (2003) reviewed 10 studies of exhibitionists using penile plethysmography and found that 9 out of the 10 studies suggested that exhibitionists in clinical settings did not have a preference for exposing themselves. Similarly, a study by Seto and Lalumiere (2001) of over 1000 child molesters using phallometric testing as a validator demonstrated that less than one-third had an underlying pedophilic arousal pattern.

Patient self-report of recurrent sexual urges and sexual fantasies are explicit indicators of a person's sexual arousal pattern that require little, if any, inferential suppositions on the part of the clinician. The relationship between a person's behavior and his sexual arousal preferences, on the other hand, requires making an inference that the behavior is driven by a paraphilic preference rather than other reasons. Making behaviors definitionally equivalent to sexual urges and sexual fantasies in criterion A leads to false positive diagnoses of paraphilias by virtue of giving the clinician permission to skip this crucial step in the diagnostic assessment process, allowing for the diagnosis to potentially be applied to individuals whose sexual offenses are motivated by non-paraphilic reasons (e.g., opportunism in a person with antisocial personality disorder, intoxication-related disinhibition in a person with substance dependence).

The DSM-5 proposal to define paraphilias in terms of behaviors alone is motivated by Workgroup members' concerns regarding potential false negatives that may result from the DSM-III-R requirement that there be intense sexual urges or sexually arousing fantasies involving the paraphilic focus. Particularly in forensic settings, individuals being evaluated for the presence of paraphilic interests have little objective motivation to be truthful regarding their sexual urges and fantasies and are thus often non-cooperative with evaluators. As noted by

Blanchard (2010) with respect to pedophilia, "offenders are not necessarily rewarded for being truthful about pedophilic impulses [and]… might experience even more severe consequences of their actions if they acknowledge being pedophiles" (p. 306). Thus, the DSM-5 proposals for the majority of the paraphilias involving non-consenting persons (i.e., Exhibitionism, Frotteurism, Sexual Sadism, and Voyeurism) will continue to allow a paraphilia diagnosis to be made based entirely on behaviors in order to "lessen the dependence of diagnosis on patients' self-reports regarding urges and fantasies," perpetuating the already identified false positive problem and its potential legal ramifications in terms of civil commitment.

Furthermore, in recognition of the fact that allowing a diagnosis of a paraphilia to be based entirely on sexual acts alone has the potential to lead to false positives (Blanchard, 2010; Långström, 2010); for those paraphilias involving non-consenting persons, the Workgroup has proposed setting a "minimum number of separate victims for diagnosing the paraphilia in uncooperative patients" (DSM-5 Paraphilia SubWorkgroup, 2010). For Pedohebephilia, at least two different child victims are required if the children are prepubescent, and three or more if at least one is pubescent; for Sexual Sadism, two or more victims on separate occasions are required, and for Exhibitionism, Frotteurism, Paraphilic Coercive Disorder, and Voyeurism, three or more non-consenting persons on separate occasions are necessary. According to the DSM-5 web site, the use of different cut-offs for different disorders is "an attempt to obtain similar rates of false positives and false negative diagnoses for all of the paraphilias" (DSM-5 Paraphilia SubWorkgroup, 2010).

Basing the diagnosis of a paraphilia entirely on exceeding a proscribed count of the number of victims is problematic for a number of reasons. First of all, although providing a precisely defined cut-off might give the appearance that these cut-off thresholds have a firm empirical basis, this is not in fact the case. As was the case with DSM-IV, one of the central requirements of the DSM-5 revision process is for recommendations to be grounded in empirical evidence (Kupfer, Regier, & Kuhl, 2008). Although the "rationale" sections on the DSM-5 web site explain why the Workgroup proposed using a victim count ("to lessen the dependence of diagnosis on patients' self-reports regarding urges and fantasies"), no empirical data were cited on the website to explain how or why these specific thresholds were set. An examination of the literature reviews conducted by the Workgroups to provide the empirical basis for the changes reveals that only the literature review for pedophilia offered any justification. In the Exhibitionsim, Frotteurism, and Voyeurism literature review, Långström (2010) stated that "to my knowledge, there are no published data that could directly advise on such behavioral determinants for the paraphilias reviewed here" (p. 322). The review justifying proposed changes for Sexual Sadism (Krueger, 2010) did not even address the cut-off issue.

The DSM-5 paraphilia literature review cites a single study by Blanchard, Klassen, Dickey, Kuban, and Blak (2001) as the justification for adopting a diagnostic threshold based on counting the number of victims, stating that "the results of Blanchard et al. (2001) show that absolute cut-off scores matter, at least up to three known offenses" (Blanchard, 2010, p. 309). Typically, when empirical data are used to justify a particular diagnostic cut-off, receiver operator characteristic (ROC) curves are calculated to determine the optimal balance of false positives and false negatives based on different diagnostic thresholds. For example, a study examining the optimal diagnostic cut-off of panic symptoms conducted as part of the MacArthur-funded DSM-IV secondary data reanalysis used ROC analysis to demonstrate that a diagnostic cut-off of four symptoms performed best using psychiatric hospitalization, ER visits, and suicide attempts as diagnostic validators (Leon, Klerman, Weissman, Fyer, & Johnson, 1992). An analogous ROC analysis to determine the optimal cut-off for the number of victims would have demonstrated that three or more victims yielded the best balance of false negatives vs. false positives based on some gold standard for a diagnosis of pedophilia, such as the individual admitting to a preferential sexual attraction to children.

The Blanchard et al. (2001) study, however, was not in any way designed to determine the diagnostic validity of a three victim diagnostic threshold. What the study actually showed was that the diagnostic sensitivity of phallometric testing for pedophilia (i.e., the percentage of men who are correctly diagnosed by the test as pedophilic) varied based on the number of victims, which were divided a priori into three groups: one victim, two victims, or three or more victims. For men with a history of offending children outside the family, diagnostic sensitivity was 30% for one victim, 42% for two victims, and 61% for three or more victims. For men with a history of offending entirely within the family, diagnostic sensitivity was 33% for one victim, 39% for two victims, and 29% for three or more victims. Blanchard et al. concluded in the discussion that "our analyses for offenders against unrelated children confirmed the expected result that men with greater numbers of victims had a greater likelihood of being diagnosed as pedophilic." They go on to say that "we were not able to demonstrate the same relation among incest offenders; however, that negative result is probably not meaningful because there were few incest offenders with multiple victims" (Blanchard et al., 2001, p. 124). The only other conclusion reached by the study was that "the more adult women with whom a patient has had sexual contact, the less likely he is to be diagnosed as pedophilic" (p. 124). It thus appears that the three victim cut-off is entirely the product of expert consensus rather than being based on any empirical footing, leaving its actual impact on rates of false positive diagnoses unknown.

Moreover, using a fixed cut-off, as is proposed for the paraphilias, introduces a significant false negative problem into

the picture, most evident in the diagnostic criteria for Pedohebephilia. Unlike the other non-consenting paraphilias, which typically involve victims picked opportunistically at random, pedophilia often involves multiple sexual offenses against one or two child victims who are known to the perpetrator (e.g., family members, children of neighbors). The proposed criteria for Pedohebephilia, which requires seeking sexual stimulation from two or more different children if both are prepubescent, or three or more different children if one or more are pubescent, would not diagnose such individuals as pedophiles.

This problem here stems from attempting to define a disorder in a way that is independent of the proper exercise of clinical judgment. Although a paraphilia diagnosis made in this way would most likely have improved reliability over the current approach, this increase in reliability would come at the expense of validity in terms of both false positives and false negatives. As noted in the introduction to the DSM, "the specific diagnostic criteria included in the DSM-IV are meant to serve as guidelines to be informed by clinical judgment and are not meant to be used in a cookbook fashion" (American Psychiatric Association, 2000, p. xxxii). As discussed above, a paraphilic arousal pattern is a mental phenomenon that may be manifested in a particular patient in a variety of ways. In those cases in which the forensic evaluator's evidentiary base is by necessity limited to the person's history of sexual offending behavior, a diagnosis of paraphilia is legitimate only if a paraphilic arousal pattern can be validly inferred from the overall pattern of behavior. Such an inference involves both ruling out other possible causes for the sexual behavior (e.g., other mental disorders such as bipolar disorder, opportunity coupled with disinhibition related to substance intoxication), and making the case that, given the nature of the pattern itself, the overwhelmingly likely motivating factor would be a paraphilic arousal pattern (e.g., involvement of a large number of victims over time coupled with the absence of sexual behavior involving non-paraphilic arousal stimuli).

In order to reduce the false positive problem inherent in allowing a paraphilia diagnosis to be made based entirely on external behavior, I recommend a return to the DSM-III-R framework which places the person's deviant sexual arousal pattern at the core of the definition and clarifying that this arousal pattern may be manifested by sexual urges, sexual fantasies, or sexual behaviors. This approach serves to clearly differentiate between the core psychopathological construct of a paraphilia (i.e., a pattern of deviant sexual arousal) and its clinical manifestations (i.e., sexual urges, fantasies, and behaviors). The current proposed wording for pedohebephilia, which separates these two constructs in its criterion A, can be used as the model for this wording change, which should be applied uniformly to the definitions of all of the paraphilias. The following wording is suggested: "Over a period of at least 6 months, recurrent and intense sexual arousal from [paraphilic focus of sexual interest], as manifested by fantasies, urges, or behaviors." Although this wording allows for the diagnosis to be made based on a consideration of the person's behavioral pattern, there is an implicit requirement that the diagnostician make the clinical judgment that the behaviors are a manifestation of the deviant sexual arousal pattern that is at the core of the paraphilia and to provide adequate explication of this in the medical record. Furthermore, I recommend rejecting the imposition of an arbitrary threshold for the number of victims and instead continue with the DSM-IV-TR requirement that "the person has acted on these sexual urges, or the sexual urges or fantasies cause marked distress or interpersonal difficulty."

A strategy for reducing diagnostic false positives that is used in the vast majority of diagnostic criteria sets in the DSM is to provide exclusionary criteria for the purpose of alerting the clinician to alternative diagnostic formulations that must be considered and ruled out (First, Frances, & Pincus, 2004). For example, the diagnostic criteria for specific phobia instruct the clinician to consider alternative diagnostic explanations for the avoidance behavior before making the diagnosis, such as a diagnosis of Obsessive–Compulsive Disorder in someone with fear of dirt related to obsessions about contamination, a diagnosis of posttraumatic stress disorder in someone who avoids stimuli associated with a severe stressor, a diagnosis of Separation Anxiety Disorder in someone who avoids going to school, etc. However, with the exception of the diagnostic criteria set for Fetishism that excludes the diagnosis if "the fetish objects are limited to articles of female clothing used in cross-dressing (as in Transvestic Fetishism) or devices designed for the purpose of tactile genital stimulation (e.g., a vibrator)," none of the paraphilia criteria sets include an exclusionary criteria.

Given the crucial need for the diagnostician to consider and rule out alternative diagnostic considerations that could account for sexual behaviors in order to prevent a false positive diagnosis of paraphilia, I recommend that new exclusionary criteria (i.e., a new criterion B) be added to the criteria sets of each of the paraphilias. This exclusionary criterion would indicate the other DSM diagnoses that could account for the sexual behavior that need to be considered and ruled out (e.g., mental retardation, substance abuse, manic episode, schizophrenia), non-psychiatric general medical conditions (e.g., severe head trauma, dementing illnesses), and non-medical non-psychiatric explanations (e.g., public urination as an alternative to a diagnosis of exhibitionism). Table 1 illustrates the proposed wording as it applies to all of the non-consent paraphilias.

Given the unique forensic implications of this diagnostic category, great care should be taken in the definition of the paraphilias to prevent false positive diagnoses. The current DSM-IV-TR criteria and the proposed DSM-5 criteria promote false positive diagnoses by allowing the diagnosis to be made without a consideration of the underlying motivation for sexual behavior which may be driven by motivations other than a paraphilic sexual arousal pattern. The proposals discussed here (i.e., placing sexual arousal at the core of the definition and

Table 1 Proposed DSM-5 criteria for paraphilias involving non-consenting victims and suggested modifications

Paraphilias involving non-consenting victims	Proposed DSM-5 criteria from www.dsm5.org	Revisions proposed to reduce potential for false positives
Exhibitionism	A. Over a period of at least 6 months, recurrent and intense sexual fantasies, sexual urges, or sexual behaviors involving the exposure of one's genitals to an unsuspecting stranger	A. Over a period of at least 6 months, recurrent and intense sexual arousal from the exposure of one's genitals to an unsuspecting stranger, as manifested by fantasies, urges, or behaviors
	B. The person is distressed or impaired by these attractions, or has sought sexual stimulation from exposing the genitals to three or more unsuspecting strangers on separate occasions	B. The behavioral manifestations are not due to the direct physiological effects of a substance (e.g., Alcohol Intoxication), a general medical condition (e.g., Alzheimer's disease) and not better accounted for another mental disorder (e.g., manic episode, Antisocial Personality Disorder) or by instances of public urination
		C. The person has acted on these sexual urges, or the sexual urges or fantasies cause marked distress or interpersonal difficulty
Frotteurism	A. Over a period of at least 6 months, recurrent and intense sexual fantasies, sexual urges, or sexual behaviors involving touching or rubbing against a nonconsenting person	A. Over a period of at least 6 months, recurrent and intense sexual arousal from touching or rubbing against a nonconsenting person, as manifested by fantasies, urges, or behaviors
	B. The person is distressed or impaired by these attractions, or has sought sexual stimulation from touching and rubbing against three or more nonconsenting persons on separate occasions	B. The behavioral manifestations are not due to the direct physiological effects of a substance (e.g., Alcohol Intoxication), a general medical condition (e.g., Alzheimer's disease) and not better accounted for any another mental disorder (e.g., manic episode, Antisocial Personality Disorder)
		C. The person has acted on these sexual urges, or the sexual urges or fantasies cause marked distress or interpersonal difficulty
Pedohebephilia[a]	A. Over a period of at least 6 months, one or both of the following, as manifested by fantasies, urges, or behaviors:	A. Over a period of at least 6 months, one or both of the following, as manifested by fantasies, urges, or behaviors:
	(1) recurrent and intense sexual arousal from prepubescent or pubescent children	(1) recurrent and intense sexual arousal from prepubescent children
	(2) equal or greater arousal from such children than from physically mature individuals	(2) equal or greater arousal from such children than from physically mature individuals
	B. One or more of the following signs or symptoms:	B. The behavioral manifestations are not due to the direct physiological effects of a substance (e.g., Alcohol Intoxication), a general medical condition (e.g., Alzheimer's disease) and not better accounted for any another mental disorder (e.g., manic episode, Antisocial Personality Disorder)
	(1) the person is distressed or impaired by sexual attraction to children	
	(2) the person has sought sexual stimulation, on separate occasions, from either of the following:	
	(a) two or more different children, if both are prepubescent	C. The person has acted on these sexual urges, or the sexual urges or fantasies cause marked distress or interpersonal difficulty
	(b) three or more different children, if one or more are pubescent	
	(3) use of child pornography in preference to adult pornography, for a period of 6 months or longer	
Sexual Sadism	A. Over a period of at least 6 months, recurrent and intense sexual fantasies, sexual urges, or sexual behaviors involving the physical or psychological suffering of another person	A. Over a period of at least 6 months, recurrent and intense sexual arousal from the physical or psychological suffering of another person, as manifested by fantasies, urges, or behaviors
	B. The person is distressed or impaired by these attractions or has sought sexual stimulation from behaviors involving the physical or psychological suffering of two or more nonconsenting persons on separate occasions	B. The behavioral manifestations are not due to the direct physiological effects of a substance (e.g., Alcohol Intoxication), a general medical condition (e.g., Alzheimer's disease) and not better accounted for any another mental disorder (e.g., manic episode, Antisocial Personality Disorder)

Table 1 continued

Paraphilias involving non-consenting victims	Proposed DSM-5 criteria from www.dsm5.org	Revisions proposed to reduce potential for false positives
		C. The person has acted on these sexual urges, or the sexual urges or fantasies cause marked distress or interpersonal difficulty
Voyeurism	A. Over a period of at least 6 months, recurrent and intense sexual fantasies, sexual urges, or sexual behaviors involving the act of observing an unsuspecting person who is naked, in the process of disrobing, or engaging in sexual activity.	A. Over a period of at least 6 months, recurrent and intense sexual arousal from observing an unsuspecting person who is naked, in the process of disrobing, or engaging in sexual activity, as manifested by fantasies, urges, or behaviors.
	B. The person is distressed or impaired by these attractions, or has sought sexual stimulation from observing three or more unsuspecting persons who are naked, disrobing, or engaging in sexual activity on separate occasions.	B. The behavioral manifestations are not due to the direct physiological effects of a substance (e.g., Alcohol Intoxication), a general medical condition (e.g., Alzheimer's disease) and not better accounted for any another mental disorder (e.g., manic episode, Antisocial Personality Disorder)
		C. The person has acted on these sexual urges, or the sexual urges or fantasies cause marked distress or interpersonal difficulty

[a] I have removed the reference to "pubescent children" from my proposed revision of the criteria as I believe this proposal exacerbates the false positive problem

adding exclusionary criteria) serve to counterbalance this risk while still accommodating the clinical reality that individuals, especially in forensic settings, are likely to falsely deny a history of deviant sexual urges and fantasies.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

Blanchard, R. (2010). The DSM diagnostic criteria for Pedophilia. *Archives of Sexual Behavior, 39*, 304–316.

Blanchard, R., Klassen, P., Dickey, R., Kuban, M., & Blak, T. (2001). Sensitivity and specificity of the phallometric test for pedophilia in nonadmitting sex offenders. *Psychological Assessment, 13*, 118–126.

DSM-5 Paraphilia SubWorkgroup. (2010). Rationale for Exhibitionism changes. http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=62#.

First, M., & Frances, A. (2008). Issues for DSM-V. Unintended consequences of small changes: The case of paraphilias. *American Journal of Psychiatry, 165*, 1240–1241.

First, M. B., Frances, A., & Pincus, H. A. (2004). *DSM-IV-TR guidebook*. Washington, DC: American Psychiatric Publishing.

First, M., & Halon, R. (2008). Use of DSM paraphilia diagnoses in sexually violent predator commitment cases. *Journal of the American Academy of Psychiatry and the Law, 36*, 443–454.

Frances, A. (1998). Problems in defining clinical significance in epidemiological studies. *Archives of General Psychiatry, 55*, 119.

Franklin, K. (2009). The public policy implications of "hebephilia". A response to Blanchard et al. [Letter to the Editor]. *Archives of Sexual Behavior, 38*, 319–320.

Green, R. (2010). Sexual preference for 14-year-olds as a mental disorder: You can't be serious!! [Letter to the Editor]. *Archives of Sexual Behavior, 39*, 585–586.

Knight, R. A. (2010). Is a diagnostic category for paraphilic coercive disorder defensible? *Archives of Sexual Behavior, 39*, 419–426.

Krueger, R. B. (2010). The DSM diagnostic criteria for Sexual Sadism. *Archives of Sexual Behavior, 39*, 325–345.

Kupfer, D. J., Regier, D. A., & Kuhl, E. A. (2008). On the road to DSM-V and ICD-11. *European Archives of Psychiatry and Clinical Neuroscience, 258*(Suppl. 5), 2–6.

Långström, N. (2010). The DSM diagnostic criteria for Exhibitionism, Voyeurism, and Frotteurism. *Archives of Sexual Behavior, 39*, 317–324.

Leon, A. C., Klerman, G. L., Weissman, M. M., Fyer, A. J., & Johnson, J. (1992). Evaluating the diagnostic criteria for panic disorder: Measures of social morbidity as criteria. *Social Psychiatry and Psychiatric Epidemiology, 27*, 180–184.

Marshall, W. L., & Fernandez, Y. (2003). *Phallometric testing with sexual offenders: Theory, research, and practice*. Brandon, VT: Safer Society Press.

Pincus, H., Frances, A., Davis, W., First, M., & Widiger, T. (1992). DSM-IV and new diagnostic categories: Holding the line on proliferation. *American Journal of Psychiatry, 149*, 112–117.

Plaud, J. J. (2009). Are there "Hebephiles" among us? A response to Blanchard et al. (2008) [Letter to the Editor]. *Archives and Sexual Behavior, 38*, 326–327.

Quinsey, V. L. (2010). Coercive Paraphilic Disorder. *Archives of Sexual Behavior, 39*, 405–410.

Seto, M. C., & Lalumiere, M. L. (2001). A brief screening scale to identify pedophilic interests among child molesters. *Sexual Abuse: A Journal of Research and Treatment, 13*, 15–25.

Tromovitch, P. (2009). Manufacturing mental disorder by pathologizing erotic age orientation: A comment on Blanchard et al. (2008) [Letter to the Editor]. *Archives of Sexual Behavior, 38*, 328.

Zander, T. K. (2009). Adult sexual attraction to early-stage adolescents: Phallometry doesn't equal pathology [Letter to the Editor]. *Archives of Sexual Behavior, 38*, 329–330.

Arch Sex Behav (2010) 39:1245–1252
DOI 10.1007/s10508-010-9659-3

LETTER TO THE EDITOR

The Specificity of Victim Count as a Diagnostic Indicator of Pedohebephilia

Ray Blanchard

Published online: 10 August 2010

In First's (2010) Letter to the Editor of this Journal, entitled "DSM-5 Proposals for Paraphilias: Suggestions for Reducing False Positives Related to Use of Behavioral Manifestations," he offers several criticisms and suggestions regarding the proposed diagnostic criteria for the paraphilic disorders in the upcoming fifth edition of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association, 2010). I am writing in response to that letter. Although I am (unavoidably) writing as a member of the Paraphilias Subworkgroup of the DSM-5 Work Group for Sexual and Gender Identity Disorders, which authored the proposed criteria, I am writing to express my own conclusions and not the considered consensus of the Subworkgroup. Any errors or omissions in this reply are mine and mine alone.

The single most important point in First's letter, as I understand it, is this: An uncooperative patient cannot be diagnosed with a specific paraphilia on the sole grounds that he has offended against *x* or more victims in a way symptomatic of that paraphilia, because such a practice would produce an excessive number of false positive diagnoses. First's view is related to the notion that behavior is a more fallible indicator of paraphilia than is self-report. First further asserts that a diagnostic criterion for uncooperative subjects based on a minimum number of individual victims would produce excessive false negatives as well as false positives, because it neglects the frequency of sexual interactions with a given victim and the time period over which repeated interactions with that victim occurred.

R. Blanchard (✉)
Law and Mental Health Program, Centre for Addiction and Mental Health, 250 College Street, Toronto, ON M5T 1R8, Canada
e-mail: ray_blanchard@camh.net

First, as one would expect, freely acknowledges that a patient's self-report may be unreliable in the adversarial context of a forensic evaluation. His response to that situation is as follows: In the absence of reliable self-report information, a diagnosis of paraphilia is possible only if two conditions are met. Firstly, there must be a *pattern* of paraphilic arousal. The term *pattern* is not defined, but it is exemplified by a large victim count plus an absence of normal sexual behavior. Secondly, other possible causes for the criminal sexual behavior must be ruled out. He includes a specific suggestion for the paraphilic disorders' diagnostic criteria to embody his second point: In order to reduce the rates of false positive diagnoses of the various paraphilias, an exclusionary criterion should be added to the criteria set of each of the paraphilias. This exclusionary criterion would list the other DSM diagnoses that should be ruled out as alternative explanations of the patient's symptoms.

First writes, in places, as if the unacceptable false positive rate that would result from diagnosing paraphilias by victim count were self-evident or already known as fact. In other places, however, he seems to entertain the possibility of treating this as an empirical question. Thus, he notes that no published study has established the minimum number of similar victims that would justify the diagnosis of a specific paraphilia in a patient who verbally denies having that paraphilia. His further statements on this head spell out not only the research design but also the precise statistical analysis that might be used to estimate the specificity and sensitivity of victim count as a diagnostic indicator of pedohebephilia. His remarks may be summarized as follows: Receiver operating characteristic (ROC) curves are typically calculated to estimate the proportions of false positive and false negative diagnoses that would result from applying a given threshold (cutting score) to a diagnostic indicator. An ROC analysis could be conducted to estimate the specificity and sensitivity

of victim count as a diagnostic indicator of pedohebephilia, using a total of three or more child victims as the cutting score, and treating men who verbally admit to an erotic preference for children as the gold standard for a diagnosis of pedohebephilia.

In what follows, I will address First's various comments and criticisms, beginning with an empirical ROC analysis along the lines he suggests. Before I begin that, however, I must clear one historical matter out of the way.

First devotes a substantial proportion of his letter to explaining the vicissitudes of *behavior* in the diagnostic criteria for the paraphilias in DSM-IV (American Psychiatric Association, 1994) and DSM-IV-TR (American Psychiatric Association, 2000). The prominence of this account might leave the reader with the impression that the current proposal to use repeated sexual offenses of the same nature as one basis for diagnosing paraphilic disorders in the DSM-5 is somehow an outgrowth or consequence of the varying uses of *behavior* in DSM-IV and DSM-IV-TR. That is, from my standpoint, not the case at all.

I suggested, in my first report to the Paraphilias Subworkgroup (Blanchard, 2010a), that a patient be diagnosed as pedohebephilic if he has offended against three or more children under the age of 15, in accordance with a long tradition of clinical research and practice, whereby men who have committed sexual crimes of a similar nature against multiple victims are diagnosed with the corresponding paraphilia, whether they acknowledge that disorder or not. This tradition might not be evident in DSM-IV or DSM-IV-TR, but its influence can clearly be seen in DSM-III (American Psychiatric Association, 1980). The DSM-III diagnostic criteria for Exhibitionism (302.40), for example, read as follows: "Repetitive acts of exposing the genitals to an unsuspecting stranger for the purpose of achieving sexual excitement, with no attempt at further sexual activity with the stranger" (p. 272). There are various clinical research studies in which men are classified into paraphilic groups on the basis of their known sexual offenses (e.g., Blanchard et al., 1999; Cantor et al., 2005; Freund, 1965).

Ad Hoc Analysis: The Specificity of Victim Count

Because First's central concern is false positive diagnoses of pedohebephilia, the most important methodological aspect of an empirical investigation is assembling a gold standard group of teleiophiles (men with an erotic preference for persons between the ages of physical maturity and physical decline). This cannot, in reality, be perfectly achieved. A teleiophilic preference is the desirable preference, in the clinic, in the courts, and in general society; and many men will falsify their self-report and manipulate their phallometric test results,[1] if they can, to obtain that diagnosis.

The closest thing to a gold standard group of teleiophiles would be a group of men whose known sexual offenses, self-report, and phallometric test results all indicated teleiophilia. In that case, however, there would be no diagnostic indicator left to test; one would have used up all one's degrees of freedom, so to speak, by employing all possible indicators for the purpose of group selection. For the following study, therefore, I selected a "silver standard" group of teleiophiles on the basis of self-report and phallometric test results. The primary limitation of such a silver standard group is that some of the men "falsely" diagnosed as pedohebephilic according to their victim count might, in fact, have been correctly diagnosed. Thus, the obtained estimate of specificity should be regarded as a minimum estimate.

Method

Subjects

Between November 1995 and October 2009, 3,166 male patients were administered the same phallometric test for erotic object (gender and age) preferences at the Kurt Freund Laboratory of the Centre for Addiction and Mental Health (Toronto, Ontario, Canada). The sources of the clinical referrals included parole and probation officers, prisons, defense lawyers, various institutions (ranging from group homes for mentally retarded persons to regulatory bodies for health or educational professionals), and physicians in private practice. As would be expected from the preponderance of criminal justice sources, the majority of patients had one or more sexual offenses against children, adults, or both. Men who had no involvement with the criminal justice system and who initiated referrals through their physicians included patients who were unsure about their sexual orientation, patients concerned about hypersexuality or "sex addiction," patients experiencing difficulties because of their excessive use of telephone sex lines or massage parlors, clinically obsessional patients with intrusive thoughts about unacceptable sexual behavior, and patients with paraphilic behaviors like masochism, fetishism, and transvestism. Subsets of these patients have been analyzed in two previous studies (Blanchard, 2010b; Blanchard et al., submitted).

There were 2,725 patients who had given informed consent for their assessment data to be used for research purposes, and whose sexual history data were complete and cross-checked at

[1] Phallometric testing is a psychophysiological technique for assessing erotic interests in males. In phallometric tests for erotic gender- and age-orientation, the individual's penile blood volume is monitored while he is presented with a standardized set of laboratory stimuli depicting male and female children, pubescents, and adults. Increases in the subject's penile blood volume (i.e., degrees of penile erection) are taken as an index of his relative attraction to different classes of persons.

the time these data were retrieved (Blanchard et al., submitted). Self-report data on sexual attractions were available for 2,715 of these. From this pool, I selected 998 subjects for the present study, using criteria explained later. The selected men had a mean age of 34.93 years ($SD = 11.98$) and a median education of high school graduation.

Materials and Measures

Sexual History

A standardized form, described in detail by Blanchard et al. (2009), was used to record the patient's history of sexual offenses. Most of that information came from objective documents that accompanied his referral (e.g., reports from probation and parole officers). The offense-history data were cross-checked against, and supplemented by, other information provided by the patient himself, including the number and nature of any additional sexual offenses that were admitted by the patient but for which he was never charged. For the present study, the total number of different children under the age of 15 with whom the patient had interacted sexually was called the *victim count*. It did not matter, for the purpose of computing this variable, if the child was male or female, consenting or coerced, related or unrelated to the patient, approached physically (e.g., fondled) or non-physically (e.g., exposed to), approached once or approached on multiple occasions.

The recorded sexual offenses included those related to the possession or (rarely) manufacture of child pornography. Detailed information on the number of images involved was often not available. This variable was therefore recorded dichotomously, as present or not present. For purposes of this study, child pornography was scored as present if the patient had charges of this nature or if he admitted to the use of child pornography (or both).

The patient was also asked to rate his sexual attraction to persons in 12 gender–age categories (e.g., females aged 17 years or older, males aged 17 years or older, females aged 15–16 years, males aged 15–16 years, and so on) using a five-point scale. The patient's information was solicited by the laboratory manager in a structured sexual history interview, which the manager conducted the same day he administered the phallometric test.

Phallometric Measurement

The Kurt Freund Laboratory is equipped for volumetric phallometry, that is, the apparatus measures penile blood volume change rather than penile circumference change. The volumetric method measures penile tumescence more accurately at low levels of response (Kuban, Barbaree, & Blanchard, 1999). A photograph and schematic drawing of the volumetric apparatus are given in Freund, Sedlacek, and Knob (1965). The major components include a glass cylinder that fits over the penis and an inflatable cuff that surrounds the base of the penis and isolates the air inside the cylinder from the outside atmosphere. A rubber tube attached to the cylinder leads to a pressure transducer, which converts air pressure changes into voltage output changes. Increases in penile volume compress the air inside the cylinder and thus produce an output signal from the transducer. The apparatus is calibrated so that known quantities of volume displacement in the cylinder correspond to known changes in transducer voltage output.

The specific test used in this study has been described in detail by Blanchard, Klassen, Dickey, Kuban, and Blak (2001). The test stimuli were audiotaped narratives presented through headphones and accompanied by slides. There were seven categories of narratives, which described sexual interactions with prepubescent girls, pubescent girls, adult women, prepubescent boys, pubescent boys, and adult men, and also solitary, nonsexual activities ("neutral" stimuli). All narratives were written in the second person and present tense and were approximately 100 words long. The narratives depicted fantasy situations in which sexual interaction with children would be relatively plausible or sexual interaction with adults would be relatively plausible as well as somatic and social attributes indicating the physical maturity of the imaginary target. Sample narratives are given in Blanchard et al. (2007). The narratives describing heterosexual interactions were recorded with a woman's voice, and those describing homosexual interactions, with a man's. Neutral stimuli were recorded with both.

Each test trial consisted of one narrative, accompanied by photographic slides on the three adjacent projection screens, which simultaneously showed the full-length front view, full-length rear view, and close-up genital region of a nude model who corresponded in age and gender to the topic of the narrative. Figure 1 in Blanchard et al. (2007) illustrates the standard pose used for the full-length front views. Each trial included three nude models, each presented for 18 s. Therefore, the total duration of a trial was 54 s, during which the examinee viewed a total of nine slides, three at a time. Neutral narratives were similarly accompanied by slides of landscapes.

The full test consisted of four blocks of seven trials, with each block including one trial of each type in fixed, pseudorandom order. Although the trial length was fixed, the intertrial interval was variable, lasting as long as necessary for penile blood volume to return to baseline. The time required to complete the test was usually about 1 h.

Two kinds of scores, representing the amount of responding and the direction of responding, were calculated from the raw phallometric data. The amount of responding was quantified with a standard measure in the Kurt Freund Laboratory, the *Output Index* or *OI* (Freund, 1967). This is the average of the three greatest responses to any stimulus category except "neutral," where penile response is expressed in cubic centimeters (cc) of blood volume increase from the start of a trial.

The assessment of response direction in the Kurt Freund Laboratory involves several steps. This process, whose calculations and rationale have been presented at length by Blanchard, Klassen, Dickey, Kuban, and Blak (2001), may be briefly explained as follows. During the stimulus trials, penile blood volume change is sampled four times per second and recorded as a curve of blood volume change over time. The examinee's response during a given trial is measured in two ways: (1) as the maximum deflection of the curve (i.e., the greatest departure from initial value occurring during the 54 s of the trial), and (2) as the area under the curve. Each examinee's 28 deflection scores are converted into standard scores, based on his own deflection data (in other words, they are converted into ipsative *z* scores), and the same operation is carried out on his area scores. Next, for each examinee, the standardized deflection and standardized area scores are averaged to yield a separate composite score for each of the 28 trials. Finally, the data are reduced to seven scores for each examinee by averaging his four composite scores in each of the seven stimulus categories. These seven *category scores* are taken as measures of the examinee's relative erotic interest in adult women, pubescent girls, prepubescent girls, and so on.

Final Gating Criteria and Assignment to Groups

Subjects were considered eligible for this study if their OI's were greater than or equal to 2.50 cc. This is slightly higher than my laboratory's customary threshold of 1.00 cc for considering a phallometric test result clinically interpretable (e.g., Blanchard et al., 2001), but it is consistent with the higher response criteria used in most other phallometric laboratories (see review in Lykins et al., 2010). An increase in penile blood volume of 2.50 cc in the Kurt Freund Laboratory corresponds roughly to 10% of a full erection for the average adult male. I chose the higher value because recent research has suggested that the reliability of phallometric diagnosis is significantly related to the magnitude of the subject's OI (Lykins et al., 2010).

A subject was selected for the Teleiophilic Group if he met both of the following criteria: (1) phallometric test results showed that the higher of his two category scores for adults (adult women and adult men) was over .25 *z* score units greater than the highest of his four category scores for children (prepubescent girls, pubescent girls, prepubescent boys, and pubescent boys), and (2) he verbally reported, in his structured sexual history interview, that his strongest sexual attractions were to women or men age 17 or older, and that he felt zero sexual attraction to girls or boys under the age of 15.

A subject was selected for the Pedohebephilic Group if he met essentially the obverse criteria: (1) the highest of his four category scores for children (prepubescent girls, pubescent girls, prepubescent boys, and pubescent boys) was over .25 *z* score units greater than the higher of his two category scores for adults (adult women and adult men),[2] and (2) he verbally reported that his strongest sexual attractions were not to women or men age 17 or older, and that he felt more than zero sexual attraction to girls or boys under the age of 15.

Using these criteria, 817 men were selected for the Teleiophilic Group, and 181 men were selected for the Pedohebephilic Group.

Results

Table 1 shows, for each group, the number of subjects who sexually offended against 0, 1, 2, 3, 4, 5, and 6 or more children under the age of 15 years. This table can be used to estimate the sensitivity and specificity of victim count as a diagnostic indicator of pedohebephilia, using different numbers of victims as the cutting score. I will use, as an illustrative example, a cutting score of three victims: Every man with fewer than three victims is classified as a teleiophile, and every man with three or more victims is classified as a pedohebephile. With that cutting score, the specificity of victim count is $(335 + 289 + 121)/817 = 91\%$, and the sensitivity is $(25 + 10 + 9 + 53)/181 = 54\%$.

The trade-off between sensitivity and specificity may be graphically represented in a receiver operating characteristic (ROC) curve. Figure 1 shows the ROC curve for the data presented in Table 1.

One measure of the adequacy of a diagnostic test is the area under the ROC curve (AUC). There is no universal agreement on what AUC values indicate a poor, fair, good, or excellent clinical test—indeed, such labeling depends partly on context—but a value around .70 might be considered "fair," and a value around .80, "good." In the present case, the AUC was .72, 95% CI [.67, .77].

There is a problem with the foregoing analysis (apart from the obvious problem that one or more of the men in the Teleiophilic Group who had offenses against children might truly have been pedohebephiles rather than "false positives," even though they claimed an erotic preference for adults and produced phallometric results consistent with that claim). The more subtle problem is that the expectation that *none* of the men in the Teleiophilic Group should have offenses against children is not matched by an equal and opposite expectation that *all* of the men in the Pedohebephilic Group should have offenses against children. A man with no criminal charges of any kind could have ended up in the Pedohebephilic Group because he came forward and requested clinical help (or because concerned caregivers requested help on his behalf), and a man with no real-life child victims could have ended up in that group via criminal charges for child pornography. Such cases would have

[2] This rule has been used for the clinical diagnosis of pedohebephilia in the Kurt Freund Laboratory for over 20 years (see Blanchard et al., 2001).

Table 1 Frequency distribution of victim count, computed on all subjects

Victim count	Group			
	Teleiophiles		Pedohebephiles	
	n	%	*n*	%
0	335	41.0	43	23.8
1	289	35.4	32	17.7
2	121	14.8	9	5.0
3	42	5.1	25	13.8
4	14	1.7	10	5.5
5	9	1.1	9	5.0
≥6	7	.9	53	29.3
Total	817	100.0	181	100.0

Table 2 Frequency distribution of victim count, computed on subjects lacking charges or self-admissions regarding child pornography use

Victim count	Group			
	Teleiophiles		Pedohebephiles	
	n	%	*n*	%
0	306	39.6	18	17.0
1	280	36.3	13	12.3
2	116	15.0	5	4.7
3	40	5.2	20	18.9
4	14	1.8	7	6.6
5	9	1.2	4	3.8
≥6	7	.9	39	36.8
Total	772	100.0	106	100.0

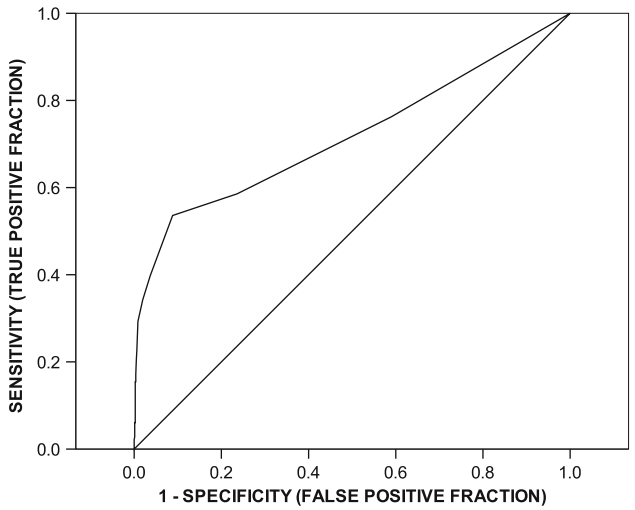


Fig. 1 ROC curve for victim count, computed on all subjects

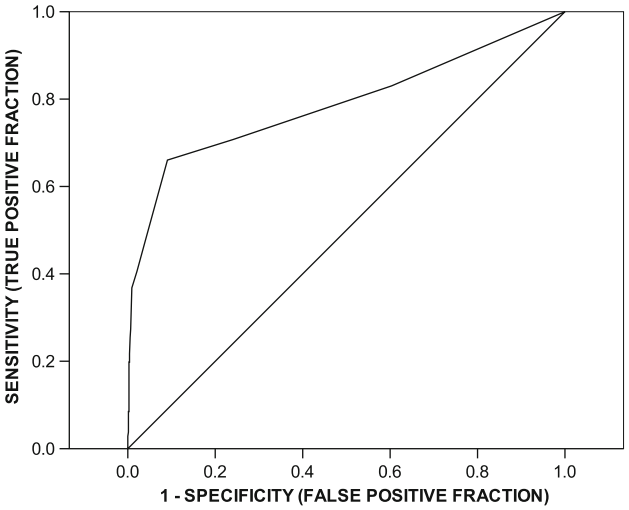


Fig. 2 ROC curve for victim count, computed on subjects lacking charges or self-admissions regarding child pornography use

the effect of lowering the computed sensitivity and the AUC of victim count as a diagnostic indicator.

I conducted two further analyses to demonstrate the problem in regarding the Pedohebephilic Group as a gold standard. Both analyses used the easily manipulated variable of child pornography. In the first analysis, I simply excluded from the whole sample any man who had been charged with, or admitted to, the use of child pornography. This reduced the Teleiophilic Group to 772 subjects, and the Pedohebephilic Group to 106 subjects.

Table 2 shows, for each group, the number of subjects who sexually offended against different numbers of children under the age of 15 years. For the same cutting score of three victims, the sensitivity increased to 66%. The specificity remained virtually unchanged, at 91%.

The manipulation also produced the expected effect on the ROC curve (Fig. 2). The AUC increased to .79, 95% CI [.73, .85].

My second method for demonstrating the effect of non-molesting routes into the Pedohebephilic Group was to treat child pornography (for diagnostic purposes) as the equivalent of child molestation. Seto, Cantor, and Blanchard (2006) found that child pornography, treated as a dichotomous variable, had a diagnostic value equal to real-life offenses against multiple victims. I therefore created a new variable, the *victim equivalency count*, by adding two "victims" to the total score for any man who had been charged with, or admitted to, the use of child pornography. In this analysis, I returned to the use of the full sample.

Table 3 shows, for each group, the number of subjects with different scores on the victim equivalency count. I used a cutting score of three victims/victim-equivalents. Remarkably, the estimates for sensitivity and specificity were almost the same as those in the previous analysis: 66 and 89%, respectively.

Table 3 Frequency distribution of victim equivalency count, computed on all subjects

Victim equivalency count	Group			
	Teleiophiles		Pedohebephiles	
	n	%	*n*	%
0	306	37.5	18	9.9
1	280	34.3	13	7.2
2	145	17.7	30	16.6
3	49	6.0	39	21.5
4	19	2.3	11	6.1
5	11	1.3	9	5.0
6	2	.2	13	7.2
7	1	.1	7	3.9
≥8	4	.5	41	22.7
Total	817	100.0	181	100.0

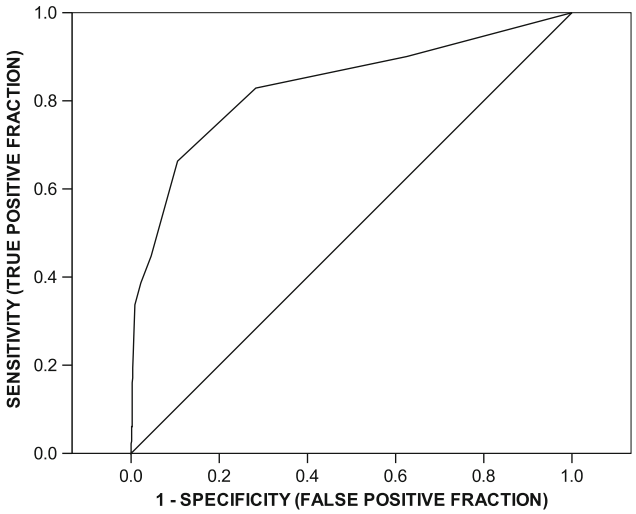


Fig. 3 ROC curve for victim equivalency count (child pornography use counted as the equivalent of two victims and added to number of real-life victims), computed on all subjects

The ROC curve was affected somewhat more (Fig. 3). The AUC increased to .83, 95% CI [.80, .87].

Discussion

First's helpful suggestion that the specificity of victim count—more specifically, the threshold score of three individual victims—should be demonstrated rather than assumed is addressed by the foregoing study. The specificity appears to be around 90%. For reasons already explained, this should be regarded as a minimum estimate of specificity. Although this study was limited to the diagnosis of pedohebephilia—the disorder for which I had the greatest quantity and highest quality of data—there is no obvious reason to assume that a cutting score in the neighborhood of three victims would produce markedly different false positive rates for other paraphilic disorders.

Responses to First's Other Major Points

Patterns versus Counts

First lays great emphasis on the notion that the diagnosis of paraphilia in uncooperative subjects must be based on a *pattern* of behaviors rather than a *count* of behaviors. In fact, he uses the word *pattern* 16 times in his letter. One would expect, given the importance of this concept to him, that he would supply two or three examples of diagnosable patterns. In fact, he supplies one: "involvement of a large number of victims over time coupled with the absence of sexual behavior involving non-paraphilic arousal stimuli."

One would also expect that if First is going to advance only one example, he would put forward his best one. That makes the limitations of this example all the more striking. Half of this "pattern" is simply victim count, with the vague term *large* substituted for a suggested minimum number of individual victims. The other half of the "pattern" is based on self-report, whose veracity is precisely the problem with uncooperative subjects in the first place. This requires some translation to be seen clearly.

In plain language, the phrase "involvement of a large number of victims over time coupled with the absence of sexual behavior involving non-paraphilic arousal stimuli" means that the patient has been performing a lot of deviant sexual activities and no normal sexual activities. If the patient has been performing some normal sexual activities then he should not, according to First, be diagnosed as paraphilic.

Now, a man's total number of normal sexual experiences (e.g., his total number of consenting adult sexual partners) is not a matter of public record in the way that his number of known sexual-offense victims is. The only way a clinician can know how many normal sexual experiences a man has had is by asking him. Thus, all an uncooperative patient has to do to avoid a diagnosis of paraphilia is to claim that he had, during the time frame in question, a number of normal sexual experiences with one or more consenting adult sexual partners, and that during all of them his mind was filled with thoughts of the situation at hand and not of fantasies that his partners were 11 years old or that he was strangling them rather than making love to them.

First's preference for patterns over counts may be based on a feeling that patterns are in the spirit of the DSM whereas counts are not. Yet, it is difficult to see how *victim counts* and *symptom counts* are completely different in form or function. Both approaches recommend a quantitative threshold for making a categorical decision. The DSM-IV-TR includes various examples of diagnoses based on a patient's showing *x* or more symp-

toms. This is especially true of the personality disorders. The core diagnostic criterion for Antisocial Personality Disorder (301.7), for example, requires the patient to have manifested three or more symptoms (p. 706). The sole diagnostic criterion for Borderline Personality Disorder (301.83) requires the patient to have shown five or more symptoms (p. 710).

Multiple Sexual Offenses Involving the Same Victim

First asserts, as established fact, that diagnosing pedohebephilia in uncooperative subjects from a minimum number of three victims would result in "a significant false negative problem." This follows from his premise that "pedophilia often involves multiple sexual offenses against one or two child victims who are known to the perpetrator." He cites no research to support this statement, nor does he attempt to define "often" as a proportion, computed on the total population of men assessed for pedohebephilia, of those men who offended against only one or two children, but on multiple occasions. In my experience, the men who get the opportunity to molest the same child on multiple occasions tend to be incest offenders, that is, men who have offended against their daughters or stepdaughters. The available evidence suggests that father–daughter incest offenders are *less* likely to be pedohebephilic than extrafamilial offenders (see original research and review by Blanchard et al., 2006).

There is a bigger problem, quite apart from the extra uncertainty in diagnosing father–daughter incest offenders, with attempting to diagnose pedohebephilia from the number of sexual interactions rather than the number of sexual victims: There is, to my knowledge, no published evidence that a man who has offended against a child on multiple occasions is more likely to be pedohebephilic than a man who has offended against a child on one occasion. There is therefore no justification for modifying the proposed criteria in order to diagnose pedohebephilia in uncooperative subjects on the grounds that they have interacted with a child on *x* or more occasions, or for *y* or more months. There is certainly no reason to drop the minimum victim criterion in order to accommodate such plausible but untested potential diagnostic signs. First himself reminds the reader that "one of the central requirements of the DSM-5 revision process is for recommendations to be grounded in empirical evidence."

Need for Formal Exclusionary Criteria

Clinicians who specialize in the assessment of sexual offenders are quite likely to see patients who, in addition to the sexual crimes that prompted the referral, present with mental retardation, antisocial personality disorder, alcohol or substance abuse, schizophrenia, (occasional) neurological disorders, and a miscellany of other conditions. The diagnostic question is essentially the same as in patients who present without obvious psychiatric disorders: Is the patient's sexual behavior probably the result of a paraphilia or of something else? The main difference between this group of patients and others is that the candidates for "something else" are more obvious.

First's comments on this head are quite brief, and it would be unfair to infer his complete views from them. It is noteworthy, however, that he does not mention the many cases in which a genuine paraphilia and some other psychiatric disorder are present in the same patient. Thus, he discusses paraphilia and schizophrenia (for example) as competing explanations for some observed sexual behavior, not as potentially comorbid conditions in which the patient's paraphilia—not his schizophrenia—may be the motivating cause of his criminal sexual behavior. In line with this perspective, he suggests that an exclusionary criterion should be added to all of the paraphilias. This exclusionary criterion would list the other DSM diagnoses that should be ruled out as alternative explanations of the patient's symptoms.

First justifies this recommendation, in part, with an example in which differential diagnosis involves disorders with relatively confusable symptoms (specific phobia versus OCD, PTSD, or separation anxiety disorder). It is worth noting that such models do not put the problem of diagnosing paraphilia in people with mental disorders in the most natural context: The symptoms of schizophrenia are not similar to the symptoms of exhibitionism, and the symptoms of mental retardation are not similar to the symptoms of pedohebephilia. Whether because he draws on such models or for some other reason, First's recommendations for the assessment of uncooperative patients emphasize differential diagnoses and ignore comorbid diagnoses, and he accordingly stresses that clinicians assessing an uncooperative patient must rule out other possible causes for sexual offending before making a diagnosis of paraphilia.

That is not my view at all. I think that the presence or absence of a paraphilia must, as much as possible, be investigated independently of any other, obvious psychiatric problem. That follows from my experience that paraphilias can be, and sometimes are, comorbid with other DSM conditions. One cannot assume that a schizophrenic patient who exposed himself to a strange woman did so *because* he has schizophrenia; he might have exhibitionism as well as schizophrenia. One certainly cannot assume that a mentally retarded patient who molested a child did so *because* he has mental retardation; there is evidence suggesting significant comorbidity of pedohebephilia and mental retardation (e.g., Blanchard et al., 1999; Rice, Harris, Lang, & Chaplin, 2008). In my opinion, adding a boilerplate exclusionary criterion to each of the paraphilia criterion sets is likely to do more harm than good, in that it would steer clinicians away from the possibility of comorbid diagnoses.

Conclusion

I stressed the importance of avoiding false positive diagnoses in my first report to the Paraphilias Subworkgroup (Blanchard,

2010a), and I explicitly suggested biasing diagnostic criteria toward making false negative diagnoses rather than false positive diagnoses. I have even been criticized for this (O'Donohue, 2010). I therefore believe that First and I are on the same side of this fundamental issue. I also believe, however, that his concerns about false positive diagnoses of paraphilia are exaggerated or misplaced, and that the changes he suggests for the proposed DSM-5 diagnostic criteria for the paraphilias are unnecessary or even undesirable.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th ed., text revision)*. Washington, DC: Author.

American Psychiatric Association. (2010). *Proposed draft revisions to DSM disorders and criteria*. http://www.dsm5.org. Retrieved July 2, 2010.

Blanchard, R. (2010a). The DSM diagnostic criteria for Pedophilia. *Archives of Sexual Behavior, 39*, 304–316.

Blanchard, R. (2010b). The fertility of hebephiles and the adaptationist argument against including hebephilia in DSM-5 [Letter to the Editor]. *Archives of Sexual Behavior, 39*, 817–818.

Blanchard, R., Klassen, P., Dickey, R., Kuban, M. E., & Blak, T. (2001). Sensitivity and specificity of the phallometric test for pedophilia in nonadmitting sex offenders. *Psychological Assessment, 13*, 118–126.

Blanchard, R., Kolla, N. J., Cantor, J. M., Klassen, P. E., Dickey, R., Kuban, M. E., et al. (2007). IQ, handedness, and pedophilia in adult male patients stratified by referral source. *Sexual Abuse: A Journal of Research and Treatment, 19*, 285–309.

Blanchard, R., Kuban, M. E., Blak, T., Cantor, J. M., Klassen, P., & Dickey, R. (2006). Phallometric comparison of pedophilic interest in nonadmitting sexual offenders against stepdaughters, biological daughters, other biologically related girls, and unrelated girls. *Sexual Abuse: A Journal of Research and Treatment, 18*, 1–14.

Blanchard, R., Kuban, M. E., Blak, T., Klassen, P. E., Dickey, R., & Cantor, J. M. (submitted). Sexual attraction to others: A comparison of two models of alloerotic responding in men. *Archives of Sexual Behavior*.

Blanchard, R., Lykins, A. D., Wherrett, D., Kuban, M. E., Cantor, J. M., Blak, T., et al. (2009). Pedophilia, hebephilia, and the DSM-V. *Archives of Sexual Behavior, 38*, 335–350.

Blanchard, R., Watson, M. S., Choy, A., Dickey, R., Klassen, P., Kuban, M., et al. (1999). Pedophiles: Mental retardation, maternal age, and sexual orientation. *Archives of Sexual Behavior, 28*, 111–127.

Cantor, J. M., Klassen, P. E., Dickey, R., Christensen, B. K., Kuban, M. E., Blak, T., et al. (2005). Handedness in pedophilia and hebephilia. *Archives of Sexual Behavior, 34*, 447–459.

First, M. B. (2010). DSM-5 proposals for paraphilias: Suggestions for reducing false positives related to use of behavioral manifestations [Letter to the Editor]. *Archives of Sexual Behavior*. doi:10.1007/s10508-010-9657-5.

Freund, K. (1965). Diagnosing heterosexual pedophilia by means of a test for sexual interest. *Behaviour Research and Therapy, 3*, 229–234.

Freund, K. (1967). Diagnosing homo- or heterosexuality and erotic age-preference by means of a psychophysiological test. *Behaviour Research and Therapy, 5*, 209–228.

Freund, K., Sedlacek, F., & Knob, K. (1965). A simple transducer for mechanical plethysmography of the male genital. *Journal of the Experimental Analysis of Behavior, 8*, 169–170.

Kuban, M., Barbaree, H. E., & Blanchard, R. (1999). A comparison of volume and circumference phallometry: Response magnitude and method agreement. *Archives of Sexual Behavior, 28*, 345–359.

Lykins, A. D., Cantor, J. M., Kuban, M. E., Blak, T., Dickey, R., Klassen, P. E., et al. (2010). The relation between peak response magnitudes and agreement in diagnoses obtained from two different phallometric tests for pedophilia. *Sexual Abuse: A Journal of Research and Treatment, 22*, 42–57.

O'Donohue, W. (2010). A critique of the proposed DSM-V diagnosis of Pedophilia [Letter to the Editor]. *Archives of Sexual Behavior, 39*, 587–590.

Rice, M. E., Harris, G. T., Lang, C., & Chaplin, T. C. (2008). Sexual preferences and recidivism of sex offenders with mental retardation. *Sexual Abuse: A Journal of Research and Treatment, 20*, 409–425.

Seto, M. C., Cantor, J. M., & Blanchard, R. (2006). Child pornography offenses are a valid diagnostic indicator of pedophilia. *Journal of Abnormal Psychology, 115*, 610–615.

Arch Sex Behav (2010) 39:1253–1260
DOI 10.1007/s10508-010-9660-x

LETTER TO THE EDITOR

Proposed Revisions to Gender Identity Disorder Diagnoses in the DSM-5

Anne A. Lawrence

Published online: 20 August 2010

In early 2010, the American Psychiatric Association (APA) announced proposed revisions to psychiatric diagnoses for the forthcoming fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5). These revisions included proposed changes to Gender Identity Disorder (GID) diagnoses in adolescents and adults (APA, 2010c), as recommended by the APA's GID Subworkgroup. The Subworkgroup members laid the groundwork for their recommendations in articles published in *Archives of Sexual Behavior* (Cohen-Kettenis & Pfäfflin, 2010; Drescher, 2010a; Meyer-Bahlburg, 2010a; Zucker, 2010). The most important elements of the proposed revisions to GID diagnoses for adolescents and adults are:

1. GIDs are conceptualized as reflecting a "marked incongruence between one's experienced/expressed gender and assigned gender" (APA, 2010c, Proposed Revision section), leading to the recommendation that the name *Gender Incongruence* (GI) replace *GID*.
2. The clinical indicators for GID/GI have been revised to reflect empirical evidence, and the proposed number of indicators required for diagnosis has been specified.
3. The presence of clinically significant distress or impairment is no longer required for the diagnosis of GID/GI but is proposed to be evaluated "separately and independently" (APA, 2010c, Rationale section, ¶ 9).
4. Persons with a disorder of sex development (DSD) have become eligible for the diagnosis of GID/GI, and the presence or absence of a DSD is used to define subtypes.
5. Subtypes based on sexual orientation have been eliminated.
6. The status of the diagnosis of Gender Identity Disorder Not Otherwise Specified (GIDNOS) is uncertain, raising the possibility that the diagnosis may not be retained.

I will comment on these elements of the proposed revisions and will then offer alternative diagnostic criteria for the diagnoses of GID/GI and GIDNOS.

Revised Conceptualization of GIDs and Proposed Name Change

The proposed revisions conceptualize the defining feature of GIDs as "marked incongruence between one's experienced/expressed gender and assigned gender" (APA, 2010c, Proposed Revision section), leading to the recommendation that the disorder be renamed Gender Incongruence (GI). The DSM-IV (APA, 1994) and DSM-IV-TR (APA, 2000), in contrast, conceptualized the defining features of GID as "strong and persistent cross-gender identification" and "persistent discomfort with [one's] sex or sense of inappropriateness in the gender role of that sex" (APA, 2000, p. 581). The newly proposed conceptualization is, in effect, a reversion to that of the DSM-III-R (APA, 1987), which described GIDs as reflecting "an incongruence between assigned sex (i.e., the sex that is recorded on the birth certificate) and gender identity" (p. 71), but with the term *gender identity* now replaced by *experienced/expressed gender* and *assigned sex* replaced by *assigned gender*. I will discuss the following points:

A. A. Lawrence
Department of Psychology, University of Lethbridge, Lethbridge, AB, Canada

A. A. Lawrence (✉)
6801 28th Ave. NE, Seattle, WA 98115, USA
e-mail: alawrence@mindspring.com

a. The term *experienced/expressed gender* is synonymous with *gender identity*, and the latter term is preferable, based on familiarity and clarity.
b. Except in some persons with a DSD, the incongruence in question is between gender identity and biologic sex, and this formulation is preferable, based on familiarity and clarity.
c. Two supposed benefits of the proposed new conceptualization are inclusion of persons with a DSD and provision of an "exit clause" for the diagnosis, but neither result would be desirable, and there would be a better way to accomplish the latter if it were desirable.
d. The proposed name Gender Incongruence is merely a euphemism for GID (except, perhaps, in persons with a DSD), and the latter name is preferable, based on continuity, familiarity, and clarity.
e. The phenomenon described by the proposed clinical indicators corresponds closely to the concept of *gender dysphoria*, suggesting that Gender Dysphoric Disorder could be an acceptable alternative name.

Experienced/Expressed Gender is Synonymous with Gender Identity

In an explanatory statement, the Subworkgroup defined *experienced/expressed gender* as "what *identity* [emphasis added] one experiences and/or expresses" (APA, 2010c, Rationale section, ¶ 2). Moreover, in the Subworkgroup's document *Dimensional Assessment for Gender Identity Disorder (Gender Incongruence) for Adolescents and Adults* (APA, 2010c, Severity section), in which the proposed clinical indicators for GI (see Table 1) are operationalized as questionnaire items, the term *gender identity* replaces *experienced/expressed gender* in the items corresponding to indicators A1 and A2. The Subworkgroup members apparently concluded that *gender identity* is not only synonymous with *experienced/expressed gender* but is better understood by patients; I believe they were correct in their conclusion. Gender identity is a widely recognized and generally well understood term; there is little justification for replacing it with a newly invented synonym.

The Incongruence in Question is Between Gender Identity and Biologic Sex

Except in persons with a DSD, in whom biologic sex may be ambiguous, the incongruence in question is between gender identity and biologic sex (reflected in assigned sex), not between gender identity and assigned gender, as the proposed language states. The Subworkgroup's own explanation concedes as much: "The term 'sex' has been replaced by assigned 'gender' in order to make the criteria applicable to individuals with a DSD" (APA, 2010c, Rationale section, ¶ 4). It is also notable that most of the proposed clinical indicators (see Table 1) reference sex, not gender. Indicators A1–A3 explicitly refer to dysphoria in relation to biologic sex characteristics, and a reference to biologic sex is also implied, if not explicitly stated, in indicators A4, A5, and perhaps A6. In any case, gender is not what is assigned at birth: Sex is assigned—or, more accurately, recognized—at birth, usually based on the morphology of the external genitalia. Birth certificates specify one's sex, not one's gender. If persons with a DSD were excluded from the diagnosis of GID/GI—a desirable outcome, as I will argue later in this Letter—then there could be no reasonable objection to framing the incongruence in relation to biologic sex, a formulation that is preferable on the basis of familiarity and clarity.

Framing the incongruence in relation to biologic sex would also prevent the diagnosis of GID/GI being given yet again to "transitioned individuals who have regrets" (APA, 2010c, Rationale section, ¶ 4), a possibility that the Subworkgroup members apparently consider desirable but that I regard as bizarre and confusing. In my opinion, it would make little sense to diagnose individuals with a mental disorder for wanting to live as members of their natal sex.

Two Supposed Benefits of the Proposed Conceptualization are not Desirable

The Subworkgroup has offered two rationales for conceptualizing GID/GI as an incongruence between experienced/expressed gender and assigned gender: to "make the criteria applicable to individuals with a DSD" (APA, 2010c, Rationale section, ¶ 4) and to provide an "exit clause" from the diagnosis, allowing "individuals who have successfully transitioned to 'lose' the diagnosis after satisfactory treatment" (APA, 2010c, Rationale section, ¶ 4).

I will argue later in this Letter that it would be inadvisable to allow persons with a DSD to be eligible for the diagnosis of GID/GI. It would also be undesirable to provide an automatic "exit clause" from the diagnosis of GID/GI, simply on the basis of having undergone nominal gender reassignment. Individuals should "lose" the diagnosis–to the extent that any psychiatric diagnosis can be lost—only if they no longer experience clinically significant gender dysphoria. The proposed diagnostic criteria for GI, however, seemingly would automatically render the diagnosis inapplicable following a sought-after change in assigned gender, because the newly assigned gender would no longer be incongruent with the patient's experienced/expressed gender. Consider a gender dysphoric adult male who for medical or financial reasons was temporarily or permanently unable to undergo cross-sex hormone therapy or vaginoplasty. If this person nevertheless underwent a change in assigned gender (say, as a result of living in the desired gender role and achieving legal recognition as a member of the desired sex), her newly assigned gender would no longer

Table 1 GID Subworkgroup's proposed clinical indicators for gender incongruence

A1	A marked incongruence between one's experienced/expressed gender and primary and/or secondary sex characteristics (or, in young adolescents, the anticipated secondary sex characteristics)
A2	A strong desire to be rid of one's primary and/or secondary sex characteristics because of a marked incongruence with one's experienced/expressed gender (or, in young adolescents, a desire to prevent the development of the anticipated secondary sex characteristics)
A3	A strong desire for the primary and/or secondary sex characteristics of the other gender
A4	A strong desire to be of the other gender (or some alternative gender different from one's assigned gender)
A5	A strong desire to be treated as the other gender (or some alternative gender different from one's assigned gender)
A6	A strong conviction that one has the typical feelings and reactions of the other gender (or some alternative gender different from one's assigned gender)

Note: Fulfillment of two or more indicators is required for diagnosis

From APA (2010c, Proposed Revision section)

be incongruent with her experienced/expressed gender, and she seemingly could no longer be diagnosed with GI, even if she were still severely gender dysphoric, due to a disparity between her gender identity and her primary and secondary sex characteristics. Surely this would not be a desirable outcome.

Moreover, to the extent that a formal diagnosis of GID/GI is necessary to justify access to clinical care, the language of the proposed criteria implies that this justification would automatically disappear once nominal gender (or sex) reassignment had occurred. Gender dysphoric patients may remain gender dysphoric and may require transition-related care for months or years after nominal gender (sex) reassignment has occurred; automatic loss of the diagnosis could jeopardize access to ongoing care.

Adding a distress/impairment criterion to the diagnosis—which would be desirable for other reasons, as I will argue later in this Letter—would be a more appropriate way to ensure that individuals who no longer experienced clinically significant gender dysphoria would not meet full criteria for the diagnosis and would no longer carry the diagnosis in the usual, unqualified sense. It is important to recognize, however, that individuals do not and should not necessarily lose all connection to a psychiatric diagnosis simply because they no longer meet full criteria for that diagnosis. The DSM-IV-TR (APA, 2000) explains that "a DSM-IV diagnosis … is not typically used to denote previous diagnoses from which the individual has recovered" (p. 1). It goes on to explain, however, that certain specifiers (i.e., Prior History, In Full Remission, and In Partial Remission) can be applied to such a diagnosis when "it may be useful to note a history of the criteria having been met for a disorder even when the individual is considered to be recovered from it" (APA, 2000, p. 1) or when the disorder is in full or partial remission.

It is hard to imagine a circumstance in which it would not be useful for a treating clinician to be aware that a patient had once met criteria for GID/GI. In my clinical experience, the fact of having once suffered from severe, persistent, clinically significant gender dysphoria, even decades earlier, remains psychologically relevant throughout life for almost all persons who have undergone sex reassignment. Therefore, retention of the diagnosis, accompanied by the specifier In Partial Remission, In Full Remission, or Prior History, would almost always be appropriate.

Gender Incongruence is Merely a Euphemism for GID

The members of the GID Subworkgroup apparently consider GI to be a genuine mental disorder: One Subworkgroup member has stated on an electronic mailing list that "no one [in the Subworkgroup] to my knowledge has recommended moving GI to a V-code" (Drescher, 2010b)—that is, classifying GI as a nondisorder "Condition That May Be a Focus of Clinical Attention" (APA, 2000, p. 739). If GI is indeed considered to be a mental disorder, how is the disorder conceptualized? As I have argued above, the proposed language conceptualizes the disorder as an incongruence between gender identity and biologic sex. In the absence of a DSD, there is no persuasive evidence that biologic sex is disordered in persons who experience this type of incongruence. Consequently, the disorder reflects disordered gender identity, which is incongruent with nondisordered biologic sex. GI, therefore, is merely a euphemism for GID.

There are good reasons to employ the more explicit term. Using the term GID would maintain continuity with current clinical usage, with previous editions of the DSM (in which all severe manifestations of gender dysphoria have been considered GIDs, even if they have received other specific names, such as Transsexualism), and with the ICD-10 (World Health Organization, 1992), in which Disorders of Gender Identity is likewise a superordinate category. The term GID is familiar to and widely understood by both professionals and laypersons. It is also clear: It specifies precisely what is disordered. The term GI, in contrast, is vague—perhaps intentionally so–concerning what, if anything, is disordered.

The Described Phenomenon is Gender Dysphoria, Which Suggests Another Possible Name

I have argued that the phenomenon described by the proposed clinical indicators is a disorder of gender identity and that there are good reasons to state this explicitly. Such a conceptualization is unacceptable to some transgender activists, however, and the Subworkgroup members understandably would prefer a conceptualization that these activists would find more

acceptable. Because the phenomenon described by indicators A1–A6 corresponds closely to the concept of gender dysphoria, Gender Dysphoric Disorder might be another appropriate name for the disorder. Subworkgroup members Cohen-Kettenis and Pfäfflin (2010) also arrived at this conclusion.

The Glossary section of the DSM-IV-TR defines gender dysphoria as "a persistent aversion toward some or all of those physical characteristics or social roles that connote one's own biological sex" (APA, 2000, p. 823). Yet another definition can be derived from Criterion B of the DSM-IV-TR diagnosis of GID: "persistent discomfort with [one's] sex or sense of inappropriateness in the gender role of that sex" (APA, 2000, p. 581). Both of these definitions capture the spirit and most of the letter of indicators A1–A6. Thus, the opening text of Criterion A for Gender Dysphoric Disorder might read like this: "A strong feeling of discomfort with one's sex or the gender role associated with one's sex, of at least 6 months duration, as manifested by two or more of the following indicators" (see Table 2).

The name Gender Dysphoric Disorder probably would not satisfy all transgender activists: Many would resist having their condition described as a disorder, albeit this would be true whether the word *disorder* appeared in the condition's name or not. A particular complaint of some activists, however, is that the name GID implies that their identities themselves are disordered; employing an alternative name would avoid emphasizing this unpleasant reality and might circumvent some complaints. Gender Incongruence would be one such alternative name, but, as I have argued above, it is less than ideal; Gender Dysphoric Disorder would be a better alternative.

Another advantage of using the name Gender Dysphoric Disorder would be that, if a distress/impairment criterion were included as an additional diagnostic criterion—as I will recommend later in this Letter—then persons who met Criterion A but not the distress/impairment criterion could be *ascertained* to have gender dysphoria, whereas persons who met full criteria could be *diagnosed* with Gender Dysphoric Disorder. The DSM-5 Paraphilias Subworkgroup recommended making a similar distinction between *paraphilias* and *paraphilic disorders* (e.g., APA, 2010a, Rationale section, ¶ 1), based on nonfulfillment versus fulfillment of a distress/impairment criterion.

Table 2 Proposed diagnostic criteria for DSM-5 GID diagnoses

Gender Identity Disorder [*or* Gender Dysphoric Disorder] in Adolescents or Adults
A. A strong feeling of discomfort with one's sex or the gender role associated with one's sex, of at least 6 months duration, as manifested by two or more of the following indicators:
1. A strong feeling of incongruence between one's gender identity and one's primary or secondary sex characteristics (or, in young adolescents, one's anticipated secondary sex characteristics)
2. A strong desire to be rid of one's primary or secondary sex characteristics (or, in young adolescents, a strong desire to prevent the development of one's anticipated secondary sex characteristics)
3. A strong desire for the primary or secondary sex characteristics of the other sex
4. A strong desire to be the other sex
5. A strong desire to be treated as the other sex
6. A strong conviction that one has the typical feelings and reactions of the other sex.
B. The condition is not concurrent with a physical intersex condition [*or* physical disorder of sex development]
C. The condition causes clinically significant distress or impairment in social, occupational, or other important areas of functioning
Subtypes:
Sexually attracted to males
Sexually attracted to females
Sexually attracted to both
Sexually attracted to neither
Gender Identity Disorder Not Otherwise Specified
This category is included for coding disorders of gender identity that are not classifiable as a specific Gender Identity Disorder and that cause clinically significant distress or impairment in social, occupational, or other important areas of functioning

Revised Clinical Indicators for GID/GI in Adolescents and Adults

The proposed changes to the clinical indicators for GID/GI represent significant improvements over the DSM-IV-TR in many respects. Based on the demonstration by Deogracias et al. (2007) that GID/gender dysphoria is best conceptualized as a unidimensional construct, the Subworkgroup has proposed that clinical indicators related to anatomic dysphoria (A1–A3; see Table 1) and cross-sex wishes or identification (A4–A6) be combined into a single, unified set and that the number of indicators required for diagnosis be explicitly specified.

The proposed indicators could be further improved in a few particulars, however. First, the proposed language for indicator A1 (see Table 1) does not make it clear whether the incongruence between experienced/expressed gender and primary or secondary sex characteristics is a sign (something the clinician observes) or a symptom (something the patient reports). Assuming the latter is intended, I would suggest inserting the words *feeling of*: "a feeling of marked incongruence between one's experienced/expressed gender [or gender identity] and primary and/or secondary sex characteristics."

Second, proposed indicators A4–A6 contain parenthetical explanations that they are applicable to persons who identify with "some alternative gender different from one's assigned gender" (see Table 1), rather than exclusively to persons who identify with "the other gender" (i.e., the other sex). There is

no empirical basis, however, for this radical expansion of the types of gender identity to which these indicators might apply. The strongest argument for the validity of the clinical indicators, at least A2–A6, is that they were derived from items in an empirically validated instrument, the Gender Identity/Gender Dysphoria Questionnaire for Adolescents and Adults (GIDYQ; Deogracias et al., 2007). The relevant GIDYQ items, however, conceptualize gender identity in binary terms. The Subworkgroup's report did not state which specific questionnaire items from the GIDYQ were used to create indicators A4–A6, but one can make reasonable guesses, assuming that items with high factor loadings were selected (the quoted items and factor loadings are from Deogracias et al., pp. 374, 377–378): (1) A4 was probably derived from GIDYQ item 16 ("the wish or desire to be a man" [for females] or "a woman" [for males]), with factor loading .93; (2) A5 was probably derived from item 6 ("better for you to live as a man" [for females] or "as a woman" [for males]), with factor loading .93; (3) A6 was probably derived from item 10 ("felt more like a man than like a woman" [for females] or "more like a woman than like a man" [for males]), with factor loading .96.

These GIDYQ items conceptualize gender identity in binary terms. By transforming them into indicators that conceptualize gender identity as a "multi-category concept or spectrum" (APA, 2010c, Rationale section, ¶ 10), the Subworkgroup has unnecessarily discarded the presumption of validity. It is also worth noting that the few items in the GIDYQ that imply other than a binary gender identity are associated with considerably lower factor loadings, albeit still high in absolute terms (the quoted items and factor loadings are again from Deogracias et al., pp. 374, 377–378): (1) Item 25, "thought of yourself as a 'transgendered person'": factor loading .78; (2) Item 2, "feeling somewhere in between a woman and a man": factor loading .58; (3) Item 9, "at times feeling more like a man and at times feeling more like a woman": factor loading .58; (4) Item 11, "felt that you did not have anything in common with either men or women": factor loading .47.

The Subworkgroup conceded that, because the wording of the GIDYQ items "is not identical to the wording of the proposed indicators, further validation work will be required during field trials" (APA, 2010c, Rationale section, ¶ 1). The parenthetical additions to indicators A4–A6, however, do not involve mere changes in wording but changes in the very way gender identity is conceptualized. This undercuts the otherwise strong empirical basis of the proposed indicators. Consequently, I recommend that the parenthetical references to "some alternative gender" in indicators A4–A6 be deleted. Gender dysphoric persons who desire to be "some alternative gender" other than male or female undoubtedly do exist (e.g., Johnson & Wassersug, 2010) but should be diagnosed with GIDNOS, not GID/GI; I will discuss this recommendation in more detail later in this Letter.

Finally, indicators A4–A6 refer to the desire to be, the desire to be treated as, or the conviction that one has the feelings and reactions of the other *gender*. I believe, however, that the reference should be to the other *sex*. As noted earlier, the Subworkgroup stated that "the term 'sex' has been replaced by assigned 'gender' in order to make the criteria applicable to individuals with a DSD" (APA, 2010c, Rationale section, ¶ 4). If persons with a DSD and persons who desire to be some sex/gender other than male or female were excluded from the diagnosis, there would be no reason not to employ the more accurate term, *sex*. Most severely gender dysphoric persons want to assume, insofar as possible, the anatomic status of the other biologic sex (not just the associated gender role) and be treated by other people as members of the other biologic sex (not just as persons who have assumed that gender role). Admittedly, the situation with respect to indicator A6 is more equivocal—*gender* is not clearly incorrect—but *sex* works equally well or better, and its use would contribute to consistency of expression.

Absence of a Distress/Impairment Criterion

GID/GI is the only major diagnosis proposed by the DSM-5 Sexual and Gender Identity Disorders Workgroup that does not incorporate a distress/impairment criterion. The absence of such a criterion is surprising: Most theoretical examinations of the concept of mental disorders undertaken in connection with the DSM-5 revision process (e.g., First & Wakefield, 2010; Stein et al., 2010; Wakefield & First, 2003) have emphasized the need to include a distress/impairment criterion to avoid problems of overdiagnosis. The absence of such a criterion is especially surprising, given that the Subworkgroup rejects the idea that a strong desire for sex reassignment necessarily indicates "inherent distress" (APA, 2010c, Rationale section, ¶ 9).

I believe the risk of overdiagnosis is significant if a distress/impairment criterion is not included in the diagnostic criteria for GID/GI. In recommending that "the GI diagnosis be given on the basis of the A criterion alone and that distress and/or impairment … be evaluated separately and independently" (APA, 2010c, Rationale section, ¶ 9), the Subworkgroup members have declared that it is appropriate to diagnose a mental disorder in gender variant people who are not significantly distressed or impaired by their gender variance. To the extent that psychiatric diagnoses are inherently stigmatizing, such a declaration is troubling, even if one believes, as I do, that some inherent distress, albeit perhaps not always clinically significant distress, can be assumed for most of the proposed clinical indicators.

What kinds of persons could be diagnosed with GID/GI on the basis of indicators A1–A6 alone, despite experiencing little or no distress or impairment? Probably they would be persons who fulfilled only indicators A1 and A6, the only

indicators that do not involve strongly felt, unmet cross-sex desires. Hypothetically, such persons might strongly identify with the other sex (and thus feel their identity to be incongruent with their anatomy) and might also believe they had the typical feelings and reactions of the other sex but might not feel any strong desire to undergo sex reassignment. Consequently, they might be neither distressed nor impaired by living as members of their natal sex. Natal females, in whom substantial sex-atypical expression is often socially tolerated, might be especially likely to experience little or no distress or impairment, despite identifying with and believing themselves similar to the other sex. I can see no justification for diagnosing GID/GI in such persons; adding a distress/impairment criterion would prevent this.

As noted earlier, if it were considered desirable to provide an "exit clause" from the diagnosis of GID/GI, a distress/impairment criterion would provide a straightforward way of accomplishing this. It would make better sense intuitively to lose a diagnosis because one no longer suffered any distress or impairment than because nominal gender reassignment had theoretically eliminated the incongruence between gender identity and assigned gender.

Eligibility of Persons with a DSD for the Diagnosis of GID/GI

The presence of a DSD, formerly called an intersex condition, was an exclusion criterion for the diagnosis of Transsexualism in the DSM-III (APA, 1980) but not in the DSM-III-R (APA, 1987). A DSD was reintroduced as an exclusion criterion for the diagnosis of GID in the DSM-IV (APA, 1994) and continued in the DSM-IV-TR (APA, 2000). The reintroduction of a DSD as an exclusion criterion was the result of an extensive review by the GID Subcommittee of the DSM-IV Task Force; one of the principal conclusions of this review was summarized by Meyer-Bahlburg (1994):

> Intersex patients with significant gender identity problems or gender change do differ from nonintersex patients with GID in prevalence, in age at onset and presentation, and in the sex ratio, and the evidence available—in spite of its methodological shortcomings—makes it very likely that the development of gender problems in intersex patients is in most cases not directly comparable to GID as it develops in nonintersex patients. The two forms of gender identity problems are unlikely to be the same disorder. (p. 33)

Meyer-Bahlburg recommended that it would be best to "exempt all intersex patients from the GID diagnosis" (p. 35)—that is, from being diagnosed with either GID or GIDNOS—but this recommendation was not adopted: In the DSM-IV and DSM-IV-TR, intersex patients were excluded from the diagnosis of GID proper but could be diagnosed with GIDNOS.

Fifteen years later, Meyer-Bahlburg (2009) expressed conclusions almost identical to his earlier ones:

> The differences between GIV [gender identity variants] in DSD individuals and GIV in non-DSD individuals with regard to phenomenon, context of presentation, epidemiology, etiology, and (pediatric) treatment settings are so large that, at this stage of our knowledge, identical diagnostic categories and treatment approaches are not justified, as has recently also been argued by Mazur, Colsman, and Sandberg (2007). (p. 231)

The review articles authored by the Subworkgroup members offered no compelling new evidence that these conclusions should be modified or disregarded: Meyer-Bahlburg (2010a) expressed no doubts about his previously stated conclusions, and Zucker (2010) recommended that a physical intersex condition continue to be an exclusion criterion for the diagnosis of GID in children. As noted by Meyer-Bahlburg (2009), Mazur et al. (2007) also concluded that, because of the many differences between GID in persons with and without intersex conditions, "it would be prudent to consider them as separate entities when initiating an evaluation" (p. 236). Accordingly, the proposed recommendation to make persons with a DSD eligible for the diagnosis of GID/GI seems ill-advised.

One could argue that the DSM-IV and DSM-IV-TR already allowed males with two very different conditions and probably different etiologies—homosexual and nonhomosexual gender dysphoria—to receive the same GID diagnosis, distinguishing the two conditions using subtypes (Lawrence, 2010). The proposal to allow gender dysphoric persons with and without a DSD to receive the same GID/GI diagnosis and to distinguish their conditions using subtypes could be seen as merely the extension of an existing principle. But extending the diagnosis of GID/GI to persons with a DSD would only unnecessarily increase variability within a diagnostic category that already encompasses substantial variability.

The Subworkgroup would have been better advised to follow Meyer-Bahlburg's (1994) original recommendation and make persons with a DSD ineligible for any GID/GI diagnosis. Gender dysphoric patients with a DSD already have a medical diagnosis that provides access to treatment, unlike gender dysphoric persons without a DSD. The stigma associated with psychiatric diagnoses is yet another reason for excluding persons with a DSD from any GID/GI diagnosis, or at least from the diagnosis of GID/GI proper.

Elimination of Subtypes Based on Sexual Orientation

Beginning in 1980, subtypes based on sexual orientation have been used in connection with GID diagnoses in every edition

of the DSM. Subtypes based on sexual orientation offer substantial descriptive, prognostic, and heuristic value; the Subworkgroup's proposal to eliminate them is ill-advised. In a review article (Lawrence, 2010), written at the request of the GID Subworkgroup (Meyer-Bahlburg, 2010b), I observed:

> The most widely used and influential typologies for transsexualism and gender identity disorder (GID) in adolescents and adults employ either sexual orientation or age of onset of GID-related symptoms as bases for categorization…. Typologies based on sexual orientation and age of onset of GID-related symptoms are roughly comparable in ease and reliability of subtype assignment. Typologies based on sexual orientation, however, employ subtypes that are less ambiguous and better suited to objective confirmation and that offer more concise, comprehensive clinical description. Typologies based on sexual orientation are also superior in their ability to predict treatment-related outcomes and comorbid psychopathology and to facilitate research. Commonly expressed objections to typologies based on sexual orientation are unpersuasive when examined closely. The DSM should continue to employ subtypes based on sexual orientation for the diagnosis of GID in Adolescents or Adults or its successor diagnosis. (p. 514)

Uncertain Status of the Diagnosis of GIDNOS

The status of the diagnosis of GIDNOS is uncertain: According to the DSM-5 website, "Changes, or lack thereof, are still under review by the work group" (APA, 2010b). This suggests that the diagnosis of GIDNOS might not be retained in the DSM-5. I believe that the diagnosis should be retained.

There should be a diagnostic category other than GID/GI for gender dysphoric persons who do not identify as members of the other sex, desire to have the primary or secondary sex characteristics of the other sex, or desire to live as members of the other sex. Too little is known about such persons to confidently place them in the same diagnostic category as persons who unequivocally identify with and want to belong to the other sex (i.e., transsexuals), a population about which a great deal is known. Although there have been Internet-based surveys of persons who claim to identify as neither men nor women, I am not aware of any systematic studies involving actual clinical patients who so identify. Absent such studies, I believe it would be premature to diagnose such patients with GID/GI.

This kind of cautious approach to nosology explains why, for the last 30 years, the DSM has always provided one or more alternative GID categories for individuals who cannot confidently be classified with transsexuals for diagnostic purposes: Atypical Gender Identity Disorder in the DSM-III; Gender Identity Disorder of Adolescence or Adulthood, Nontranssexual Type and GIDNOS in the DSM-III-R; and GIDNOS in the DSM-IV and DSM-IV-TR. In my opinion, such an alternative diagnostic category (i.e., GIDNOS) should also exist in the DSM-5, and persons with gender dysphoria or gender identity problems who identify as neither men nor women and wish to live as neither men nor women should receive a diagnosis of GIDNOS, not GID/GI.

Proposed Revised Language for GID Diagnoses in the DSM-5

Table 2 displays the alternative diagnostic criteria I propose for GID diagnoses in the DSM-5, based on the foregoing analysis.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., revised). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

American Psychiatric Association. (2010a). *Proposed revision: 302.3, Transvestic Fetishism*. Retrieved from http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=189.

American Psychiatric Association. (2010b). *Proposed revision: 302.6, Gender Identity Disorder Not Otherwise Specified*. Retrieved from http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=194.

American Psychiatric Association. (2010c). *Proposed revision: 302.85, Gender Identity Disorder in Adolescents or Adults*. Retrieved from http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=482.

Cohen-Kettenis, P. T., & Pfäfflin, F. (2010). The DSM diagnostic criteria for gender identity disorder in adolescents and adults. *Archives of Sexual Behavior, 39*, 499–513.

Deogracias, J. J., Johnson, L. L., Meyer-Bahlburg, H. F. L., Kessler, S. J., Schober, J. M., & Zucker, K. J. (2007). The Gender Identity/Gender Dysphoria Questionnaire for adolescents and adults. *Journal of Sex Research, 44*, 370–379.

Drescher, J. (2010a). Queer diagnoses: Parallels and contrasts in the history of homosexuality, gender variance, and the diagnostic and statistical manual. *Archives of Sexual Behavior, 39*, 427–460.

Drescher, J. (2010b, February 20). Re: Hyperbole [Electronic mailing list message]. *Sexnet* listserv.

First, M. B., & Wakefield, J. C. (2010). Defining 'mental disorder' in DSM-V. *Psychological Medicine*. doi:10.1017/S0033291709992339.

Johnson, T. W., & Wassersug, R. J. (2010). Gender identity disorder outside the binary: When gender identity disorder-not otherwise specified is not good enough [Letter to the Editor]. *Archives of Sexual Behavior, 39*, 597–598.

Lawrence, A. A. (2010). Sexual orientation versus age of onset as bases for typologies (subtypes) for gender identity disorder in adolescents and adults. *Archives of Sexual Behavior, 39*, 514–545.

Mazur, T., Colsman, M., & Sandberg, D. E. (2007). Intersex: Definitions, examples, gender stability, and the case against merging with transsexualism. In R. Ettner, S. Monstrey, & A. E. Eyler (Eds.), *Principles of transgender medicine and surgery* (pp. 235–259). Binghamton, NY: Haworth Press.

Meyer-Bahlburg, H. F. L. (1994). Intersexuality and the diagnosis of gender identity disorder. *Archives of Sexual Behavior, 23*, 21–40.

Meyer-Bahlburg, H. F. L. (2009). Variants of gender differentiation in somatic disorders of sex development: Recommendations for version 7 of the World Professional Association for Transgender Health's Standards of Care. *International Journal of Transgenderism, 11*, 226–237.

Meyer-Bahlburg, H. F. L. (2010a). From mental disorder to iatrogenic hypogonadism: Dilemmas in conceptualizing gender identity variants as psychiatric conditions. *Archives of Sexual Behavior, 39*, 461–476.

Meyer-Bahlburg, H. F. L. (2010b, February 21). Re: DSM-5 GID issues—some clarifications [Electronic mailing list message]. *Sexnet* listserv.

Stein, D. J., Phillips, K. A., Bolton, D., Fulford, K. W. M., Sadler, J. Z., & Kendler, K. S. (2010). What is a mental/psychiatric disorder? From DSM-IV to DSM-V. *Psychological Medicine*. doi:10.1017/S0033291709992261.

Wakefield, J. C., & First, M. B. (2003). Clarifying the distinction between disorder and nondisorder: Confronting the overdiagnosis (false positives) problem in DSM-V. In K. A. Phillips, M. B. First, & H. A. Pincus (Eds.), *Advancing DSM: Dilemmas in psychiatric diagnosis* (pp. 23–55). Arlington, VA: American Psychiatric Association.

World Health Organization. (1992). *International statistical classification of diseases and related health problems* (10th rev., Vol. 1). Geneva, Switzerland: Author.

Zucker, K. J. (2010). The DSM diagnostic criteria for gender identity disorder in children. *Archives of Sexual Behavior, 39*, 477–498.

Arch Sex Behav (2010) 39:1261–1273
DOI 10.1007/s10508-010-9618-z

ORIGINAL PAPER

Infants' Preferences for Toys, Colors, and Shapes: Sex Differences and Similarities

Vasanti Jadva · Melissa Hines · Susan Golombok

Received: 3 September 2008 / Revised: 16 February 2010 / Accepted: 18 February 2010 / Published online: 16 March 2010

Abstract Girls and boys differ in their preferences for toys such as dolls and trucks. These sex differences are present in infants, are seen in non-human primates, and relate, in part, to prenatal androgen exposure. This evidence of inborn influences on sex-typed toy preferences has led to suggestions that object features, such as the color or the shape of toys, may be of intrinsically different interest to males and females. We used a preferential looking task to examine preferences for different toys, colors, and shapes in 120 infants, ages 12, 18, or 24 months. Girls looked at dolls significantly more than boys did and boys looked at cars significantly more than girls did, irrespective of color, particularly when brightness was controlled. These outcomes did not vary with age. There were no significant sex differences in infants' preferences for different colors or shapes. Instead, both girls and boys preferred reddish colors over blue and rounded over angular shapes. These findings augment prior evidence of sex-typed toy preferences in infants, but suggest that color and shape do not determine these sex differences. In fact, the direction of influence could be the opposite. Girls may learn to prefer pink, for instance, because the toys that they enjoy playing with are often colored pink. Regarding within sex differences, as opposed to differences between boys and girls, both boys and girls preferred dolls to cars at age 12-months. The preference of young boys for dolls over cars suggests that older boys' avoidance of dolls may be acquired. Similarly, the sex similarities in infants' preferences for colors and shapes suggest that any subsequent sex differences in these preferences may arise from socialization or cognitive gender development rather than inborn factors.

Keywords Sex · Gender · Infants · Toy preference · Color preference · Shape preference

V. Jadva (✉) · S. Golombok
Centre for Family Research, Faculty of Social and Political Sciences, University of Cambridge, Cambridge CB2 3RF, UK
e-mail: vj227@cam.ac.uk

M. Hines
Faculty of Social and Political Sciences, University of Cambridge, Cambridge, UK

Introduction

Children show clear sex-typed toy preferences, with girls showing more interest than boys do in dolls and boys showing more interest than girls do in vehicles (Alexander & Hines, 1994; Pasterski et al., 2005; Serbin, Poulin-Dubois, Colburne, Sen, & Eichstedt, 2001). In addition to these differences between the sexes, within sex analyses show that boys play more with masculine toys, like vehicles and weapons, than with feminine toys, like dolls and tea sets. In contrast, although girls play more with feminine toys than with masculine toys when the feminine toys are sufficiently interesting (e.g., Hines & Alexander, 2008; Pasterski et al., 2005), they sometimes show no significant preference for feminine over masculine toys (e.g., Berenbaum & Hines, 1992; Servin, Nordenstrom, Larsson, & Bohlin, 2003). The strong male preference for same-sex toys has sometimes been described as boys avoiding girls' toys (Maccoby & Jacklin, 1974).

Sex-typed toy preferences have been seen in infants (Alexander, Wilcox, & Woods, 2009; Campbell, Shirley, Heywood, & Crook, 2000; Serbin et al., 2001; Snow, Jacklin, & Maccoby, 1983), grow larger as childhood progresses (Golombok & Hines, 2002), and have been reported into young adulthood (Alexander, 2006). These sex differences have been documented using an array of research methodologies, including inventories of children's toys at home, observation of children's toy contact in a playroom, parental interviews and questionnaires, and visual preferences and eye-tracking.

Perspectives on the acquisition of sex-typed play, including toy preferences, can be categorized broadly into social learning theories, cognitive theories, and hormonal theories. Social learning theories posit that children are socialized into different gender role behaviors, including toy play (Bandura, 1977; Mischel, 1966). Boys are reinforced for engaging with male-typical toys and girls for engaging with female-typical toys. Opposite sex-typed behavior is punished or not rewarded, which leads to extinction. Children can also learn which behaviors to adopt by modeling individuals of the same sex as themselves or by complying with labels identifying behaviors as appropriate for children of one sex or the other.

From a social learning perspective, infants' preferences for sex-typed toys would suggest that the differential treatment of boys and girls begins at an early age. In support of this view, Snow et al. (1983) found that fathers of 12-month-old infants were less likely to give dolls to their sons than to their daughters. In addition, Caldera and Sciaraffa (1998) observed parents and their 18- to 23-month-old infants playing with either a doll or a clown. They found that parents of boys called their sons' attention to the clown more than the doll and parents of girls called their daughters' attention to the doll more than the clown. Similarly, Pomerleau, Bolduc, Malcuit, and Cossette (1990) found that parents of infants aged 5–25 months created different home environments for boys and girls. Boys had more sports equipment, tools and vehicles, and girls had more dolls and fictional characters. Thus, socialization of very young infants may be occurring not only through parents' interactions with their sons and daughters, but also in the way in which they design their infants' home environments.

Cognitive theories include cognitive developmental theories (Kohlberg, 1966) and gender schema theories (Bem, 1981; Martin & Halverson, 1981; Martin, Ruble, & Szkrybalo, 2002). According to cognitive developmental theories, gender role acquisition involves three stages: gender labeling, gender stability, and gender constancy. It is at this last stage, where the child understands that gender remains the same across different situations, that sex-typed preferences were originally thought to emerge. More recently, researchers have suggested that gender constancy is not a prerequisite for gender-typed behavior and, indeed, young children show sex-typed preferences before gender constancy is attained (Ruble, Martin, & Berenbaum, 2006). Gender schema theorists posit that children develop gender schemas to organize and structure gender-related information from their environment. The process of gender typing is thought to begin once the child is able to categorize him/herself as belonging to a particular gender. For cognitive theorists, sex-typed behavior follows from a child knowing his or her own gender and becoming aware of the stereotypes that exist in the social environment.

From the hormonal perspective, sex differences arise, in part, from early hormonal differences between boys and girls (Hines, 2004). In particular, sex differences in the prenatal hormone environment are thought to produce differences in neural organization, such that high concentrations of androgens, hormones typically produced in large amounts by the male fetus, lead to brain masculinization and increased male-typical behavior. One approach to understanding the effects of sex hormones has been to study children with congenital adrenal hyperplasia (CAH), a genetic condition where the female fetus is exposed to abnormally high concentrations of androgens. These studies have shown that girls with CAH spend more time playing with masculine toys and less time playing with feminine toys compared to control group girls (Berenbaum & Hines, 1992; Meyer-Bahlburg et al., 2004; Nordenstrom, Servin, Bohlin, Larsson, & Wedell, 2002; Pasterski et al., 2005). The suggestion that this may result from parents encouraging male-typical toy play in girls with CAH (Quadagno, Briscoe, & Quadagno, 1977) has not been supported by research finding that parents encourage feminine toy play, not masculine toy play, more in their daughters with CAH than in their unaffected daughters (Pasterski et al., 2005). Normal variability in androgen exposure prenatally also relates to male-typical childhood behavior (Auyeung et al., 2009; Hines et al., 2002), suggesting that the findings for girls with CAH relate to their androgen exposure, not to other aspects of the disorder. Sex-typed toy preferences similar to those seen in children have also been reported in two species of non-human primates, vervet monkeys and rhesus monkeys (Alexander & Hines, 2002; Hassett, Siebert, & Wallen, 2008), providing additional evidence of some innate contribution.

Given the evidence that sex differences in toy preferences emerge early in life and appear to relate, in part, to hormonal or other inborn influences, some researchers have begun to ask what properties of sex-typed toys differentially attract boys and girls (Alexander & Hines, 2002; Campbell et al., 2000). For example, are boys attracted to wheels and motion, and girls to faces and imaginary role-play? Moller and Serbin (1996) argued that toy preferences result from what the toy can do, rather than from children knowing that a toy is appropriate for their own gender. Similarly, Alexander (2003) suggested that sex-typed toy preferences may result from a preference for different object features, including color, movement, or form.

Sex-typed Toys and Color

Toys for boys and girls tend to differ in many ways. One of the most obvious is color. Pennell (1994) found that girls' toys tended to be colored in pastel shades, especially pink and lavender, and boys' toys tended to be colored in intense colors, such as red, blue, and black. These colors are also differentially preferred by girls and boys. For example, Picariello, Greenberg, and Pillemer (1990) asked 3- to 7-year-old children to choose their favorite felt pig from a choice of pigs colored in either

stereotypically masculine colors (navy blue, brown, maroon) or stereotypically feminine colors (light pink, bright pink, lavender), and found that they were likely to choose a pig in a color stereotyped as for their own sex. Similarly, Chiu et al. (2006) asked girls and boys aged 3–12 years to choose their three favorite colors from a color chart, and found that boys preferred blue to pink/purple, and girls preferred pink/purple to blue. There also are sex differences in children's use of color in drawings (Iijima, Arisaka, Minamoto, & Arai, 2001). Girls use more "warm" colors, including pink, than boys, whereas boys use more "cold" colors, such as gray and blue, compared to girls. Sex-typed color preferences appear to persist into adulthood; Hurlbert and Ling (2007) examined the color preferences of adults using a forced choice color picking task and found females to prefer reddish purple and males to prefer blue-green.

Few studies have examined the color preferences of children below the age of 3 years, and none have looked at sex differences in infants' preferences. However, babies as young as 3 months can see color (Bornstein, 1985; Cohen, DeLoache, & Strauss, 1979; Franklin, Pitchford, Mahony, & Davies, 2006) and both male and female infants between the ages of 3 and 5 months appear to like red most and green least (Adams, 1987; Bornstein, 1975). It is not known, however, if infants display sex-typed color preferences similar to those of older children and adults or, if so, when these sex differences emerge.

Sex-typed Toys and Shape

In addition to color, sex-typed toys differ in their shape. For instance, cars and other vehicles tend to be angular, whereas dolls tend to be rounded. Although research has not examined preferences for different shaped toys per se, some studies have examined the content of drawings, finding sex differences in images produced by adults, as well as children. Franck and Rosen (1949) found that men tend to "close off" stimuli, to enlarge images (mainly by extending the image upwards), and to emphasize sharp or angular lines, while women tend to leave the stimulus areas "open", to elaborate the drawing within the confines of the presented lines and to blunt or round off any angular lines. Among children, girls are more likely than boys to draw flowers, butterflies, the sun, and human motifs, whereas boys are more likely than girls to draw mobile objects, such as vehicles, trains, aircraft, and rockets (Iijima et al., 2001).

The present study examined toy preferences, as well as color and shape preferences, in infants ages 12, 18, and 24 months. We evaluated the hypotheses that these young children show preferences for sex-typical toys and colors, for sex-typed toys in sex-typed colors, and for angular versus rounded shapes. Infants across a range of ages were studied in anticipation of determining not only infants' sex-typed preferences, but also the age at which any such preferences emerge.

Method

Participants

Parents of infants were contacted through nurseries and mother and baby groups in London, UK. Infants were recruited into three age categories: 12 months (for boys: $M = 54.51$ weeks, $SD = 5.20$; for girls: $M = 53.66$ weeks, $SD = 4.76$), 18 months (for boys: $M = 80.45$ weeks, $SD = 3.42$; for girls, $M = 81.26$ weeks, $SD = 5.24$), and 24 months (for boys: $M = 106.58$ weeks, $SD = 6.36$; for girls, $M = 105.85$ weeks, $SD = 5.28$). Each age category consisted of 20 boys and 20 girls. Most infants ($N = 116$) participated with their mothers; four infants participated with their fathers. Each parent–infant pair was paid £10 sterling (about $20) for taking part in the study.

The majority of mothers (72, 60%) and fathers (81, 67.5%) had a professional occupation, as defined by the modified version of the Registrar Generals classification (OPCS and Employment Department Group, 1991) and 94 (78.3%) mothers and 96 (80%) fathers held a university degree. Sixty-six (55%) of the mothers were not working at the time of study, 13 (10.8%) worked full-time, and 41 (34.2%) worked part-time. Ninety-eight (81.7%) infants were Caucasian according to the Commission for Racial Equality classification for ethnicity.

Measures

We used a preferential looking task, whereby two images were shown simultaneously to the infant in a darkened room. Each image in each stimulus pair was mounted in a square, colored in gray (hue = 160, saturation = 0, luminance = 202). The infant's face was recorded by videotape and later coded for the length of time that the infant looked at each image. The stimuli used for the preferential looking task were chosen to test specific hypotheses, and these stimuli, and the hypotheses they were chosen to assess, are described below.

Color Stimuli

Four pairs of stimuli were used to evaluate infants' preferences for colors on their own. These stimulus pairs examined the hypotheses that boys prefer blue and girls prefer pink, as well as that infants show these sex-typed color preferences when brightness is controlled. Two pairs of stimuli compared pink (hue = 234, saturation = 235, luminance = 191) and blue (hue = 146, saturation = 240, luminance = 115). To ensure that the color of the stimuli matched the shades of pink and blue of existing toys, two toys (a doll's dress and a building block) were scanned directly into the computer and their shades of pink and blue were recorded. Because pink and blue are made up of different brightness (luminance) levels, with pink being brighter than blue, and because differences in the brightness levels of colors have been

shown to modify infants' color preferences (Cohen et al., 1979), two additional stimulus pairs were used to control for brightness. The pink was matched for brightness with the blue to produce red (hue = 234, saturation = 235, luminance = 115), and the blue was matched for brightness with the pink to produce pale blue (hue = 146, saturation = 240, luminance = 191). Thus, there were four pairings: pink/blue; red/pale blue; pink/pale blue; and red/blue.

Toy Stimuli

Two sex-typed toys (a doll and a car) provided the toy stimulus pairings. Simple line drawings of a doll and a car were scanned into a computer to create the stimuli (Fig. 1). To allow assessment of relationships between toy and color, as well as toy preferences on their own, the car and the doll were colored in the same four colors used for the color stimuli (pink, blue, red, pale blue).

The stimuli were paired to examine specific hypotheses. To test the hypothesis that boys and girls prefer sex-typed toys in sex-typed colors, we compared the doll to the car when colored in sex congruent colors, i.e., pink doll/blue car. To test the hypothesis that the preference for sex-typed toys would be weaker when they are colored in cross sex-typed colors, we compared the doll to the car when colored in sex incongruent colors, i.e., blue doll/pink car. To examine the same hypotheses with brightness controlled, we paired the doll to the car when colored in sex congruent colors and sex incongruent colors controlling for the difference in brightness levels of pink and blue. As all possible color combinations were included, this resulted in four pairings: two pairings of toys colored in sex congruent colors (i.e., red doll/blue car and pink doll/pale blue car) and two parings of toys colored in sex incongruent colors (i.e., red car/blue doll and pale blue doll/pink car).

We also tested the hypothesis that boys and girls differ in their preference for the car and doll when both toys were of the same color or no color, by pairing the doll with the car of the same color, i.e., pink car/pink doll, blue doll/blue car, and by pairing a colorless car with a colorless doll. Finally, to test the hypothesis that boys and girls differ in preferences for the colors pink and blue, we paired pink to blue with the toy held constant: blue doll/pink doll and pink car/blue car.

Shape Stimuli

Three pairs of stimuli tested the hypothesis that boys and girls differ in their preferences for angular shapes versus rounded shapes: an angular triangle paired with a triangle with rounded edges (rounded triangle), an angular star paired with a star with rounded edges (rounded star), and an overlapping square and rectangle (rectangles) paired with an overlapping circle and oval (circles). The shapes were colored in white (Fig. 2.)

Procedure

Procedures were conducted at City University in London. On arrival, parents and infants were taken into a reception room where they were informed about the procedure for the study and parental consent was obtained. They were then taken into the laboratory where parents were asked to seat their infants in their laps. In front of them, at a distance of 2 m, was a large white screen onto which the prepared images were projected. Hidden behind the screen was a stand holding a video camera and speakers. Only the lens of the video camera, which protruded from a hole cut out of the screen, was visible from the front of the screen. Parents were advised not to direct their child to a particular stimulus, either verbally or physically. They were also told that they could stop the testing procedure at any time by getting up from their seat. The experimenter sat in the observation room, separated from the laboratory by a one-way mirror.

As in other preferential looking studies (Campbell et al., 2000; Serbin et al., 2001), two stimuli were presented simultaneously, one on either side of the child's central gaze. The stimuli measured 45 × 45 cm and were located approximately 45 cm apart when projected onto the screen. The experimenter waited for the child to have a central gaze before showing each pair of stimuli. The infant could also be encouraged to look centrally at the screen by projecting a red spot onto the central point of the

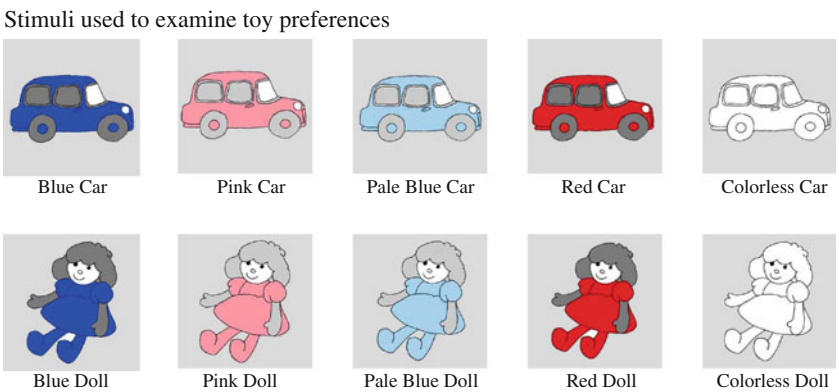


Fig. 1 Stimuli used to examine toy preferences

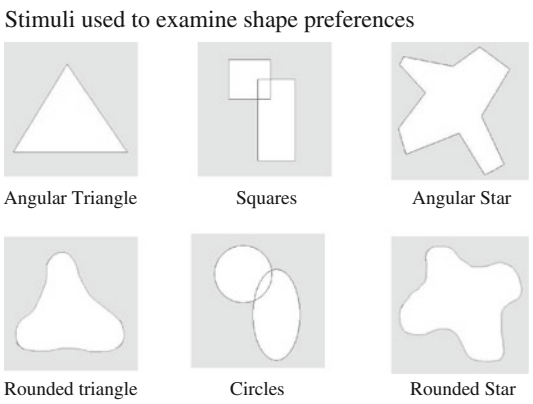


Fig. 2 Stimuli used to examine shape preferences

screen (used when the infant was looking in the direction of the screen) or by playing a sound (used when the infant was looking away from the screen area or was being especially fidgety). Generally, these devices were only required before the first pair of stimuli were presented.

The first sets of stimuli shown were the four pairs of color stimuli combined with the 11 pairs of toy stimuli. To ensure counterbalancing, each pairing was shown twice, with each stimulus within a pair appearing once on the left and once on the right side of the child's gaze. Thus, 30 pairs of stimuli were shown for 5 s each. The shape stimuli were shown after the color and toy stimuli. The three pairs of shape stimuli were counterbalanced producing six pairs shown for 5 s each. Order of presentation was randomized within each of the two groups of stimuli.

Data Analysis

Coding of the videotapes from the toy, color, and shape presentations was carried out by playing the tape on a VHS videorecorder and freezing the initial image. The frame advance function was then used to move the picture frame by frame. Data were coded directly onto a spreadsheet where it was noted whether the infant was looking at the left hand image, the right hand image or neither image during each frame. There were a total of 25 frames per second. As the images were counterbalanced by showing each stimulus per pair on each side of the child's central gaze, the total looking times for both images of a pair were added together. This meant that the final score ranged from 0 to 10 s (or 0 to 250 frames). To assess inter-rater reliability, a randomly selected sample of the videotapes was coded by two scorers. Pearson correlation coefficients for the pairings, calculated using the combined raw scores for each pair of stimuli, ranged from .80 to .99 with an average correlation of .95.

Data were lost for some subjects during the preferential looking task due to infant fussiness. As counterbalancing was achieved by showing the same pairings twice, some infants saw only one of the two versions of a pairing. When this happened, the data for that infant for both versions of that particular pairing were deleted. During the first session 107 of the 120 infants saw all the pairings. Fewer infants participated in the shape preference procedure due to fussiness, and 90 infants saw all the pairings for the shape preference procedure.

Some infants looked longer at the pairings than others. To adjust for these differences, scores were converted into the proportion of time spent looking at one stimulus over the total looking time for both stimuli. Proportions were transformed into percentage values; thus, an infant looking at a particular stimulus for 50% of the time meant that no preference was shown. All subsequent analyses were conducted using these percentage values.

Results

Mean proportions of time that infants looked at each of the color, toy, and shape stimulus pairings broken down by sex and age are shown in Table 1. Sex and age differences and their interaction were evaluated using analysis of variance for each of the pairings (See Table 2).

Color Stimuli

No main effects of sex were found for any of the four color pairings. A significant main effect of age was found only for the red/pale blue pairing, $F(2, 101) = 4.94$, $p < .01$ with 12-month-olds looking significantly longer at red compared to 24-month-olds ($p < .01$). There were no significant interaction effects. A composite score was computed to examine sex and age influences on preferences for pink/red versus blue/pale blue collapsed across all four pairings. There were no significant main effects of sex or age and no interaction between sex and age.

Toy Stimuli

Five of the 11 pairings designed to test specific hypotheses were significant, and, contrary to the expectation that sex-typed toys would be of most interest when of sex-typed colors, findings suggested that infants preferred looking at sex-typed toys whether or not they were of sex-typed colors, but only when the brightness of colors was matched (Table 1). In addition, means, even when not significant, were in the direction consistent with a preference for sex-typed toys, regardless of their color.

Shape Stimuli

No main effects of sex or age and no interactions were found for any of the three shape pairings comparing rounded shapes to angular shapes. An overall score, collapsed across all three pairings, also showed no main effects of sex or age and no interaction between sex and age.

Table 1 The mean proportion (%) of looking time for individual toy, color, and shape pairings by sex and by age

	Girls			Boys		
	12 months *M* % (*SD*)	18 months *M* % (*SD*)	24 months *M* % (*SD*)	12 months *M* % (*SD*)	18 months *M* % (*SD*)	24 months *M* % (*SD*)
Color						
Pairings comparing pink to blue						
Pink	45.8 (15.5)	50.1 (11.2)	46.9 (11.9)	47.6 (14.3)	46.3 (15.2)	52.5 (16.5)
Blue	54.2 (15.5)	49.9 (11.2)	53.1 (11.9)	52.4 (14.3)	53.7 (15.2)	47.5 (16.5)
Red	65.1 (14.2)	59.1 (16.0)	55.8 (16.9)	66.4 (15.2)	59.8 (15.5)	54.5 (12.5)
Pale blue	34.9 (14.2)	40.9 (16.0)	44.2 (16.9)	33.6 (15.2)	40.2 (15.5)	45.5 (12.5)
Pairings comparing pink to blue with brightness controlled						
Pink	53.7 (19.3)	46.7 (17.9)	53.7 (11.9)	61.4 (20.2)	49.1 (22.8)	53.1 (15.8)
Pale blue	46.3 (19.3)	53.3 (17.9)	46.3 (11.9)	38.6 (20.2)	50.9 (22.8)	46.9 (15.8)
Red	55.7 (23.5)	54.1 (11.6)	60.0 (17.1)	54.2 (15.2)	58.1 (12.3)	51.2 (11.0)
Blue	44.3 (23.5)	45.9 (11.6)	40.0 (17.1)	45.8 (15.2)	41.9 (12.3)	48.8 (11.0)
Toys						
Pairings comparing doll to car of sex-congruent and sex-incongruent colors						
Pink doll	55.2 (20.4)	51.0 (17.2)	44.5 (20.1)	49.5 (12.8)	47.0 (20.8)	43.7 (17.3)
Blue car	44.8 (20.4)	49.0 (17.2)	55.5 (20.1)	50.5 (12.8)	53.0 (20.7)	56.3 (17.3)
Blue doll	65.2 (17.7)	54.5 (21.7)	53.4 (16.3)	60.6 (21.2)	45.5 (20.9)	47.8 (17.5)
Pink car	34.8 (17.7)	45.5 (21.7)	46.6 (16.3)	39.4 (21.2)	54.5 (20.9)	52.2 (17.5)
Pairings comparing doll to car of sex-congruent and sex-incongruent colors (brightness controlled)						
Red doll	66.3 (14.3)	57.8 (15.6)	61.3 (20.2)	61.9 (15.4)	44.1 (19.6)	47.1 (17.1)
Blue car	33.7 (14.3)	42.2 (15.6)	38.7 (20.2)	38.1 (15.4)	55.9 (19.6)	52.9 (17.1)
Blue doll	52.5 (15.9)	56.5 (11.7)	49.3 (14.8)	53.6 (14.3)	44.5 (20.2)	39.5 (12.7)
Red car	47.5 (15.9)	43.5 (11.7)	50.7 (14.8)	46.4 (14.3)	55.5 (20.2)	60.5 (12.7)
Pink doll	61.6 (18.7)	60.5 (18.9)	55.9 (14.2)	63.9 (16.8)	43.6 (15.2)	49.3 (11.4)
Pale blue car	38.4 (18.7)	39.5 (18.9)	44.1 (14.2)	36.1 (16.8)	56.4 (15.2)	50.7 (11.4)
Pale blue doll	56.7 (15.7)	57.8 (13.9)	54.4 (16.8)	54.7 (24.1)	47.1 (20.2)	43.5 (14.1)
Pink car	43.3 (15.7)	42.2 (13.9)	45.6 (16.8)	45.3 (24.1)	52.9 (20.2)	56.5 (14.1)
Pairings comparing doll to car with color held constant						
Pink doll	57.7 (16.5)	54.1 (16.2)	51.1 (21.3)	58.3 (16.7)	53.9 (25.8)	48.8 (13.5)
Pink car	42.3 (16.5)	45.9 (16.2)	48.9 (21.3)	41.7 (16.7)	46.1 (25.8)	51.2 (13.5)
Blue doll	56.3 (23.3)	54.3 (11.8)	55.9 (13.8)	56.9 (13.4)	45.5 (21.7)	48.7 (17.3)
Blue car	43.7 (23.3)	45.7 (11.8)	44.1 (13.8)	43.1 (13.4)	54.5 (21.7)	51.3 (17.3)
Neutral doll	57.3 (14.0)	50.4 (11.3)	50.7 (15.9)	58.8 (17.4)	44.5 (14.9)	45.5 (11.7)
Neutral car	42.7 (14.0)	49.6 (11.3)	49.3 (15.9)	41.2 (17.4)	55.5 (14.9)	54.5 (11.8)
Pairings comparing pink to blue with toy held constant						
Pink doll	45.2 (18.8)	56.9 (16.7)	48.6 (9.6)	48.1 (13.5)	41.8 (12.9)	44.0 (9.4)
Blue doll	54.8 (18.8)	43.1 (16.7)	51.4 (9.6)	51.9 (13.5)	58.2 (12.9)	56.0 (9.4)
Pink car	49.9 (17.3)	54.0 (16.5)	50.6 (11.6)	49.4 (18.0)	47.4 (13.5)	46.0 (10.0)
Blue car	50.1 (17.3)	46.0 (16.5)	49.4 (11.6)	50.6 (18.0)	52.6 (13.5)	54.0 (10.0)

Table 1 The mean proportion (%) of looking time for individual toy, color, and shape pairings by sex and by age

	Girls			Boys		
	12 months *M* % (*SD*)	18 months *M* % (*SD*)	24 months *M* % (*SD*)	12 months *M* % (*SD*)	18 months *M* % (*SD*)	24 months *M* % (*SD*)
Shape						
Pairings comparing rounded to angular images						
Circles	63.5 (22.8)	54.1 (19.1)	58.5 (13.2)	60.3 (29.1)	64.3 (9.9)	58.1 (15.4)
Squares	36.5 (22.8)	45.9 (19.1)	41.5 (13.2)	39.7 (29.1)	35.7 (9.9)	41.9 (15.4)
Rounded star	51.0 (14.9)	52.4 (11.6)	47.6 (10.8)	57.6 (15.4)	49.0 (15.2)	46.3 (19.2)
Angular Star	49.0 (14.9)	47.6 (11.6)	52.4 (10.8)	42.4 (15.4)	51.0 (15.2)	53.7 (19.2)
Rounded triangle	58.5 (17.8)	52.3 (15.4)	53.9 (15.2)	66.8 (18.0)	60.4 (22.9)	51.4 (17.4)
Triangle	41.5 (17.8)	47.7 (15.4)	46.1 (15.2)	33.2 (18.0)	39.6 (22.9)	48.6 (17.4)

Composite Stimuli

We next combined stimulus pairs to provide more reliable, composite estimates of children's preferences for sex-typed toys and sex-typed colors. Preferences for sex-typed toys were assessed by computing infants' average scores for looking at the doll versus the car, irrespective of color, across all pairings. Preferences for sex-typed colors were assessed by computing infants' average scores for looking at pink/red versus blue/pale blue, across all color pairings and, irrespective of the toy, across all toy pairings. Results are summarized in Tables 3 and 4.

The combined analysis of toy type revealed a main effect of sex. Girls looked longer at the doll than boys did, and boys looked longer at the car than girls did, $F(1, 101) = 7.68$, $p < .01$. There also was a main effect of age, $F(2, 101) = 6.84$, $p < .01$. Infants looked significantly longer at the doll at 12 months of age than at either 18 months ($p < .01$) or 24 months ($p < .01$), but 18- and 24-month-olds did not differ. There was no significant interaction between sex and age. The combined analysis of color preferences across all toys and color pairings showed no significant main effects of sex or age and no significant interaction.

Using the composite scores, we also looked at within sex preferences for same sex-typed toys over other sex-typed toys (i.e., boys' preferences for cars over dolls and girls' preferences for dolls over cars) in each age group (Fig. 3). Girls showed a significant preference for the doll over the car at ages 12 months, $t(18) = 4.99$, $p < .001$, and 18 months, $t(16) = 3.40$, $p < .01$, but this difference, though in the same direction, was not statistically significant at 24 months. Boys also showed a significant preference for the doll over the car at 12 months, $t(16) = 3.75$, $p < .01$. At ages 18 and 24 months, boys no longer showed a preference for the doll, and, although they looked longer at the car than the doll at these later ages, their preference for the car was not statistically significant.

Finally, at the suggestion of a reviewer, we analyzed difference scores, obtained by subtracting percentage looking time at the colorless doll and car from percentage looking time at the same stimuli when colored. These analyses also suggested that infants did not show sex typed color preferences. There were no significant main or interaction effects for the composite color difference scores, and only one main effect and no interactions for any of the individual pairings.

Infant Preferences Regardless of Sex and Age

Because no sex or age differences emerged for the color or shape stimuli, we examined color and shape preferences, irrespective of sex and age (see Table 5). For the color stimuli, infants looked longer at red than blue and longer at red than pale blue. For the shape stimuli, infants looked longer at circles than squares and longer at rounded triangles than triangles. For the four color stimuli combined, infants looked longer at pink/red than blue/pale blue, $t(112) = 5.67$, $p < .001$. For the three shape stimuli combined, infants looked longer at rounded images than angular images, $t(88) = 5.14$, $p < .001$.

Discussion

Our results found both sex differences and sex similarities in infants' toy, color, and shape preferences. We saw the expected sex differences in toy preferences, with girls showing more interest than boys in dolls, and boys showing more interest than girls in cars. These results did not interact with age. The differences were most apparent in stimulus pairings when colors were controlled for brightness. Contrary to prediction, however, sex-typed toy preferences were not stronger when toys were of sex-typed colors. In addition, infants did not show the predicted sex differences in color or shape preferences. Instead, we saw sex similarities in these areas. Both boys and girls preferred reddish colors to blue colors, and rounded shapes to angular shapes. There was also an age effect for interest in the doll. Both boys and girls looked longer at the doll at age 12 months, than at 18 or 24 months.

Table 2 Analysis of variance for color, toy, and shape stimuli

Source	df	*F*	*p*
Color			
Pairings comparing pink to blue			
Pink vs. Blue			
Sex	1	<1	ns
Age	2	<1	ns
Sex × Age	2	1.06	ns
Error	110		
Red vs. Pale blue			
Sex	1	<1	ns
Age	2	4.94	<.01
Sex × Age	2	<1	ns
Error	111		
Pairings comparing pink to blue with brightness controlled			
Pink vs. Pale blue			
Sex	1	<1	ns
Age	2	2.68	.07
Sex × Age	2	<1	ns
Error	109		
Red vs. Blue			
Sex	1	<1	ns
Age	2	<1	ns
Sex × Age	2	1.63	ns
Error	111		
Toys			
Pairings comparing doll to car of sex-congruent and sex-incongruent colors			
Pink doll vs. Blue car			
Sex	1	1.05	ns
Age	2	1.90	ns
Sex × Age	2	<1	ns
Error	109		
Blue doll vs. Pink car			
Sex	1	3.15	.08
Age	2	5.50	<.01
Sex × Age	2	<1	ns
Error	109		
Pairings comparing doll to car of sex-congruent and sex-incongruent colors (brightness controlled)			
Red doll vs. Blue car			
Sex	1	11.38	<.01
Age	2	6.23	<.01
Sex × Age	2	1.01	ns
Error	111		
Blue doll vs. Red car			
Sex	1	5.89	.02
Age	2	3.31	.04
Sex × Age	2	2.05	ns
Error	110		
Pink doll vs. Pale blue car			
Sex	1	5.75	.02
Age	2	5.44	<.01
Sex × Age	2	3.43	.04
Error	110		
Pale blue doll vs. Pink car			
Sex	1	5.47	.02
Age	2	1.33	ns
Sex × Age	2	<1	ns
Error	108		
Pairings comparing doll to car with color held constant			
Pink doll vs. Pink car			
Sex	1	<1	ns
Age	2	1.82	ns
Sex × Age	2	<1	ns
Error	112		
Blue doll vs. Blue car			
Sex	1	2.46	ns
Age	2	1.46	ns
Sex × Age	2	<1	ns
Error	109		
Neutral doll vs. Neutral car			
Sex	1	1.45	ns
Age	2	6.76	<.01
Sex × Age	2	<1	ns
Error	111		
Pairings comparing pink to blue with toy held constant			
Pink doll vs. Blue doll			
Sex	1	4.80	.03
Age	2	<1	ns
Sex × Age	2	4.17	.02
Error	112		
Pink car vs. Blue car			
Sex	1	2.01	ns
Age	2	<1	ns
Sex × Age	2	<1	ns
Error	110		
Shape			
Pairings comparing rounded to angular images			
Circles vs. Squares			
Sex	1	<1	ns
Age	2	<1	ns
Sex × Age	2	1.01	ns
Error	85		
Rounded star vs. Angular star			
Sex	1	<1	ns
Age	2	1.86	ns

Table 2 continued

Source	df	*F*	*p*
Sex × Age	2	<1	ns
Error	87		
Rounded triangle vs. Triangle			
Sex	1	1.55	ns
Age	2	2.42	.10
Sex × Age	2	<1	ns
Error	111		

Controlling the brightness of colors was a novel aspect of the current study and, given that sex differences in toy preferences were most obvious when brightness was controlled, this could be a useful design feature for future studies. Controlling brightness may be particularly important in studies such as ours, which present images in a darkened room, allowing the brightness of a color, as well as its hue or other characteristics to influence its attractiveness.

Our observation that 12- to 24-month-old boys show more interest than girls do in cars, and that girls of this age show more interest than boys do in dolls, resemble observations of sex differences in toy preferences in older children, and add to evidence that these sex differences emerge at a very young age. Such early sex differences could reflect inborn tendencies for girls and boys to prefer different toys. This interpretation is consistent with findings linking prenatal androgen exposure to toy preferences in children (Hines, 2004) and with findings of similar sex differences in toy preferences in non-human primates (Alexander & Hines, 2002; Hassett et al., 2008). Additionally, early socialization could contribute to sex differences in infants, since they have already been provided with sex-typed toys (Pomerleau et al., 1990). Thus, their looking preferences may reflect the type of toys that they have been exposed to in their environment. This interpretation would suggest that children learn sex-typed behaviors at a very young age.

Cognitive developmental processes related to gender are not likely to explain sex-typed toy preferences in 12- to 24-month-old infants. At this age, many infants would not have reached even the first stage of gender acquisition (gender labeling). In addition, although there is evidence that female infants may display some understanding of gender by the age of 18 months, this is apparently not the case for boys (Serbin et al., 2001). Serbin et al. concluded that the role of gender identity in the acquisition of gender role learning needs to be re-evaluated, because toy preferences are found in male infants, even though they do not appear to be aware of their gender identity. Our findings also argue for reconsidering the role of cognitive understanding of gender, at least in the initial phase of children's acquisition of sex-typed toy preferences. Cognitive factors may play a role in later years, however, as sex-typed toy preferences become increasingly evident (Golombok et al., 2008).

We did not see sex differences in preferences for pink or reddish colors over blue, nor did we see sex differences in preferences for angular versus rounded shapes. Therefore, our findings did not support Alexander's (2003) suggestion that differences in color or shape preferences explain sex differences in toy preferences, at least at this early stage of development. Indeed, the causal relationships may be the opposite. Sex differences in toy preferences may contribute to sex differences in preferences for colors or shapes. For example, girls may learn to like pink because many of the toys they play with are pink. Alternatively, or additionally, they may learn this color preference through social or cognitive mechanisms. For example, girls may learn to prefer pink through modeling older girls who like pink, or through cultural labeling of pink as for girls. Similar mechanisms could explain sex differences in shape preferences. In addition to suggesting that the different colors of sex-typed toys could drive boys and girls differential interest in them, Alexander (2003) has suggested that females and males may have evolved to prefer pink and blue, respectively, a suggestion that has been reiterated by others (e.g., Hurlbert & Ling, 2007). Our findings argue against these suggestions as well.

Our observation that boys at 12 months of age, like girls, prefer the doll to the car is similar to that of Serbin et al. (2001), who found that children of both sexes prefer to look at dolls over trucks at 12 months of age. Both findings argue against suggestions that boys' strong preference for masculine toys or avoidance of feminine toys, such as dolls, is inborn (Hassett et al.,

Table 3 The mean proportion (%) of composite scores by sex and by age

	Girls			Boys		
	12 months *M* % (*SD*)	18 months *M* % (*SD*)	24 months *M* % (*SD*)	12 months *M* % (*SD*)	18 months *M* % (*SD*)	24 months *M* % (*SD*)
Pairings comparing dolls to cars combined across all stimulus pairs						
Doll	57.2 (6.3)	54.6 (5.5)	52.7 (9.3)	56.4 (7.0)	46.8 (12.6)	47.9 (6.5)
Car	42.8 (6.3)	45.4 (5.5)	47.3 (9.3)	43.6 (7.0)	53.2 (12.6)	52.1 (6.5)
Pairings comparing pink/red to blue/pale blue combined across all stimulus pairs						
Pink/red	53.5 (3.6)	51.7 (4.3)	51.9 (4.6)	53.3 (4.9)	50.7 (4.0)	52.1 (4.4)
Blue/pale blue	46.5 (3.6)	48.3 (4.3)	48.1 (4.6)	46.7 (4.9)	49.3 (4.0)	47.5 (4.2)

Table 4 Analysis of variance for composite scores

Source	df	*F*	*p*
Pairings comparing dolls to cars combined across all stimulus pairs			
Dolls vs. Cars			
Sex	1	7.68	<.01
Age	2	6.84	<.01
Sex × Age	2	1.57	ns
Error	101		
Pairings comparing pink/red to blue/pale blue combined across all stimulus pairs			
Pink/red vs. Blue/pale blue			
Sex	1	<1	ns
Age	2	2.31	.11
Sex × Age	2	<1	ns
Error	101		

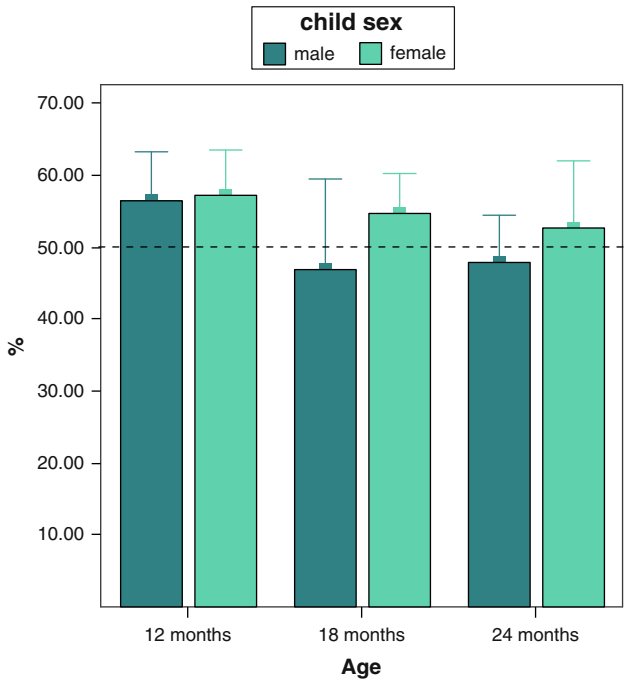


Fig. 3 Infants overall looking times at the doll when shown with the car (collapsed across all pairings) by sex and age

Table 5 Mean proportion (%) of time spent looking at stimuli for all infants irrespective of sex and age for color and shape stimuli

Stimuli pairings	*N*	%	*SD*	*t*	*p*
Color					
Pairings comparing pink to blue					
Blue	116	51.85	14.14	1.41	ns
Pink		48.15	14.14		
Red	117	60.15	15.39	7.14	<.001
Pale blue		39.84	15.39		
Pairings comparing pink to blue with brightness controlled					
Red	117	55.59	15.72	3.85	<.001
Blue		44.41	15.72		
Pink		52.97	18.52	1.72	ns
Pale blue	115	47.03	18.52		
Shape					
Pairings comparing rounded to angular images					
Circles	91	59.63	19.29	4.76	<.001
Squares		40.38	19.29		
Rounded star	93	50.84	14.64	<1	ns
Angular star		49.15	14.64		
Rounded triangle	93	57.44	18.24	3.93	<.001
Triangle		42.57	18.24		

2008; Williams & Pleil, 2008), and argue instead for the importance of social learning or cognitive developmental processes in the development of this particular aspect of sex-typed toy preferences. Consistent with this argument, boys avoidance of feminine toys has been found to increase with age, and to be stronger when an observer is present (Hartup, Moore, & Sager, 1963). Boys also receive stronger reinforcement than girls do to avoid cross sex toy play (Fagot, 1977; Lytton & Romney, 1991; Pasterski et al., 2005), and they are more likely than girls are to imitate the behavior of same sex models (Perry & Bussey, 1979). Thus, reinforcement and modeling could play an important role in boys' eventual strong preference for masculine toys or avoidance of feminine toys.

Instead of providing evidence of sex differences in infants' visual preference for pink and blue, our findings suggest that infants prefer red, irrespective of their sex. Other studies also have reported that infants from as young as 2 months of age look longer at red than at other colors (Adams, 1987; Bornstein, 1975; Franklin et al., 2006). Similarly, both girls and boys, ages 2.5–5 years, have been found to indicate that their favorite colors are red and pink (Zentner, 2001). The absence of sex differences in infants' and young children's color preferences, coupled with findings that older children display sex-typed color choices (Chiu et al., 2006; Picariello et al., 1990), suggests that children learn these preferences. The timing of the emergence of sex-typed color preference (after age two, or maybe even five, years) is also

consistent with cognitive developmental perspectives, which suggest that sex differences in children's behavior emerge as children develop a cognitive understanding of their gender and its stability and constancy, a process that continues after the age of two until as late as age seven years or older (Ruble et al., 2007).

Our results also suggest sex similarity rather than difference in infants' shape preferences; irrespective of sex, infants looked longer at rounded shapes (circles, rounded triangles) than at angular shapes (squares, triangles). The preference for rounded over angular shapes could relate to the emotional responses that different shapes elicit. A study asking college students to rate their emotional response to stimuli consisting of either an ellipse or a straight line found that roundedness conveyed warmth and acute angles conveyed threat (Aronoff, Woike, & Hyman, 1992). Bar and Neta (2006) showed adults a series of stimuli consisting of sharp angled images or curved images, and found that they liked the curved objects and disliked the sharp objects. It was suggested that sharp angles convey a sense of threat which results in a negative bias. It also has been suggested that the visual properties of angularity could reflect the facial attributes of an angry face and roundedness could reflect the facial attributes of a happy face (Aronoff, Barclay, & Stevenson, 1988). Further studies could examine if these associations that have been found in adults are also seen in children.

The question of male and female shape preferences in general also would benefit from additional research. As noted above, Iijima et al. (2001) reported that boys tend to draw angular shapes and girls tend to draw rounded shapes, and Franck and Rosen (1949) reported similar findings for the drawings of adults, with men tending to sharp, angular lines and women to rounded lines. Both these studies looked at shapes that people draw and not what they prefer if given a choice, however. Studies involving choice have not produced clear sex differences for shape preferences. For instance, Munroe, Munroe, and Lansky (1976) presented children with two containers holding either spherical or cubed sweets, and found that although both sexes chose the spherical sweets more than the cubed sweets, this preference was only significant for girls.

In addition to seeing unexpected sex similarities in the color and shape preferences of infants, we saw an unpredicted effect of age. Regardless of sex, infants looked longer at the doll at age 12 months than at later ages. The interest of 12-month-old infants of both sexes in dolls was also noted by Serbin et al. (2001), who suggested that it might relate to infants' interest in faces (Morton & Johnson, 1991). If so, our results suggest that this interest is more pronounced in younger infants than in older infants.

The preferential looking paradigm is a widely used method for assessing the preferences of pre-verbal infants. However, it assumes that infants will look longer at images that they prefer. Future studies using observational methods could see if similar sex differences to those we report using preferential looking are found for play with actual toys, colors and shapes. Also, it could be argued that larger samples might have produced evidence of sex differences in color or shape preferences. However, our study used a large number of infants than other studies of this type ($n = 30$ in Alexander et al., 2009, $n = 60$ in Campbell et al., 2000, and $n = 77$ in Serbin et al., 2001), and provided 83% and 76% power to detect moderate-sized or larger sex differences in color and shape preferences, respectively. Effect sizes also suggest that if any sex differences exist in these areas, they are small to negligible, and, for shape, in the direction opposite to that predicted ($d = 0.06$ for color preferences, and $d = -34$ for shape preferences). Finally, although we suggest that the absence of certain sex differences (e.g., sex-typed color preferences) in 12- to 24-month-old infants argues that inborn factors do not play a role in their emergence, we can not rule out the possibility that inborn factors, such as prenatal androgen exposure, contribute to sex differences that appear only later in life. Although it seems more likely that socialization or cognitive developmental processes are the primary cause of differences that are present only in older children, additional research, looking, for example, at color preferences in individuals exposed to high levels of androgens prenatally or in non-human primates, could help resolve this issue.

The current study adds to growing evidence that infants younger than 2 years of age display sex-typed toy preferences, with boys showing more interest than girls do in cars, and girls showing more interest than boys do in dolls. Within sex analyses found that the female preference for dolls over cars begins as early as 12 months of age, whereas boys of this age also prefer dolls to cars. The male preference for cars over dolls, or avoidance of dolls, emerges later, suggesting that socialization or cognitive development, rather than inborn factors, causes the male avoidance of feminine toys. Similarly the lack of sex differences in color or shape preferences in infants suggests that sex differences in these areas emerge later, perhaps also under the influence of socialization or cognitive developmental processes. In addition to seeing sex differences in infants, we also observed sex similarities. Infants of both sexes preferred reddish colors to blue and rounded shapes to angular shapes. One implication of our findings is that sex differences in toy preferences in infancy are not driven by sex-linked preferences for different colors or shapes, since sex differences in these areas are not yet present. Instead, the direction of influence could be the opposite. Girls may learn to prefer pink, for instance, because the toys that they enjoy playing with are often colored pink. Finally, our results suggest that different types of factors influence different aspects of children's sex-typed preferences. Inborn factors, such as the prenatal testosterone surge in male fetuses, may be particularly important for boys' greater interest than girls in vehicles and girls' greater interest than boys in dolls. In contrast, sex-typed color and shape preferences, and the male avoidance of girls' toys, which appear to emerge later in life, may depend more extensively on sex-related differences in socialization or cognitive developmental processes.

References

Adams, R. J. (1987). An evaluation of color preference in early infancy. *Infant Behavior and Development, 10*, 143–150.

Alexander, G. M. (2003). An evolutionary perspective of sex-typed toy preferences: Pink, blue and the brain. *Archives of Sexual Behavior, 32*, 7–14.

Alexander, G. M. (2006). Associations among gender-linked toy preferences, spatial ability, and digit ratios: Evidence from an eye-tracking analysis. *Archives of Sexual Behavior, 35*, 699–709.

Alexander, G. M., & Hines, M. (1994). Gender labels and play styles: Their relative contribution to children's selection of playmates. *Child Development, 65*, 869–879.

Alexander, G. M., & Hines, M. (2002). Sex differences in response to children's toys in nonhuman primates (*Cercopithecus aethiops sabaeus*). *Evolution and Human Behavior, 23*, 467–479.

Alexander, G. M., Wilcox, R., & Woods, R. (2009). Sex differences in infants' visual interest in toys. *Archives of Sexual Behavior, 38*, 427–433.

Aronoff, J., Barclay, A. M., & Stevenson, L. A. (1988). The recognition of threatening facial stimuli. *Journal of Personality and Social Psychology, 54*, 647–655.

Aronoff, J., Woike, B. A., & Hyman, L. M. (1992). Which are the stimuli in facial displays of anger and happiness? Configurational bases of emotion recognition. *Journal of Personality and Social Psychology, 62*, 1050–1066.

Auyeung, B., Baron-Cohen, S., Ashwin, E., Knickmeyer, R., Taylor, K., Hackett, G., & Hines, M. (2009). Fetal testosterone predicts sexually differentiated childhood behavior in girls and in boys. *Psychological Science, 20*, 144–148.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Bar, M., & Neta, M. (2006). Humans prefer curved visual objects. *Psychological Science, 17*, 645–648.

Bem, S. L. (1981). Gender-schema theory: A cognitive account of sex-typing. *Psychological Review, 88*, 354–364.

Berenbaum, S. A., & Hines, M. (1992). Early androgens are related to childhood sex-typed toy preferences. *Psychological Science, 3*, 203–206.

Bornstein, M. H. (1975). Qualities of color vision in infancy. *Journal of Experimental Child Psychology, 19*, 401–419.

Bornstein, M. H. (1985). Human infant color-vision and color perception. *Infant Behavior and Development, 8*, 109–113.

Caldera, Y. M., & Sciaraffa, M. A. (1998). Parent-toddler play with feminine toys: Are all dolls the same? *Sex Roles, 3*, 657–668.

Campbell, A., Shirley, L., Heywood, C., & Crook, C. (2000). Infants' visual preference for sex-congruent babies, children, toys and activities: A longitudinal study. *British Journal of Developmental Psychology, 18*, 479–498.

Chiu, S. W., Gervan, S., Fairbrother, C., Johnson, L. L., Owen-Anderson, A. F. H., Bradley, S. J., & Zucker, K. J. (2006). Sex-dimorphic color preference in children with gender identity disorder: A comparison to clinical and community controls. *Sex Roles, 55*, 385–395.

Cohen, L. B., DeLoache, J. S., & Strauss, M. S. (1979). Infant visual perception. In J. D. Osofsky (Ed.), *Handbook of infant development* (pp. 393–438). New York: John Wiley and Sons.

Fagot, B. I. (1977). Consequences of moderate cross-gender behavior in preschool children. *Child Development, 48*, 902–907.

Franck, K., & Rosen, E. (1949). A projective test of masculinity-femininity. *Journal of Consulting Psychology, 13*, 247–256.

Franklin, A., Pitchford, N., Mahony, L., & Davies, I. R. (2006, June). *Color preferences in infancy*. Paper presented at the annual meeting of the XVth Biennial International Conference on Infant Studies, Westin Miyako, Kyoto, Japan.

Golombok, S., & Hines, M. (2002). Sex differences in social behavior. In P. K. Smith & C. H. Hart (Eds.), *Blackwell handbook of childhood social development* (pp. 117–136). Malden, MA: Blackwell Publishing.

Golombok, S., Rust, J., Zervoulis, K., Croudace, T., Golding, J., & Hines, M. (2008). Developmental trajectories of sex-typed behavior in boys and girls: A longitudinal general population study of children aged 2.5–8 years. *Child Development, 79*, 1583–1593.

Hartup, W. W., Moore, S. G., & Sager, G. (1963). Avoidance of inappropriate sex-typing by young children. *Journal of Consulting Psychology, 27*, 467–473.

Hassett, J. M., Siebert, E. R., & Wallen, K. (2008). Sex differences in rhesus monkey toy preferences parallel those of children. *Hormones and Behavior, 54*, 359–364.

Hines, M. (2004). *Brain gender*. New York: Oxford University Press.

Hines, M., & Alexander, G. M. (2008). Monkeys, girls, boys and toys: A confirmation [Letter to the editor]. *Hormones and Behavior, 54*, 478–479.

Hines, M., Golombok, S., Rust, J., Johnston, K. J., Golding, J., & The ALSPAC Study Team. (2002). Testosterone during pregnancy and gender role behavior of preschool children: A longitudinal, population study. *Child Development, 73*, 1678–1687.

Hurlbert, A. C., & Ling, Y. (2007). Biological components of sex differences in color preference. *Current Biology, 17*, 623–625.

Iijima, M., Arisaka, O., Minamoto, F., & Arai, Y. (2001). Sex differences in children's free drawings: A study of girls with congenital adrenal hyperplasia. *Hormones and Behaviour, 40*, 99–104.

Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In E. E. Maccoby (Ed.), *The development of sex differences* (pp. 82–173). Stanford, CA: Stanford University Press.

Lytton, H., & Romney, D. M. (1991). Parents' differential socialization of boys and girls: A meta-analysis. *Psychological Bulletin, 109*, 267–296.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Martin, C. L., & Halverson, C. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development, 52*, 1119–1134.

Martin, C. L., Ruble, D. N., & Szkrybalo, J. (2002). Cognitive theories of early gender development. *Psychological Bulletin, 128*, 903–933.

Meyer-Bahlburg, H. F. L., Dolezal, C., Baker, S. W., Carlson, A. D., Obeid, J. S., & New, M. I. (2004). Prenatal androgenization affects gender-related behavior but not gender identity in 5-12-year-old girls with congenital adrenal hyperplasia. *Archives of Sexual Behavior, 33*, 97–104.

Mischel, W. (1966). A social learning view of sex differences in behaviour. In E. E. Maccoby (Ed.), *The development of sex differences* (pp. 58–81). Stanford, CA: Stanford University Press.

Moller, L. C., & Serbin, L. A. (1996). Antecedents of toddler gender segregation: Cognitive consonance, gender-typed toy preferences and behavioral compatibility. *Sex Roles, 35*, 445–460.

Morton, J., & Johnson, M. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review, 98*, 164–181.

Munroe, R. H., Munroe, R. L., & Lansky, L. M. (1976). A sex difference in shape preference. *Journal of Social Psychology, 98*, 139–140.

Nordenstrom, A., Servin, A., Bohlin, G., Larsson, A., & Wedell, A. (2002). Sex-typed toy play behavior correlates with degree of prenatal androgen exposure assessed by CYP21 genotype in girls with congenital adrenal hyperplasia. *Journal of Clinical Endocrinology and Metabolism, 87*, 5119–5124.

OPCS and Employment Department Group. (1991). *Standard classification of occupations*. London: Her Majesty's Stationery Office.

Pasterski, V. L., Geffner, M. E., Brain, C., Hindmarsh, P., Brook, C., & Hines, M. (2005). Prenatal hormones and postnatal socialization by parents as determinants of male-typical toy play in girls with congenital adrenal hyperplasia. *Child Development, 76*, 264–278.

Pennell, G. E. (1994). Babes in toyland: Learning an ideology of gender. *Advances in Consumer Research, 21*, 359–364.

Perry, L. C., & Bussey, K. (1979). The social learning theory of sex difference: Imitation is alive and well. *Journal of Personality and Social Psychology, 37*, 1699–1712.

Picariello, M. L., Greenberg, D. N., & Pillemer, D. (1990). Children's sex related stereotyping of colors. *Child Development, 61*, 1453–1460.

Pomerleau, A., Bolduc, D., Malcuit, G., & Cossette, L. (1990). Pink or blue: Environmental gender stereotypes in the first 2 years of life. *Sex Roles, 22*, 359–367.

Quadagno, D. M., Briscoe, R., & Quadagno, J. S. (1977). Effects of perinatal gonadal hormones on selected nonsexual behavior patterns: A critical assessment of the nonhuman and human literature. *Psychological Bulletin, 84*, 62–80.

Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (2006). Gender development. In N. Eisenberg (Ed.), *Handbook of child psychology* (Vol. 3, pp. 858–932). New York: Wiley.

Ruble, D. N., Taylor, L. J., Cyphers, L., Greulich, F. K., Lurye, L. E., & Shrout, P. E. (2007). The role of gender constancy in early gender development. *Child Development, 78*, 1121–1136.

Serbin, L. A., Poulin-Dubois, D., Colburne, K. A., Sen, M. G., & Eichstedt, J. A. (2001). Gender stereotyping in infancy: Visual preferences for and knowledge of gender-stereotyped toys in the second year. *International Journal of Behavioral Development, 25*, 7–15.

Servin, A., Nordenstrom, A., Larsson, A., & Bohlin, G. (2003). Prenatal androgens and gender-typed behavior: A study of girls with mild and severe forms of congenital adrenal hyperplasia. *Developmental Psychology, 39*, 440–450.

Snow, M. E., Jacklin, C. N., & Maccoby, E. E. (1983). Sex-of-child differences in father-child interaction at one year of age. *Child Development, 49*, 227–232.

Williams, C. L., & Pleil, K. E. (2008). Toy story: Why do monkey and human males prefer trucks? *Hormones and Behavior, 54*, 355–358.

Zentner, M. R. (2001). Preferences for colours and colour emotion combinations in early childhood. *Developmental Science, 4*, 389–398.

Arch Sex Behav (2010) 39:1275–1288
DOI 10.1007/s10508-009-9595-2

ORIGINAL PAPER

Viewing Time Effects Revisited: Prolonged Response Latencies for Sexually Attractive Targets Under Restricted Task Conditions

Roland Imhoff · Alexander F. Schmidt · Uta Nordsiek · Charlotte Luzar · Andrew W. Young · Rainer Banse

Received: 24 October 2008 / Revised: 10 December 2009 / Accepted: 10 December 2009 / Published online: 3 March 2010

Abstract Sexually attractive stimuli are watched longer than unattractive stimuli. The processes underlying this robust and reliable viewing time effect are presently not well understood. In the present research comprising four experiments (total $N = 250$), four classes of potential explanations are proposed and the derived implications were experimentally tested. Contrary to explanations based on either deliberate delay or attentional adhesion to sexually attractive stimuli, prolonged response latencies were also found under restricted task conditions. Sexually preferred targets elicited longer response latencies in a self-paced evaluation task when stimulus pictures were presented for 750 ms (Experiment 1) or for 500 ms and followed by a pattern mask (Experiment 2). Prolonged latencies for sexually preferred targets were also observed when sexual attractiveness was rated in a speeded binary decision task with a response window of 1000 ms (Experiment 3). Eventually, it was shown that the response latency effect in the speeded binary choice task was still preserved when only the heads of target individuals were presented instead of the bodies (Experiment 4). Mate identification and schematic processes are discussed as the remaining plausible mechanisms for prolonged response latencies for sexually attractive targets under restricted conditions.

Keywords Viewing time · Sexual preference · Sexual interest · Visual reaction times · Indirect measures

R. Imhoff (✉) · A. F. Schmidt · U. Nordsiek · C. Luzar · R. Banse
Social and Legal Psychology, University of Bonn, Kaiser-Karl-Ring 9, 53111 Bonn, Germany
e-mail: rimhoff@uni-bonn.de

A. W. Young
Department of Psychology, University of York, York, UK

Introduction

Over the last decade, the indirect assessment of sexual preferences has received growing attention. Direct measures of sexual preference, such as questionnaires and clinical interviews, rely on participants' willingness and ability to accurately report information about their sexual interest for assessing sexual preference. Therefore, the usefulness of such direct methods is particularly questionable in the forensic context where denial and dissimulation of deviant sexual interest can be expected if assessments are part of legal proceedings. As an alternative to self-report methods, a number of indirect measures have been proposed, such as penile plethysmography (PPG; e.g., Freund, 1963), the Implicit Association Tests (IAT; e.g., Gray, Brown, MacCulloch, Smith, & Snowden, 2005), the Choice Reaction Task (CRT; e.g., Wright & Adams, 1994), and viewing time measures (VT; e.g., Harris, Rice, Quinsey, & Chaplin, 1996).

These instruments infer sexual preference from objective measures either based on physiological indicators of sexual arousal or response latencies. Both approaches have in common that reactions elicited by stimuli showing target persons belonging to groups of sexual interest (men versus women, children versus adults) are recorded. The measurement rationale relies on the fact that certain stimulus categories induce more sexual arousal as indicated by stronger tumescence (PPG), that certain classes of target individuals are more strongly semantically associated with the concept of sex or sexual interest than others (IAT), or that sexually preferred stimuli function as distracters that interfere with performance in a primary task (CRT).

Viewing Time Measures

Since the seminal work of Rosenzweig (1942), it is well established that pictures of sexually attractive persons are

watched longer than pictures of sexually unattractive persons when sexuality is salient. This basic effect is so reliable and robust that it is used for the indirect assessment of sexual preferences in forensic settings (e.g., Abel, Jordan, Hand, Holland, & Phipps, 2001). However, whereas the underlying processes of most other indirect measures of sexual preference are relatively well established, surprisingly little is known about the mechanisms underlying viewing time effects. In the standard viewing time procedure, participants are asked to evaluate pictures of target individuals on a graded scale of sexual attractiveness. The response latency of this judgment is unobtrusively measured. Across studies, there is a very robust finding that the response latency is longer for sexually attractive as compared to sexually unattractive targets and, in turn, viewing time measures can be used to discriminate between participants with respect to sexual preference (Flak, Beech, & Fisher, 2007; Kalmus & Beech, 2005; Laws & Gress, 2004), including homosexual and heterosexual men (Zamansky, 1956), heterosexual men and women (Israel & Strassberg, 2009; Quinsey, Ketsetzis, Earls, & Karamanouikan, 1996), and child sex offenders and non-offenders (e.g., Banse, Schmidt, & Clabour, 2010; Gress, 2005; Harris et al., 1996). The discriminatory ability of viewing time measures has been claimed to be similar or even superior to phallometric measures (e.g., Abel, Huffman, Warberg, & Holland, 1998) but in a recent critical review of the VT literature Sachsenmaier and Gress (2009) propose that "studies have yet to determine, whether the measure is at least as accurate as or perhaps more so than PPG" (p. 55). However, recent evidence suggests that VT measures outperform other indirect measures like the IAT (Banse et al., 2010).

Despite the robustness of the viewing time effect, there seems to be virtually no empirical research on the underlying mechanisms. Also, theoretical accounts are sparse. In review articles, the viewing time effect is commonly introduced at the descriptive level without further theorizing. For example, Laws and Gress (2004) stated that "the rationale underlying the test is that clients will look longer at pictures they find sexually attractive" (p. 184). Others (e.g., Flak et al., 2007; Kalmus & Beech, 2005) categorize viewing time measures as "attentional techniques" and argue that "assessments measuring viewing time assume that individuals will look longer at images they consider attractive than they would view unattractive or neutral images," explicitly distinguishing it from other techniques that "discriminate the effect of increased attention upon information processing tasks" (Kalmus & Beech, 2005, p. 208).

As a background theory, scholars in the area of forensic research often refer to Singer's (1984) model of sexual arousal (e.g., Flak et al., 2007; Kalmus & Beech, 2005), which consists of three consecutive phases. The first reaction is termed an *aesthetic response*, a "hedonic feeling in response to a sexual stimulus" that "develops into a more active orientation toward the sexual stimulus," followed by an approach step, and then a third step of physiological genital response (Singer, 1984). The attentional process described as an aesthetic response is generally believed to cause the viewing time effect (e.g., Kalmus & Beech, 2005). Others take an evolutionary psychology perspective by postulating that longer viewing time may be adaptive for mate seeking because "it reflects the initial stage of courtship, locating and evaluating an appropriate partner" (Quinsey et al., 1996). Sexual attraction is closely related to reproduction; therefore, it seems plausible that the cognitive system has adapted to directing attention to potential sexual mates, i.e., sexually preferred individuals (Redouté et al., 2000). However, in evolutionary psychology, a satisfactory explanation of a phenomenon requires that there is an understanding not only of ultimate but also proximal mechanisms that cause the observed behavior. Thus, although both Singer's theory of sexual arousal and the evolutionary psychology approach offer a starting point, very little is known about the actual psychological mechanisms that cause the effect of prolonged viewing times for sexually attractive targets. It is the aim of the current research to specify four alternative accounts of the underlying processes and to test them empirically.

Deliberate Delay

First, the most parsimonious explanation of why judgment of sexually highly attractive stimuli is prolonged is that watching those stimuli is rewarding and that terminating this by any response is therefore deliberately delayed. This hypothesis is corroborated by neurophysiological evidence. Watching sexually attractive stimuli elicits neuronal activities in brain areas commonly associated with the human reward system (e.g., Ishai, 2007; Karama et al., 2002; Mouras et al., 2003; Ponseti et al., 2006; Redouté et al., 2000; Safron et al., 2007; Stoléru et al., 1999). Humans are assumed to be motivated by hedonism (e.g., Epstein, 1990) in that they try to reach and maintain positive affective states. Singer (1984) argued that it is the "hedonic feeling in response to a sexual stimulus" that motivates the individual to keep the sexually attractive object in view. In addition, even in the absence of stimuli, participants might be motivated to delay responding to prolong rewarding reminiscence of the stimuli or sexual fantasies. In short, viewing time effects could result from the controlled and intentional delay to keep a sexually pleasant stimulus in view or keep a sexually pleasant internal representation.

Attentional Adhesion to Sexual Stimuli

Second, the delayed responding could be mediated by the automatic process of attention direction toward presented sexually attractive stimuli. It can be argued that sexually attractive stimuli automatically bind attention and distract participants from their actual task to rate the persons' sexual attractiveness. Responses are, therefore, delayed. A very similar assumption underlies the rationale of the CRT (Santtila et al., 2009; Wright

& Adams, 1994) and recent research suggests that sexual arousal can indeed increase attentional adhesion to attractive opposite-sex targets (Maner, Gailliot, Rouby, & Miller, 2007). Although deliberate delay and attentional adhesion are clearly distinct, they have not been well differentiated in the literature so far. It should be noted that both explanations are not mutually exclusive. It is conceivable that sexually attractive stimuli could automatically attract and bind attention; the visual processing could then elicit positive affect, which subsequently causes deliberate prolonged viewing and a delayed judgment to maintain the pleasurable state.

Sexual Content Induced Delay (SCID)

In the SCID literature (Geer & Bellard, 1996; Geer & Melton, 1997), it has been found that the presentation of erotic stimuli induces hesitancy in decision making. Spiering, Everaerd, and Elzinga (2002) provided data to support their interpretation of SCID as an evolutionary adaptive activation of conscious regulation modules. To the degree that stimuli presented in viewing time tasks are sexually explicit (as compared to neutral or mildly erotic; Spiering, Everaerd, & Laan, 2004) viewing time effects may be a special case of the general SCID phenomenon.

Mate Identification

A fourth class of explanations postulates internal processes that are automatically triggered by either sexually attractive stimuli as such or the specific task to rate their sexual attractiveness. Briefly presented stimuli could trigger internal attentional processes to erotic cues as well as expectancies and/or schematic concepts (Wiegel, Scepkowski, & Barlow, 2007). Finally, the effect could also emerge as a result of the task commonly connected to viewing time measures. It is conceivable that denying sexual attractiveness is on average faster than confirming it and responses for sexually attractive stimuli are thus prolonged.

The Present Research

As we have seen so far, there are at least four classes of plausible explanations for prolonged response latencies of sexually attractive stimuli. However, to the best of our knowledge, these plausible explanations have never been empirically tested, and it was the aim of the present research to start to close this gap. The deliberate delay hypothesis can be tested by constraining the viewing conditions of sexually attractive targets in a way that it becomes unlikely that the observed latencies are caused by a controlled delay. If the deliberate delay hypothesis is true, prolonged response latencies should be reduced or eliminated if the response can only be given in the absence of the hedonically rewarding stimulus (i.e., after the stimulus has disappeared). Likewise, attentional adhesion necessarily requires the presence of the stimulus and should be eliminated in its absence. The first two experiments tested the effect of stimulus presence on viewing time effects and provide first evidence that these effects also emerge in the absence of stimuli. To further elucidate the nature of the underlying process we tested the boundary conditions of the viewing time effect in Experiments 3 and 4 by drastically reducing the time frame in which a response can be given and by presenting only the heads of the target stimuli (Experiment 4).

Experiment 1

Prolonged response latencies for sexually attractive stimuli due to deliberate delay crucially depend on the presence of the sexually attractive stimuli. If the sexually attractive stimuli are removed before participants give their attractiveness rating, there is no longer a reason to deliberately delay the attractiveness rating because it will not prolong the time participants can watch the stimuli. Furthermore, an absent stimulus is unlikely to hold attention and distract from the actual rating task. To test these predictions, the presentation time of target stimuli was experimentally manipulated in Experiment 1. To the extent that prolonged RTs for sexually attractive targets are caused by deliberate delay or attentional adhesion to the stimulus, prolonged RTs for sexually preferred targets were expected in the standard viewing time task, but these should vanish under restricted presentation conditions. To the extent that internal processes cause delayed responding to sexually attractive stimuli these should occur even in the absence of stimuli. Heterosexual and homosexual men were recruited as groups with contrasting sexual preferences that are not confounded by sex differences in response to visual sexual stimuli (Rupp & Wallen, 2008).

Method

Participants

A sample of 35 heterosexual and 24 homosexual men was recruited by posters and via online forums for a study on attractiveness. Participants were informed that the experiment would entail direct and indirect measures of their sexual interest in men, women, boys, and girls. Participants were further informed that they could withdraw from the experiment at any time without disadvantage, and that all data were collected and stored in fully anonymous form. Written consent was obtained prior to the experiment. Participants received 5 Euro for participation. The mean age of participants was 24.8 years. Their age was independent of sexual orientation, $t(57) = 1.52$. Sexual orientation was checked by an Explicit Sexual Interest Questionnaire (ESIQ; Banse et al., 2010). Out of ten possible sexual behaviors or fantasies regarding women, heterosexual men reported an average of $M = 9.8$, $SD = .47$, whereas homosexual men reported $M = 1.5$, $SD = 1.10$. For sexual behaviors and fantasies regarding men, this pattern reversed: heterosexual

men reported an average of $M = .4$, $SD = .92$ and homosexuals an average of $M = 9.8$ ($SD = .41$).

Stimuli and Materials

The target stimuli were 40 computer-constructed photographs of male and female individuals taken from the Not Real People (NRP) picture Set (Subset B, Pacific Psychological Assessment Corporation, 2004; for examples, see Laws & Gress, 2004), featuring male and female individuals belonging to five categories of sexual maturation (corresponding to Tanner, 1978). The Tanner categories 1 to 3 depict pictures of prepubescent children of increasing maturity, Tanner category 4 adolescents, and Tanner category 5 young adults. All individuals were shown in bathing clothes of different colors.

Viewing Time Measure

The subset of 40 target pictures was divided into two halves with two target individuals per age × sex combination. The assignment of these two sets to the experimental conditions was counterbalanced. In the first experimental condition (the standard viewing time procedure), target pictures and a rating scale (1 = "sexually not attractive" to 5 = "sexually very attractive") were presented simultaneously. Both the picture and the scale were presented until the response was given and confirmed by pressing the enter key. In the second condition (the restricted condition), the target picture was presented for 750 ms, followed by the response scale. The response scale remained visible on the screen until a response was entered and confirmed. The response time (RT) in the attractiveness rating task served as the dependent variable. Recording of RT started with the presentation of the Likert scale.

Procedure

Upon their arrival in the laboratory, participants were informed about the aims and procedures of the experiment. After giving their informed consent, four different indirect assessment tasks were performed in the following order: a Sexual Misattribution Procedure (Imhoff, Schmidt, Bernhardt, Dierksmeier, & Banse, 2010), the viewing time task, an Implicit Association Test, and a Choice Reaction Task. Only the viewing time task was of interest here, so the results of the other measures will not be reported in this article. At the end of the experiment, participants completed a sexual preferences questionnaire, and were then debriefed, paid for their participation, and thanked.

Design

The factorial design was a 2 (Participant Sexual Orientation: Heterosexual versus Homosexual) × 2 (Target Sex: Male versus Female) × 5 (Target Age: Tanner Categories 1–5) × 2 (Display Condition: Standard versus Restricted), with one between-subjects factor and three within-subjects factors.

Results

The response latencies for each of the 2 × 5 target categories were averaged. A 2 (Participants' Sexual Orientation) × 2 (Target Sex) × 5 (Target Age) × 2 (Display Condition) repeated measures analysis of variance (ANOVA) revealed that a viewing time effect emerged, as indicated by a significant three-way Participants' Sexual Orientation × Target Age × Target Sex interaction, $F(4, 54) = 10.46$, $p < .001$, $\eta_p^2 = .44$. Figure 1 shows the mean latencies as a function of Participants' Sexual Orientation, Target Sex and Target Age for both display conditions separately. Means show that adults generally elicited the longest RTs, $F(4, 54) = 29.27$, $p < .001$, $\eta_p^2 = .68$, but that this effect was more pronounced for the preferred sex (Participant Sexual Orientation × Target Sex). However, contrary to the predictions derived from the deliberate delay and attentional adhesion hypotheses, this effect was independent of display condition as indicated by the non-significant four-way interaction, $F(4, 54) = 1.03$.

Table 1 shows the results in a condensed design of only two target age categories: prepubescent (Tanner 1–3) versus postpubescent (Tanner 4–5). Both homosexual and heterosexual men showed the longest RTs for postpubescent stimuli of the preferred sex and the shortest RTs for prepubescent stimuli of the non-preferred sex. Standardized preference scores for postpubescent males over females almost perfectly predicted participants' sexual orientation in a ROC analysis, $AUC = .95$, $p < .001$.

Discussion

The results of Experiment 1 showed that sexually more attractive targets elicited longer response latencies than sexually less attractive targets in a rating task of sexual attractiveness by homosexual and heterosexual men. Contrary to the prediction derived from the deliberate delay or attentional adhesion accounts, this effect was obtained not only for the standard viewing time condition, but also under conditions of restricted stimulus display. Although the stimuli were presented for only 750 ms, and the judgment of sexual attractiveness was performed after the target picture had disappeared, prolonged latencies for sexually preferred stimuli emerged. Additionally, neither effect sizes nor the criterion validity of measures showed any substantial difference between the standard and restricted presentation conditions. Thus, Experiment 1 did not support the claim that prolonged viewing time effects under conditions of restricted stimulus display can be explained by a deliberate delay of the response to keep a sexual stimulus in view or an attentional adhesion to a sexually attractive stimulus. This result raises the question of whether the label "Viewing Time" for the

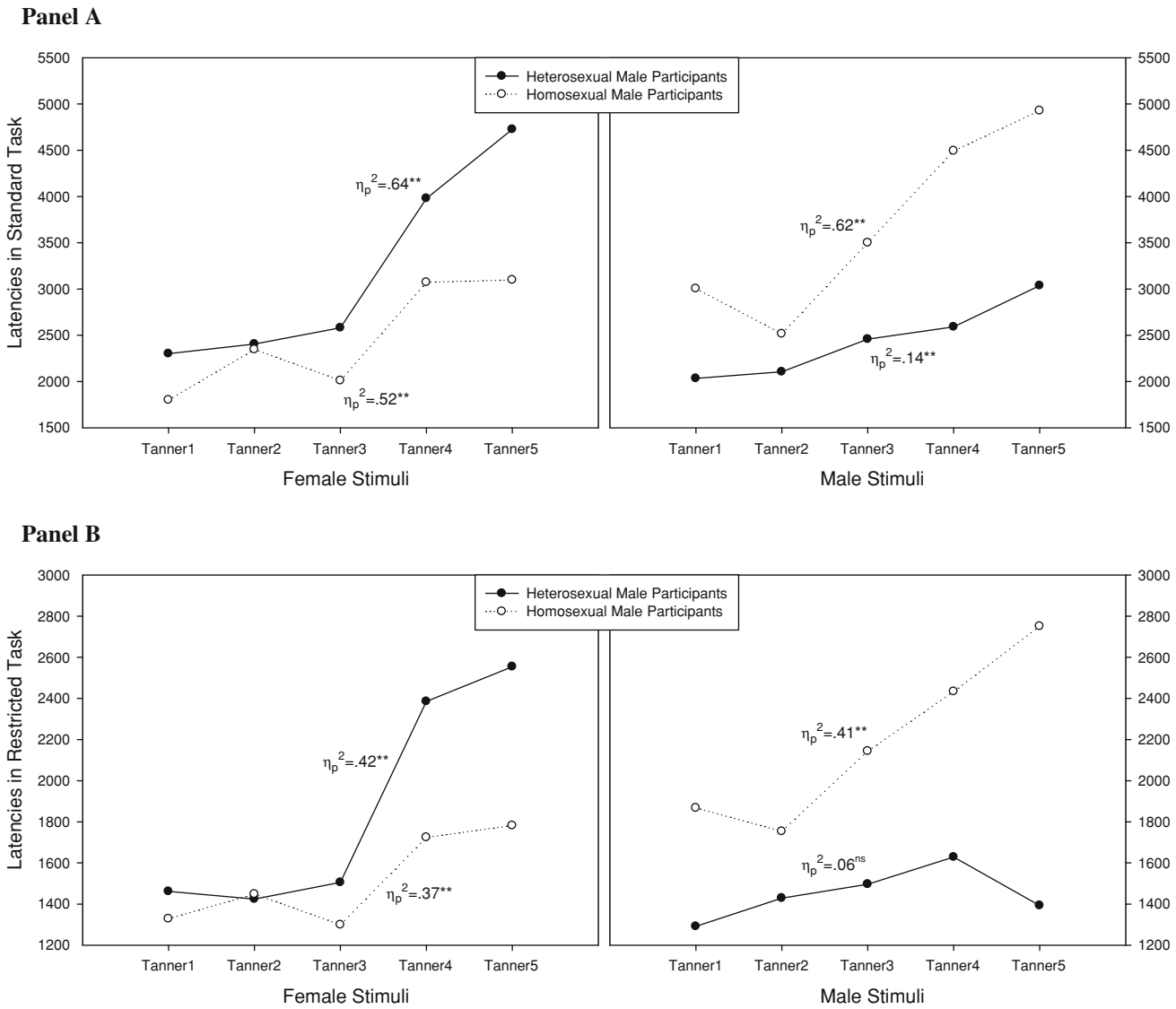


Fig. 1 Response latencies (ms) as a function of Target Age (Tanner Category) and Target Sex (female versus male) for heterosexual and homosexual male participants in an unrestricted standard viewing time (**a**) and a restricted display task (stimulus presentation 750 ms) (**b**) in Experiment 1. Effect sizes for linear within-subject contrasts for Target Age, ** $p < .01$, * $p < .05$

observed effect is a misnomer. In fact, participants under restricted conditions viewed all stimuli for the same amount of time but still differed in their latencies. The effects could thus be better described as prolonged response latencies for sexually attractive targets (PRELSAT).

It could be argued, however, that even in the absence of the stimulus pictures participants may experience afterimages of the targets on their retina. Thus, it cannot be excluded that this (retinal) afterimage was distracting and/or rewarding and therefore responsible for the effect in the restricted condition. In order to eliminate this possibility, a second experiment was conducted in which target pictures were displayed for an even shorter time of 500 ms, and then masked before the rating could be given. For Experiment 2, we chose heterosexual men and women as a sample. Recent research suggests that women are generally less specific in their sexual response than men, showing subjective and genital sexual arousal to pictures of both men and women (Chivers, Rieger, Latty, & Bailey, 2004; Chivers, Seto, & Blanchard, 2007). However, prior research on the viewing time effect has provided mixed results. Although standard viewing time effects emerged also for women, these effects were usually smaller than those for heterosexual males (Israel & Strassberg, 2009; Quinsey et al., 1996). To explore this gender specificity effect also for the restricted display variation, we recruited a community sample of men and women. By excluding not clearly heterosexual individuals, we created groups in which participant sex served as a proxy for contrasting sexual preferences.

Table 1 Criterion validity and mean latencies (ms) as a function of stimulus maturity (prepubescent versus postpubescent) and stimulus sex (male versus female) for standard viewing time and restricted display procedures with heterosexual and homosexual men (Experiment 1) and heterosexual men and women (Experiment 2)

	Female stimuli				Male stimuli				Effect size	Criterion validity	
	Prepubescent		Postpubescent		Prepubescent		Postpubescent				
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*	*AUC*	Correct classifications (%)
Experiment 1: standard viewing time											
Heterosexual men ($n = 35$)	2429^{ab}	1150	4353^{c}	1614	2199^{a}	1004	2814^{b}	1523	1.09	.95	85.7
Homosexual men ($n = 24$)	2052^{a}	890	3085^{b}	1401	3007^{b}	1769	4711^{c}	1907	−0.86		83.3
Experiment 1: restricted display											
Heterosexual men ($n = 35$)	1464^{a}	616	2469^{b}	986	1405^{a}	612	1511^{a}	587	1.00	.96	85.7
Homosexual men ($n = 24$)	1358^{a}	599	1753^{b}	907	1922^{b}	1208	2592^{c}	1016	−0.83		79.2
Experiment 2: standard viewing time											
Heterosexual men ($n = 33$)	1934^{ab}	683	3737^{c}	1676	1891^{a}	770	2385^{b}	1596	0.62	.80	63.6
Heterosexual women ($n = 36$)	1812^{a}	547	2954^{b}	1492	1839^{a}	436	2838^{b}	1015	−0.17		72.2
Experiment 2: restricted display											
Heterosexual men ($n = 33$)	1181^{a}	288	2063^{b}	985	1323^{a}	585	1532^{a}	697	0.53	.86	78.8
Heterosexual women ($n = 36$)	1256^{a}	296	1617^{b}	567	1263^{a}	315	1795^{b}	582	−0.25		77.8

Note: Different index letters in one row indicate significant differences in simple tests (Bonferroni-corrected $\alpha = .008$). Effect sizes for the different latencies are based on male versus female postpubescent stimuli. All predictions are based on difference scores (male versus female postpubescent stimuli), $p < .001$

Experiment 2

The setup of Experiment 2 was similar to Experiment 1, except that the number of trials was increased and the target stimuli were presented for 500 ms. Most importantly, target stimuli were immediately followed by a pattern mask to overwrite any afterimages on the retina. With regard to the sample, heterosexual men and women were selected as known groups with contrasting sexual preference.

Method

Participants

The sample consisted of 69 participants (33 men, 36 women) with a mean age of 26.2 years, and contrasting sexual preferences (sexual interest in men: $M = .7$, $SD = 1.32$, for males and $M = 9.4$, $SD = .83$, for females; sexual interest in women $M = 9.7$, $SD = .74$, for men and $M = 1.9$, $SD = 1.78$, for women). Participants were informed that the study investigated sexual interest towards children and adults of both sexes and written consent was obtained.

Stimuli and Materials

As compared to Experiment 1, the number of trials was doubled to 40 in each condition to achieve an even more reliable measure. In the restricted condition, presentation time was further reduced to 500 ms, followed by a pattern mask that was displayed for 250 ms.

Procedure

After the viewing-time task, participants completed the ESIQ as an explicit measure of sexual preference, were debriefed, and thanked.

Results

A 2 (Participant Sex) × 2 (Target Sex) × 5 (Target Age) × 2 (Display Condition) repeated measures ANOVA was calculated. As in Experiment 1, the PRELSAT effect emerged independently of display condition. The interaction of Participant Sex × Target Age × Target Sex was significant, $F(4, 64) = 6.77$, $p < .001$, $\eta_p^2 = .30$, and was not qualified by a four-way interaction with display condition, $F < 1$. A general increase of latencies with increasing target age was stronger for targets of the preferred sex (Fig. 2). Whereas heterosexual men showed a specific pattern of longer RTs in both display conditions, heterosexual women showed a non-specific pattern of increased RTs for both male and female postpubescents (Table 1). The criterion validity was thus not as high as in Experiment 1, but still substantial and significant in the standard condition, $AUC = .80$, $p < .001$, as well as in the restricted condition, $AUC = .86$, $p < .001$.

Discussion

Replicating Experiment 1, the results provided evidence that the PRELSAT effect did not vanish under even more restricted conditions. Masking the stimuli before presenting the scale ruled

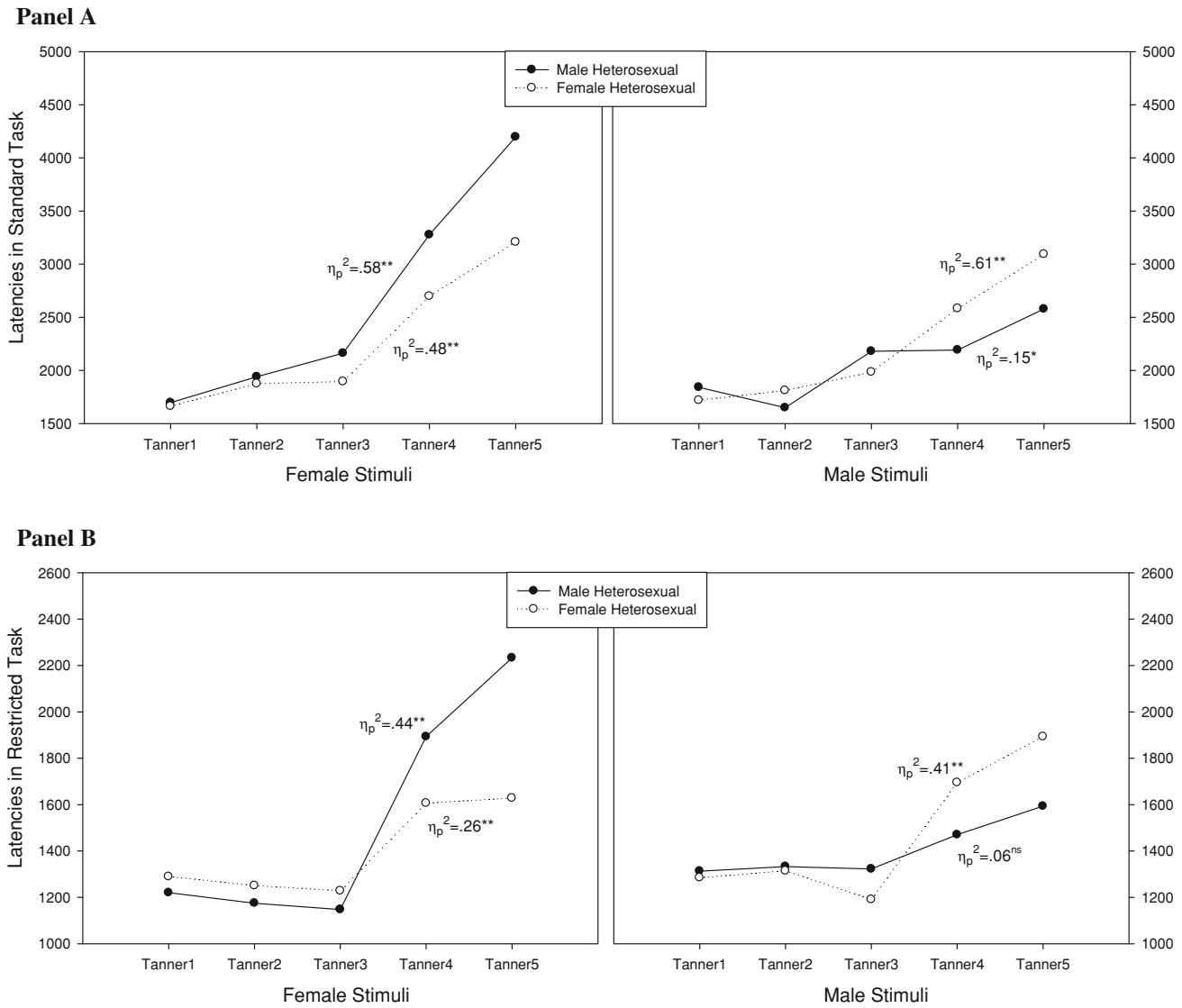


Fig. 2 Response latencies (ms) as a function of Target Age (Tanner Category) and Target Sex (female versus male) for heterosexual men and women in an unrestricted standard viewing time (**a**) and a restricted display task (stimulus presentation 500 ms, masked) (**b**) in Experiment 2. Effect sizes for linear within-subject contrasts for Target Age, ** $p < .01$, * $p < .05$

out an explanation based on afterimage effects. Thus, stimulus visibility was not required to produce PRELSAT effects.

An inspection of the absolute latencies suggests that, despite identical patterns, in the standard viewing time condition it took participants roughly twice as long to rate the sexual attractiveness of the presented stimuli as in the restricted condition. This difference can be partially explained by the time to actually see and perceive the stimuli that were included in the RT under standard conditions, but not under restricted conditions. However, it seems plausible that, in the absence of any time pressure, participants spontaneously engaged in behavior induced by sexually attractive stimuli. For example, they might engage in processes of social comparisons—checking the physique of the stimulus as compared to their own or their partners. In fact, they may have enjoyed watching sexually attractive stimuli and thus deliberately delayed their response. However, the results of the two experiments provided evidence that the prolonged response latencies did not depend on such a mechanism. Furthermore, by allowing participants to take more time, it seems likely that while potentially tapping into this hedonic pleasure effect, additionally more noise was recorded that blurred the measure.

Having established the fact that processes independent of stimulus presence were sufficient to produce PRELSAT effects, we were interested in further narrowing down the plausible explanations. It is conceivable that deliberate delay is not motivated by watching visually pleasant stimuli but rather by an internal reminiscence initially activated by such a stimulus (Wiegel et al., 2007). Hence, constraining the sensory input does not preclude the possibility that the pictures of sexually attractive targets elicit sexual thoughts, fantasies, and/or expectancies that

are pleasurable and thus are maintained before responding to the task. Therefore, a third experiment was conducted to restrict the participants' behavior after exposure to sexually attractive stimuli. This was achieved by changing the nature of the task from an evaluation task to a speeded performance measure.

Experiment 3

In an attempt to constrain fantasizing and imagery following the presentation of targets, the self-paced evaluation of sexual attractiveness was replaced by a simple binary decision task in which participants had to decide whether a target person was either a "potential sexual partner" or "not a potential sexual partner" for them. Participants were instructed to make their decision as fast as possible within 1000 ms. The speeded task should eliminate deliberative processes leading to deliberately delayed responding.

Method

Participants

The sample consisted of 58 heterosexual participants (29 men, 29 women), after excluding two bisexual participants on base of their ESIQ scores. The mean age was 29.0 years, and there was no significant age difference between men and women, $t(56) < 1$. All participants gave their informed, written consent to participate in a study on sexual preferences.

Procedure

For the speeded viewing-time trials, participants were asked to classify as quickly as possible the randomly presented targets by pressing the left (no potential sexual partner) or the right response button (potential sexual partner). Participants then completed a filler task, the standard viewing time procedure, and the ESIQ before taking again the same speeded response task to assess the stability of the measure.

Stimuli and Materials

The standard viewing time condition was identical to the one in Experiment 2. In the speeded condition, both categories were anchored in black letters next to the top right and left corner of the picture, respectively. After exceeding a response time of 1000 ms, an error message "too slow!" appeared above the picture stimulus.

Design and Specific Hypotheses

The experiment followed a 2 (Participant Sex) × 2 (Target Sex) × 5 (Target Age) × 3 (Response Condition: Standard versus Speeded 1 versus Speeded 2) mixed factorial design with one between-subjects factor and three within-subjects factors.

Fig. 3 Response latencies (ms) as a function of Target Age (Tanner Category) and Target Sex (female versus male) for male and female heterosexual participants in standard viewing time (**a**) and two speeded response tasks (1000 ms response window, two assessments) (**b**, **c**) in Experiment 3. Effect sizes for linear within-subject contrasts for Target Age, ** $p < .01$, * $p < .05$

Under standard viewing conditions, we expected prolonged RTs for sexually preferred targets. Female participants were expected to show the longest RTs for adult male targets, whereas male participants were expected to show the longest reaction times for adult female targets. Under conditions of speeded responding, deliberate delay due to internal reminiscence should be strongly reduced or eliminated, and thus prolonged RTs for sexually preferred targets under speeded condition would suggest a different process.

Results

For the speeded response measures, all trials with RTs greater than 1000 ms were defined as errors and discarded. One participant's results were excluded from the analysis of the second speeded response task as his error rate was more than 2 SD above the mean.

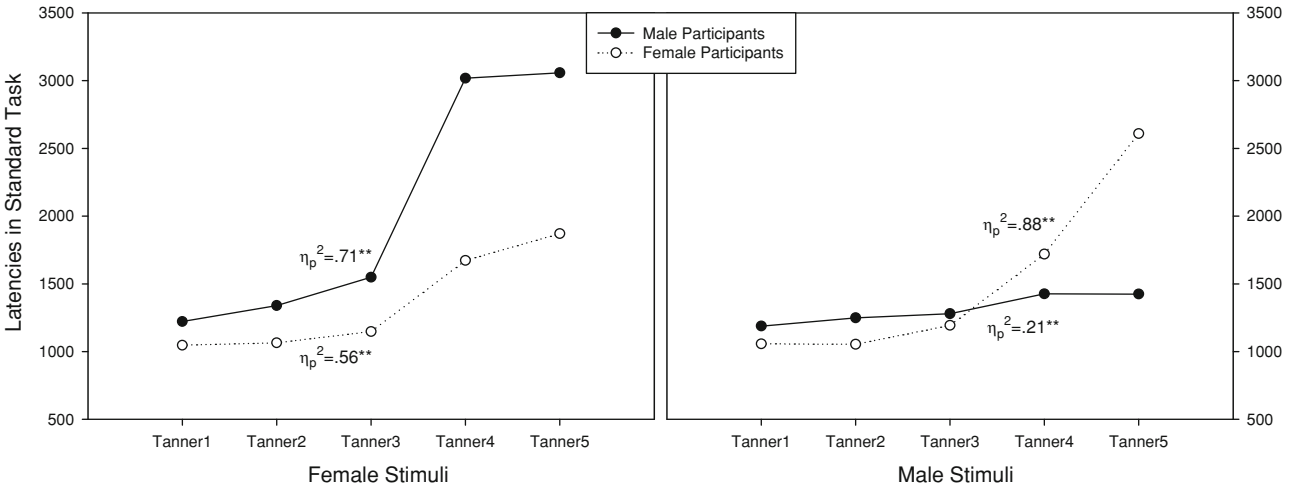
A 2 (Participant Sex) × 2 (Target Sex) × 5 (Target Age) × 3 (Response Condition) ANOVA yielded a three-way Participant Sex × Target Sex × Target Age interaction, $F(4, 52) = 23.31$, $p < .001$, $\eta_p^2 = .62$. Contrary to Experiments 1 and 2, it was qualified by a significant four-way interaction with response condition, $F(8, 48) = 8.94$, $p < .001$, $\eta_p^2 = .60$, indicating that the size and/or direction of the three-way interaction depended on response condition. Separate analyses of all three conditions showed this was due to the fact that the hypothesized three-way interaction was significant and comparable for all three conditions but more pronounced in the second speeded condition. In the second speeded condition, the effect was somewhat larger, $F(4, 52) = 36.91$, $p < .001$, $\eta_p^2 = .74$, than in the standard viewing time procedure, $F(4, 53) = 16.85$, $p < .001$, $\eta_p^2 = .56$, and the first speeded procedure, $F(4, 52) = 13.71$, $p < .001$, $\eta_p^2 = .51$. Figure 3 shows comparable and hypothesized patterns of the means for all three conditions: the effect of longer RTs for older targets was stronger for the preferred sex. In contrast to Experiment 2, in all three measures heterosexual women showed an effect of target sex for postpubescent stimuli (Table 2).

The psychometric quality of the speeded response measure can be inferred from an *AUC* between .98 and 1.00, resulting in correct classifications of up to 97% of the participants as well as a retest reliability of $r_{tt} = .86$, $p < .001$, for the standardized difference score for postpubescent stimuli.

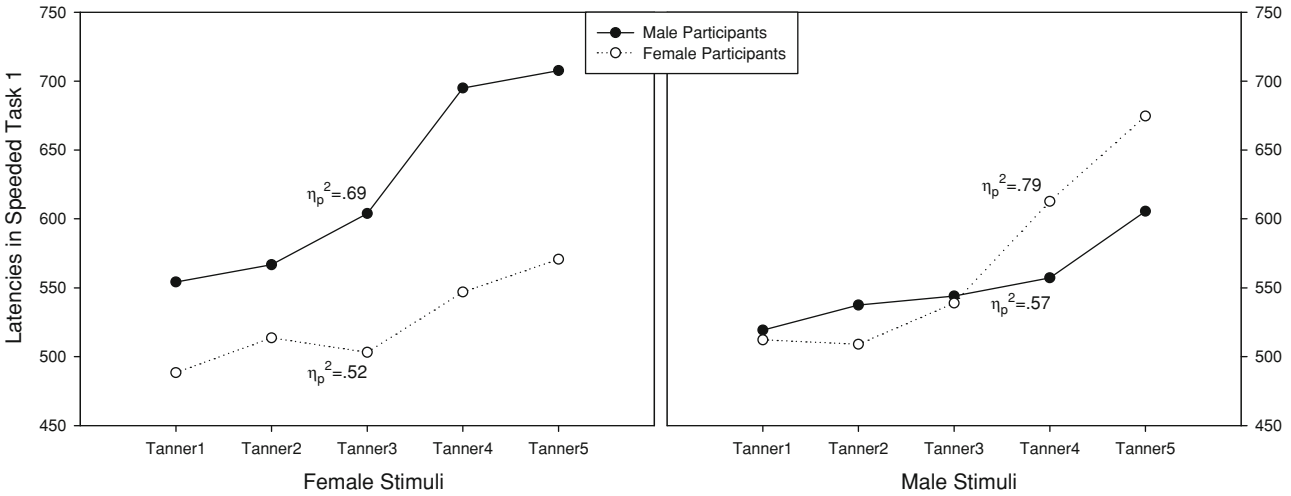
Discussion

The results of Experiment 3 provided evidence that the effect of prolonged RTs for sexually preferred targets persisted even

Panel A



Panel B



Panel C

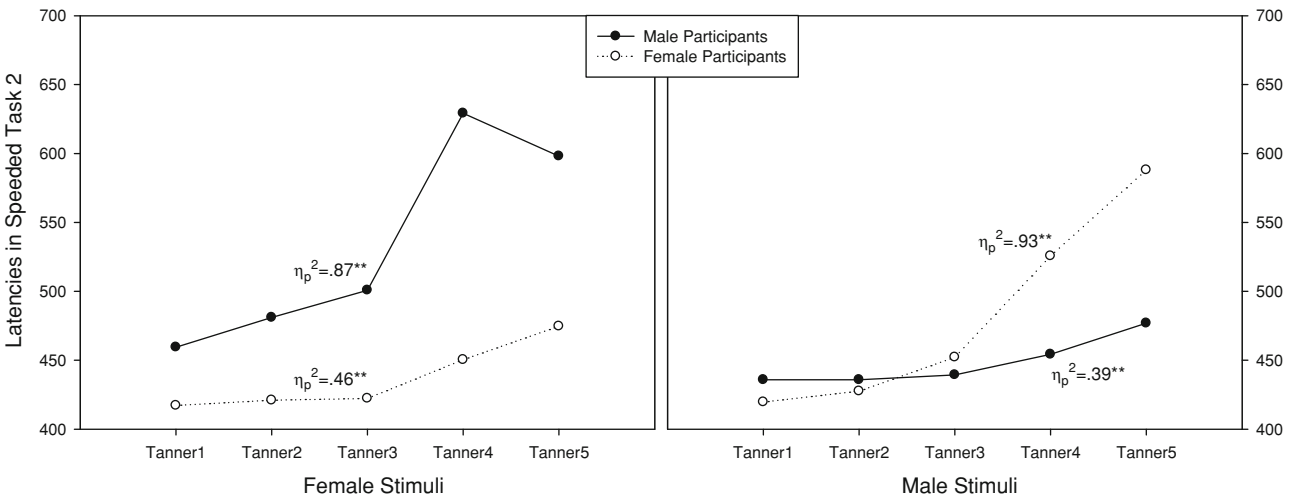


Table 2 Criterion validity and mean latencies (ms) as a function of stimulus maturity (prepubescent vs. postpubescent) and stimulus sex (male vs. female) for standard viewing time and speeded response procedures with heterosexual men and women in Experiments 3 and 4

	Female stimuli				Male stimuli				Effect size	Criterion validity	
	Prepubescent		Postpubescent		Prepubescent		Postpubescent				
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*	*AUC*	Correct classifications (%)
Experiment 3: standard viewing time											
Heterosexual men ($n = 29$)	1370^{a}	733	3038^{b}	1348	1239^{a}	576	1425^{a}	848	1.26	.96	93.1
Heterosexual women ($n = 29$)	1086^{a}	276	1772^{b}	745	1101^{a}	260	2164^{c}	568	−0.62		93.1
Experiment 3: speeded response 1											
Heterosexual men ($n = 29$)	572^{a}	105	701^{c}	105	531^{b}	108	579^{a}	89	1.69	.98	93.1
Heterosexual women ($n = 29$)	502^{a}	79	558^{c}	98	519^{b}	83	640^{d}	68	−1.13		89.7
Experiment 3: speeded response 2											
Heterosexual men ($n = 28$)	481^{b}	92	612^{c}	90	437^{a}	83	465^{b}	95	2.35	1.00	96.4
Heterosexual women ($n = 29$)	420^{a}	63	463^{b}	99	433^{ab}	58	555^{c}	68	−1.36		96.6
Experiment 4: speeded response											
Heterosexual men ($n = 36$)	513^{b}	92	583^{c}	88	471^{a}	93	515^{b}	103	1.04	.96	97.2
Heterosexual women ($n = 28$)	450^{a}	86	475^{b}	105	476^{b}	94	569^{c}	94	−1.13		89.3

Note: Different index letters in one row indicate significant differences in simple tests (Bonferroni-corrected $\alpha = .008$). Effect sizes for the different latencies are based on male vs. female postpubescent stimuli. All predictions are based on difference scores (male vs. female postpubescent stimuli), $p < .001$

under conditions of a speeded binary response task. This result rules out deliberative reminiscence as the cause of prolonged RTs. The speeded response conditions essentially yielded the same results as the standard condition, the effect sizes were even somewhat larger, and the classifications as correct as in the standard viewing time condition. It seems plausible that performance under time pressure might be a clearer reflection of the automatic process underlying implicit sexual preference. The large effect sizes, together with the more automatic nature and the fully satisfactory retest reliability, may make this paradigm an attractive alternative to the standard viewing time procedure for diagnostic purposes. Finally, to further narrow down plausible explanations of PRELSAT effects, we reduced the potential erotic content of stimuli Experiment 4.

Experiment 4

In the SCID literature, it has been reported that these delay effects appear only after presenting sexually explicit content (pornographic images), whereas no SCID was found after presenting mere erotic images or pictures of nude models (Spiering et al., 2004). Most viewing time research relies on portrayal of nude or only partly clothed persons, sometimes displaying the genital area (Harris et al., 1996; for the use of nude stimuli, see Brown, Amoroso, Ware, Pruesse, & Pilkey, 1973; Love, Sloan, & Schmidt, 1976; Quinsey et al., 1996; Ware, Brown, Amoroso, Pilkey, & Pruesse, 1972). In these experiments, PRELSAT effects could be considered a special case of a more general SCID Effect.

In Experiments 1 to 3, the target persons were clothed in bathing suits. Although unlikely to produce the same effect as explicit sexual imagery, particularly for participants with a corresponding sexual preference, men wearing tight swimming trunks and women wearing bikinis might function as mildly sexually arousing stimuli. To investigate whether this (mildly) erotic content is a necessary condition to produce the PRELSAT effect, heads-only pictures were edited from the full body pictures used in the previous experiments. If the PRELSAT effect is merely a specification of the SCID effect, no latency differences between sexually preferred and non-preferred targets should be expected.

Method

Participants

A total of 28 female (sexual interest in men $M = 9.2$, $SD = .94$ and women $M = 1.8$, $SD = 1.57$) and 36 male (sexual interest in men $M = .7$, $SD = 1.37$ and women $M = 9.3$, $SD = 1.23$) heterosexual participants were, on average, 25.6 years old, independent of sex, $t(62) = 1.32$. They gave their informed consent for participating in a study that dealt with sexual attractiveness of faces.

Procedure

After completing the speeded response procedure, participants completed the ESIQ as a manipulation check, were debriefed, and thanked.

Stimuli and Materials

The 80 items pictures used as items before were digitally cropped to show only the heads of the target individuals. Except for that, the speeded response task was similar to the speeded variations used in Experiment 3.

Design

The design was the same as used for the different response conditions in Experiment 3.

Results

As in Experiment 3, only trials with RTs below 1000 ms were used for data analysis. A 2 (Participant Sex) × 2 (Target sex) × 5 (Target age) repeated measures ANOVA was conducted. Contrary to the SCID hypothesis, the Participant Sex × Target Sex × Target Age three-way-interaction was significant, $F(4, 59) = 13.62, p < .001, \eta_p^2 = .48$. As illustrated in Fig. 4, and as in Experiments 1 to 3, a linear effect of target age was most pronounced for stimuli of the preferred sex. Both heterosexual men and women showed the longest latencies for postpubescent targets of the preferred sex, resulting in a high criterion validity of the standardized difference measure, $AUC = .96$, $p < .001$ (Table 2).

Discussion

Participants showed longer RTs to rate members of a sexually preferred category as a potential sexual partner compared to non-preferred target categories. This pattern emerged under the time pressure of a response window of 1000 ms and in the absence of any primary and most secondary sexual characteristics, as only heads of targets were presented. Thus, it is unlikely that sexually arousing content and resulting SCID was responsible for the effect replicated here.

General Discussion

Across four experiments, the present study provided consistent evidence that participants showed prolonged response latencies for rating the sexual attractiveness of targets belonging to the sexually preferred category even under conditions of strongly restricted stimulus presentation and speeded responding. Prolonged response latencies emerged (1) in the absence of target pictures (Experiments 1 and 2), (2) under speeded responding (Experiments 3 and 4) and (3) with target stimuli devoid of any primary and most secondary sexual characteristics (Experiment 4). A comparison between the psychometric properties of the standard viewing time and restricted presentation or speeded response conditions generally indicated somewhat improved performance under the restricted conditions.

The findings shed light on the processes underlying the PRELSAT effect. Four plausible explanations were introduced. Deliberate delay to keep a stimulus in view as well as attentional adhesion both require the presence of the stimulus, as is the case in standard viewing time procedures. The results of the present study challenge these explanations. If the PRELSAT effect is based on deliberate watching of arousing stimuli and intentional delay of the response, it should be strongly reduced under restricted conditions, i.e., in the absence of the stimuli, under time pressure, or if no sexually arousing stimuli (faces) are used. A deliberate delay of

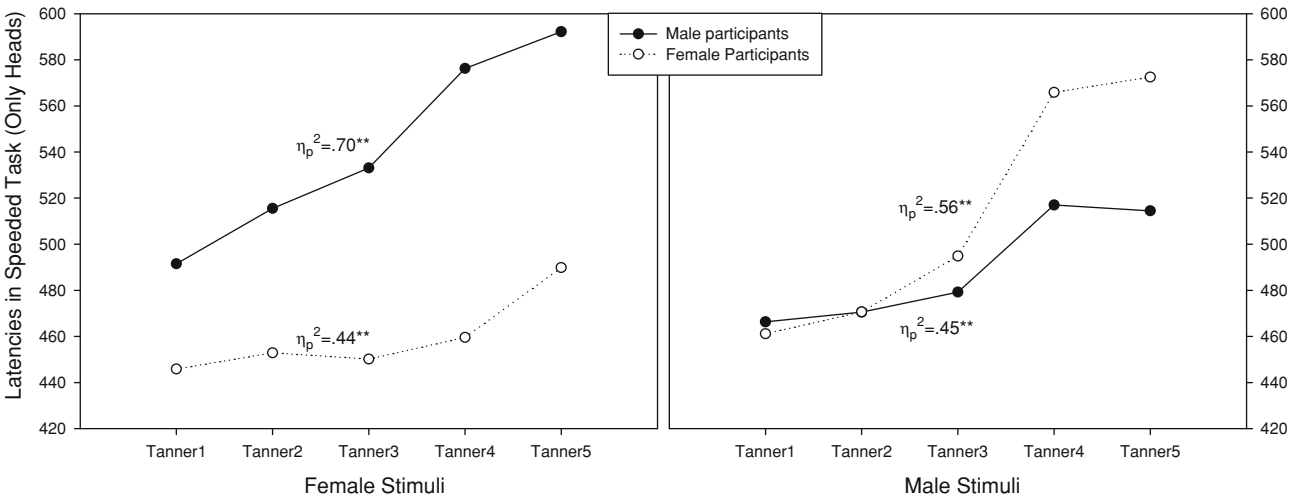


Fig. 4 Response latencies (ms) as a function of Target Age (Tanner Category) and Target Sex (female versus male) for male and female heterosexual participants in a speeded response task with portrait pictures in Experiment 4. Effect sizes for linear within-subject contrasts for Target Age, ** $p < .01$, * $p < .05$

the response due to hedonically rewarding internal images, scripts or schemas connected to sexual content was ruled out by drastically reducing the response window to 1000 ms. However, across all four experiments, the effect of prolonged response latencies for sexually preferred stimuli remained stable and large, and correctly discriminated between participant groups according to their sexual preference.

If it is the case that a sexually attractive stimulus differentially *distracts attention* from the rating task, the effect should be strongly reduced in the absence of stimuli. The attentional adhesion hypothesis is, therefore, difficult to reconcile with the results of restricted stimulus presentation conditions in Experiments 1 and 2, unless one ascribes an attention-grabbing power to an absent stimulus. PRELSAT effects emerged even in the absence of stimuli, leaving some kind of (highly automatic) internal processes resulting from either the brief presentation of sexual stimuli or the specific task to rate sexual attractiveness as the only remaining explanation. SCID, i.e., hesitancy in decision making after priming with sexually explicit images, was ruled out in Experiment 4 by using stimuli devoid of any primary sexual characteristics. Thus, findings from the four experiments suggest that the underlying process functions rather quickly, and also in the absence of stimuli. These are most likely processes of mate identification and potentially resulting schematic processes. We propose two processes that can be reconciled with the results.

Task-Specific Cognitive Processing

It is conceivable that PRELSAT effects are based on cognitive processes resulting from the identification of potential sexual partners, i.e., structural demands of the task. Rating the sexual attractiveness or acceptability of a sexual partner requires to correctly classify the stimulus regarding age and sex before judging the attractiveness. For a positive answer, a heterosexual man first has to confirm that the target person is a woman, i.e., has the appropriate sex (female) and age (postpuberty). Thus, to reach a decision regarding the sexual attractiveness or suitability of the target as a sexual partner, participants need to integrate the three criteria of sex, age, and attractiveness. Participants may test these criteria sequentially ("Is this person at an age appropriate for sexual attractiveness?"; "Is this person of the sex I find attractive?"; "Are the physical features sexually attractive to me?"). The process can be stopped as soon as one feature check results in a negative response. Thus, for all non-preferred targets, a negative outcome of any of the three checks is sufficient to give a low sexual attractiveness score or to reach the decision "no potential sexual partner." Conversely, as long as the outcome of sequential checks is positive, it is necessary to continue target scrutiny until all three criteria are evaluated. Such a systematic processing can account for longer response latencies for sexually attractive targets, as positive identification of sexual mates always require the evaluation of all three criteria, whereas negative decisions require the evaluation of one, two, or three criteria. As the actual decisions participants made in Experiments 3 and 4 showed that not all preferred adults were identified as potential sexual partners (roughly 50–80%), participants had apparently also evaluated the individual attractiveness of targets.

Stimulus-Specific Schematic Processing

Sexually preferred stimuli can only have an effect after they are identified as such. However, as an alternative to the mere identification process, prolonged latencies could also follow from internal processes automatically triggered by the identification of an object as a potential sexual partner. In fact, the previous literature suggested that the underlying processes of viewing time effects are "attentional" (e.g., Kalmus & Beech, 2005), implying processes elicited by sexually attractive stimuli, not by the task. Although attentional adhesion to present stimuli was ruled out as an explanation in Experiments 1 and 2, automatic attention could be directed toward internal representation (e.g., sexual fantasies, scripts, or schemata).

Age Effects for Individuals of the Non-Preferred Sex

These two plausible accounts are both reconcilable with the observed pattern of increasing response latencies, even for adults of the non-preferred sex. However, both accounts would imply two different explanations. From a mere identification perspective, we would expect a target age effect also for the non-preferred target sex (i.e., heterosexual men should be faster to discard male children than adult men), if target age can be identified faster than target sex. Such a main effect of target age for the non-preferred sex is exactly what we find almost across all experiments (for a similar effect for male participants, see Quinsey et al., 1996). From a perspective of schematic processing, such an effect could be explained if adults of the non-preferred sex are more associated with the concept of sexuality and related schemata than children (i.e., the concept of sexuality in heterosexual men is less associated with male children than adult men).

Based on our data neither of the two processes—task-specific processing or stimulus-specific processing—can be ruled out. Future research will have to separate task from stimulus effects to empirically test these two accounts. This was beyond the scope of the present study.

Speeded Response Variant

The present research has replicated that sexually preferred targets elicit longer latencies than non-preferred targets under unrestricted conditions. Contrary to predictions derived from two plausible hypotheses regarding the underlying processes (deliberate delay and attentional adhesion), prolonged response

latencies for sexually preferred targets were not reduced under conditions of restricted stimulus presentation or speeded responding. We therefore argue that, at least under these restricted conditions, other processes cause prolonged response latencies for sexually preferred targets. The processes that cause differential effects for sexually preferred and non-preferred targets take place within approximately 700 ms after the presentation of target stimuli. Both above-mentioned processes might be involved in causing prolonged reaction latencies between 700 ms and the typical latency of up to 5000 or 6000 ms in the standard viewing time paradigm. However, the effect sizes and the criterion validity in the standard version were not superior to the restricted versions. It thus appears that the proportion of viewing time variance that is a valid indicator of sexual preference is rather confined to the early phases of processing. The variance due to latencies beyond 700 ms seems to be largely blurred by noise and non-specific behavior that does not add to the diagnostic value of viewing time. Although our results cannot rule out that deliberate delay does indeed lead to an increase in latencies under standard conditions (which would imply different mechanisms involved in the different variations), the data suggest that this additional time due to delay does not turn Viewing Time into a more valid measure.

This reasoning implies that speeded variants of the viewing time paradigm may be preferable to the standard viewing time measure. Time constraints turn the task into a performance measure that is generally more likely to tap into automatic processes. These may be harder to control and thus less prone to faking compared to controlled processes. Although indirect measures are generally assumed to be more immune to faking than self-reports, PPG has been widely criticized for being liable to deliberate faking by suppression of an erection (Kalmus & Beech, 2005; Konopansky & Konopansky, 2000), most successfully by cognitive deflection. Despite the fact that viewing time measures have been shown to successfully detect even socially sanctioned sexual interest, it has to be established empirically to what extent speeded and standard viewing time procedures are robust against faking.

Female Non-Specificity

Our results also shed some light on the intriguing questions of sex differences in gender specificity of sexual arousal. Previous studies have shown that women generally show less specific reactions to male stimuli compared to female stimuli than men do. Our findings confirm this general pattern in Experiment 2 (women showed no specificity at all) and in Experiment 3 (women showed less specificity than men). When the stimuli were reduced to heads (Experiment 4), this difference vanished and women's latencies differed between male and female stimuli to the same degree as men's. Whether this is an effect of the stimulus reduction or a characteristic of the specific sample is open to future research.

The present study has also shown that faces are sufficient to elicit prolonged response latencies for sexually preferred targets. This result might be particularly important for forensic contexts where it is highly desirable to use non-erotic content for the assessment of deviant sexual preference, for ethical as well as legal reasons (Abel et al., 2001). Future research will have to provide evidence as to whether the restricted and reduced viewing time variants introduced in this study are suitable also for forensic samples (e.g., that pedophiles show longer latencies for faces of children).

References

Abel, G. G., Huffman, J., Warberg, B., & Holland, C. L. (1998). Visual reaction time and plethysmography as measures of sexual interest in child molesters. *Sexual Abuse: A Journal of Research and Treatment, 10*, 81–95.

Abel, G. G., Jordan, A., Hand, C. G., Holland, L. A., & Phipps, A. (2001). Classification models of child molesters utilizing the Abel Assessment for Sexual Interest™. *Child Abuse and Neglect, 25*, 703–718.

Banse, R., Schmidt, A. F., & Clabour, J. (2010). Indirect measures of sexual interest in child sex offenders: A multi-method approach. *Criminal Justice and Behavior, 37*, 319–335.

Brown, M., Amoroso, D. M., Ware, E. E., Pruesse, M., & Pilkey, D. W. (1973). Factors affecting viewing time of pornography. *Journal of Social Psychology, 90*, 125–135.

Chivers, M. L., Rieger, G., Latty, E., & Bailey, J. M. (2004). A sex difference in the specificity of sexual arousal. *Psychological Science, 15*, 736–744.

Chivers, M. L., Seto, M. C., & Blanchard, R. (2007). Gender and sexual orientation differences in sexual response to sexual activities versus gender of actors in sexual films. *Journal of Personality and Social Psychology, 93*, 1108–1121.

Epstein, S. (1990). Cognitive-experiential self-theory. In L. A. Pervin (Ed.), *Handbook of personality theory and research* (pp. 165–192). New York: Guilford.

Flak, V., Beech, A., & Fisher, D. (2007). Forensic assessment of deviant sexual interests: The current position. *Issues in Forensic Psychology, 6*, 70–83.

Freund, K. (1963). A laboratory method of diagnosing predominance of homo- and hetero-erotic interest in the male. *Behaviour Research and Therapy, 1*, 85–93.

Geer, J. H., & Bellard, H. S. (1996). Sexual content induced delays in unprimed lexical decisions: Gender and context effects. *Archives of Sexual Behavior, 25*, 91–107.

Geer, J. H., & Melton, J. S. (1997). Sexual content induced delay with double-entendre words. *Archives of Sexual Behavior, 26*, 295–316.

Gray, N. S., Brown, A. S., MacCulloch, M. J., Smith, J., & Snowden, R. J. (2005). An implicit test of the associations between children and sex in pedophiles. *Journal of Abnormal Psychology, 114*, 304–308.

Gress, L. Z. C. (2005). Viewing time and sexual interest: Another piece in the puzzle. *Journal of Sexual Aggression, 11*, 117–125.

Harris, G. T., Rice, M. E., Quinsey, V. L., & Chaplin, T. C. (1996). Viewing time as a measure of sexual interest among child molesters and heterosexual men. *Behaviour Research and Therapy, 34*, 389–394.

Imhoff, R., Schmidt, A. F., Bernhardt, J., Dierksmeier, A., & Banse, R. (2010). An inkblot for male sexual preference: A semantic variant of the affect misattribution procedure. Manuscript submitted for publication.

Ishai, A. (2007). Sex, beauty and the orbifrontal cortex. *International Journal of Psychophysiology, 63*, 181–185.

Israel, E., & Strassberg, D. S. (2009). Viewing time as an objective measure of sexual interest in heterosexual men and women. *Archives of Sexual Behavior, 38*, 551–558.

Kalmus, E., & Beech, A. R. (2005). Forensic assessments of sexual interest: A review. *Aggression and Violent Behavior, 10*, 193–217.

Karama, S., Lecours, A. R., Leroux, J.-M., Bourgoin, P., Beaudoin, G., Joubert, S., et al. (2002). Areas of brain activation in males and females during viewing of erotic film excerpts. *Human Brain Mapping, 16*, 1–13.

Konopansky, R. J., & Konopansky, A. W. B. (2000). Remaking penile plethysmography. In D. R. Laws, S. M. Hudson, & T. Ward (Eds.), *Remaking relapse prevention with sex offenders* (pp. 257–284). Thousand Oaks, CA: Sage.

Laws, D. R., & Gress, L. Z. C. (2004). Seeing things differently: The viewing time alternative to penile plethysmography. *Legal and Criminological Psychology, 9*, 183–196.

Love, R. E., Sloan, L. R., & Schmidt, M. J. (1976). Viewing pornography and sex guilt: the priggish, the prudent, and the profligate. *Journal of Consulting and Clinical Psychology, 76*, 624–629.

Maner, J. K., Gailliot, M. T., Rouby, D. A., & Miller, S. L. (2007). Can't take my eyes off you: Attentional adhesion to mates and rivals. *Journal of Personality and Social Psychology, 93*, 389–401.

Mouras, H., Stoléru, S., Bittoun, J., Glutron, D., Pélégrini-Issac, M., Paradis, A.-L., et al. (2003). Brain processing of visual sexual stimuli in healthy men: A functional magnetic resonance imaging study. *NeuroImage, 20*, 855–869.

Pacific Psychological Assessment Corporation (PPAC). (2004). *The NRP (Not Real People) stimulus set for assessment of sexual interest*. Victoria, BC: Author.

Ponseti, J., Bosinski, H. A., Wolff, S., Peller, M., Jansen, O., Mehdorn, H. M., et al. (2006). A functional endophenotype for sexual orientation in humans. *NeuroImage, 23*, 825–833.

Quinsey, V., Ketsetzis, M., Earls, C., & Karamanouikan, A. (1996). Viewing time as a measure of sexual interest. *Ethology and Sociobiology, 17*, 341–354.

Redouté, J., Stoléru, S., Grégoire, M.-C., Costes, N., Cinotti, L., Lavenne, F., et al. (2000). Brain processing of visual sexual stimuli in human males. *Human Brain Mapping, 11*, 162–177.

Rosenzweig, S. (1942). The photoscope as an objective device for evaluating sexual interest. *Psychosomatic Medicine, 4*, 150–157.

Rupp, H. A., & Wallen, K. (2008). Sex differences in response to visual sexual stimuli: A review. *Archives of Sexual Behavior, 37*, 206–218.

Sachsenmaier, S. J., & Gress, C. L. Z. (2009). The Abel Assessment for Sexual Interests-2: A critical review. In D. Thornton & D. R. Laws (Eds.), *Cognitive approaches to the assessment of sexual interest in sexual offenders* (pp. 31–57). Chichester, UK: Wiley.

Safron, A., Barch, B., Bailey, J. M., Gitelman, D. R., Parrish, T. B., & Reber, P. J. (2007). Neural correlates of sexual arousal in homosexual and heterosexual men. *Behavioral Neuroscience, 121*, 237–248.

Santtila, P., Mokros, A., Viljanen, K., Koivisto, M., Sandnabba, N. K., & Osterheider, M. (2009). Assessment of sexual interest using a choice reaction time task and priming: A feasibility study. *Legal and Criminological Psychology, 14*, 65–82.

Singer, B. (1984). Conceptualizing sexual arousal and attraction. *Journal of Sex Research, 20*, 230–240.

Spiering, M., Everaerd, W., & Elzinga, B. (2002). Conscious processing of sexual information: Interference caused by sexual primes. *Archives of Sexual Behavior, 31*, 159–164.

Spiering, M., Everaerd, W., & Laan, E. (2004). Conscious processing of sexual information: Mechanisms of appraisal. *Archives of Sexual Behavior, 33*, 369–380.

Stoléru, S., Grégoire, M.-C., Gérard, D., Decety, J., Lafarge, E., Cinotti, L., et al. (1999). Neuroanatomical correlates of visually evoked sexual arousal in human males. *Archives of Sexual Behavior, 28*, 1–21.

Tanner, J. (1978). *Foetus into man: Physical growth from conception to maturity*. Cambridge, MA: Harvard University Press.

Ware, E. E., Brown, M., Amoroso, D. M., Pilkey, D. W., & Pruesse, M. (1972). The semantic meaning of pornographic stimuli for college males. *Canadian Journal of Behavioural Science, 4*, 204–209.

Wiegel, M., Scepkowski, L. A., & Barlow, D. H. (2007). Cognitive-affective processes in sexual arousal and sexual dysfunction. In E. Janssen (Ed.), *The psychophysiology of sex* (pp. 143–165). Bloomington, IN: Indiana University Press.

Wright, L., & Adams, H. (1994). Assessment of sexual preference using choice reaction time task. *Journal of Psychopathology and Behavioral Assessment, 16*, 221–231.

Zamansky, H. S. (1956). A technique for measuring homosexual tendencies. *Journal of Personality, 24*, 436–448.

Arch Sex Behav (2010) 39:1289–1296
DOI 10.1007/s10508-009-9559-6

ORIGINAL PAPER

Sex-Dimorphic Face Shape Preference in Heterosexual and Homosexual Men and Women

Aaron N. Glassenberg · David R. Feinberg · Benedict C. Jones · Anthony C. Little · Lisa M. DeBruine

Received: 4 November 2008 / Revised: 4 September 2009 / Accepted: 5 September 2009 / Published online: 15 October 2009

Abstract Studies have used manipulated faces to test the preferences of heterosexual individuals for sexually dimorphic facial cues. In contrast to previous studies, which have generally excluded homosexual participants, we directly compared homosexual and heterosexual male and female preferences for manipulated sexual dimorphism in faces (homosexual males: $n = 311$; heterosexual males: $n = 215$; homosexual females: $n = 159$; heterosexual females: $n = 218$). Prior studies on sexual orientation and preferences for faces that were paired with masculine and feminine behavioral descriptors suggest that homosexual men prefer more masculine men and that homosexual women demonstrate no preference for either masculinity or femininity in women. In our study, we tested for similarities and differences among heterosexual and homosexual males and females with regard to their preferences for a more specific aspect of faces: sexual dimorphism of face shape. Homosexual men demonstrated stronger preferences for masculinity in male faces than did all of the other groups. Homosexual women demonstrated stronger preferences for masculinity in female faces than did heterosexual women. These results suggest attractiveness judgments of same-sex faces made by homosexual individuals are not a mirror image of those made by heterosexual individuals of the opposite sex. Our data suggest that face preferences of homosexual individuals reflect a system of biologically and socially guided preferences at least as complex as those found among heterosexual individuals.

Keywords Face · Attractiveness · Homosexual · Heterosexual · Masculinity · Femininity · Sociosexual Orientation Inventory

A. N. Glassenberg
Department of Psychology, Harvard University, Cambridge, MA, USA

D. R. Feinberg (✉)
Department of Psychology, Neuroscience, and Behaviour, McMaster University, Hamilton, ON L8S 4K1, Canada
e-mail: feinberg@mcmaster.ca

B. C. Jones · L. M. DeBruine
School of Psychology, Aberdeen University, Aberdeen, Scotland, UK

A. C. Little
Department of Psychology, Stirling University, Stirling, Scotland, UK

Introduction

Studies of preferences for sexual dimorphism in faces among heterosexual raters suggest preferences evolved to maximize potential benefits of mate choices (for reviews, see Feinberg, 2008; Fink & Penton-Voak, 2002; Jones et al., 2008; Little & Perrett, 2002). Similar interpretations of preferences for sexual dimorphism in voices (for review, see Feinberg, 2008) and bodies (Little, Jones, & Burriss, 2007) have also been proposed.

Masculine facial characteristics are positively associated with indices of male health (Rhodes, Chan, Zebrowitz, & Simmons, 2003; Thornhill & Gangestad, 2006) and dominance (Mueller & Mazur, 1996). While masculinity in male faces may cue dominance and health, masculine men are less likely to invest in offspring and relationships than are relatively feminine men (Boothroyd, Jones, Burt, DeBruine, & Perrett, 2008; Burnham et al., 2003; Gray, 2003; Gray et al., 2004; Gray, Kahlenberg, Barrett, Lipson, & Ellison, 2002). Moreover, masculine men are more interested in pursuing short-term relationships than are relatively feminine men (Boothroyd et al., 2008; Rhodes, Simmons, & Peters, 2005). Female preferences for masculinity in male faces are modulated by a

large number of factors, including menstrual cycle phase and state hormone levels (Frost, 1994; Johnston, Hagel, Franklin, Fink, & Grammer, 2001; Jones, Little et al., 2005; Penton-Voak et al., 1999; Welling et al., 2007), hormonal contraceptives (Cornwell et al., 2004; Feinberg, DeBruine, Jones, & Little, 2008; Little, Jones, Penton-Voak, Burt, & Perrett, 2002), body morphology (Penton-Voak et al., 2003; Scarbrough & Johnston, 2005), self-perceived attractiveness (Little, Burt, Penton-Voak, & Perrett, 2001; Little & Mannion, 2006), and sex drive (Welling, Jones, & DeBruine, 2008). Variations in female preferences for masculine men (i.e., those with masculine face and body shapes and those with low vocal frequencies) may function to maximize the benefits of female mate choices (e.g., by increasing offspring health or reducing mate search time) (Feinberg, 2008; Jones et al., 2008).

Potential adaptations are also evident in male preferences for female facial femininity (Jones, DeBruine, Little, & Feinberg, 2007; Perrett et al., 1998; Welling et al., 2008). Facial femininity is positively associated with indices of female reproductive health (Law-Smith et al., 2006; Penton-Voak et al., 2003). Although male preferences for sexually dimorphic face cues are relatively stable in comparison to female preferences, men alter their face preferences based on social feedback (Jones, DeBruine, Little, Conway et al., 2007; Jones et al., 2008; Little, Burriss, Jones, DeBruine, & Caldwell, 2008). Individual differences in male preferences may maximize the benefits of choices by reducing mate search time or by increasing the efficiency with which men allocate their mating resources (Jones, DeBruine, Little, Conway et al., 2007; Jones et al., 2008; Little, Burriss et al., 2008).

Studies of dating advertisements suggest that homosexual men prefer masculine men more than do heterosexual women, and that homosexual women prefer feminine women more than do heterosexual men (Bailey, Kim, Hills, & Linsenmeier, 1997; Child, Low, McCormick, & Cocciarella, 1996; Lippa, 2007). Also, among homosexual and heterosexual men and women, neural reward systems associated with viewing attractive faces (Kampe, Frith, Dolan, & Frith, 2001; O'Doherty et al., 2003) were more active when viewing preferred-sex faces than when viewing non-preferred-sex faces (Ishai, 2007; Kranz & Ishai, 2006). Thus, preferences for facial sexual dimorphism may differ between heterosexual and homosexual individuals. To the best of our knowledge, however, only one previous study has investigated similarities and differences between homosexual and heterosexual male and female preferences for masculinity, assessing preferences for faces paired with vignettes illustrating whether the person was behaviorally masculine or feminine (Bailey et al., 1997). This study found that homosexual men preferred men who were described to be masculine more than they preferred men who were described to be feminine. Homosexual women, on the other hand, showed no consistent bias in their preferences for masculine versus feminine women.

A potential limitation of Bailey et al. (1997) is that the study used vignettes describing behavioral masculinity rather than objective manipulations of sexually dimorphic characteristics. Therefore, it is unclear if the effects reported by Bailey et al. extend to face preferences. Thus, we compared homosexual and heterosexual male and female preferences for masculinized and feminized male and female faces. Following previous studies of preferences for facial masculinity (Little & Hancock, 2002; Little & Mannion, 2006; Penton-Voak, Perrett, & Peirce, 1999; Perrett et al., 1998), face stimuli were manufactured using computer graphics methods that objectively and systematically manipulate sexual dimorphism of two-dimensional shape cues in facial images. Following Bailey et al. (1997), we hypothesized that aggregate analyses of homosexual men would show preferences for male facial masculinity and that homosexual women may not show preferences for facial masculinity or femininity.

Levels of sociosexuality vary between heterosexual and homosexual individuals (Bailey, Gaulin, Agyei, & Gladue, 1994). Among heterosexual individuals, less restricted women (i.e., individuals who are more open to short-term relationships) prefer masculine male bodies (Provost, Kormos, Kosakoski, & Quinsey, 2006) and faces (Waynforth, Delwadia, & Camm, 2005) more than do more restricted individuals (i.e., individuals who are less open to short-term relationships), but these findings do not necessarily extend to female preferences for male facial masculinity. To our knowledge, no previous studies have tested for possible relationships between sociosexuality and male preferences for feminine women. Nor have any studies tested for possible relationships between sociosexuality and preferences for sexual dimorphism among homosexual individuals. Men with increased masculinity and women with increased femininity tend to report less restricted sexual behavior than those with decreased sexual dimorphism (Boothroyd et al., 2008; Rhodes et al., 2005). Therefore, individuals with less restricted sociosexual profiles may prefer increased sexual dimorphism. To determine if people with less restricted sociosexual profiles prefer increased sexual dimorphism, we tested for associations between sociosexuality and preferences for sexual dimorphism in each group of participants (homosexual and heterosexual men and women). We predicted that individuals with less restricted sociosexual profiles would show stronger preferences for masculinity among male faces and femininity among female faces as compared to individuals with more restricted sociosexual profiles.

Method

Participants

A total of 903 participants were recruited online (M age = 32.0 years, SD = 11.3 years). Of these, 215 men (M age =

32.33 years, $SD = 11.11$) and 218 women (M age = 30.55 years, $SD = 9.27$) identified themselves as heterosexual, and 311 men (M age = 34.85 years, $SD = 12.30$) and 159 women (M age = 28.91 years, $SD = 9.85$) identified themselves as homosexual. Sexual orientation was determined by asking participants to select one of seven statements that best described their sexual orientation. The seven statements provided in the survey were taken from the Kinsey scale (Kinsey, Pomeroy, & Martin, 1948).

To separate our population into its four groups (homosexual and heterosexual male and female), we coded the Kinsey scale data such that individuals who rated themselves as "exclusively homosexual" and "predominantly homosexual, only incidentally heterosexual" were classified as homosexual, whereas individuals who rated themselves as "exclusively heterosexual" and "predominantly heterosexual, only incidentally homosexual" were classified as heterosexual. Bisexual people were not the focus of this study, and people who scored along the other two points of the Kinsey scale were not brought into our data analysis.

Our participants came from a number of online sources where we posted advertisements asking for men and women who were interested in helping with a 20-min study on facial attractiveness. A total of 10.7% of participants were recruited through Harvard University and Harvard Business School research websites (of that group, 2.3% were undergraduate students at Harvard University and 8.4% of participants were non-students who were recruited by Harvard Business School to participate in online surveys). A total of 72.5% of our participants found our study while browsing online sites not associated with Harvard. Of those, 56% found the link to our study on the volunteer section of Craigslist, 19.7% of participants came to the study by clicking on Google advertisements, and the remainder of participants came from eight other websites that primarily involved social networking and gay and lesbian discussion (<8% per website). A total of 16.8% of our participants declined to specify where they found out about the study. A preliminary analysis showed that ethnicity, menstrual cycle, relationship context, and hormonal contraceptive use did not have any effect on the dependent measures, so these variables are not discussed further.

Measures and Procedure

We used standard computer graphics methods to manipulate the sexual dimorphism in facial images (Benson & Perrett, 1993; Rowland & Perrett, 1995; Tiddeman, Burt, & Perrett, 2001). Each face was delineated with 179 landmark points. To create endpoints of the manipulations, the shape, color, and texture of the faces of 50 men and 50 women faces were averaged. Next, we created 10 base-faces upon which the manipulations would be applied; each was comprised of three faces of random individuals from a different set of images than that which was used to create the endpoints. By using averages of three individuals, we were able to make each face more representative of the average male or female face, while still maintaining separate identities (Little & Hancock, 2002; Penton-Voak et al., 1999; Perrett et al., 1998). Next, we added ±50% of the linear differences in 2D shape between the male and female prototypes to each base face to manufacture masculinized and feminized versions. This process created 10 pairs of images in total (each pair consisted of a masculinized and feminized version of the same base face, which produced five male pairs and five female pairs). Finally, each individual face was averaged with its mirror image to control for symmetry (Little, DeBruine, Jones, & Feinberg, 2008; Little, Jones et al., 2008). Images were aligned to the average interpupillary distance. In prior studies, these manipulations have been shown to alter masculinity ratings in the predicted way (DeBruine et al., 2006; Penton-Voak, Jacobson, & Trivers, 2004; Penton-Voak & Perrett, 2000; Perrett et al., 1998; Rhodes, Hickford, & Jeffery, 2000). For an example of the face stimuli used in this study, see Fig. 1.

The study was conducted online (www.surveymonkey.com). To recruit participants, advertisements were placed on various websites described above. Research has demonstrated that laboratory-based and internet-based research on facial attractiveness yield consistent results (Feinberg et al., 2005; Jones, DeBruine, Little, Conway et al., 2007; Jones, Perrett et al., 2005; Welling et al., 2008; Wilson & Daly, 2004).

Participants indicated their sexual orientation and answered basic demographic questions. Each participant was shown the 10 pairs of faces sequentially (each pair consisting of a masculinized and feminized version of the same base face as shown in Fig. 1), and asked "Which face do you consider more attractive?" for each pair. The 10 pairs of faces were presented in a fully counterbalanced order. Further, the location of the masculinized and feminized face within each pair was also counterbalanced, meaning that masculinized faces were sometimes displayed on the left and sometimes were displayed on the right. Last, for approximately half of our participants, male face pairs were presented first. Female face pairs were presented first to the remaining participants. Participants who left the study without rating all faces were not included in the analyses.

Following face preference ratings, participants completed the Sociosexual Orientation Inventory (SOI) (Boothroyd et al., 2008; Simpson & Gangestad, 1991). The SOI consists of seven questions that assess tendencies to be inclined towards short- versus long-term relationships. It includes questions concerning both behavior and attitudes (see Table 1 for a list of questions in this scale along with a principal components analysis). Finally, we also asked female participants to answer questions about their use of birth control and the timing of their last period.

Fig. 1 Display of an example of masculinized (*right*) and feminized (*left*) versions of male (*top*) and female (*bottom*) faces used in this study. Although this is a global manipulation of the differences between male and female faces, key features that relate to sexual dimorphism are eyebrow size, jaw shape, and eye size, among other features (Perrett et al., 1998; Perrett, May, & Yoshikawa, 1994)

Results

We calculated the proportion of masculine faces chosen as more attractive than feminine faces separately for male and female faces (preference scores >.5 indicate preferences for masculine faces; preference scores <.5 indicate preferences for feminine faces; preference scores = .5 indicate no preference). We tested whether each group of participants preferred masculinized or feminized face images more than chance when judging the attractiveness of the faces of men and women. One-sample *t*-tests (chance = 0.5) indicated that both heterosexual and homosexual men preferred feminized versions over masculinized versions of the faces of women: heterosexual: $t(214) = -12.65$, $p < .001$; homosexual: $t(306) = -3.53$, $p < .001$. Heterosexual men preferred the faces of feminine men over the faces of masculine men, $t(214) = -3.78$, $p < .001$, whereas homosexual men preferred the faces of men that were masculinized over the faces of men that were feminized, $t(214) = 7.04$, $p < .001$. See Fig. 2 for a comparison between all groups.

One-sample *t*-tests also indicated that both heterosexual and homosexual women preferred the faces of women that were feminized to those that were masculinized: heterosexual: $t(217) = -15.08$, $p < .001$; homosexual: $t(157) = -7.27$, $p < .001$. Heterosexual women showed no preference for either masculinized or feminized male faces, $t(217) < 1$, whereas homosexual women preferred the faces of men that were feminized to the faces of men that were masculinized, $t(157) = -9.33$, $p < .001$.

To compare the strength of facial masculinity preferences, we conducted a 2 (Sex of Rater) × 2 (Sexual Orientation) × 2 (Sex of Stimuli) analysis of variance (ANOVA). The ANOVA yielded a significant three-way interaction between sex of rater, sexual orientation, and sex of stimuli on preferences for masculinized versus feminized faces, $F(1, 894) = 40.1$, $p < .001$.

Independent sample *t*-tests for male and female raters were carried out to interpret the three-way interaction. All Levine's tests for equality of variances were non-significant, with the exception of preferences of heterosexual women for masculinity in the faces of women, $F = 6.79$, $p = .01$. In the case where equal variances were not assumed, Brown–Forsythe *t*-tests with adjusted degrees of freedom were used. In all other cases, actual degrees of freedom were used.

We compared the facial masculinity preferences of heterosexual and homosexual women. These independent samples *t*-tests showed that heterosexual women demonstrated stronger preferences for male facial masculinity faces than did homosexual women, $t(375) = 6.77$, $p < .001$, and that homosexual women demonstrated significantly stronger preferences for female facial masculinity than did heterosexual women, Brown–Forsyth $t(303.38) = -2.92$, $p < .01$.

We compared the facial masculinity preferences of heterosexual and homosexual men. Homosexual men demonstrated stronger facial masculinity preferences than did heterosexual men for both male, $t(520) = -7.42$, $p < .001$, and female faces, $t(520) = -6.72$, $p < .001$.

We compared the preferences for facial masculinity among homosexual men and heterosexual women and compared preferences facial masculinity preferences among homosexual women and heterosexual men. Homosexual men showed stronger preferences for female facial masculinity, Levine's test for inequality of variance: $F = 6.89$, $p < .001$; Brown–Forsyth $t(510.13) = 8.15$, $p < .001$, and male facial masculinity, $t(524) = 4.37$, $p < .001$, than did heterosexual women. Homosexual women preferred male facial masculinity less than did heterosexual men, $t(372) = 4.37$, $p < .001$, and showed stronger preferences for female facial masculinity than did heterosexual men, $t(371) = -2.05$, $p < .05$.

Table 1 Rotated principal components for Sociosexual Orientation Inventory (Simpson & Gangestad, 1991)

Question	Attitudes	Behavior
How many sexual relationships have you had in the last 5 years?	.47	.60
With how many partners have you had sexual relationships in the last 5 years?	.58	.66
How often do you fantasize about having sex?	.47	−.13
How many one night stands have you had?	.42	.58
Sex without love is OK.	.77	−.39
I can imagine myself being comfortable and enjoying "casual sex" with different partners.	.82	−.35
I would have to be closely attached to someone (both emotionally and psychologically) before I could feel comfortable and fully enjoy having sex with him or her?	.76	−.33

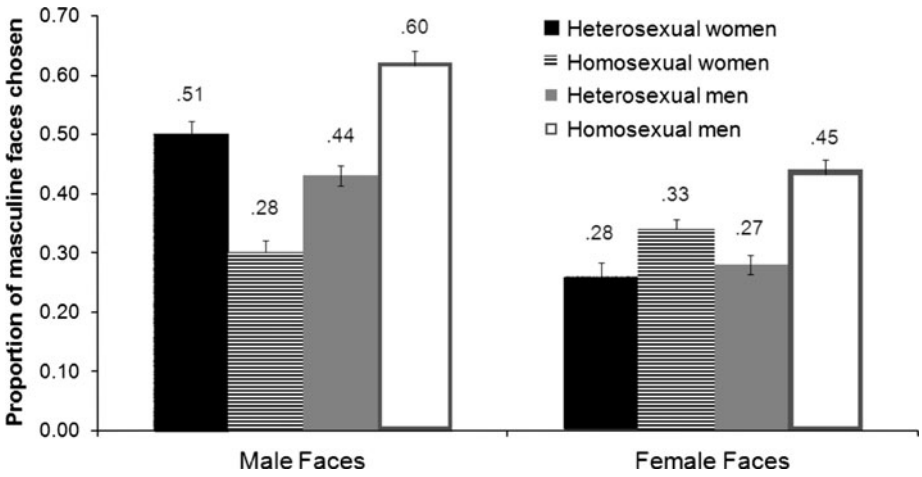


Fig. 2 Differences among heterosexual and homosexual male and female preferences for manipulated sexual dimorphism in male and female faces. *Bars* show means and SEMs

Sociosexual Orientation Inventory and Facial Preferences

Following Webster and Bryan (2007), we entered all SOI variables into a principal components analysis with varimax rotation. The PCA revealed two factors with eigenvalues >1. The first factor (explaining ∼40% of the variance) related most strongly to sexual attitudes, whereas the second factor (explaining ∼22% of the variance) related most strongly to sexual behaviors. These factors were used in subsequent analyses. See Table 1 for factor loadings.

Among heterosexual men, the attitude factor of the SOI was positively correlated with preferences for female facial femininity, $R(125) = .20, p < .05$. Those individuals who were less restricted preferred feminine female faces more than did restricted individuals. Among homosexual men, the sociosexuality attitude factor was also positively correlated with preferences for male facial masculinity, $R(259) = .17, p < .001$. All other correlations (i.e., those among women, and those using the behavior score of the SOI) were non-significant (all $R < .12$, all $p > .075$, which corresponds to sociosexual behavior marginally predicting the proportion of male faces where masculinity was chosen most attractive to heterosexual men).

Discussion

Heterosexual and homosexual men and women demonstrated similarities and differences in their preferences for sexual dimorphism in face shape. Heterosexual and homosexual men and women preferred female facial femininity to masculinity. Regarding ratings of male faces, heterosexual men preferred femininity, whereas homosexual men preferred masculinity. Heterosexual women demonstrated no preference for femininity or masculinity in male faces, whereas homosexual women preferred femininity in male faces. Homosexual men demonstrated stronger preferences for male facial masculinity than did heterosexual men and women. The preferences for female facial femininity among homosexual men were also significantly weaker than the preferences of both heterosexual men and women. Homosexual women preferred female facial femininity more than did homosexual men and

heterosexual women. These findings suggest that while homosexual preferences for same-sex faces and heterosexual preferences for opposite-sex faces are directionally similar, the strengths of these preferences vary depending on sex of the rater, sexual orientation, and the sex of the face being rated.

Bailey et al. (1997) suggested a number of hypotheses that relate to why homosexual masculinity preferences may be similar or different than heterosexual preferences. The first hypothesis is that homosexual people tend to be more open to sex-atypical behavior and, because they are more accepting of sex-atypical sexual partners, may not prefer sexual dimorphism in partners. Neither our data nor those reported by Bailey et al. supported this hypothesis. We found that homosexual men had stronger preferences for male facial masculinity and homosexual women had stronger preferences for female facial femininity than would be expected by chance, indicating homosexual men and women preferred exaggerated sex-typical appearance in potential partners' faces.

The second hypothesis proposed by Bailey et al. (1997) was that homosexual male preferences mirror heterosexual female preferences, whereas homosexual female preferences mirror heterosexual male preferences. This hypothesis appears to be the most relevant to our data. We found that although the direction of preferences was similar between the aforementioned groups, the strength of preferences was not always as predicted by Bailey et al. Homosexual men had stronger preferences for male facial masculinity than did heterosexual women. It may be that this indicates that homosexual men truly prefer male masculinity more than do heterosexual women. However, this interpretation of our results requires careful consideration. Indeed, heterosexual female preferences for male facial masculinity are highly variable, depending on a number of trait, state, physical, and psychological variables described in the introduction. Thus, if generalized preferences for sexual dimorphism among women categorize the individual sample from which they are drawn more than reflecting preferences in general, then there may be no a priori preferences from heterosexual women with which to compare to those from homosexual men or women. Similarly, homosexual women preferred female facial femininity less than did heterosexual men. These results could be qualified by a number of factors related to individual differences in female preferences. Indeed, while we found sociosexuality predicted male but not female preferences, there could be a number of factors that may qualify our results. We discuss this further below.

Another hypothesis proposed by Bailey et al. (1997) is that homosexual male and female preferences for sexual dimorphism in same-sex individuals are solely dictated by their own masculinity. Although the data reported here did not explicitly test this hypothesis, we did investigate how sociosexuality may predict preferences for facial sexual dimorphism. This may serve as an indirect indicator of how individual masculinity might predict facial masculinity preferences because masculine individuals tend to have less restricted sociosexual profiles than do feminine individuals (Boothroyd et al., 2008). Nevertheless, among heterosexual men, those with less restricted sociosexual preferences preferred femininity in both male and female faces to a greater degree than did more restricted men. Among homosexual men, those with less restricted sociosexual preferences preferred male facial masculinity more than did men with more restricted sociosexual preferences. Thus, regardless of sexual orientation, less sociosexually restricted men have stronger preferences for sexual dimorphism in the sex they prefer than more restricted men do. That we found no association between sociosexuality and female preferences for facial sexual dimorphism is consistent with other studies that also found no association between female sociosexuality and preferences for masculinity in male faces (Provost et al., 2006; cf. Waynforth et al., 2005). Given the diversity of sources of individual variation in female preferences for sexual dimorphism (for review, see Feinberg et al., 2008), we encourage future study on individual differences in homosexual and heterosexual preferences for facial masculinity.

In summary, this is the first study to address homosexual individuals' preferences for sexual dimorphism in faces using manipulated stimuli. It appears that preferences of homosexual men are not a reflection of the preferences of heterosexual women and the preferences of homosexual women are not a reflection of the preferences of heterosexual men. Our data suggest that the preferences of homosexual men are more comparable to the preferences of heterosexual men and the preferences of homosexual women are more comparable to the preferences of heterosexual women with regards to sexual dimorphism in faces because of the manner in which individual differences in preferences function. In other words, if homosexual preferences for same-sex faces are reflections of heterosexual preferences for opposite-sex faces, then we would have expected the sociosexual attitudes of homosexual men *not* to predict their preferences for facial masculinity. Instead, we found that the sociosexual attitudes of homosexual men *did* predict their preferences for masculinity in the faces of men. Thus, within each sex, homosexual and heterosexual preferences may be more similar than once thought.

Acknowledgments David R. Feinberg is supported by Canada Foundation for Innovation, The Ministry of Research and Innovation of Ontario: Project # 17515, and The Social Sciences and Humanities Research Council of Canada: Grant #410-2009-2924. Anthony Little is supported by a Royal Society University Research Fellowship.

References

Bailey, J. M., Gaulin, S., Agyei, Y., & Gladue, B. A. (1994). Effects of gender and sexual orientation on evolutionary relevant aspects of human mating psychology. *Journal of Personality and Social Psychology, 66*, 1081–1093.

Bailey, J. M., Kim, P. Y., Hills, A., & Linsenmeier, J. A. W. (1997). Butch, femme, or straight acting? Partner preferences of gay men and lesbians. *Journal of Personality and Social Psychology, 73*, 960–973.

Benson, P. J., & Perrett, D. I. (1993). Extracting prototypical facial images from exemplars. *Perception, 22*, 257–262.

Boothroyd, L. G., Jones, B. C., Burt, D. M., DeBruine, L. M., & Perrett, D. I. (2008). Facial correlates of sociosexuality. *Evolution and Human Behavior, 29*, 211–218.

Burnham, T. C., Chapman, J. F., Gray, P. B., McIntyre, M. H., Lipson, S. F., & Ellison, P. T. (2003). Men in committed, romantic relationships have lower testosterone. *Hormones and Behavior, 44*, 119–122.

Child, M., Low, K. G., McCormick, C. M., & Cocciarella, A. (1996). Personal advertisements of male-to-female transsexuals, homosexual men, and heterosexuals. *Sex Roles, 34*, 447–455.

Cornwell, R. E., Boothroyd, L., Burt, D. M., Feinberg, D. R., Jones, B. C., Little, A. C., et al. (2004). Concordant preferences for opposite-sex signals? Human pheromones and facial characteristics. *Proceedings of the Royal Society of London Series B-Biological Sciences, 271*, 635–640.

DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., et al. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proceedings of the Royal Society of London Series B-Biological Sciences, 273*, 1355–1360.

Feinberg, D. R. (2008). Are human faces and voices ornaments signaling common underlying cues to mate value? *Evolutionary Anthropology, 17*, 112–118.

Feinberg, D., DeBruine, L. M., Jones, B. C., & Little, A. C. (2008). Correlated preferences for men's facial and vocal masculinity. *Evolution and Human Behavior, 29*, 233–241.

Feinberg, D. R., Jones, B. C., DeBruine, L. M., Moore, F. R., Law Smith, M. J., Cornwell, R. E., et al. (2005). The voice and face of woman: One ornament that signals quality? *Evolution and Human Behavior, 26*, 398–408.

Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science, 11*, 154–158.

Frost, P. (1994). Preference for darker faces in photographs at different phases of the menstrual cycle: Preliminary assessment of evidence for a hormonal relationship. *Perceptual and Motor Skills, 79*, 507–514.

Gray, P. B. (2003). Marriage, parenting, and testosterone variation among Kenyan Swahili men. *American Journal of Physical Anthropology, 122*, 279–286.

Gray, P. B., Chapman, J. F., Burnham, T. C., McIntyre, M. H., Lipson, S. F., & Ellison, P. T. (2004). Human male pair bonding and testosterone. *Human Nature, 15*, 119–131.

Gray, P. B., Kahlenberg, S. M., Barrett, E. S., Lipson, S. F., & Ellison, P. T. (2002). Marriage and fatherhood are associated with lower testosterone in males. *Evolution and Human Behavior, 23*, 193–201.

Ishai, A. (2007). Sex, beauty and the orbitofrontal cortex. *International Journal of Psychophysiology, 63*, 181–185.

Johnston, V. S., Hagel, R., Franklin, M., Fink, B., & Grammer, K. (2001). Male facial attractiveness: Evidence for a hormone-mediated adaptive design. *Evolution and Human Behavior, 22*, 251–267.

Jones, B. C., DeBruine, L. M., Little, A. C., Conway, C. A., Welling, L. L. M., & Smith, F. (2007). Sensation seeking and men's face preferences. *Evolution and Human Behavior, 28*, 439–446.

Jones, B. C., DeBruine, L. M., Little, A. C., & Feinberg, D. R. (2007). The valence of experience with faces influences generalized preferences. *Journal of Cultural and Evolutionary Psychology, 5*, 1–11.

Jones, B. C., DeBruine, L. M., Perrett, D. I., Little, A. C., Feinberg, D. R., & Law Smith, M. J. (2008). Effects of menstrual cycle phase on face preferences. *Archives of Sexual Behavior, 37*, 78–84.

Jones, B. C., Little, A. C., Boothroyd, L. G., DeBruine, L. M., Feinberg, D. R., Law Smith, M. J., et al. (2005). Commitment to relationships and preferences for femininity and apparent health in faces are strongest on days of the menstrual cycle when progesterone level is high. *Hormones and Behavior, 48*, 283–290.

Jones, B. C., Perrett, D. I., Little, A. C., Boothroyd, L. G., Cornwell, R. E., Feinberg, D. R., et al. (2005). Menstrual cycle, pregnancy and oral contraceptive use alter attraction to apparent health in faces. *Proceedings of the Royal Society of London Series B-Biological Sciences, 272*, 347–354.

Kampe, K. K. W., Frith, C. D., Dolan, R. J., & Frith, U. (2001). Reward value of attractiveness and gaze—Making eye contact enhances the appeal of a pleasing face, irrespective of gender. *Nature, 413*, 589.

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

Kranz, F., & Ishai, A. (2006). Face perception is modulated by sexual preference. *Current Biology, 16*, 63–68.

Law-Smith, M. J., Perrett, D. I., Jones, B. C., Cornwell, R. E., Moore, F. R., Feinbert, D. R., et al. (2006). Facial appearance is a cue to oestrogen levels in women. *Proceedings of the Royal Society of London Series B-Biological Sciences, 273*, 135–140.

Lippa, R. A. (2007). The preferred traits of mates in a cross-national study of heterosexual and homosexual men and women: An examination of biological and cultural influences. *Archives of Sexual Behavior, 36*, 193–208.

Little, A. C., Burriss, R. P., Jones, B. C., DeBruine, L. M., & Caldwell, C. A. (2008). Social influence in human face preference: Men and women are influenced for long-term but not short-term attractiveness decisions. *Evolution and Human Behavior, 27*, 140–146.

Little, A. C., Burt, D. M., Penton-Voak, I. S., & Perrett, D. I. (2001). Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry in male faces. *Proceedings of the Royal Society of London Series B-Biological Sciences, 268*, 39–44.

Little, A. C., DeBruine, L. M., Jones, B. C., & Feinberg, D. R. (2008). Symmetry and sexual-dimorphism in human faces: Interrelationships in preference suggest a common advert. *Behavioural Ecology, 19*, 902–908.

Little, A. C., & Hancock, P. J. B. (2002). The role of masculinity and distinctiveness in judgments of human male facial attractiveness. *British Journal of Psychology, 93*, 451–464.

Little, A. C., Jones, B. C., & Burriss, R. P. (2007). Preferences for masculinity in male bodies change across the menstrual cycle. *Hormones and Behavior, 51*, 633–639.

Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M., & Perrett, D. I. (2002). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proceedings of the Royal Society of London Series B-Biological Sciences, 269*, 1095–1103.

Little, A. C., Jones, B. C., Waitt, C., Tiddeman, B. P., Feinberg, D. R., Perrett, D. I., et al. (2008). Symmetry is related to sexual dimorphism in faces: Data across culture and species. *PLoS One, 3*, 2106.

Little, A. C., & Mannion, H. (2006). Viewing attractive or unattractive same-sex individuals changes self-rated attractiveness and face preferences in women. *Animal Behaviour, 72*, 981–987.

Little, A., & Perrett, D. I. (2002). Putting beauty back in the eye of the beholder. *Psychologist, 15*, 28–32.

Mueller, U., & Mazur, A. (1996). Facial dominance of West Point cadets as a predictor of later military rank. *Social Forces, 74*, 823–850.

O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: The role of medial orbito-

frontal cortex in facial attractiveness. *Neuropsychologia, 41*, 147–155.

Penton-Voak, I. S., Jacobson, A., & Trivers, R. (2004). Population differences in attractiveness judgements of male and female faces: Comparing British and Jamaican samples. *Evolution and Human Behavior, 25*, 355–370.

Penton-Voak, I. S., Little, A. C., Jones, B. C., Burt, D. M., Tiddeman, B. P., & Perrett, D. I. (2003). Female condition influences preferences for sexual dimorphism in faces of male humans (Homo sapiens). *Journal of Comparative Psychology, 117*, 264–271.

Penton-Voak, I., & Perrett, D. I. (2000). Consistency and individual differences in facial attractiveness judgements: An evolutionary perspective. *Social Research, 67*, 219–244.

Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., et al. (1999). Menstrual cycle alters face preference. *Nature, 399*, 741–742.

Penton-Voak, I. S., Perrett, D. I., & Peirce, J. W. (1999). Computer graphic studies of the role of facial similarity in judgments of attractiveness. *Current Psychology, 18*, 104–117.

Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394*, 884–887.

Perrett, D. I., May, K. A., & Yoshikawa, S. (1994). Facial shape and judgments of female attractiveness. *Nature, 368*, 239–242.

Provost, M. P., Kormos, C., Kosakoski, G., & Quinsey, V. L. (2006). Sociosexuality in women and preference for facial masculinization and somatotype in men. *Archives of Sexual Behavior, 35*, 305–312.

Rhodes, G., Chan, J., Zebrowitz, L. A., & Simmons, L. W. (2003). Does sexual dimorphism in human faces signal health? *Proceedings of the Royal Society of London Series B-Biological Sciences, 270*, S93–S95.

Rhodes, G., Hickford, C., & Jeffery, L. (2000). Sex-typicality and attractiveness: Are supermale and superfemale faces super-attractive? *British Journal of Psychology, 91*, 125–140.

Rhodes, G., Simmons, L. W., & Peters, M. (2005). Attractiveness and sexual behavior: Does attractiveness enhance mating success? *Evolution and Human Behavior, 26*, 186–201.

Rowland, D. A., & Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications, 15*, 70–76.

Scarbrough, P. S., & Johnston, V. S. (2005). Individual differences in women's facial preferences as a function of digit ratio and mental rotation ability. *Evolution and Human Behavior, 26*, 509–526.

Simpson, J. A., & Gangestad, S. W. (1991). Individual differences in sociosexuality: Evidence for convergent and discriminant validity. *Journal of Personality and Social Psychology, 60*, 870–883.

Thornhill, R., & Gangestad, S. W. (2006). Facial sexual dimorphism, developmental stability, and susceptibility to disease in men and women. *Evolution and Human Behavior, 27*, 131–144.

Tiddeman, B. P., Burt, D. M., & Perrett, D. I. (2001). Prototyping and transforming facial texture for perception research. *IEEE Computer Graphics Applications, 21*, 42–50.

Waynforth, D., Delwadia, S., & Camm, M. (2005). The influence of women's mating strategies on preference for masculine facial architecture. *Evolution and Human Behavior, 26*, 409–416.

Webster, G. D., & Bryan, A. (2007). Sociosexual attitudes and behaviors: Why two factors are better than one. *Journal of Research in Personality, 41*, 917–922.

Welling, L. L. M., Jones, B. C., & DeBruine, L. M. (2008). Sex drive is positively associated with women's preferences for sexual dimorphism in men's and women's faces. *Personality and Individual Differences, 44*, 161–170.

Welling, L. L. M., Jones, B. C., DeBruine, L. M., Conway, C. A., Smith, M. J. L., Little, A. C., et al. (2007). Raised salivary testosterone in women is associated with increased attraction to masculine faces. *Hormones and Behavior, 52*, 156–161.

Wilson, M., & Daly, M. (2004). Do pretty women inspire men to discount the future? *Proceedings of the Royal Society of London Series B-Biological Sciences, 271*, S177–S179.

Arch Sex Behav (2010) 39:1297–1304
DOI 10.1007/s10508-009-9527-1

ORIGINAL PAPER

Women's Probability of Conception Is Associated with their Preference for Flirtatious but not Masculine Facial Movement

Edward R. Morrison · Andrew P. Clark · Lisa Gralewski · Neill Campbell · Ian S. Penton-Voak

Received: 3 December 2008 / Revised: 16 June 2009 / Accepted: 26 June 2009 / Published online: 18 August 2009

Abstract Women's preferences for facial structure vary over the menstrual cycle. Little is known, however, as to how preferences for behavior may be influenced by hormonal factors. Here, we demonstrate that social properties of facial motion influence attractiveness judgments in the absence of other cues, and that women's preferences for these displays vary over the menstrual cycle, as has been demonstrated for structural traits of men's faces in static stimuli. We produced shape-standardized facial models that were animated with male movement and assessed for flirtatiousness by 16 women and attractiveness by 47 women. In fertile phases of the menstrual cycle, women showed stronger preferences for flirtatious movement, but not for absolute movement. These data show that women (1) recognize specific mating-relevant social cues in male facial movement and (2) are differentially influenced by these cues at different phases of the menstrual cycle. This preference for flirtatiousness may promote the adaptive allocation of mating effort towards men who are, in turn, likely to respond positively.

Keywords Facial attractiveness · Facial movement · Dynamic stimuli · Mate choice · Menstrual cycle · Flirting

E. R. Morrison
Department of Psychology, University of Portsmouth, Portsmouth, UK

A. P. Clark · I. S. Penton-Voak (✉)
Department of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol BS8 1TU, UK
e-mail: I.S.Penton-Voak@bristol.ac.uk

L. Gralewski · N. Campbell
Department of Computer Science, University of Bristol, Bristol, UK

Introduction

In humans, sexual intercourse can only result in pregnancy in the late follicular phase around ovulation. Several changes have been documented in women's behaviors and preferences over the menstrual cycle that, functionally, seem adaptive in light of this short window of fertility. During the follicular phase, women are more likely to go out to places where they can meet men (Haselton & Gangestad, 2006) and make themselves more attractive through grooming, ornamentation, and dress (Haselton, Mortezaie, Pillsworth, Bleske-Rechek, & Frederick, 2007). Women's gait also changes over the menstrual cycle (Provost, Quinsey, & Troje, 2008a). Furthermore, at these times, women particularly prefer taller men (Pawlowski & Jasienska, 2005), deeper male voices (Feinberg et al., 2006), and the odor of more symmetrical men (Gangestad & Thornhill, 1998) (for a review, see Gangestad & Thornhill, 2008). Sexually dimorphic or symmetrical traits are recognized throughout nature as signaling heritable quality in males (Andersson, 1994; Møller & Thornhill, 1998) and, by preferring them when fertile, females increase the chances of passing them on to their offspring hence increasing their own reproductive success. Cyclic shifts have also been documented for women's preferences for male faces. During the fertile window, women prefer more masculine faces than they do at other times of the month (Johnston, Hagel, Franklin, Fink, & Grammer, 2001; Jones et al., 2005b; Penton-Voak et al., 1999; Penton-Voak & Perrett, 2000; Scarbrough & Johnston, 2005). Additionally, one recent study has found a cyclic preference for facial symmetry in humans, which may be a correlate of facial masculinity (Little, Jones, Burt, & Perrett, 2007).

To date, most studies of menstrual cycle shifts in preferences have concentrated on traits that are sexually dimorphic (height, voice pitch, facial structure) and hence potentially indicate heritable biologically quality. All three of these traits

are associated with testosterone (Dabbs & Mallinger, 1999; Heald et al., 2003; Penton-Voak & Chen, 2004), which is itself thought to be an honest signal of immunocompetence (Folstad & Karter, 1992), although this link lacks direct empirical support to date. The logic behind this so-called "good genes" argument is that preferences evolve for traits that honestly signal mate quality. This quality is to some extent heritable, and can be passed on to offspring, thereby improving reproductive success. In humans, mate quality has many components, but considerable research has concentrated on the possibility that attractiveness may be associated with developmental health, a key aspect of biological quality. There is some evidence for this link (e.g., Hume & Montgomerie, 2001; Langlois et al., 2000; Zebrowitz & Rhodes, 2004). Interestingly, men with attractive faces are more heterozygous at tested loci within the major histocompatibility complex, which is important for immune function (Roberts et al., 2005). Signals to good genes, however, may indicate costs as well as benefits. Although facial masculinity may indicate biological quality, it also drives negative personality attributions, such as coldness and dishonesty (Perrett et al., 1998), and signals less willingness to invest in offspring (Roney, Hanson, Durante, & Maestripieri, 2006). A masculine male may, therefore, represent a high-quality but high-risk mate who offers good genes but less investment than a less masculine male. In the context of these conflicting cues, ovulatory shifts in preferences are often proposed to result from females "trading off" traits desirable in long-term and short-term partners (paternal investment and good genes, respectively) depending on the chance of conception (Gangestad & Thornhill, 2008).

Women, however, may also profit from attending to traits in addition to those associated with "good genes." Cues or signals that indicate a likely return on mating effort may also be preferred, irrespective of the relationship between such cues and traits associated with good genes. Mating effort is costly in terms of time, energy, and resources, and individuals should minimize wasted effort if they are limited by any of these factors. Flirting may signal to both men and women an increased probability of reward for mating effort invested (Mishra, Clark, & Daly, 2007). Although females are usually construed as choosy and males as indiscriminate throughout nature, humans are a largely monogamous species which may select for substantial male choice (Johnstone, Reynolds, & Deutsch, 1996; Kenrick, Sadalla, Groth, & Trost, 1990). A flirting man may, therefore, be signaling interest to a woman, and this cue would be especially valuable to women time-constrained by windows of fertility.

Behavioral cues, such as flirtatiousness, are an important component of attractiveness, but few studies have examined them. Although dynamic mating displays are well-studied in other species (e.g., Bakker, Künzler, & Mazzi, 1999; Maynard Smith, 1956; Rowland, 1995), most facial attractiveness research in humans has relied on photographs or photographic composites. Flirtatious displays are difficult, if not impossible, to study with static stimuli because they are often dynamic by nature, and recent studies indicate that the attractiveness of the same face in a static photograph and a dynamic video do not necessarily correlate (Lander, 2008; Penton-Voak & Chang, 2008; Rubenstein, 2005). Flirtatious displays may also be linked to good genes in humans, although this idea is speculative. Simpson, Gangestad, and Biek (1993) found that symmetrical men engaged in different courtship tactics than asymmetrical men. It is worth noting that testosterone is associated with both anatomical and behavioral displays in a wide variety of species (Adkins-Regan, 2005). In birds, for example, testosterone treatment can restore castrated cockerels' crowing and sexual behavior and can even bring female canaries into song (for a review, see Fusani, 2008). The possibility that preferences for dynamic behavioral traits are influenced by menstrual cycle phase has received little research attention. Gangestad, Simpson, Cousins, Garver-Apgar, and Christensen (2004) did examine preferences for videotaped behavioral displays, showing that women's preferences for directly competitive courtship behavior were heightened during fertile phases of the menstrual cycle when the men were being judged as potential short-term, but not long-term, partners. Here, directly competitive courtship behavior was defined from factor analysis of men's behavioral tactics and nonverbal behaviors to include things like derogation of a male competitor, asserting their own superiority, and lack of laughing. That study, however, was not concerned with examining the dynamic aspects of the behavior alone, as video footage contained both dynamic and structural cues (although there was an attempt to statistically control for static face attractiveness in this study, there were potential interactions between dynamic and static cues that could not be controlled by these means).

In order to investigate dynamic cues, it is useful to isolate them from potential confounding variables. Motion capture is an important tool in this respect and can provide a clear insight into the importance of motion itself in attractiveness judgments in humans. For example, Brown et al. (2005) used motion capture to demonstrate an association between body symmetry and dance quality. Morrison, Gralewski, Campbell, and Penton-Voak (2007) used a form of motion capture to examine facial dynamics and showed that sex-typical facial motion in females was rated attractive by males and females. In this study, however, sex typical (i.e., masculine) facial movement was not related to attractiveness for male targets, yet both male and female raters agreed on the male faces that were attractive based on their movement, suggesting that some (as yet unidentified) quality or qualities in male facial movement drive attractiveness judgments.

In the current study, we investigated whether women's preferences for two properties of facial movement varied over the menstrual cycle. First, we assessed whether preferences for sex-typicality in facial movement varied in a similar way to the well demonstrated shifting preferences for structural masculinity. Such an effect has recently been demonstrated by Provost, Troje, and Quinsey (2008b), who isolated the biological motion of walkers, showing that fertile women particularly preferred masculinity in shape-controlled point-light displays. The large anatomical differences in, for example, pelvic structure, lead to easily observable sex differences in walking style. Although anatomical constraints that lead to sex differences in facial movement are less obvious than in gait, there are nevertheless reliable cues to sex in facial movement alone (Hill & Johnston, 2001; Morrison et al., 2007). If preferences for masculinity in facial movement are analogous to preferences for masculinity in the structural properties of faces and the dynamic properties of gait, we should expect preferences for masculine movement to be highest in the follicular phase of the menstrual cycle.

Secondly, we assessed whether preferences for flirtatious facial movement, which signals likely reward for mating effort, is more important to women when their conception probability is high. To test these ideas, we measured women's preferences for flirtatious dynamic facial displays as a function of their likelihood to conceive following intercourse, using actuarial data from an earlier study (Wilcox, Dunson, Weinberg, Trussell, & Baird, 2001). Conception probability is a useful variable when investigating the adaptiveness of menstrual shifts in behavior and preferences. The mechanism behind such shifts is presumed to be hormonal, and there is evidence that women's facial preference shifts are associated with days of the menstrual cycle when levels of estrogen (Jones et al., 2005a) and progesterone (DeBruine, Jones, & Perrett, 2005) are predicted to be high. We specifically investigated women's preferences for short-term attractiveness judgments, as menstrual cycle effects are particularly predicted in this context (e.g., Jones et al., 2005b; Penton-Voak et al., 1999).

Method

Participants

Facial Animations

Facial animations were the same as those in Morrison et al. (2007). Briefly, we obtained facial motion data from 1-min video sequences of five men in conversation with an interviewer about their upcoming Christmas holidays. These men were young, Caucasian undergraduates recruited at Stirling University.

Flirtatiousness Ratings

Participants were 16 psychology students who received course credit for participation.

Attractiveness Ratings

A further set of 47 women (M age = 19.8 years), none of whom were taking the contraceptive pill, rated the same animations.

Procedure

Facial Animations

Facial landmark points were labeled semi-automatically with 30 points per frame. The movement of these points was then transferred to an androgynous line face and then split into sequences of 10 s, resulting in 30 clips. For further technical details of these computer graphic techniques, see Gralewski, Campbell, Morrison, and Penton-Voak (2006). For each 10 s clip, the total amount of movement was estimated by calculating the distance travelled by the centroid point (the imaginary point representing the average spatial position of each landmark) across each frame of the sequence (M = 261 pixels, SD = 50.1; Morrison et al., 2007). This earlier study also provided a masculinity score for each clip, which was the number of times a given clip was correctly identified as male in the previous study. Here, 108 raters (68 female, 40 male) viewed the 30 male clips in the current study and 30 female clips constructed using identical techniques in random order, and decided whether each was male or female (M of male clips = 65.8/108 possible male identifications, SD = 7.5, no sex of rater effects, see Morrison et al., 2007, for full details of this study). The study was approved by the ethics committee of the Department of Experimental Psychology, University of Bristol.

Flirtatiousness Ratings

Participants viewed the animations in random order and rated the flirtatiousness of each clip from 1 (not very flirtatious) to 7 (very flirtatious).

Attractiveness Ratings

Participants viewed the same clips in random order and rated their short-term attractiveness on a 1 to 7 scale. Participants were given the following definition: "By 'short-term attractiveness,' we refer to the type of person who would be attractive to you in a short-term relationship. This implies that the relationship may not last a long time. Examples of this type of relationship would include a single date accepted on the spur

of the moment, an affair within a long-term relationship, or the possibility of a one-night stand." All participants were aware that they were viewing male faces.

Women were asked how many days it had been since the onset of their last period, and from this we calculated each participant's probability of conception from a single act of unprotected intercourse using benchmark data from Wilcox et al. (2001), who provided likelihood of conception from a single act of intercourse for each day of the menstrual cycle based on a study of 221 women who were attempting to conceive.

Measures

Inter-rater agreement for the flirtatiousness ratings was high (Cronbach's $\alpha = 0.79$). We therefore averaged the ratings, giving each clip a single flirtatiousness score ($M = 4$, $SD = 0.95$). We were interested in the relationship among flirtatiousness, masculinity, and attractiveness of the clips, and the influence of the menstrual cycle on this relationship, and performed simple correlations to assess these associations. To assess individual participant's preferences for flirtatiousness, we performed a median split on the animations to divide them into the 15 most flirtatious and the 15 least flirtatious. We then averaged each woman's attractiveness ratings for the animations in these groups to give her two scores. The difference between each woman's average for the high flirtatiousness and low flirtatiousness groups represented her "preference for flirtatiousness." We then repeated this procedure for a median split on masculinity to calculate each woman's "preference for masculinity." For use as a control measure, we also calculated a "preference for movement" score in the same way using a median split on the total amount of movement in each of the 30 clips.

One potential weakness with simple correlations is that data must be aggregated to create the dependent variables. We therefore performed another analysis involving hierarchical linear modeling, which allows each individual rating to be included by nesting each rating within each female rater. This approach also allowed us to consider the effects of flirtatiousness, masculinity, and movement simultaneously because all three are included in a regression equation for each rater. Another advantage is that hierarchical modeling can test for main effects of conception likelihood, i.e., do women give higher or lower ratings generally when fertile? Hierarchical linear modeling is a form of regression in which some data at one level share error terms and hence are not independent. This shared variance can be taken into account by organizing the related Level 1 data under Level 2 units of analysis. At Level 1 of our model was the relationship between flirtatiousness, total movement, and masculinity of each clip as independent variables and the individual woman's attractiveness rating of that clip as the dependent variable. Thus, there was a population of regression slopes each representing one woman's ratings. Level 2 of the model examined the effect of the rater variable—each woman's probability of conception—on the intercept and slope of the regression lines. All p-values are two-tailed.

Results

Conception Probabilities

Using data from Wilcox et al. (2001; Table 1, column 1, estimated mean probability of pregnancy of following a single act of unprotected intercourse), we estimated that the mean probability of conception of participants when tested in our study was 2.7% ($SD = 3.0\%$; range, 0–8.6%). In the late follicular phase (days 7–15), this value was 4.8% ($SD = 2.9\%$, $n = 19$), and the combined conception probability during menses (days 1–6, $n = 14$) and the luteal phase (days 16–29, $n = 14$) was 1.3% ($SD = 2.1\%$). These values were significantly different $t(45) = -4.9$, $p < .001$. These data are as would be expected from a random sample of naturally cycling women. Wilcox et al. (2001) also split their sample to provide conception probability data for women with regular and irregular cycles. Substituting either of these additional data sets for the combined data in our analyses had no influence on the pattern or significance of any effects.

First Order Correlational Analysis

To assess whether conception probability was associated with preferences for flirtatiousness, masculinity, and total movement, we performed two Spearman's nonparametric correlations. The correlation between flirtatiousness preference and conception likelihood was positive and significant ($r_s = .44$, $p = .002$, $n = 47$). There was no significant association between the probability of conception and preference for masculinity ($r_s = -.19, n = 47$) or preference for overall movement in the clips ($r_s = .22, n = 47$). See Fig. 1 for scatterplots of these relationships.

Hierarchical Regression Analysis

In our hierarchical model, the slopes of the Level 1 relationship represented each woman's preference for flirtatious movement, masculine movement, or total amount of movement. The steepness of the flirtatiousness slopes was positively associated with each woman's probability of conception, $\gamma = 0.10$, $t(45) = 2.01$, $p = .043$, meaning that when more fertile, women found flirtatious faces especially attractive. The intercepts of these slopes were not affected by probability of conception, $\gamma = 0.10, t(45) = 1.42, p = .16$, meaning that women did not give higher or lower ratings overall

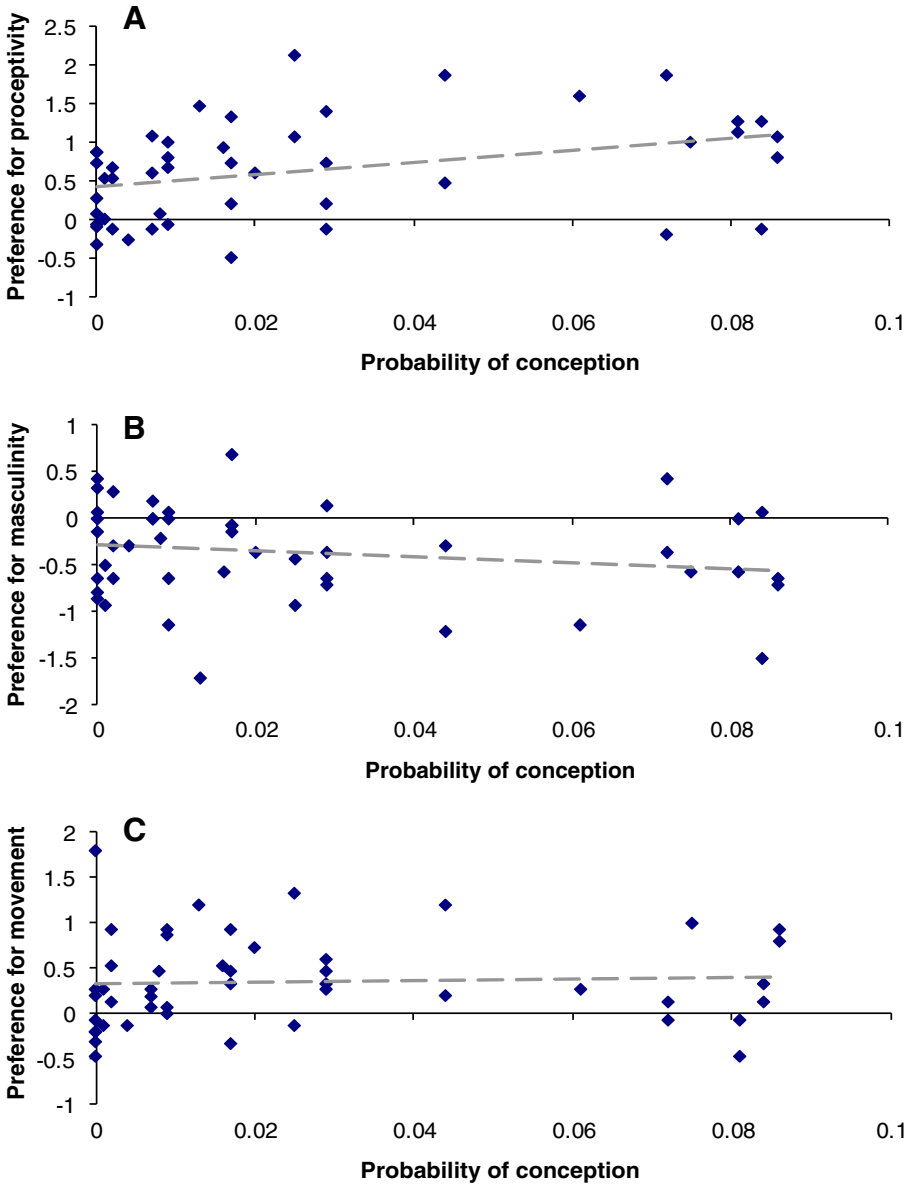


Fig. 1 Associations between the conception probability for flirtatious movement (Panel **A**), masculine movement (Panel **B**), and total movement in dynamic male faces (Panel **C**)

when fertile. There were no such effects for women's preference for masculinity ($\gamma = -0.02, t(45) < 1$), or movement ($\gamma = -0.05, t(45) = -1.24$).

Flirtatious animations were attractive generally, as evidenced by a main effect at Level 1 in the model, $\gamma = 0.24$, $t(45) = 6.10, p < .001$. This result would be unsurprising if it were just the result of a halo effect whereby those animations that moved more were rated more attractive (as in Morrison et al., 2007) and more flirtatious. But this was not the case, since the model also included and controlled for the amount of movement in each animation. Masculinity was negatively associated with flirtatiousness ($r_p = 0.40$, $n = 30$, $p = .03$), but this relationship disappeared when controlling for the amount of movement in each clip (partial correlation $r = -0.10$, $df = 27$). The intraclass correlation coefficient for the null model was 0.23, meaning that 23% of the variances in attractiveness ratings could be attributed to differences between the raters.

Discussion

Our data showed that women's preference for flirtatious facial movement increased at times of the menstrual cycle when they are more likely to conceive following intercourse. This finding was consistent with a growing body of evidence that female preferences for men's traits undergo menstrual cycle shifts, but was the first to examine facial movement independently of structural cues. Interestingly, the same menstrual shift was not apparent in terms of women's preference for the amount of movement in faces. Women did not seem to be merely responding to a low level perceptual variable like

amount of movement. Instead, participants were differentially responding to the social value of the signal.

We found no evidence that women's preferences shifted towards masculine motion when they are fertile. This is at odds with findings using both static faces (e.g., Penton-Voak et al., 1999) and whole-body motion (Provost et al., 2008b). With static faces, masculinity is defined in a variety of ways, but all relate to either measured or perceived sex differences in facial shape (for a discussion of these methods, see DeBruine et al., 2006), which are in turn related to hormone levels (Penton-Voak & Chen, 2004; Pound, Penton-Voak, & Surridge, 2009). For the Provost et al. (2008b) stimuli, masculinity denoted sex differences in motion identified from Fourier analysis of point-light walkers and principal components analysis. Hormones affect body morphology and hence presumably gait, although this link is not directly established. In our case, masculinity was operationalized simply by identifying how often each dynamic clip was identified as male in a forced choice decision task, and no direct link between endocrine activity and facial movement has been hypothesized or established. It is possible that this difference in definition accounts for our lack of a finding, but in this context it is worth noting that most alternative methods of estimating static face masculinity show strong relationships to each other and outcome variables of interest (DeBruine et al., 2006). A more likely explanation may lie in the differences in social attributions made to static and dynamic masculinity. Although masculine facial structure is often attributed with negative personality attributions of coldness and dominance, it is also associated with some stereotypically attractive features, such as a prominent jaw line. Unlike masculinity in static faces, masculine movement is not readily associated with any positive features, as the major sex difference in facial movement is one of degree, with males moving their faces less than women overall (Morrison et al., 2007). This idea was supported by our finding that flirtatious and masculine movement were negatively correlated, but this relationship was mediated by the amount of movement. The relative lack of absolute movement found in our "masculine" dynamic displays (in the operationalization of masculinity we used here) may not signal biological quality so much as a social disengagement, which is unlikely to be considered attractive in any context.

Our definition of masculinity was based on social perception rather than rated masculinity (asking participants how masculine each clip looked) or the mathematical differences between male and female movement (e.g., based on Fourier analysis). This is to allow comparison with the results of Berry (1991) and Hill and Johnston (2001). Rated masculinity is also subjective and may be prone to halo effects based on attractiveness, and although a Fourier-based analysis of sex differences in movement would be interesting, a key issue is whether viewers can accurately perceive sex in dynamic facial displays. Therefore, ratings from an objective sex-identification task are a logical first step in investigating masculinity cues that lead to perceptual differences in stimuli. Correct identification of the sex of the male clips was 61%. This is not dissimilar to the values of 68% from Berry (1991) and about 60% from Hill and Johnston (2001), and reflects the difficulty of the task. This is not surprising given the apparent lack of anatomical constraints on facial movement; men and women can move their faces in any way they want, so it is interesting that facial movement was nevertheless diagnostic of sex.

There were some methodological considerations worth noting in our study. We used the forward counting method to estimate each woman's position in the menstrual cycle from the onset of last menstruation. This method is not as accurate as the reverse counting method because of the variability in the length of the follicular compared to the luteal phase. However, any inaccuracy will only have added noise to our data rather than systematic bias, and it is therefore noteworthy that we nevertheless found a robust effect with this relatively crude method. Furthermore, although the chance of conception may have predicted preferences for flirtatious facial movement, it is clearly not the mechanism underlying the preference shifts. After all, markedly different hormonal profiles can underlie the same mean conception probability. Recent studies examining preferences for masculinity in male faces have attempted to explicitly examine the endocrine correlates of preferences for artificially manipulated masculinity in faces (Welling et al., 2007) and preferences for male faces from men with measured testosterone (Roney & Simmons, 2008). These studies are somewhat contradictory, with the earlier study implicating women's testosterone levels in shifting preferences, while the latter suggests a key role for estradiol. Nevertheless, both studies suggest that cyclic shifts in hormone levels are closely associated with shifting preferences. Future research could search explicitly for associated hormonal fluctuations by measuring levels directly.

Determining what benefit women may actually gain by increasing their attraction to flirtatiousness during fertile peaks remains unresolved. Flirtatiousness may signal both direct benefits, via increased probability of reward for invested mating effort, and indirect benefits (i.e., good genes), via the costs involved in developing the necessary resources for producing skilled social behavior (cf. links between testosterone and mating display in other species). Menstrual cycle shifts in preferences are often described as favoring cues for indirect benefits, because conception is the only time when these benefits will actually be realized (although see Jones et al., 2005b for an alternative, although not necessarily incompatible exception, to this generalization). In this case, however, the direct benefits of conserving mating effort when faced with cues indicating that such effort will be wasted will also be of particular value to mid-cycle women, as the time

available during the fertile window is limited. Neither form of benefit can be excluded by our results, and it should be noted that they do not represent mutually exclusive alternatives; menstrual cycle shifts could serve both adaptive goals simultaneously (see Jones et al., 2008, for a related argument).

Studies of facial motion may be particularly important for menstrual cycle research, because the dynamics of a face can change quickly, unlike the static properties. For example, women can take advantage of a momentary opportunity when they see a man who is signaling interest by flirting. Improvements in motion capture technology mean that dynamic signals in the face can be investigated. Future investigations into cyclic changes in women's preferences may benefit from the consideration of such transient traits which may be important for potential exploitation of strategic mate choice opportunities.

Acknowledgements ERM is supported by an ESRC PhD studentship. This work was partly funded by ESRC grant RES-062-23-0358.

References

Adkins-Regan, E. (2005). *Hormones and animal social behavior*. Princeton, NJ: Princeton University Press.

Andersson, M. (1994). *Sexual selection*. Princeton, NJ: Princeton University Press.

Bakker, T. C. M., Künzler, R., & Mazzi, D. (1999). Condition-related mate choice in sticklebacks. *Nature, 401*, 234.

Berry, D. S. (1991). Child and adult sensitivity to gender information in patterns of facial motion. *Ecological Psychology, 3*, 349–366.

Brown, W. M., Cronk, L., Grochow, K., Jacobson, A., Liu, C. K., Popovic, Z., et al. (2005). Dance reveals symmetry especially in young men. *Nature, 438*, 1148–1150.

Dabbs, J. M., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences, 27*, 801–804.

DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., et al. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proceedings of the Royal Society B, 273*, 1355–1360.

DeBruine, L. M., Jones, B. B., & Perrett, D. I. (2005). Women's attractiveness judgments of self-resembling faces change across the menstrual cycle. *Hormones and Behavior, 47*, 379–383.

Feinberg, D. R., Jones, B. C., Law Smith, M. J. L., Moore, F. R., DeBruine, L. M., Cornwell, R. E., et al. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior, 49*, 215–222.

Folstad, I., & Karter, A. J. (1992). Parasites, bright males, and the immunocompetence handicap. *American Naturalist, 139*, 603–622.

Fusani, L. (2008). Testosterone control of male courtship in birds. *Hormones and Behavior, 54*, 227–233.

Gangestad, S. W., Simpson, J. A., Cousins, A. J., Garver-Apgar, C. E., & Christensen, P. N. (2004). Women's preferences for male behavioral displays change across the menstrual cycle. *Psychological Science, 15*, 203–207.

Gangestad, S. W., & Thornhill, R. (1998). Menstrual cycle variation in women's preferences for the scent of symmetrical men. *Proceedings of the Royal Society B, 265*, 927–933.

Gangestad, S. W., & Thornhill, R. (2008). Human oestrus. *Proceedings of the Royal Society B, 275*, 991–1000.

Gralewski, L., Campbell, N., Morrison, E., & Penton-Voak, I. (2006). Analysis of facial dynamics using a tensor framework. *Journal of Multimedia, 6*, 10–21.

Haselton, M. G., & Gangestad, S. W. (2006). Conditional expression of women's desires and men's mate guarding across the ovulatory cycle. *Hormones and Behavior, 49*, 509–518.

Haselton, M. G., Mortezaie, M., Pillsworth, E. G., Bleske-Rechek, A., & Frederick, D. A. (2007). Ovulatory shifts in human female ornamentation: Near ovulation, women dress to impress. *Hormones and Behavior, 51*, 40–45.

Heald, A. H., Ivison, F., Anderson, S. G., Cruickshank, K., Laing, I., & Gibson, J. M. (2003). Significant ethnic variation in total and free testosterone concentration. *Clinical Endocrinology, 58*, 262–266.

Hill, H., & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology, 11*, 880–885.

Hume, D. K., & Montgomerie, R. (2001). Facial attractiveness signals different aspects of "quality" in women and men. *Evolution and Human Behavior, 22*, 93–112.

Johnston, V. S., Hagel, R., Franklin, M., Fink, B., & Grammer, K. (2001). Male facial attractiveness: Evidence for hormone-mediated adaptive design. *Evolution and Human Behavior, 22*, 251–267.

Johnstone, R. A., Reynolds, J. D., & Deutsch, J. C. (1996). Mutual mate choice and sex differences in choosiness. *Evolution, 50*, 1382–1391.

Jones, B. C., DeBruine, L. M., Perrett, D. I., Little, A. C., Feinberg, D. R., & Law Smith, M. J. (2008). Effects of menstrual cycle phase on face preferences. *Archives of Sexual Behavior, 37*, 78–84.

Jones, B. C., Little, A. C., Boothroyd, L., DeBruine, L. M., Feinberg, D. R., Law Smith, M. J., et al. (2005a). Commitment to relationships and preferences for femininity and apparent health in faces are strongest on days of the menstrual cycle when progesterone level is high. *Hormones and Behavior, 48*, 283–290.

Jones, B. C., Perrett, D. I., Little, A. C., Boothroyd, L., Cornwell, R. E., Feinberg, D. E., et al. (2005b). Menstrual cycle, pregnancy and oral contraceptive use alter attraction to apparent health in faces. *Proceedings of the Royal Society B, 272*, 347–354.

Kenrick, D. T., Sadalla, E. K., Groth, G., & Trost, M. R. (1990). Evolution, traits, and the stages of human courtship: Qualifying the parental investment model. *Journal of Personality, 58*, 97–116.

Lander, K. (2008). Relating vocal and visual attractiveness for moving and static faces. *Animal Behaviour, 75*, 817–822.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin, 126*, 390–423.

Little, A. C., Jones, B. C., Burt, D. M., & Perrett, D. I. (2007). Preferences for symmetry in faces change across the menstrual cycle. *Biological Psychology, 76*, 209–216.

Maynard Smith, J. (1956). Fertility, mating behaviour and sexual selection in *Drosophila subobscura*. *Journal of Genetics, 54*, 261–279.

Mishra, S., Clark, A. P., & Daly, M. (2007). One woman's behavior affects the attractiveness of others. *Evolution and Human Behavior, 28*, 145–149.

Møller, A. P., & Thornhill, R. (1998). Bilateral symmetry and sexual selection: A meta-analysis. *American Naturalist, 151*, 174–192.

Morrison, E. R., Gralewski, L., Campbell, N., & Penton-Voak, I. S. (2007). Facial movement varies by sex and is related to attractiveness. *Evolution and Human Behavior, 28*, 186–192.

Pawlowski, B., & Jasienska, G. (2005). Women's preferences for sexual dimorphism in height depend on menstrual cycle phase and expected duration of relationship. *Biological Psychology, 70*, 38–43.

Penton-Voak, I. S., & Chang, H. (2008). Attractiveness judgments of individuals vary across emotional expression and movement conditions. *Journal of Cultural and Evolutionary Psychology, 6*, 89–100.

Penton-Voak, I. S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior, 25*, 229–241.

Penton-Voak, I. S., & Perrett, D. I. (2000). Female preference for male faces changes cyclically: Further evidence. *Evolution and Human Behavior, 21*, 39–48.

Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., et al. (1999). Menstrual cycle alters face preference. *Nature, 399*, 741–742.

Perrett, D. I., Lee, K. J., Penton-Voak, I. S., Rowland, D. R., Yoshikawa, S., Burt, D. M., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394*, 884–887.

Pound, N., Penton-Voak, I. S., & Surridge, A. K. (2009). Testosterone responses to competition in men are related to facial masculinity. *Proceedings of the Royal Society B, 276*, 153–159.

Provost, M. P., Quinsey, V. L., & Troje, N. F. (2008a). Differences in gait across the menstrual cycle and their attractiveness to men. *Archives of Sexual Behavior, 37*, 598–604.

Provost, M. P., Troje, N. F., & Quinsey, V. L. (2008b). Short-term mating strategies and attraction to masculinity in point-light walkers. *Evolution and Human Behavior, 29*, 65–69.

Roberts, S. C., Little, A. C., Gosling, L. M., Perrett, D. I., Jones, B. C., Carter, V., et al. (2005). MHC-heterozygosity and human facial attractiveness. *Evolution and Human Behavior, 26*, 213–226.

Roney, J. R., Hanson, K. N., Durante, K. M., & Maestripieri, D. (2006). Reading men's faces: Women's mate attractiveness judgments track men's testosterone and interest in infants. *Proceedings of the Royal Society B, 273*, 2169–2175.

Roney, J. R., & Simmons, Z. L. (2008). Women's estradiol predicts preference for facial cues of men's testosterone. *Hormones and Behavior, 53*, 14–19.

Rowland, W. J. (1995). Do female sticklebacks care about male courtship vigour? Manipulation of display tempo using video playback. *Behaviour, 132*, 951–961.

Rubenstein, A. J. (2005). Variation in perceived attractiveness. *Psychological Science, 16*, 759–762.

Scarbrough, P. S., & Johnston, V. S. (2005). Individual differences in women's facial preferences as a function of digit ratio and mental rotation ability. *Evolution and Human Behavior, 26*, 509–526.

Simpson, J. A., Gangestad, S. W., & Biek, M. (1993). Personality and nonverbal social behaviour: An ethological perspective of relationship initiation. *Journal of Experimental Social Psychology, 29*, 434–461.

Welling, L. L. M., Jones, B. C., DeBruine, L. M., Conway, C. A., Law Smith, M. J., Little, A. C., et al. (2007). Raised salivary testosterone in women is associated with increased attraction to masculine faces. *Hormones and Behavior, 52*, 156–161.

Wilcox, A. J., Dunson, D. B., Weinberg, C. R., Trussell, J., & Baird, D. D. (2001). Likelihood of conception with a single act of intercourse: Providing benchmark rates for assessment of post-coital contraceptives. *Contraception, 63*, 211–215.

Zebrowitz, L. A., & Rhodes, G. (2004). Sensitivity to "bad genes" and the anomalous face overgeneralization effect: Cue validity, cue utilization, and accuracy in judging intelligence and health. *Journal of Nonverbal Behavior, 28*, 167–185.

Arch Sex Behav (2010) 39:1305–1319
DOI 10.1007/s10508-009-9498-2

ORIGINAL PAPER

Latina Mother–Daughter Dyads: Relations Between Attachment and Sexual Behavior Under the Influence of Alcohol or Drugs

Mario De La Rosa · Frank R. Dillon · Patria Rojas · Seth J. Schwartz · Rui Duan

Received: 17 September 2008 / Revised: 5 March 2009 / Accepted: 23 March 2009 / Published online: 28 April 2009

Abstract Associations among mother–daughter attachment, mother and daughter substance abuse, and daughter's sexual behavior under the influence of drugs and alcohol were investigated among 158 adult U.S. Latina daughters. Latina daughters were sampled from four mother–daughter dyad types: substance abusing mother and daughter, substance abusing mother only, substance abusing daughter only, and non-substance-abusing mother and daughter. Substance abusing daughters with substance abusing mothers, and daughters who were less strongly attached to their mothers, reported more sex under the influence of drugs. Age, marital status, substance abuse, and mother's substance abuse all influenced the daughter's sex under the influence of alcohol. An unexpected positive association between attachment and sex under the influence of alcohol was found for daughters who were more closely attached to a substance abusing mother. Implications for future research, and HIV/AIDS and drug prevention and treatment programs for Latinas are discussed.

Keywords Mother–daughter attachment · Alcohol and drug intoxication · Substance abuse · HIV risk behaviors · Latinas

M. De La Rosa (✉) · F. R. Dillon · P. Rojas · R. Duan
Center for Research on U.S. Latino HIV/AIDS and Drug Abuse (CRUSADA), Florida International University, 11200 SW 8th St., PCA 356, Miami, FL 33199, USA
e-mail: delarosa@fiu.edu

S. J. Schwartz
Department of Epidemiology and Public Health, University of Miami, Miami, FL, USA

Introduction

During the past decade, there has been an increasing interest in documenting the effects of substance use on sexual risk behaviors in Latinas (Castañeda & Gómez Bastidas, 2005; Kaplan, Erickson, & Juarez-Reyes, 2002; Matos et al., 2004; Paxton, Myers, Hall, & Javanbakht, 2004). This interest is likely due to the disproportionate rate of HIV infection affecting Latinas in the United States (Centers for Disease Control and Prevention, 2007). HIV/AIDS rates are increasing for women in general, as well as for individuals of Latin American descent. In 1992, women constituted 14% of adults and adolescents diagnosed with HIV/AIDS; however, by 2003 this increased to 22% (CDC, 2004), and increased again to 27% in 2005 (CDC, 2007). Latinas have an especially high risk for infection with HIV. In 2005, the HIV/AIDS contraction rate for Latina women was five times the rate for non-Latina White women (CDC, 2007).

Extant research has found that use of alcohol and other drugs places Latina and non-Latina teens and young adults at risk for unsafe sexual practices (Youth Risk Behavior Surveillance Survey [YRBSS] 2008). In 2007, 17% of sexually active adolescent Latinas reported being under the influence of alcohol or drugs during their most recent sexual experience (YRBSS, 2008). White females report the highest prevalence of sex under the influence of drugs or alcohol (19.8%), and black females report the least (12.9%). Studies have found that young adult Latinas who frequently use alcohol (Castañeda & Gómez Bastidas, 2005; Matos et al., 2004) or illicit drugs (Kaplan et al., 2002; Paxton et al., 2004) were more likely to engage in unprotected sex than those who used alcohol less frequently or not at all. However, because most of the research in this area has been conducted with adolescents and young adults, it is unclear whether the association between substance use and sexual risk taking persists into adulthood.

An important issue that has not received much attention in the literature on Latina sexual risk behavior is engaging in sexual behavior while under the influence of alcohol or drugs. Although preliminary studies linking drug and alcohol use to unsafe sex suggest that substance-using Latinas are more likely to engage in unsafe sexual practices, these analyses often do not focus specifically on sexual behaviors while the woman is intoxicated. Such behaviors are especially risky due to the effects of drugs and alcohol on cognition and judgment in substance users (Logan, Cole, & Leukefeld, 2002; Volkow, 2006). Specifically, intoxication increases the likelihood of impulsive and risky decisions and decreases inhibition (Bellis et al., 2008; George et al., 2008; Stoner, George, Peters, & Norris, 2007), which may, in turn, increase the likelihood of unsafe sex.

A second important issue concerns parental influences on Latinas' sexual risk behaviors into adulthood. Parental influences can be subdivided into at least two categories: relationships with the parents and the parents' own substance use. In samples of U.S. Latina/os, studies have identified parental substance use as a risk for substance use in offspring (Barrett & Turner, 2006) and have suggested that closeness to parents serves to protect Latina/o adults against substance use and HIV risk behavior (Wang, Matthew, Chiu, Fan, & Bellamy, 2007). However, as previously noted, the vast majority of studies on familial influences on Latina substance use and HIV risk behavior have focused on adolescents. To our knowledge, no extant research has examined familial factors, particularly attachment to mother, as risks and protective mechanisms vis-à-vis Latina adults' engaging in unsafe sexual behavior while under the influence of alcohol or drugs. More research is needed on adult Latinas because of the cultural value placed on strong family relationships well into adulthood (i.e., familism; De La Rosa & White, 2001; Muñoz-Laboy, 2008; Santisteban, Muir-Malcolm, Mitrani, & Szapocznik, 2002). Given these research needs, the present study was designed to examine substance abuse and familial risks for engaging in sexual risk behaviors while under the influence of alcohol or drugs in a community sample of adult Latinas.

The mother traditionally plays a key role in the Latino family (Gomez & Marin, 1996). Latina women—and especially mothers—have traditionally been viewed as owing the family a duty of self-sacrifice (Galanti, 2003). This duty has been termed *marianismo*. Marianismo is the female counterpart to *machismo*, which refers to the exaggerated masculine gender roles to which Latino men are expected to adhere within the family (Galanti, 2003). To be considered the *buena ama de casa* (good housewife and mother), women are encouraged to be selfless in their dedication to the family. This tradition of sacrifice and its accompanying servitude often extends to daughters in accordance with traditional cultural values. Daughters are often socialized to treat their mothers with high levels of respect so they can be considered a *buena hija* (good daughter). This often includes complying with the mother's wishes and not shaming the family by engaging in deviant behaviors (e.g., substance abuse, premarital sex). As a result of traditional cultural values, Latina mothers and daughters often forge deep lifelong bonds, in which they rely heavily on one another for emotional, financial, and other support over the lifespan (Delgado, 1998; Garcia-Preto, 1998; Vega, 1990; Villanueva Dixon, Graber, & Brooks-Gunn, 2008). Although *marianismo* is often questioned by some younger Latinas born in the U.S., it generally remains an important component of Latino cultural expectations. Thus, traditional Latino cultural values are theorized to foster mother–daughter relationships characterized by high levels of attachment, abstention from substance, daughters' premarital chastity, and avoidance of health risk behaviors like sex under the influence of alcohol or drugs. In particular, the daughter's attachment to her mother is hypothesized to protect against health risk behaviors.

In addition to considering the culturally based, family dynamics affecting Latina mothers and daughters, the present study was guided by Bogenschneider's (1996) ecological risk/protective model. This model is based on Bronfenbrenner's (1986) ecological theory of human development and on Lerner's (1995) theory of developmental contextualism. Similar to the concept of familism, these theoretical perspectives also suggest that the family represents the primary context for human development over the lifespan (Steinberg, 2001; Szapocznik & Coatsworth, 1999). Parents and other family members have the longest history with the individual. They play major roles in shaping patterns of development. These influences extend well beyond childhood and adolescence, such that family influences continue to be important in adulthood (Fergusson, Boden, & Horwood, 2007; Overbeek, Stattin, Vermulst, Ha, & Engels, 2007). Adult children and their parents—especially mothers and daughters—have been shown to influence one another's behavior considerably throughout the lifespan (Fingerman, 1995; Fingerman, Chen, Hay, Cichy, & Lefkowitz, 2006). As previously noted, the familism often endorsed by Latino families, along with the *marianismo* (and accompanying family ties) attached to the female gender role in many of these families, may render the adult daughter-mother relationship even more important (Galanti, 2003).

Ecological theories of human development also have been increasingly attentive to cultural concerns (e.g., Thurston & Vissandjée, 2005). Specifically, within the context of immigration, family relationships may be transformed as individual family members adapt to a new cultural context. Family relationships may operate differently depending on how long family members have been in the U.S. (Szapocznik & Kurtines, 1993). Therefore, nativity and length of time in the U.S. are considered salient contextual factors in the lives of the women in the present study and are included in the analytic plan.

The Present Study

The purpose of this article was to report the findings from a study investigating the roles of mother–daughter attachment, and maternal and daughter substance use, in adult daughters' engagement in sexual behavior under the influence of substances, while statistically controlling for the daughters' nativity and length of time in the U.S. To most effectively examine the interplay between maternal and daughter substance abuse on the daughter's engagement of sex under the influence of alcohol or drugs, we recruited four types of mother–daughter dyads: mother–daughter both drug or alcohol abusers (Dyad 1); mother-abuser and daughter non-abuser (Dyad 2); mother-non-abuser and daughter-abuser (Dyad 3); and mother–daughter both non-abusers (Dyad 4).

The study was designed to address four hypotheses: First, as a manipulation check, daughters who were not drug abusers (and whose mothers were not drug abusers—those daughters categorized into Dyad 4) were hypothesized to report the least sex under the influence of alcohol or drugs. Hypothesis 1 also allows us to explore for differences between two types of non-substance abusing daughters—those with a substance abusing mother (Dyad 2) and those without a substance abusing mother (Dyad 4). Second, we anticipated that daughter reports of attachment to mother would be inversely associated with sex under the influence of alcohol or drugs regardless of dyad type. Third, we hypothesized that being born in the U.S., and having resided in the U.S. for longer periods of time among immigrant daughters, would be positively associated with the daughter's risk for sex under the influence of alcohol or drugs regardless of dyad type. Finally, given that younger, unmarried women are at higher risk for substance use (Delva et al., 2005), it was expected that unmarried and younger daughters would be more likely to engage in sex under the influence of alcohol or drugs regardless of dyad type.

Method

Participants

Participants were 158 self-identified Latina mother–daughter dyads (total $N = 316$) across four drug use categories: Dyad 1 = drug-abusing mothers and daughters ($n = 80$); Dyad 2 = drug-abusing mothers and non-abusing daughters ($n = 38$); Dyad 3 = non-abusing mothers and drug-abusing daughters ($n = 102$); and Dyad 4 = non-abusing mothers and daughters ($n = 96$).

Approximately 46% of daughters ($n = 72$) were born in the United States. About 15% of daughters ($n = 24$) were born in Cuba. The next largest subgroup was 9.5% of daughters ($n = 15$), who were born in Colombia. The remainder of daughters reported the following birth countries: Peru ($n = 8$, 5.1%), Nicaragua ($n = 6$, 3.8%), Honduras ($n = 6$, 3.8%), Venezuela ($n = 6$, 3.8%), Dominican Republic ($n = 5$, 3.2%), Puerto Rico ($n = 4$, 2.5%), and Mexico ($n = 4$, 2.5%). One daughter (0.6%) reported being born from each of the following countries: Argentina, Bolivia, Brazil, and Chile. Table 1 summarizes the daughters' other demographic variables (age; receipt of financial, emotional, or legal assistance with immigration; family-of-origin socioeconomic status; personal income; educational level; employment status; marital status; and length of time in U.S.) across dyad type.

Procedure

Data for the present analyses were taken from a study of intergenerational transmission of drug abuse between Latina mothers and daughters in Miami-Dade County, Florida. The study was approved by, and conducted in compliance with, the Institutional Review Board at Florida International University. The general criteria for inclusion of mothers and daughters in the parent study were: consenting to be interviewed for at least 2–3 h; 18 years old or older; self-identifying as Latina; living in Miami-Dade County, Florida; and willing to provide two telephone numbers to researchers for correspondence during participation in the study.

Mothers and daughters were classified into the four dyads based on whether they qualified as abusers of alcohol, marijuana, cocaine, heroin, ecstasy, and/or prescription drugs. For the purpose of the current study, alcohol abusers were defined in terms of at least one binge drinking episode per month (adapted from Naimi et al., 2003)—at least 4–5 glasses of wine, 3–4 cans/bottles of beer, or 3–4 four-ounce drinks of hard liquor per occasion-during the 12 months prior to assessment. Illicit drug abusers were defined as participants who reported at least 3 days per week of marijuana use, 2 days per week of cocaine use, one or more occasions of heroin use per week, and/or least 3 ecstasy use occasions per month during the 12 months prior to assessment. Abuse of prescribed medication was measured by asking participants whether they had taken medicine without a doctor's authorization, in larger amounts than prescribed, or for longer periods than prescribed, in the 12 months prior to assessment (adapted from Turner et al., 2001).

All study participants were recruited for a one time face-to-face interview by using a snowball sampling (chain referral) method (McCracken et al., 1997). Given the fairly low representation of drug and alcohol abusers in the general population, different strategies were used to recruit substance abusing and non-abusing participants. Non-abusing Latinas were recruited to participate in the study through community health fairs, health clinics, radio announcements on local Spanish-language stations, and advertisements on local television channels. Substance abusing participants were recruited through substance abuse support groups such as Narcotics Anonymous (NA) and Alcoholics Anonymous

Table 1 Summary of demographics, sexually transmitted infections, and unprotected sex rates for daughters across dyads

Demographics	Dyad 1: mother and daughter, abusers ($n = 40$)	Dyad 2: mother abuser, daughter non-abuser ($n = 19$)	Dyad 3: mother non-abuser, daughter abuser ($n = 51$)	Dyad 4: mother and daughter, non-abusers ($n = 48$)	Group comparisons
Age (years)					
M (*SD*)	23.42 (5.96)	24.58 (7.34)	28.86 (9.45)	30.67 (9.81)	$F(3, 156) = 6.25, p < .001, \eta^2 = .11$ Dyad 1 < Dyad 3, $p < .01$ Dyad 1 < Dyad 4, $p < .001$ Dyad 2 < Dyad 4, $p = .01$
Daughter's nativity					
Caribbean	4	4	10	14	$\chi^2(9, N = 158) = 31.27, p < .001, \eta^2 = .15$
Central America	9	0	6	3	
South America	2	4	10	19	
U.S.	25	1	25	12	
Assistance with immigration					
Yes	4	3	4	11	$\chi^2(3, N = 158) = 2.49, p = .48, \eta^2 < .01$
No	12	5	22	25	
Family of origin socioeconomic status[a]					
M (*SD*)	1.98 (0.53)	1.89 (0.57)	2.06 (0.54)	2.04 (0.50)	$F(3, 157) = 0.47, p = .71, \eta^2 < .01$
Personal income-past year[b]					
M (*SD*)	4.03 (3.02)	2.42 (1.68)	3.35 (2.44)	3.50 (2.68)	$F(3, 157) = 1.67, p = .18, \eta^2 = .03$
Education level[c]					
M (*SD*)	2.28 (1.26)	2.47 (0.96)	2.63 (1.18)	2.65 (0.91)	$F(3, 157) = 1.03, p = .38, \eta^2 = .02$
Employment status					
Employed	19	13	30	31	$\chi^2(3, N = 158) = 3.45, p = .32, \eta^2 = .02$
Unemployed	21	6	21	17	
Marital status					
Married	6	3	8	16	$\chi^2(3, N = 158) = 6.43, p = .09, \eta^2 = .04$
Unmarried	34	16	43	32	
Length of life in U.S. (%)					
M (*SD*)	73.70 (36.36)	75.37 (32.61)	68.94 (36.02)	49.03 (35.37)	$F(3, 154) = 4.78, p < .01, \eta^2 = .02$ Dyad 1 > Dyad 4, $p = .001$ Dyad 2 > Dyad 4, $p = .007$ Dyad 3 > Dyad 4, $p = .006$
Total number of sexually transmitted infections					
M (*SD*)	0.38 (0.67)	0.32 (0.75)	0.32 (0.62)	0.33 (0.81)	$F(3, 157) = 0.06, p = .98, \eta^2 < .01$
Condom use ratio					
Oral sex ($n = 122$)					
M (*SD*)	0.07% (0.36)	No condom use reported	No condom use reported	2.92% (16.66)	$F(3, 121) = 0.86, p = .46, \eta^2 = .02$

Table 1 continued

Demographics	Dyad 1: mother and daughter, abusers ($n = 40$)	Dyad 2: mother abuser, daughter non-abuser ($n = 19$)	Dyad 3: mother non-abuser, daughter abuser ($n = 51$)	Dyad 4: mother and daughter, non-abusers ($n = 48$)	Group comparisons
Anal sex ($n = 31$)					
M (*SD*)	66.67% (57.74)	50.00% (70.71)	11.77% (33.21)	33.33% (50.00)	$F(3, 30) = 1.83, p = .17, \eta^2 = .17$
Vaginal sex ($n = 140$)					
M (*SD*)	37.73% (42.61)	53.96% (49.37)	35.51% (40.98)	30.84% (41.47)	$F(3, 139) = 1.16, p = .33, \eta^2 = .03$

[a] 1 = Poorer than most, 2 = About the same as others, 3 = Richer than most

[b] 1 = $0 to 4999 to 11 = $50,000 or more

[c] 1 = less than high school to 5 = graduate degree

(AA) meetings and by advertising in a local alternative newspaper, on an FM radio station, and through announcements posted at local drug court programs.

Interviews were conducted in either Spanish or English by trained interviewers using a structured questionnaire. This questionnaire consisted primarily of existing and widely used measures. All measures were pilot-tested to ensure that they were culturally and linguistically sensitive to Latino culture and to women. Many of these measures were already available in both English and Spanish and had previously been used with Latinas. Measures not available in Spanish were back-translated into Spanish in accordance with guidelines established by the Institutional Review Board and Office of Sponsored Research Administration at Florida International University. The questionnaire took 2–3 h to complete, and took place at locations convenient to participants.

Measures

Demographics Variables

A demographic form asked participants to report their age, nativity, the receipt of financial, emotion, or legal assistance with immigration, family of origin socioeconomic status, annual personal income, educational level, employment status, marital status, and length of time in U.S. Length of time in U.S. was defined as percent of the participant's life that had been spent in the U.S. This percentage was calculated by dividing the number of years a participant reported living in the U.S. by her age. Because Latino families often may move back and forth between the United States and their countries of origin (Congress, 2004), even U.S.-born daughters may have spent some time living outside of the United States.

HIV and Other Sexually Transmitted Infections (STI)

Each participant was asked if she ever had been informed by a doctor or other health professional that she had gonorrhea, syphilis, herpes, chlamydia, genital warts, Hepatitis C, HIV/AIDS, vaginitis, or pelvic inflammatory disease. These questions were adapted from a survey designed for a study of substance use and sexual risk behavior trajectories of young adults (Life Course and Health Research Center, 1997). Participant responses to the STI items were summed to compute a total number of STI score. Mean total number of STI scores are reported by dyad type in Table 1.

Unprotected Sex

Each participant was asked the number of times she had engaged in oral, anal, and vaginal sex during the 12 months prior to assessment. Participants were subsequently asked, "How many of those times did you or your partner use a condom

during oral, anal, and vaginal sex?" These items also were derived from the survey designed for a study of substance use and sexual risk behavior trajectories of young adults (Life Course and Health Research Center, 1997). From these responses, a condom use to sexual behavior ratio was calculated for each participant. Mean participant condom use ratios by sexual behavior and dyad type are reported in Table 1.

Attachment

A 25-item version of the Inventory of Parent and Peer Attachment (IPPA; Armsden & Greenberg, 1987) was used to assess attachment between mothers and daughters. Data from daughters rating their attachment relationships with their mothers were used in the present analyses. Daughters were asked to indicate their agreement with each item using a 5-point Likert scale ranging from *Almost never or never true* to *Almost always or always true*. In prior research, internal consistency and test–retest reliability coefficients for the IPPA have been acceptable (Adam & Chase-Lansdale, 2002). However, because the IPPA was originally developed with adolescent and young adult samples, we conducted a confirmatory factor analysis to ensure that the previously reported IPPA subscales/factors (i.e., communication, trust, and alienation) were appropriate for the present sample of adult Latina daughters. The confirmatory factor analysis of the three factor model produced an inadequate fit to the data (see Kline, 2005, for a review of fit statistics and acceptable values). The chi-square value (272 *df*) was 578.67 ($p < .0001$), the comparative fit index (CFI) value was .85, and the root mean square error of approximation (RMSEA) value was .08. One of the original subscales (alienation) yielded inadequate evidence of internal consistency in the current sample [e.g., Cronbach's alpha coefficient = .68, indicating the standard error of measurement will be over half (0.57) a standard deviation and suggesting a large amount of error or inconsistency in the measure; Nunnally & Bernstein, 1994]. It was, therefore, possible that the scoring algorithm for the measure was not appropriate for this population. As a result, we analyzed a factor solution where all of the items patterned onto a single factor. This solution provided an adequate fit to the data, $\chi^2(261) = 443.44$, $p < .0001$; CFI = .91; RMSEA = .07. Therefore, a total attachment scale score was used in the current study. We calculated a total attachment scale score by first reverse scoring the alienation scale items and summing these reverse-scored items with the items from the other subscales (communication and trust). The Cronbach's alpha coefficient for the total scale score was .93 for the daughter sample in the current study.

Alcohol Use

The alcohol consumption items from the Health and Daily Living Form (Billings, Cronkite, & Moos, 1983) were used to assess participant alcohol use frequency during the 12 months prior to assessment and, in part, to determine dyad membership (based on use rates during the 12 months prior to assessment). The questions used in the current study assessed how frequently participants consumed alcohol (i.e., alcohol frequency) in the 3 and 12 months prior to assessment, using an 8-point Likert scale ranging from 1 (never uses) to 8 (uses every day).

Drug Use

The Drug Use Frequency measure (DUF; O'Farrell, Fals-Stewart, & Murphy, 2003) was used to assess the frequency of drug use among the participants during the 12 months prior to assessment and, in part, to determine dyad membership (based on use rates during the 12 months prior to assessment). The DUF also was used to assess illicit drug use and non-medical use of prescription drugs (e.g., sedatives) in the 3 and 12 months prior to assessment. The DUF measure assesses frequency of use for each separate illicit and prescription drug (sedatives, hypnotics or tranquilizers; cannabis; stimulants; heroin; opioids; cocaine; PCP; hallucinogens; inhalants; others) using an 8-point Likert scale ranging from 1 (never uses) to 8 (uses every day). Past research (O'Farrell et al., 2003) indicates that DUF scores correlate closely ($r = .87$) with collateral reports of drug use frequency and with self-reports on other drug and alcohol use measures (r values ranged from .83 to .98).

Sex Under the Influence of Substances

Sexual behavior under the influence of substances was measured using six items taken from a survey designed for a study of substance use and sexual risk behavior trajectories of young adults (Life Course and Health Research Center, 1997). Participants' responses to the six selected items were used to create a sex under the influence of substance risk index (SUISRI). The six items assessed the frequency and effects of substance use by both the participant and her sexual partner(s) during the past 12 months. The first item assessed how often the respondent (or her partner) drank alcohol before or during sex during the past 12 months. The next item asked how often the respondent (or her partner) used drugs before or during sex during the past 12 months. For these two items, participants were asked to indicate their frequency on the following scale for each item: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Usually, 5 = Always. The third item assessed, on average, how strongly the participant was affected (i.e., intoxicated) by alcohol use prior to or during sex. The fourth item asked how strongly the participant's sexual partner was affected (i.e., intoxicated) by alcohol use prior to or during sex. Similarly, the fifth item assessed on average, how strongly the participant was affected (i.e., intoxicated)

by drug use prior to or during sex, and the sixth item asked how strongly the participant's sexual partner was affected (i.e., intoxicated) by drug use prior to sex. For these last two items, participants were asked to indicate their response on a scale of 1 = not at all, 2 = somewhat, or 3 = very strongly. The sex under the influence of alcohol risk index (SUIARI) score was calculated by summing the three items referring to alcohol. The sex under the influence of drugs risk index (SUIDRI) score was calculated by summing the three items referring to drugs.

Data Analysis Plan

The preliminary data analyses included calculating frequency distributions for continuous variables to determine whether they violated the assumption of normality. To explore for potential confounding variables, analyses of variance (ANOVA) and chi-square tests were performed to investigate for differences in demographic variables across the four dyads. We also explored total number of sexually transmitted infections, as well as condom use rates across dyads. For variables where the omnibus ANOVA was significant, pair-wise comparisons were conducted to identify significant differences between and among dyads.

To test our primary hypotheses, we conducted three hierarchical regression analyses. First, a bivariate correlation matrix was created for all variables involved in the hierarchical regression analyses to assess for potential multicollinearity. Next, we examined the hypothesized predictors of sex under the influence of any substance (alcohol or drugs). We then estimated two planned post hoc decomposition models—one with sex under the influence of drugs (including both illicit and prescription drugs) as the dependent variable, and another with sex under the influence of alcohol as the dependent variable. Predictors were entered sequentially into each equation in four blocks. Daughter's age and marital status were entered first, followed by nativity and time lived in the United States, followed by dyad type, and finally by attachment to mother. Entering age and marital status in the first block allowed us to control for their effects when examining the contributions of the predictor variables. Dyad type was entered in the third block using dummy coded variables for Dyads 1–3, with Dyad 4 as the reference group.

Results

Frequency distributions for continuous variables were examined to determine whether they violated the assumption of normality. All study variables appeared to be reasonably normally distributed with the exception of condom use during oral sex. Too few participants reported using condoms during oral sex, causing the variable to be too positively skewed for analyses. No significant differences were found among dyads for total number of sexually transmitted infections, or condom use ratios for anal or vaginal sex.

Demographic Differences Across Dyads

Significant dyad differences were found in age, percentage of life spent in the U.S., and nativity of daughters (described below). No significant differences were found among dyads for daughter's assistance with immigration, socioeconomic status of family of origin, current personal income, education level, employment status, and marital status.

Daughter Age, Nativity, and Years in U.S. by Dyad Type

As reported in Table 1, daughters from the "Mother, Daughter, Substance Abusers" dyad (Dyad 1) were significantly younger than daughters from the "Mother Non-Abuser, Daughter Abuser" dyad (Dyad 3) and from the "Mother, Daughter, Non-Abusers" dyad (Dyad 4). Daughters from the "Mother Abuser, Daughter Non-Abuser" dyad (Dyad 2) also were significantly younger than daughters from the "Mother, Daughter, Non-Abusers" dyad (Dyad 4). Results also showed that the percentage of U.S.-born versus immigrant daughters also differed by dyad (see Table 1). Daughters in substance-abusing-daughter dyads were more likely to be U.S.-born, compared to other dyads. Finally, daughters from the "Mother, Daughter, Non-Abusers" dyad (Dyad 4) reported significantly less time in the U.S. than all other dyad types.

Hierarchical Regression Analyses

A bivariate correlation matrix, including variables involved in regression analyses, is shown in Table 2. Thirteen participants were dropped from the regression analysis because they reported no sexual activity in the 12 months prior to assessment. The remaining 145 participants were used in subsequent analyses. Approximately 61% of the sample reported sex under the influence of alcohol, and 27.6% reported sex under the influence of drugs. Table 3 provides descriptive data for attachment, alcohol use frequency, drug use frequency, and sex under the influence of drugs and alcohol by dyad type. Results of the regression analyses are shown in Tables 4, 5 and 6, and are described below.

Sex Under the Influence of Any Substance

Results indicated that the overall set of predictors was related to sex under the influence of substances, accounting for 27% of variability, $F(8, 144) = 6.35$, $p < .001$. Of the 8 hypothesized predictors of sex under the influence of drugs or alcohol, 2 were statistically significant. These effects included Dyad 1 (drug-abusing mothers and daughters) versus

Table 2 Correlations among predictor and criterion variables

Variable	1	2	3	4	5	6	7	8	9	10	11
1. SUISRI	–										
2. SUIDRI	.88**	–									
3. SUIARI	.73**	.34**	–								
4. Age	−.09	−.11	−.02	–							
5. Marital status[a]	−.22**	−.16	−.29**	.27**	–						
6. Length of life in U.S. (%)	.21*	.21*	.16	−.28**	−.11	–					
7. Nativity[b]	−.22**	−.19*	−.17*	.31**	.10	−.83**	–				
8. Dyad 1	.35**	.31**	.22**	−.27**	−.08	.14	−.19*	–			
9. Dyad 2	−.07	−.07	.02	−.12	−.05	.11	−.09	−.22**	–		
10. Dyad 3	.14	.03	.21*	.10	−.09	.08	−.04	−.40**	−.26**	–	
11. Dyad 4	−.43**	−.28**	−.44**	.23**	.20*	−.29**	.28**	−.39**	−.24**	−.46**	–
12. Attachment	−.19*	−.30**	.05	−.06	.08	−.16*	.14	−.29**	.09	.07	.14

Note: *SUISRI* sex under the influence of substances risk index, *SUIDRI* sex under the influence of drugs risk index, *SUIARI* sex under the influence of alcohol risk index. Dyad 1 = drug-abusing mothers and daughters; Dyad 2 = drug-abusing mothers and non-abusing daughters; Dyad 3 = non-abusing mothers and drug-abusing daughters; Dyad 4 = non-abusing mothers and daughters

* $p < .05$; ** $p < .01$

[a] 0 = not married, 1 = married

[b] 0 = U.S. born, 1 = non-U.S. born

Table 3 Descriptive statistics for predictor and criterion variables by dyad type

Variables	Dyad 1 *M* (*SD*)	Dyad 2 *M* (*SD*)	Dyad 3 *M* (*SD*)	Dyad 4 *M* (*SD*)
Attachment	80.51 (19.56)	94.68 (20.94)	91.80 (19.01)	94.00 (15.74)
Alcohol use—past 12 months[a]	2.77 (0.86)	2.96 (1.47)	2.90 (0.93)	2.04 (1.12)
Illicit drug use—past 12 months				
Sedatives[a]	1.83 (1.96)	1.06 (0.25)	1.84 (2.13)	1.47 (1.72)
Cannabis[a]	4.03 (2.99)	1.35 (1.00)	2.42 (2.30)	1.14 (0.65)
Stimulants[a]	1.78 (1.64)	1.00 (0.00)	1.17 (0.63)	1.16 (0.91)
Heroin[a]	1.03 (0.16)	1.00 (0.00)	1.12 (0.86)	1.00 (0.00)
Opiods[a]	1.08 (0.36)	1.00 (0.00)	1.06 (0.31)	1.16 (1.06)
Cocaine[a]	1.86 (1.88)	1.00 (0.00)	1.43 (1.36)	1.00 (0.00)
PCP[a]	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
Hallucinogens[a]	1.00 (0.00)	1.00 (0.00)	1.02 (0.14)	1.00 (0.00)
Inhalants[a]	1.03 (0.16)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
SUISRI	7.79 (4.56)	4.88 (2.83)	6.31 (3.43)	3.00 (2.01)
SUIARI	5.05 (2.81)	4.19 (2.69)	4.85 (2.52)	2.19 (1.89)
SUIDRI	3.89 (3.15)	2.00 (2.29)	2.65 (2.72)	1.33 (1.29)

Note: *SUISRI* sex under the influence of substances risk index, *SUIDRI* sex under the influence of drugs risk index, *SUIARI* sex under the influence of alcohol risk index. Dyad 1 = drug-abusing mothers and daughters; Dyad 2 = drug-abusing mothers and non-abusing daughters; Dyad 3 = non-abusing mothers and drug-abusing daughters; Dyad 4 = non-abusing mothers and daughters

[a] 1 = Never to 8 = Everyday

Dyad 4 (non-abusing mothers and daughters) membership ($\beta = .50, p < .001$), and Dyad 3 (non-abusing mothers and drug-abusing daughters) versus Dyad 4 membership ($\beta = .35, p < .001$). Drug abusing daughters reported more sex under the influence of substances. Attachment to mother did not emerge as a significant predictor in this model.

Sex Under the Influence of Drugs

Results indicated that the overall set of predictors was related to sex under the influence of drugs, accounting for 20% of variability, $F(8, 144) = 4.12, p < .001$. Of the 8 hypothesized predictors of sex under the influence of drugs, 2 were

Table 4 Predictors of sex under the influence of substances risk index: results of a hierarchical regression model ($n = 145$)

Variable	*B*	*SE B*	β	*t*	R^2	ΔR^2	*F*	*df*
Step 1								
Age	−0.01	0.04	−.02	0.08				
Marital status	−1.98	0.78	−.22	−2.54*	.05	.05	3.76*	2, 144
Step 2								
Age	0.02	0.04	.04	0.46				
Marital status	−1.90	0.77	−.21	−3.12*				
Length of life in U.S. (%)	0.01	0.02	.08	0.52				
Nativity	−1.13	1.10	−.15	−1.03	.09	.04	3.53**	4, 144
Step 3								
Age	0.05	0.04	.10	1.26				
Marital status	1.33	0.71	−.15	−1.87				
Length of life in U.S. (%)	0.01	0.01	.05	0.41				
Nativity	−0.60	1.00	−.08	−0.60				
Dyad 1 versus Dyad 4	4.47	0.81	.52	5.55**				
Dyad 2 versus Dyad 4	1.57	0.99	.13	1.58				
Dyad 3 versus Dyad 4	2.84	0.74	.35	3.84**	.27	.18	7.22**	7, 144
Step 4								
Age	0.04	0.04	.09	1.12				
Marital status	−1.28	0.72	−.14	−1.79				
Length of life in U.S. (%)	0.01	0.01	.04	0.34				
Nativity	−0.61	1.00	−.08	−0.61				
Dyad 1 versus Dyad 4	4.33	0.84	.50	5.18**				
Dyad 2 versus Dyad 4	1.58	1.00	.13	1.59				
Dyad 3 versus Dyad 4	2.83	0.74	.35	3.83**				
Attachment	−0.01	0.02	−.05	−0.64	.27	.00	6.35**	8, 144

Note: Dyad 1 = drug-abusing mothers and daughters; Dyad 2 = drug-abusing mothers and non-abusing daughters; Dyad 3 = non-abusing mothers and drug-abusing daughters; Dyad 4 = non-abusing mothers and daughters

* $p < .05$; ** $p < .01$

statistically significant in the final step of the model, including attachment ($\beta = -.21$, $p < .05$) and Dyad 1 (drug-abusing mothers and daughters) versus Dyad 4 (non-abusing mothers and daughters) membership ($\beta = .32, p < .01$). Drug abusing daughters with a drug abusing mother, and those reporting less attachment to their mother, reported more sex under the influence of drugs. Marital status may have failed to explain a significant portion of the variability in sex under the influence of drugs due to low statistical power resulting from the small number of married (15.2%) and unmarried (31.2%) daughters reporting sex under the influence of drugs.

Sex Under the Influence of Alcohol

Results indicated that the overall set of predictors was related to sex under the influence of alcohol, accounting for 29% of variability, $F(8, 144) = 6.91, p < .001$. Of the 8 hypothesized predictors of sex under the influence of alcohol, 6 were statistically significant in the final step of the model, including daughter age ($\beta = .20$, $p < .05$), marital status ($\beta = -.24$, $p < .01$), attachment ($\beta = .19, p < .05$), Dyad 1 (drug-abusing mothers and daughters) versus Dyad 4 (non-abusing mothers and daughters) membership ($\beta = .50$, $p < .001$), Dyad 2 (drug-abusing mothers and non-abusing daughters) versus Dyad 4 membership ($\beta = .21$, $p < .05$), and Dyad 3 (non-abusing mothers and drug-abusing daughters) versus Dyad 4 membership ($\beta = .40, p < .001$). Older or unmarried daughters reported more sex under the influence of alcohol. In contrast to sex under the influence of alcohol, larger percentages of married (39.4%) and unmarried (67.9%) daughters reported engaging in sex under the influence of alcohol, thereby, increasing the power to detect marital status effects. Daughters with drug abusing mothers also report more sex under the influence of alcohol. Surprisingly, daughters with closer attachments to their mothers reported more sex under the influence of alcohol. Because this finding was counterintuitive and contrary to our hypothesis, we conducted a series of post hoc regression analyses to explore it.

Table 5 Predictors of sex under the influence of drugs risk index: results of a hierarchical regression model ($n = 145$)

Variable	*B*	*SE B*	β	*t*	R^2	ΔR^2	*F*	*df*
Step 1								
Age	−0.02	0.03	−.07	−0.68				
Marital status	−0.88	0.55	−.14	−1.61	.03	.03	2.14	2, 144
Step 2								
Age	−0.01	0.03	−.02	−0.17				
Marital status	−0.82	0.54	−.13	−1.51				
Length of life in U.S. (%)	0.01	0.01	.15	1.03				
Nativity	−0.22	0.80	−.04	−0.28	.06	.03	2.32	4, 144
Step 3								
Age	0.01	0.03	.05	0.51				
Marital status	−0.59	0.53	−.10	−1.12				
Length of life in U.S. (%)	0.01	0.01	.14	0.99				
Nativity	0.01	0.77	.01	0.01				
Dyad 1 versus Dyad 4	2.28	0.61	.38	3.76**				
Dyad 2 versus Dyad 4	0.38	0.75	.05	0.52				
Dyad 3 versus Dyad 4	1.00	0.56	.18	1.79	.16	.10	3.68*	7, 144
Step 4								
Age	0.01	0.03	.01	0.04				
Marital status	−0.45	0.52	−.07	−0.86				
Length of life in U.S. (%)	0.01	0.01	.10	0.71				
Nativity	−0.05	0.75	−.01	−0.07				
Dyad 1 versus Dyad 4	1.90	0.61	.32	3.10**				
Dyad 2 versus Dyad 4	0.43	0.73	.05	0.59				
Dyad 3 versus Dyad 4	1.01	0.55	.18	1.85				
Attachment	−0.03	0.01	−.21	−2.46*	.20	.15	4.12	8, 144

Note: Dyad 1 = drug-abusing mothers and daughters; Dyad 2 = drug-abusing mothers and non-abusing daughters; Dyad 3 = non-abusing mothers and drug-abusing daughters; Dyad 4 = non-abusing mothers and daughters

* $p < .05$; ** $p < .01$

Post Hoc Hierarchical Regression Analyses

We conducted two separate post hoc hierarchical regression analyses to investigate the positive association between attachment to mother and sex under the influence of alcohol. The analyses were conducted with two subsamples of daughters. The first subsample consisted of daughters with a drug abusing mother (Dyads 1 and 2), and the second subsample involved daughters with a non-drug abusing mother (Dyads 3 and 4). The aim of these analyses was to determine whether the positive association between attachment and sex under the influence of alcohol risk was unique to the daughters with a drug abusing mother—possibly reflecting transmission of the mother's risk behavior (Wilder & Watt, 2002). That is, we examined whether attachment to a drug abusing mother was associated with greater risk for sex under the influence of alcohol. The same sequence of variables and blocks used in the primary regression analysis was also used in these post hoc analyses. Because only two dyad types were included in each model, a single variable was used to represent dyad type.

Daughters with Drug Abusing Mothers

Fifty-four daughters with drug abusing mothers were used in the first post hoc analysis. Results indicated that the overall set of predictors was related to sex under the influence of alcohol, accounting for 23% of variability, $F(6, 53) = 2.35$, $p < .05$. Of the 6 hypothesized predictors of sex under the influence of drugs, 4 were statistically significant in the final step of the model, including daughter age ($\beta = .37, p < .05$), marital status ($\beta = -.35$, $p < .05$), attachment ($\beta = .37$, $p < .05$), and dyad membership ($\beta = .30$, $p < .05$). Once again, older or unmarried daughters reported more sex under the influence of alcohol. Drug abusing daughters with a drug abusing mother (Dyad 1) reported more sex under the influence of alcohol compared to non-drug abusing daughters with a drug abusing mother (Dyad 2). Daughters with greater

Table 6 Predictors of sex under the influence of alcohol risk index: results of a hierarchical regression model ($n = 145$)

Variable	*B*	*SE B*	β	*t*	R^2	ΔR^2	*F*	*df*
Step 1								
Age	0.02	0.03	.07	0.80				
Marital status	−1.97	0.54	−.31	−3.62**	.09	.09	6.59**	2, 144
Step 2								
Age	0.04	0.03	.11	1.31				
Marital status	−1.92	0.54	−.30	−3.54**				
Length of life in U.S.	0.01	0.01	.07	0.46				
Nativity	−0.58	0.77	−.11	−0.75	.11	.03	4.31**	4, 144
Step 3								
Age	0.05	0.03	.16	1.86				
Marital status	−1.47	0.51	−.23	−2.88**				
Length of life in U.S.	0.01	0.01	.02	0.16				
Nativity	−0.32	0.72	−.06	−0.45				
Dyad 1 versus Dyad 4	2.72	0.58	.44	4.71**				
Dyad 2 versus Dyad 4	1.82	0.73	.21	2.50*				
Dyad 3 versus Dyad 4	2.29	0.53	.40	4.34**	.26	.15	7.22**	7, 144
Step 4								
Age	0.06	0.03	.20	2.33*				
Marital status	−1.59	0.50	−.25	−3.17**				
Length of life in U.S.	0.01	0.01	.05	0.42				
Nativity	−0.31	0.71	−.06	−0.44				
Dyad 1 versus Dyad 4	3.10	0.59	.50	5.28**				
Dyad 2 versus Dyad 4	1.79	0.71	.21	2.50*				
Dyad 3 versus Dyad 4	2.30	0.52	.40	4.43**				
Attachment	0.03	0.01	.19	2.49*	.29	.03	6.29	8, 144

Note: Dyad 1 = drug-abusing mothers and daughters; Dyad 2 = drug-abusing mothers and non-abusing daughters; Dyad 3 = non-abusing mothers and drug-abusing daughters; Dyad 4 = non-abusing mothers and daughters

* $p < .05$; ** $p < .01$

attachment to their drug abusing mother reported more sex under the influence of alcohol.

Daughters with Non-Drug Abusing Mothers

Ninety daughters with non-drug abusing mothers were used in the second subsample analysis. Results indicated that the overall set of predictors was related to the sex under the influence of alcohol, accounting for 35% of variability, $F(6, 89) = 7.60$, $p < .0001$. Of the 6 hypothesized predictors of sex under the influence of drugs, 2 were statistically significant in the final step of the model, including marital status ($\beta = -.22, p < .05$) and Dyad 3 versus Dyad 4 membership ($\beta = .42$, $p < .001$). Again, unmarried daughters reported more sex under the influence of alcohol. Drug abusing daughters with a non-drug abusing mother (Dyad 3) reported more sex under the influence of alcohol risk compared to non-drug abusing daughters with a non-drug abusing mother (Dyad 4). Attachment was not related to sex under the influence of alcohol ($\beta = .14, p = .14$).

Discussion

The present study examined predictors of sexual behavior under the influence of alcohol and drugs among a community sample of adult Latinas. The results offer new knowledge concerning the association between adult Latina mother–daughter relations, substance abuse, and sexual behavior under the influence of substances, a rarely studied HIV risk behavior, especially in this population. Demographics (marital status, age), daughter and maternal substance abuse, and daughter attachment to mother each explained significant variance in sexual behavior under the influence of alcohol and/or drugs. Although this study has begun to identify salient familial and substance abuse related variables associated with sex under the influence of substances, future studies are needed to identify additional determinants of this hazardous risk behavior to help reduce health disparities in HIV contraction among adult Latinas in the U.S.

Age and marital status explained a significant portion of the variability in sex under the influence of alcohol. The

finding that older daughters were more likely to engage in sexual behavior under the influence of alcohol, controlling for other predictors, is consistent with studies indicating that the highest rate of drinking among women occurs between the ages of 25 and 44 (National Institutes of Health, 2000). For unmarried Latinas, sexual activity may be more likely to occur in the context of their or their sexual partners' heavy drinking in comparison to married women. This suggests that alcohol abuse may particularly increase HIV risk among unmarried adult Latinas. This finding is consistent with others suggesting that alcohol and drug use decreases following marriage (Bachman et al., 2002). More statistical power may have been necessary to detect associations for sex under the influence of drugs. Given that sex under the influences of substances is associated with impaired decision making (Volkow, 2006) and with lower likelihood of condom use (Maisto, Carey, Carey, Gordon, & Schum, 2004), it is important for future research to use larger samples to explore the effects of marital status on sexual behavior under the influence of drugs.

Daughter nativity and length of time in the U.S. did not predict sex under the influence of substances, although it is important to note that daughters from the substance abusing dyads reported more time in the U.S. and were more often U.S. born. Nativity, immigration status, and language use have been found to be associated with sexual risk behaviors in Latina/o adolescents and young adults (Abraido-Lanza, Armbrister, Flórez, & Aguirre, 2006; Adam, McGuire, Walsh, Basta, & LeCroy, 2005; Afable-Munsuz & Brindis, 2006). Latina immigrants often struggle with balancing traditional Latino norms and values (e.g., forbidding premarital sex) with conflicting sets of sexual expectations and values espoused by their U.S. born peers and partners (Raffaelli & Ontai, 2001). Latino cultural directives or values, such as those encouraged by *marianismo*, often socialize Latinas to be submissive to their male partners and may impair Latinas' decision-making ability concerning the initiation of sexual intercourse resulting in greater sexual risk behaviors (Weeks, Schensul, Williams, Singer, & Grier, 1995). However, an association between nativity and length of time in the U.S. and sex under the influences did not emerge in the present study.

The absence of associations of nativity and time lived in the U.S. to sexual behavior under the influence of alcohol and drugs may be due, in part, to our use of a different sexual outcome than has been assessed in prior work. Variables related to length of time in the U.S. have been suggested as alternately protective or risk enhancing, based on the sexual outcome studied (Deardorff, Tschann, & Flores, 2008). For example, although greater acculturation to American cultural practices is related to earlier sexual activity in Latinas, once sexual activity has been initiated, low levels of acculturation to American practices are thought to place Latinas at risk for reduced condom use (Marin, Gomez, Tschann, & Gregorich, 1997). The null finding obtained in the present study may have resulted from aggregating these different sets of relationships. Additionally, nativity and years spent in the U.S. may affect behavioral outcomes differently depending on the cultural context in which the person resides (Schwartz, Pantin, Prado, Sullivan, & Szapocznik, 2006). In Miami-Dade County, which has been a Latino ethnic enclave for two generations, individuals may not acculturate to American cultural practices with increasing time in the U.S. Years spent in the U.S. may, therefore, not be associated with cultural change for some immigrants to Miami and, perhaps, other ethnic enclaves as well.

Dyad Type

Not surprisingly, substance abusing daughters engaged in more sex under the influence of substances. This finding is consistent with literature citing the association of substance abuse with sexual risk behaviors among Latina women (Matos et al., 2004; Paxton et al., 2004). It is noteworthy that only substance-abusing daughters with a substance-abusing mother differed from non-abusing daughters with non-abusing mothers (the reference group) in sex under the influence of drugs. Unexpectedly, substance abusing daughters with a non-abusing mother were not statistically different from the reference group in terms of sex under the influence of drugs. In contrast, not only did substance-abusing daughters with substance-abusing mothers differ from the reference group in sex under the influence of alcohol, but surprisingly, non-abusing daughters with abusing mothers also reported more sex under the influence of alcohol when compared to the reference group. This complex pattern of findings suggests a particular risky effect of mother's substance abuse on daughter's sex under the influence of alcohol. Sex under the influence of alcohol was far more prevalent than sex under the influence of drugs. Thus, not only was there more statistical power to detect effects on sex under the influence of alcohol, but alcohol use is legal and more socially accepted than drug use. Nevertheless, given the present findings, alcohol abuse immediately prior to sexual behavior may be an important target for preventive efforts to reduce HIV contraction rates among adult Latinas.

Attachment to Mother

When alcohol and drugs were combined into a single *substance use* variable, attachment to mother failed to explain a significant portion of the variability in sex under the influence of substances. However, contrasting patterns emerged when sex under the influence of each substance was analyzed separately. As hypothesized, closer attachment to mother was related to less sex under the influence of drugs. Thus, attachment with mother appeared to serve as a protective mechanism against sex under the influence of drugs. However,

attachment to a drug abusing mother was identified as a *risk* for sex under the influence of alcohol. Both of these findings may be viewed as consistent with Bogenschneider's (1996) ecological risk/protective theoretical framework and with Lerner's (1995) developmental-contextual perspective. As outlined earlier, the family represents the primary context for human development over the lifespan (Steinberg, 2001; Szapocznik & Coatsworth, 1999). Attachment to the mother may help to protect Latinas from obtaining drugs (or affiliating with drug-using sexual partners). However, because alcohol is widely available to adults, attachment to a substance using mother appeared to increase risk. Many substance abusing mothers identified alcohol as their drug of choice (second only to non-prescribed sedatives). It is possible, then, that risk for consumption of a specific substance—in this case alcohol—was responsible for the counter-intuitive finding. Attachment to an alcohol-abusing mother was associated with risk for heavy alcohol use prior to sexual activity in adult Latina women.

Limitations

The present findings should be interpreted in light of several important limitations. First, the cross-sectional design used in the present study does not permit causal inferences to be drawn. Future longitudinal research is needed to study the ways in which Latina mothers' substance abuse, as well as their daughters' substance abuse, affects the daughter's sexual behavior under the influence of drugs and alcohol over the lifespan. Furthermore, it is also important to examine other parent–child pairings—such as father–daughter, father–son, and mother–son—in terms of sex under the influence of substances in Latina/os. Second, although it helps to recruit relatively small and dispersed populations, snowball sampling increases the likelihood of obtaining a nonrepresentative sample. To counter this limitation, efforts were undertaken to include participants from major Latino subgroups, such as Cuban Americans and Puerto Ricans. However, some groups (e.g., Mexican American and Puerto Rican) were not well represented due to their underrepresentation in the Miami-Dade County area in general. According to 2003 U.S. Census data (U.S. Census Bureau, 2007), Hispanic subgroups in Miami-Dade County constituted 60.97% of the total county population. The percentage of each subgroup was estimated as 30.98% Cuban, 23.35% Other Hispanic or Latino, 4.07% Puerto Rican, and 2.57% Mexican. Thus, the current study's sample was representative of Latinas living in Miami-Dade County, but not of the larger U.S. Latino population. Future studies are needed with nationally representative samples to enhance the generalizability of the present results. Third, sexual risk behaviors and substance abuse are highly stigmatized among Latinas, which often inhibits their discussing and revealing personal information with others (Larkey et al., 2002). Although assessors were trained to detect inconsistencies in participants' responses and to note such inconsistencies in a clinically and culturally sensitive manner (De La Rosa, Rahill, Rojas, & Pinto, 2007), socially desirable responding may have occurred. Finally, given that we attempted to recruit substance abusing participants through self-help support groups and announcements posted at local drug court programs (among other methods), the substance abusing participants may have been seeking help—or in trouble with the law—more often than would be obtained within a random sample.

Despite these limitations, the results from the current study can serve as a platform for future research investigating sexual decision making and sex under the influence of substances among Latinas. It is important for future research, as well as for substance abuse and HIV interventionists, to assess sexual behavior under the influence of alcohol and drugs among Latinas in the U.S. Interventionists and researchers should consider the particular risks for unmarried, older, and substance abusing Latinas, as well as those with a substance abusing mother. Focusing on these characteristics of risk may improve the efficacy of interventions with Latina participants and clients. Such culturally informed research and interventions also may help to identify additional contextual and intrapersonal factors affecting sex under the influence of alcohol and drugs among Latinas. Continued research and intervention development may help to facilitate the genesis and refinement of empirically based HIV and drug prevention and treatment programs for Latinas.

Acknowledgments This study was supported by award number P20MD002288 from the National Center for Minority and Health Disparity and award number R24DA014260 from the National Institute on Drug Abuse. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Center on Minority Health and Health Disparities, the National Institute on Drug Abuse, or the National Institutes of Health.

References

Abraido-Lanza, A. F., Armbrister, A. N., Flórez, K. R., & Aguirre, A. N. (2006). Toward a theory-driven model of acculturation in public health research. *American Journal of Public Health, 96*, 1342–1346.

Adam, E. K., & Chase-Lansdale, P. L. (2002). Home sweet home(s): Parental separations, residential moves, and adjustment problems in low income adolescent girls. *Developmental Psychology, 38*, 792–805.

Adam, M. B., McGuire, J. K., Walsh, M., Basta, J., & LeCroy, C. (2005). Acculturation as a predictor of the onset of sexual intercourse among Latino and white teens. *Archives of Pediatrics and Adolescent Medicine, 159*, 261–265.

Afable-Munsuz, A., & Brindis, C. D. (2006). Acculturation and the sexual and reproductive health of Latino youth in the United States: A literature review. *Perspectives on Sexual and Reproductive Health, 38*, 208–219.

Armsden, G. C., & Greenberg, M. T. (1987). The Inventory of Parent and Peer Attachment: Individual differences and their relationship to

psychological well-being in adolescence. *Journal of Youth and Adolescence, 16*, 427–454.

Bachman, J. G., O'Malley, P. M., Schulenberg, J., Johnston, L. D., Bryant, A. L., & Merline, A. C. (2002). *The decline of substance use in young adulthood: Changes in social activities, roles, and beliefs*. Mahwah, NJ: Lawrence Erlbaum Associates.

Barrett, A. E., & Turner, R. J. (2006). Family structure and substance use problems in adolescence and early adulthood: Examining explanations for the relationship. *Addiction, 101*, 109–120.

Bellis, M. A., Hughes, K., Calafat, A., Juan, M., Ramon, A., Rodriguez, J. A., et al. (2008). Sexual uses of alcohol and drugs and the associated health risks: A cross sectional study of young people in nine European cities. *BMC Public Health, 8*, 155.

Billings, A. G., Cronkite, R. C., & Moos, R. H. (1983). Social-environmental factors in unipolar depression: Comparisons of depressed patients and nondepressed controls. *Journal of Abnormal Psychology, 92*, 119–133.

Bogenschneider, K. (1996). An ecological risk/protective theory for building prevention programs, policies, and community capacity to support youth. *Family Relations, 45*, 127–138.

Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: Research perspectives. *Developmental Psychology, 22*, 723–742.

Castañeda, D., & Gómez Bastidas, E. (2005). HIV/AIDS risk behavior of Mexican women working in a maquiladora. *Interamerican Journal of Psychology, 39*, 267–274.

Center for Disease Control and Prevention. (2004). Cases of HIV infection and AIDS in the United States, 2003. In *HIV/AIDS surveillance report* (Vol. 15). Atlanta, GA: U.S. Department of Health and Human Services.

Center for Disease Control and Prevention. (2007). Cases of HIV infection and AIDS in the United States and dependent areas, 2005. In *HIV/AIDS surveillance report* (rev. ed., Vol. 17). Atlanta, GA: U.S. Department of Health and Human Services.

Congress, E. P. (2004). Cultural and ethical issues in working with culturally diverse patients and their families: The use of the culturagram to promote cultural competent practice in health care settings. *Social Work in Health Care, 39*, 249–262.

Deardorff, J., Tschann, J. M., & Flores, E. (2008). Sexual values among Latino youth: Measurement development using a culturally based approach. *Cultural Diversity and Ethnic Minority Psychology, 14*, 138–146.

De La Rosa, M. R., Rahill, G. J., Rojas, P., & Pinto, E. (2007). Cultural adaptations in data collection: Field experiences. *Journal of Ethnicity in Substance Abuse, 6*, 163–180.

De La Rosa, M. R., & White, M. (2001). A review of the role of social support systems in the drug use behavior of Hispanics. *Journal of Psychoactive Drugs, 33*, 233–241.

Delgado, M. (1998). *Social services in Latino communities: Research and strategies*. New York: Haworth Press.

Delva, J., Wallace, J. M., O'Malley, P. M., Bachman, J., Johnston, L., & Schulenberg, J. E. (2005). The epidemiology of alcohol, marijuana, and cocaine use among Mexican American, Puerto Rican, Cuban American, and other Latin American eighth-grade students in the United States: 1991–2002. *American Journal of Public Health, 95*, 696–702.

Fergusson, D. M., Boden, M., & Horwood, L. J. (2007). Exposure to single parenthood in childhood and later mental health, educational, economic, and criminal behavior outcomes. *Archives of General Psychiatry, 64*, 1089–1095.

Fingerman, K. L. (1995). Aging mothers' and their adult daughters' perceptions of conflict behaviors. *Psychology and Aging, 10*, 639–650.

Fingerman, K. L., Chen, P. C., Hay, E. L., Cichy, K. E., & Lefkowitz, E. S. (2006). Ambivalent reactions in the parent and offspring relationship. *Journals of Gerontology: Psychological Sciences, 61*, 152–160.

Galanti, G. A. (2003). The Hispanic family and male-female relationships: An overview. *Journal of Transcultural Nursing, 14*, 180–185.

Garcia-Preto, N. (1998). *Latinas in the United States: Bridging two worlds*. New York: Guilford Press.

George, W. H., Cue Davis, K., Norris, J., Heiman, J. R., Stoner, S. A., Schacht, R. L., et al. (2008). Indirect effects of acute alcohol intoxication on sexual risk-taking: The roles of subjective and physiological sexual arousal. *Archives of Sexual Behavior*. doi: 10.1007/s10508-008-9346-9.

Gomez, C. A., & Marin, B. V. (1996). Gender, culture, and power: Barriers to HIV-prevention strategies for women. *Journal of Sex Research, 33*, 355–362.

Kaplan, C. P., Erickson, P. I., & Juarez-Reyes, M. (2002). Acculturation, gender role orientation, and reproductive risk-taking behavior among Latina adolescent family planning clients. *Journal of Adolescent Research, 17*, 103–121.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Larkey, L. K., Staten, L. K., Ritenbaugh, C., Hall, R. A., Buller, D. B., Bassford, T., et al. (2002). Recruitment of Hispanic women to the Women's Health Initiative: The case of embajadoras in Arizona. *Controlled Clinical Trials, 23*, 289–298.

Lerner, R. M. (1995). *America's youth in crisis: Challenges and options for programs and policies*. Thousand Oaks, CA: Sage Publications.

Life Course and Health Research Center. (1997). *Transitions: A study of stress and well being in young adulthood. Wave I questionnaire*. Title of project: Drug Use Trajectories—Ethnic/Racial Comparisons. NIDA Grant Number DA10772. Florida International University, Miami, FL. Retrieved on August 26, 2002 (available at http://www.fiu.edu/~lchrc/transitions.htm).

Logan, T. K., Cole, J., & Leukefeld, C. (2002). Women, sex, and HIV: Social and contextual factors, meta-analysis of published interventions, and implications for practice and research. *Psychological Bulletin, 128*, 851–885.

Maisto, S. A., Carey, M. P., Carey, K. B., Gordon, C. M., & Schum, J. L. (2004). Effects of alcohol and expectancies on HIV-related risk perception and behavioral skills in heterosexual women. *Experimental and Clinical Psychopharmacology, 12*, 288–297.

Marin, B. V., Gomez, C. A., Tschann, J. M., & Gregorich, S. E. (1997). Condom use in unmarried Latino men: A test of cultural constructs. *Health Psychology, 16*, 458–467.

Matos, T. D., Robles, R. R., Sahai, H., Colón, H. M., Reyes, J. C., Amalia Marrero, C., et al. (2004). HIV risk behaviors and alcohol intoxication among injection drug users in Puerto Rico. *Drug and Alcohol Dependence, 76*, 229–234.

McCracken, C. F. M., Boneham, M. A., Copeland, J. R. M., Williams, K. E., Wilson, K., Scott, A., et al. (1997). Prevalence of dementia and depression among elderly people in black and ethnic minorities. *British Journal of Psychiatry, 171*, 269–273.

Muñoz-Laboy, M. A. (2008). Familism and sexual regulation among bisexual Latino men. *Archives of Sexual Behavior, 37*, 773–782.

Naimi, T. S., Brewer, R. D., Mokdad, A. H., Denny, C., Serdula, M., & Marks, J. S. (2003). Binge drinking among US adults. *Journal of the American Medical Association, 289*, 70–75.

National Institutes of Health. (2000). *Women of color health data book*. Office of the Director, National Institutes of Health.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

O'Farrell, T. J., Fals-Stewart, W., & Murphy, M. (2003). Concurrent validity of a brief self-report drug use frequency measure. *Addictive Behaviors, 28*, 327–337.

Overbeek, G., Stattin, H., Vermulst, A., Ha, T., & Engels, R. C. M. E. (2007). Parent–child relationships, partner relationships, and

emotional adjustment: A birth-to-maturity prospective study. *Developmental Psychology, 43*, 429–437.

Paxton, K. C., Myers, H. F., Hall, N. M., & Javanbakht, M. (2004). Ethnicity, serostatus, and psychosocial differences in sexual risk behavior among HIV-seropositive and HIV-seronegative women. *AIDS and Behavior, 8*, 405–415.

Raffaelli, M., & Ontai, L. L. (2001). "She's sixteen years old and there's boys calling over to the house": An exploratory study of sexual socialization in Latino families. *Culture, Health, and Sexuality, 3*, 295–310.

Santisteban, D. A., Muir-Malcolm, J. A., Mitrani, V. B., & Szapocznik, J. (2002). *Integrating the study of ethnic culture and family psychology intervention science*. Washington, DC: American Psychological Association.

Schwartz, S. J., Pantin, H., Prado, G., Sullivan, S., & Szapocznik, J. (2006). Nativity and years in the receiving culture as markers of acculturation in ethnic enclaves. *Journal of Cross-Cultural Psychology, 37*, 345–353.

Steinberg, L. (2001). Adolescent development. *Annual Review of Psychology, 52*, 83–110.

Stoner, S. A., George, W. H., Peters, L. M., & Norris, J. (2007). Liquid courage: Alcohol fosters risky sexual decision-making in individuals with sexual fears. *AIDS and Behavior, 11*, 227–237.

Szapocznik, J., & Coatsworth, J. D. (1999). An ecodevelopmental framework for organizing risk and protection for drug abuse: A developmental model of risk and protection. In M. Glantz & C. R. Hartel (Eds.), *Drug abuse: Origins and interventions* (pp. 331–366). Washington, DC: American Psychological Association.

Szapocznik, J., & Kurtines, W. M. (1993). Family psychology and cultural diversity: Opportunities for theory, research and application. *American Psychologist, 48*, 400–407.

Thurston, W. E., & Vissandjée, B. (2005). An ecological model for understanding culture as a determinant of women's health. *Critical Public Health, 15*, 229–242.

Turner, B. J., Fleishman, J. A., Wenger, N., London, A. S., Burnam, M. A., Shapiro, M. F., et al. (2001). Effects of drug abuse and mental disorders on use and type of antiretroviral therapy in HIV-infected persons. *Journal of General Internal Medicine, 16*, 625–633.

U.S. Census Bureau. (2007). *American community survey 2003 multi-year profile for Miami-Dade County*. Retrieved September 10, 2008, from http://www.census.gov/acs/www/Products/Profiles/Chg/2003/ACS/Tabular/050/05000US120861.htm.

Vega, W. A. (1990). Hispanic families in the 1980s: A decade of research. *Journal of Marriage and the Family, 52*, 1015–1024.

Villanueva Dixon, S., Graber, J. A., & Brooks-Gunn, J. (2008). The roles of respect for parental authority and parenting practices in parent–child conflict among African American, Latino, and European American families. *Journal of Family Psychology, 22*, 1–10.

Volkow, N. (2006). Hispanic drug abuse research: Challenges and opportunities. *Drug and Alcohol Dependence, 84S*, S4–S7.

Wang, M. Q., Matthew, R. F., Chiu, Y. W., Fan, Y., & Bellamy, N. D. (2007). Latent model analysis of substance use and HIV risk behaviors among high-risk minority adults. *Journal of Alcohol & Drug Education, 51*, 35–62.

Weeks, M. R., Schensul, J. J., Williams, S. S., Singer, M., & Grier, M. (1995). AIDS prevention for African-American and Latina women: Building culturally and gender-appropriate intervention. *AIDS Education and Prevention, 7*, 251–263.

Wilder, E. I., & Watt, T. T. (2002). Risky parental behavior and adolescent sexual activity at first coitus. *The Milbank Quarterly, 80*, 481–524.

Youth Risk Behavior Surveillance Survey. (2008). Youth online: Comprehensive results. Accessed July 1, 2008 at http://apps.nccd.cdc.gov/yrbss/QuestYearTable.asp?cat=4&Quest=Q62&Loc=XX&Year=2007&compval=&Graphval=yes&path=byHT&loc2=&colval=2007&rowval1=Sex&rowval2=Race&ByVar=CI&Submit2=GO.

Arch Sex Behav (2010) 39:1321–1330
DOI 10.1007/s10508-009-9510-x

ORIGINAL PAPER

Partner Selection among Latino Immigrant Men Who Have Sex with Men

Fernanda T. Bianchi · Michele G. Shedlin · Kelly D. Brooks · Marcelo Montes Penha · Carol A. Reisen · Maria Cecilia Zea · Paul J. Poppen

Received: 22 October 2008 / Revised: 26 May 2009 / Accepted: 26 May 2009 / Published online: 18 August 2009

Abstract This qualitative study explored partner selection in a sample of immigrant Latino men who have sex with men (MSM). In-depth interviews were conducted with men living in the greater New York metropolitan area who had been born in Brazil ($n = 10$), Colombia ($n = 14$), or the Dominican Republic ($n = 9$). One focus group was conducted with MSM from each of the three countries (9 Brazilian, 11 Colombian, and 5 Dominican participants). A grounded theory approach revealed three main themes relating to partner selection. The first concerned stereotypes of how Latino and Anglo-American men tend to behave in their sexual encounters and relationships. The participants perceived Latinos to be more affectionate and passionate, whereas they saw Anglo-American men as more independent and practical. These cultural discrepancies sometimes resulted in a preference for Latino partners. A second theme concerned stereotypes of the national groups, including expectations that Brazilians would be sexy and sensual and that Dominicans would have large penises. As found in other research on MSM of color, ethnic and national stereotypes were associated with experiences of sexual objectification. The third theme addressed the importance of masculine characteristics in sexual attraction and partner selection. Negative feelings towards effeminate men who did not conform to normative male physical or behavioral presentation reflect a stigma found inside and outside of the gay community. These findings suggest that gender and ethnic stereotypes play an important role in shaping partner choice and have implications for sexual risk and relationship formation.

Keywords Latino · MSM · Sexual partner · Stereotypes · Masculinity

F. T. Bianchi (✉) · C. A. Reisen · M. C. Zea · P. J. Poppen
Department of Psychology, George Washington University, 2125 G Street NW, Washington, DC 20052, USA
e-mail: fbianchi@gwu.edu

M. G. Shedlin
College of Nursing, New York University, New York, NY, USA

K. D. Brooks
Department of Psychology, Illinois Wesleyan University, Bloomington, IL, USA

M. Montes Penha
Department of American Studies Program, New York University, New York, NY, USA

Introduction

Research on the selection of sexual partners has frequently focused on factors influencing the choice of long-term mates by heterosexual men and women, whereas the process of partner selection among men who have sex with men (MSM) has received much less attention. The literature has shown that individuals often seek mates who are similar to themselves in intelligence, values, personality attributes, and physical characteristics (Buss, 1985). In addition, certain traits in potential partners, including intelligence, sense of humor, honesty, kindness, good looks, facial attractiveness, and dependability, have been reported as important by men and women—both gay and straight (Bailey, Gaulin, Agyei, & Gladue, 1994; Gómez, Mason, & Alvarado, 2005; Lippa, 2007). Age of partners has also been reported as relevant, with a preference for younger aged partners professed by gay and straight men (Silverthorne & Quinsey, 2000).

This qualitative study examined partner selection in a sample of immigrant Latino men who have sex with men. Little has been reported about partner selection among MSM, and even less is known about the process among Latino immigrant MSM. This group is at high risk for HIV (CDC, 2008) and, therefore, it is important to gain further understanding of their sexual partnering.

Theoretical approaches to the issue of partner selection have often taken an evolutionary perspective, emphasizing the survival value of strategies of mate selection (Buss, Shackelford, Kirkpatrick, & Larsen, 2001), including when explaining differences in preferred traits for partners in long-term relationships and short-term sexual encounters (Buss & Schmitt, 1993). Although MSM grow up within a dominant heterosexual culture that emphasizes the link between sex and reproduction, it has been argued that among MSM the type of relationship desired is more relevant than evolutionary factors in determining the importance of various traits (Gobrogge et al., 2007). In a study of internet personal advertisements, gay men sought short-term sexual encounters more frequently than straight men, but both gay and straight men reported a wider acceptable age range for short-term than long-term partners. In studying MSM, it is important to examine partner selection for short-term sexual encounters, as well as longer term relationships.

The process of partner selection can also be shaped by issues stemming from ethnicity and culture (Díaz & Ayala, 1999; Gómez et al., 2005). In qualitative interviews with participants in the Seropositive Urban Men's Study, race and ethnicity were often cited when describing an ideal partner, with some men preferring partners of their own ethnicity and others preferring those with different backgrounds (Gómez et al., 2005). Racial, ethnic, and national stereotypes influenced not only participants' preferences for certain partners, but also potential partners' perceptions of participants. Many men of color have reported experiences of sexual stereotyping and objectification (Gómez et al., 2005); for example, African American men have described being approached solely because of the stereotyped view that they have large penises (Teunis, 2007).

Partner selection among MSM also can be influenced by gender expectations. Gómez et al. (2005) noted that the majority of their participants valued masculine and devalued feminine appearance in their partners, but that this pattern was especially evident among the Latino and African American men. For these men, masculine appearance and comportment were typically viewed as essential in a partner.

The preference for physically masculine partners may be shaped by a gay cultural context, as well as more general Western patriarchal values that accord greater status to men. The development in the gay community of an ideal of a physically masculine man has been described as a reaction against earlier conceptions of gay men as effeminate (Connell, 2005), as well as a response to AIDS and its accompanying loss of body mass strength (Halkitis, 2001). Furthermore, it has been argued that the stigma associated with gender-nonconformity in the larger society is internalized by gay men and results in a preference for masculine partners and an antipathy toward effeminate men (Taywaditep, 2001).

The desire for partners who behave in a masculine manner can also be viewed within a Latino cultural context. Appropriate masculine behavior in Latino cultures has sometimes been described with the construct of machismo. Although there has been disagreement about the definition and current applicability of the term, machismo was historically depicted as hypermasculinity (Gutmann, 1996; Mirandé, 1997). Characteristics seen as male (e.g., activity, strength, dominance, sexual prowess, aggression) were prescribed, and those seen as female (e.g., passivity, weakness, and vulnerability) were proscribed for men (Mirandé, 1997).

The enactment of masculinities in Latino cultures, however, has been shown to be more complex and varied than this definition implies (Gutmann, 1996; Mirandé, 1997). Two recent studies of machismo provided evidence of multiple dimensions of Latino masculinity, including emotional expressiveness and affiliation (Arciniega, Anderson, Tovar-Blank, & Tracey, 2008; Torres, Solberg, & Carlstrom, 2002). Furthermore, differences in gender stereotypes for expressing emotions (e.g., love, sadness, embarrassment) have been shown to be smaller for Latinos in the U.S. than for European Americans, thus suggesting greater cultural acceptance for emotional expression among Latino than among Anglo men (Durik et al., 2006). However, among Latino MSM, as among other groups of MSM, other aspects of femininity, such as mannerism and appearance, are devalued (Sandfort, Melendez, & Díaz, 2007; Taywaditep, 2001).

This qualitative study investigated partner selection—the process of choosing and being chosen—among Brazilian, Colombian, and Dominican MSM in the New York area. Although representation of these three groups in the U.S. has substantially increased in recent decades, they have received little research attention (Logan, 2001; Luiz, 2005). It has been argued that more research should address heterogeneity among Latinos (Torres et al., 2002), and the inclusion of Latino immigrants from three countries enabled us to examine both commonalities and discrepancies in the experience of Latino MSM. In this article, we explored three broad issues that arose repeatedly in the conversation of the men in this study. First, we explored how stereotypes of Latino and Anglo-American men impacted their sexual encounters and relationships. Second, we examined how ethnic and national stereotypes influenced partner selection. Third, we investigated how conceptions of masculinity shaped partner choice.

Method

Participants and Procedure

The research for this article was conducted in the greater New York City area (Manhattan, Queens, Newark) among immigrant Brazilian, Colombian, and Dominican MSM. Participants were recruited between autumn 2004 and spring 2005 as part of a larger research project on contextual influences of sexual risk behavior among immigrant Latino MSM (see Zea, Reisen, Poppen, & Bianchi, in press). Data reported in this paper came from focus groups and in-depth interviews conducted in the qualitative phase of the research project. Recruitment methods included snowball sampling, advertisements in gay publications and on internet websites, flyers, and referrals from health and prevention programs serving the Latino gay community.

Three focus groups were conducted, one for participants from each of the three countries of origin, in the participants' native language (i.e., Spanish or Portuguese). There were two facilitators in each focus group, at least one of whom was a native speaker and at least one of whom was a gay man (sometimes the same person). Focus group guides were used to structure the conversation in a manner that ensured relevant issues were covered, but facilitated spontaneous discussion and inclusion of unanticipated topics. Topics covered included immigrant gay community, sexual partners and activities, sexual identity, drug and alcohol use, and HIV and STI risk. Individuals of the three national groups were asked about ways in which their nationality affected their sex lives within the gay community in New York City. Focus groups lasted approximately 2 h. Participants were reimbursed $50 for their time and given $5 to cover the cost of transportation.

Eligibility criteria included having been born in Brazil, Colombia, or the Dominican Republic, being 18 years of age or older, and having had sex with a man in the last three months. Individuals who met eligibility requirements could participate in either the focus groups or in-depth interviews, but not both. The number of participants in each of the focus groups was as follows: 9 in the Brazilian group, 11 in the Colombian group, and 5 in the Dominican group. The ages of focus group participants ranged from 20 to 57 years, with a mean of 37 years. The average length of time in the USA was approximately 10 years, with a range of 4 months to 24 years. About 60% of the participants in focus groups reported that they were in a relationship with a primary male partner.

Semi-structured in-depth interviews were conducted in the language of preference of the participant. Interviewers were native speakers of Spanish or Portuguese, but also fluent speakers of English. Interview guides were developed to loosely structure the conversations and to address a variety of topics, including migration, sexual behavior, sexual partners, condom use, and drugs and alcohol. Questions were sufficiently open-ended to allow unanticipated topics and themes to emerge and be discussed. Interviews lasted approximately 90 min, and participants received $65 in reimbursement for time and transportation costs. Questions that addressed the topics for this paper most directly were: "Do you think that your race, ethnicity or national origin impacts your sexual life?" and "Tell me a little about whom you are sexually attracted to. Do you have any preferences in terms of ethnicity, style, age, gender roles, etc.?"

Participants in the in-depth interviews met the same inclusion criteria as focus group participants. Fourteen Colombian, 12 Dominican, and 10 Brazilian men participated in the in-depth interviews, but three Dominican participants were later eliminated because they revealed in the interviews that they were not immigrants, but rather had been born in the U.S. The remaining 33 interviewees comprised the final sample. Table 1 shows the demographic characteristics of the participants from the in-depth interviews, who are listed there and in the text below with pseudonyms. We do not use any name for the focus group participants.

The average age of those interviewed was 35 years, with a range from 18 to 58. The median length of time in the U.S. for interview participants was 10 years, with a range from one to 26 years. About 40% of the participants reported having a primary partner at the time of the interview. The Dominican participants differed from the Colombian and Brazilian participants in that they were slightly younger and less educated. Even though the majority of the sample had some college education, many of our participants were economically disadvantaged because of their immigrant status.

All focus groups and interviews were tape-recorded and data were transcribed for analysis using Atlas.ti 5.0. Interviews were coded in the original languages by a team of four researchers, three of whom were native speakers of Portuguese or Spanish. There were at least two coders for each interview, one of whom was a native speaker of the language of the interview. The coding of the first four interviews was done simultaneously by three coders in order to establish criteria for coding. Thereafter, the majority of interviews were coded by two coders and discrepancies in coding were resolved through discussion and consensus as suggested by methodology guidelines provided by Frieze (2008). The analysis of the data was guided by the principles of grounded theory (Glaser, 1992), such that themes and relationships were identified and then modified, accepted, or rejected as additional information and insight emerged over the course of the analysis. Categories were developed within each theme and continuous coding and use of memos were applied to allow for ongoing revision of categories. The final scheme of themes and categories was reviewed and checked against the coded data.

Table 1 Demographic characteristics of participants for in-depth interviews ($N = 33$)

Name[a]	Age	Years in U.S.	Education	Relationship status	HIV status
Brazilians					
Antônio	43	19	Some college	Widower	Positive
Jorge	52	10	College	Single	Positive
Francisco	52	23	Some college	Partnered	Positive
João	34	5	College	Single	Positive
Gilberto	36	7	College	Single	Positive
Joaquim	39	19	College	Partnered	Unknown
Guiomar	25	5	Some college	Partnered	Unknown
Bruno	28	1	Some college	Partnered	Unknown
Edson	28	2	College	Girlfriend	Negative
Oswaldo	31	8	High school	Single	Negative
Colombians					
Rodolfo	50	16	High school	Partnered	Unknown
Edgar	45	24	Some college	Single	Unknown
Luis	58	24	Graduate school	Single	Positive
Gabriel	48	3	College	Single	Negative
Eduardo	35	9	College	Partnered	Negative
Esteban	34	10	College	Single	Negative
Rubén	35	15	Some college	Partnered	Negative
Gonzalo	42	3	Graduate school	Partnered	Positive
Mario	38	7	Some college	Single	Negative
Carlos	23	15	High school	Partnered	Negative
Jairo	28	10	College	Partnered	Negative
Luciano	34	9	High school	Single	Negative
Mauricio	29	16	College	Single	Negative
Álvaro	33	2	Graduate school	Single	Negative
Dominicans					
Fabio	45	20	High school	Single	Unknown
Guillermo	25	14	High school	Single	Unknown
Leopoldo	18	18	Some high school	Single	Negative
Samuel	42	12	Some college	Single	Positive
Pedro	33	4	Some college	Partnered	Unknown
Pablo	23	6	Some college	Single	Negative
Rafael	31	26	Graduate school	Single	Unknown
Miguel	28	28	High school	Single	Negative
Felix	32	16	Some college	Partnered	Negative

[a] All names in this table are fictitious

Results

Three themes concerning partner selection emerged in the interviews and focus groups: (1) stereotypes of Latino and Anglo-American men; (2) stereotypes of the specific immigrant groups; and (3) masculinity in sexual preferences and roles. The topics covered within each of these themes are presented in Table 2, as is the number of respondents from each group who contributed to each category. Because some topics emerged spontaneously within interviews and focus groups, not all topics were discussed by all respondents.

Stereotypes of Latino and Anglo-American Men

The participants reported a consistent difference in how Latino and Anglo-American men approached sexual encounters and romantic relationships. Sex was described as being something more "cerebral" and practical for Anglo-Americans, but more emotional and passionate for Latinos. The perception that Anglo-American men are less affectionate and more pragmatic than Latinos sometimes resulted in feelings of emptiness or lack of intimacy experienced by participants during sex. A Brazilian man (Jorge) described

Table 2 Number of interviews and focus groups contributing to major themes and categories by ethnic group

Major themes and categories	Brazilian	Dominican	Colombian
Stereotypes of Latino and Anglo-American men			
Latinos—more passionate and lack control over sexual situations	1 IDI[a]	–	2 IDI
Anglo-Americans—more cold, practical, independent in relationships	3 IDI, FG[b]	–	2 IDI, FG
Idealized image of the "gringo" boyfriend	1 IDI, FG	FG	FG
Stereotypes of the specific immigrant groups			
Sexual objectification of Brazilian and Dominican men	7 IDI, FG	4 IDI	–
Positive consequences on sexual stereotypes	1 IDI	1 IDI, FG	FG
Negative consequences on sexual stereotypes	3 IDI, FG	1 IDI, FG	1 IDI
Masculinity in sexual preferences and roles			
Preference for masculine men and/or devaluation of effeminacy	3 IDI, FG	3 IDI, FG	7 IDI
Sexual roles influenced by masculinity/femininity of partners	2 IDI	–	5 IDI

[a] In-depth interviews (IDI), $N = 33$

[b] Focus group (FG), Brazilian ($N = 9$), Dominican ($N = 5$), Colombian ($N = 11$)

sexual relations in the U.S. as devoid of feelings and mechanical:

> [In the U.S.] you are like a robot, a sex machine… you have sex, you get up, you are done, you leave, you clean yourself, and you don't even know who you had sex with. With me, it is different, it is not like that. Feelings must be involved, one must feel something to be able to have sex with another person, which the majority here don't do.

In addition, the men in this study expressed the belief that Latinos are passionate, more spontaneous, and less likely to plan or prepare for a sexual encounter—qualities that could potentially lead to greater HIV risk. For example, this situation was described by a Colombian participant (Edgar):

> I think it is like one's nature, no? That is, I don't know if [Latinos] are hotter or more passionate, or what it is. But it's like one loses control over taking care of oneself. On the other hand, not an American. An American… it has happened to me that we go to bed and the first thing he does is open the night table and put the condom on me for when the moment comes. Not the Latino. In fact, one says to him, "Do you have a condom?" "No, I don't have a condom." "Well then, there is no condom, let's go."

A lack of feeling of closeness was an especially salient theme that emerged throughout discussions relating to Anglo-American partners, especially when participants were seeking an enduring relationship. Some of the participants (e.g., Gabriel) believed that "it is very difficult to find a [stable] partner in the U.S." As João, a Brazilian man put it, "Here, it is very hard to have relationships, I have tried several times…people only want to have sex." The difficulty in finding stable relationships in the U.S. was partially attributed to cultural differences in norms concerning affiliation, relationships, and emotional attachment. A participant in the Colombian focus group made the following observation:

> When you first start seeing someone, you are in love, and you want to be with that person all the time, you want to make plans with that person, you want to share good things with that person. In contrast to an American, who makes plan on his own, [a Colombian] makes plans with his family and his partner. Therefore, there are these cultural clashes…

The dissatisfaction with relationships in the U.S. differed starkly from participants' expectations upon coming to the U.S. Many initially anticipated dating or forming long-term relationships with Anglo-American men. The desire of Latino MSM to explore something new—to date Anglo-Americans rather than other Latinos—was discussed in both the Brazilian and Dominican focus groups. One participant explained: "I believe that each one of us when you first set foot in this country that the last thing you thought was that you would have a relationship with another Brazilian gay man, [but] with a Gringo, yes. Am I lying?"

Not only were Anglo-American men new, appealing, and different, they represented a social status unavailable to many newly arrived immigrant Latinos. A Colombian focus group participant described an idealized image of the "blue-eyed and white" Anglo-American man, a partner preference arising from white men's dominant status within the U.S. Moreover, Latino men who became romantically involved with American men were often able to experience some of the privilege associated with their partner's social status. A Brazilian man (Bruno) described the experiences of some immigrant MSM with Anglo-American partners in this manner:

> I think that the Brazilians who become involved with Americans do fine for a while and they live a fantasy life that is not theirs, because many Americans have a good

> social position…And the Brazilians who arrive here and get involved with them get taken in by these conditions—of living well, of having a car available, and in the majority of times of having easy money through the boyfriend….It is not like they [the Brazilians] get closer [to Americans] out of self-interest, but rather the person offers them these things and they let themselves be swayed by it.

Some participants described how their initial desire to have a relationship with Anglo-Americans dissipated over time. They found themselves gravitating back towards other Latino partners. As a Dominican focus group participant noted, he felt "better off with my own race [ethnic group] because I know it better and get along well with my own race." A participant in the Colombian focus groups explained it in this way:

> So, you try to have a relationship with an American, in general it is one of the first relationships you have, and even if you speak English very well, with you being Latino, there will always be a time when, when you get frustrated. You don't understand, for example, what your partner is trying to say, you don't understand what he is feeling, what he is experiencing…and these are things that keep repeating themselves, keep repeating …and then you say, "Oh well, there is a cultural difference here, and I prefer to be with a person who has customs similar to mine". Even if this other [Latino] person is jealous or possessive, they are not as individualistic as the Americans are….and we are not as cold as the Americans are. We are used to having a person who is more attached by the bone to us. And that is why I think I prefer a Latino or Colombian.

Stereotypes of the Immigrant Groups

In addition to talking about their own preferences for Latino and Anglo-American partners, participants discussed how perceived characteristics of men from their national groups affected potential partners' reactions to them. Brazilians and Dominicans had strong beliefs about their group's sexual reputations, whereas Colombians did not.

There was a widely shared opinion among Brazilian participants that Brazil's image as an exotic and sexually charged country was responsible for an image of Brazilians as highly sensual and sexual by nature. According to the participants, Brazilian MSM were regarded as very desirable in the U.S. because of the expectation that they were good lovers, who were "thinking of, talking about, and looking for sex" all the time. A Brazilian man (Gilberto) described other MSM's reaction after finding out his nationality in this way:

> When you say that you are Brazilian, man, it sounds like it is something out of this world. It seems that they are winning the lottery…there is a certain frenzy, it is like a shot of energy.

In a similar vein, many potential sexual partners' attraction to Dominicans was attributed to the reputation that Dominicans have large penises. Dominican participants also reported that others frequently assumed they were sexually virile and always ready for sex. They noted that with non-Dominican partners, they were often expected to be the active and insertive partner. A Dominican participant (Felix) described the reputation of Dominican gay men in the U.S. in this way: "I think that being Dominican gives me the stamp of being hot and it's expected that… I be more "macho," more active, and that I enjoy penetrating, that I have a big penis."

In contrast to the Brazilian and Dominican participants, Colombians did not report any particular sexual stereotypes or reputation that would set them apart from other Latinos. Some Colombians described themselves as being passionate and affectionate in their sexual interactions, similar to other Latinos. In addition, some Colombian participants emphasized their manner of speaking and level of education as features that made them attractive to Anglo-American men.

The data indicated both positive and negative ramifications of the sexual stereotypes of Latino men. Although the reputations of Latinos in general or Brazilians and Dominicans in particular made attracting casual partners easier, many participants felt that these generalizations contributed to sexual objectification. For example, Alvaro, a Colombian, expressed such feelings when he stated that "[Americans think] 'Ah Latinos, to have sex with a Latino must be very good.' But, they really are not interested in anything else. And I am tired of that."

Similarly, a Brazilian, Oswaldo, noted his experience of sexual objectification due to his nationality being equated with sex in the following way:

> Many people attack me [sexually] because I am Brazilian. As soon as you mention that you are Brazilian, they want sex. Do you understand? This is very offensive, you know? You probably have heard about this [before]. You mention being Brazilian and everybody [says], "Wow! My gosh, Brazilian! Hmmm, Good lay!" This happens a lot… being Brazilian means that one is great for sex.

Pedro, a Dominican participant, described the process of becoming disillusioned with the objectifying treatment he experienced:

> That was a label that they put on me, "The sexy one," sexy everywhere I turned, and in the beginning it seemed to be quite flattering, but after a while I realized

> that it was, or at least to my mind it seemed to be, something quite offensive to tell you that you are sexy, because what they mean was that you are good to fuck with, nothing else.

Participants also described ways in which sexual objectification sometimes resulted in pressure to perform sexually or to conform to expectations. A Brazilian focus group participant noted: "One has to prove that one is good [in bed] since one is Brazilian, one must show the national product…it's like people are saying to you, 'Come on, you Brazilian, prove that you are a Brazilian for real!" Francisco noted that objectification led potential sexual partners to assume that as a Brazilian, he would be willing to engage in a wide range of sexual behaviors: "I don't like [this reputation] because sometimes it is taken to the level of vulgarity or that anything can happen and I believe in individuality; not all Brazilians will have that sensuality or that sexuality."

Stereotypes may be particularly difficult for those men who fail to conform to the expectations placed on their group, which can even lead to distancing one's self from the group. João explained: "Since my sexual situation is not very active right now, I avoid saying that I am Brazilian because I want to avoid causing a certain disappointment….I tend to say that I am Latino or South American."

Masculinity in Sexual Preferences and Roles

Beliefs about masculinity and femininity were widely discussed by participants in relation to partner selection. Masculine characteristics were viewed as important criteria for sexual attraction and several participants indicated that they would never become interested in a man who was effeminate. A Brazilian participant (Oswaldo) described his preference for sexual partners in this way:

> I like men who are very masculine. If the person is effeminate, then it is a real turn off. The person can be homosexual, as long as he is not flamboyant….[if he is] I am not attracted, but when he is a man, a real man then I am attracted.

Similarly, a Dominican focus group participant remarked, "I go to bed with guys who are masculine. Be it to fuck or be fucked, but an effeminate man, no."

The acceptability of certain female characteristics varied for different participants. Rubén, a Colombian, reported that his partner had a small body size, like a woman, but was not feminine. This physical characteristic made Rubén feel more masculine and in control, as if he were the "Papi." Other participants noted that ways of dressing or grooming traditionally associated with women (e.g., tweezing one's eyebrows, wearing cologne) were "too gay" and a deterrent to pursuing a relationship.

The definition of what constitutes masculinity was broader, however, than might be commonly found in Anglo-American culture in that it included a greater acceptance of emotionality. The Latino men in this study did not associate emotional expressiveness with femininity. As one Colombian focus group participant explained, "We want to feel not feminine, but rather possessive and passionate." Thus, this type of emotional behavior was not viewed as incompatible with male gender expectations.

The stigma associated with gender non-conformity in both the larger society and gay sub-culture provided an additional motive for the choice of masculine partners. Bruno noted his distaste for partners who display feminine traits:

> If I wanted a woman I would grab a real one. So, I like a person who carries himself as a man, who talks in a normal way, who dresses like a man, and that he be a person that when we are seen together on the street people would not say, "Here comes two faggots."

One participant, Luciano, described being conflicted about his feminine tendencies and attempting to conform to a masculine ideal. He described himself as being effeminate his entire life and talked about his internal struggles to become more masculine in this manner:

> I was very effeminate while growing up and I suffered a lot with that…so I want to be masculine. That is what I want and that is what I like in other men, so I also want to be masculine. So, I started to study the things that I did, how I saw myself, the things that I did unconsciously that were feminine-like, that were considered feminine by society, and I started, little by little [changing] and I am still working at it.

Partner selection for MSM could also involve the issue of sexual role, and sometimes in the Latino cultural context, insertive partners are considered masculine and receptive partners are considered feminine. As one Colombian participant (Jairo) put it, "We associate being the bottom as taking the role of the woman, of being passive…of one who is willing to receive the man and letting the man penetrate." A preference for masculine-seeming receptive partners was also reported, however, which is consistent with the general preference for masculine partners. A Colombian participant (Gabriel) noted that he enjoyed having partners who were receptive, but "act straight, in other words like men."

Most men in this study reported being versatile in their role for anal intercourse (i.e., taking insertive and receptive roles), but concerns about gender implications of sexual roles influenced their behavior. For example, Alvaro remarked that he preferred to be the top and was only willing to be the receptive partner when, "the person [is] very virile, someone who awakens my desire to be 'bottom'….very, very masculine." Another participant, Bruno, reported taking the insertive role

at the beginning of the sexual encounter, and thereby establishing his masculinity: "I don't have a problem in being bottom or top….because it is an exchange. But, in the beginning I get pleasure from being the top like a man."

Discussion

Cultural expectations and stereotypes influenced the process of partner selection among the Brazilian, Colombian, and Dominican immigrant MSM in ways that were relevant for both short- and long-term relationships. Common perceptions of Anglo-American men as cold, practical, and independent contrasted with self-perceptions of Latino men as passionate, emotionally expressive, loyal, and desirous of close relationships. Although one would expect such perceived differences to be salient in ongoing relationships, it was also an issue in brief encounters.

The more collectivist orientation of Latino participants was evident in their desire for greater emotional contact with partners, even in short-term sexual encounters, where they sometimes perceived their Anglo-American partners as behaving in an impersonal manner. In addition, the Latino men's expectation that being in love involved having an intense relationship, spending time together and with family, and sharing all aspects of one's life conflicted with the more independent approach sometimes taken by their Anglo-American partners, which at times resulted in Latino men feeling lonely and dissatisfied in longer term relationships. It is interesting to note that there was a tendency in some men toward decreased desire for Anglo-American partners with increased exposure and time in the U.S. due to these cultural conflicts. Although the discrepancy in emotional style from that perceived as typical of Anglo-American men was a reported theme, it was not universally experienced by the Latino men in this study.

The cultural expectations and stereotypes described had potential implications for HIV risk. Sexual scripts for Latino men included passion and spontaneity—characteristics that many described with pride. These characteristics, however, were thought to contribute to Latinos' lack of preparation for and loss of control during sexual encounters, which ultimately hindered consistent condom use. Previous research has noted that loss of sexual control was a common reason given by Latino gay men to explain failure to engage in safe sex (Díaz & Ayala, 1999). Some Latino men in the current study interpreted the preparation, planning, and self-control that their Anglo-American partners brought to sexual encounters as antithetical to passion. It is important that HIV prevention efforts for Latinos frame safer sex techniques as compatible with caring, passion, and emotional intimacy.

Despite the cultural differences, many Latino immigrant MSM reported being drawn to Anglo-American partners, particularly in the early period after arriving in the U.S. Anglo-American partners tended to have higher social status and greater financial resources, and often provided their Latino partners with access to privilege and money. Power imbalances between partners, particularly when one partner is receiving material goods or support from the other, can diminish the dependent partner's agency in sexual situations and therefore may increase his HIV risk (Díaz & Ayala, 1999). Although this may have been the case for the low-income participants in the current study, more affluent Latino immigrants would not be expected to experience this type of power differential in their relationships with Anglo-American men.

Sexual stereotyping of racial, ethnic, and national groups was an evident influence in the process of choosing and being chosen as a sexual partner. Such stereotyping involves generalizations about potential partners based on characteristics attributed to their groups. Although there may be some actual differences among groups, the process of stereotyping fails to differentiate among individuals within groups. Stereotypes of Latino men as passionate lovers, of Brazilians as highly sensual and sexual, and of Dominicans as having large penises were repeatedly reported. Although such images could be interpreted as describing positive attributes in a lover, they contributed to sexual objectification and dehumanization. Some participants in this study felt that potential partners reduced them to their national stereotypes, failed to see them as whole people, and valued them only for their assumed sexual characteristics. Although the Colombians in this study perceived their own educational achievements as attractive to potential partners, they and the Brazilian or Dominican men made little mention of other personality characteristics.

The findings of this study were consistent with other research showing objectification as a common experience among gay men of color (Drummond, 2005; Gómez et al., 2005; Teunis, 2007). Sexual stereotyping and objectification may be especially common in partner selection for anonymous or brief sexual encounters. White men's dominant status, racist views of men of color, and perceptions of the foreign as exotic or sexually exciting all contribute to sexual objectification of men of color within the gay community (Teunis, 2007). Research on sexual objectification has highlighted the negative impact on women's mental health and well-being resulting from objectification (e.g., Fredrickson, Roberts, Noll, Quinn, & Twenge, 1998; McKinley, 1999). Recently, studies have begun to document similar effects of sexual objectification among gay men (Martins, Tiggeman, & Kirkbride, 2007).

The Brazilian, Colombian, and Dominican men in this study expressed a preference for masculine sexual partners almost exclusively, a finding that is consistent with previous research in multi-ethnic samples (Bailey, Kim, Hills, & Linsenmeier, 1997; Gómez et al., 2005). Moreover, they believed

that their own masculine appearance and behavior were important in their success in attracting partners themselves. This partiality for masculinity may reflect an implicit—and largely unconscious—endorsement of patriarchal views, which value characteristics and behaviors ascribed to men (e.g., activity, strength, aggression) over those ascribed to women (e.g., passivity, weakness, vulnerability). For Latino men, gender-conforming masculine behavior also includes emotional expression, nurturance, and family orientation (Arciniega et al., 2008; Torres et al., 2002); therefore, the greater desire for warmer partner relationships than sometimes encountered with Anglo-American men, as discussed above, is not inconsistent with enactment of masculinity.

In Latino cultures, sexual roles during anal intercourse have sometimes been associated with gender, such that insertive (active) partners are viewed as masculine and receptive (passive) partners are viewed as feminine. We found some support for this perception in our study, as the relative masculinity of partners sometimes determined sexual role, with more masculine partners tending to take the insertive role. Preference for a specific sexual role was rarely mentioned however, as an influence in partner choice; it was clearly secondary to masculine characteristics in steering partner selection.

In addition to the expressed preference for masculinity, we also found an antipathy toward effeminate men as potential partners. Attitudes stigmatizing effeminate men are evident in the larger culture, as well as in gay and Latino cultures (Connell, 2005; Mirandé, 1997; Taywaditep, 2001). This distancing from or rejection of men who are not masculine may reflect internalized homonegativity or self-hatred among the MSM themselves: for some participants, masculinity was associated with heterosexuality and valued, whereas femininity was associated with homosexuality and shunned. A number of authors have pointed out the internalized homonegativity and shame inherent in such a belief system, which is not limited to Latino men (Bergling, 2001; Taywaditep, 2001).

Because gay men have traditionally been stereotyped as displaying effeminate characteristics, distancing themselves from effeminacy and men who appear effeminate may serve to protect some MSM from the stigma associated with homosexuality. They can preserve their own self-esteem by linking sexual-orientation prejudice to gender atypicality, and not to homosexual behavior per se. In addition, because MSM continue to be targets of prejudice and discrimination based on the stigma associated with homosexuality (Herek, 2007), some MSM avoid being seen with effeminate men. In this way, they can lessen the probability of being identified as gay in public, thereby shielding themselves from being a target of homophobic behavior.

Unfortunately, the implications of these attitudes for the well-being of effeminate men can be grave. Gender nonconforming gay men tend to report more anti-gay discrimination than gender-conforming men and experience poorer mental health outcomes (D'Augelli, Grossman, & Starks, 2006; Sandfort, Melendez, & Diaz, 2007; Taywaditep, 2001). Potential rejection not only by the larger society, but also by the gay community, can lead to negative self-concept and low self-esteem among effeminate men. Moreover, the lack of attraction to effeminate men among many MSM may result in the disempowerment of such men in their sexual relationships. Effeminate men's compromised power position relative to their partners could diminish their ability to negotiate condom use and insist on safer sex practices. This situation, in combination with an expectation that effeminate men would take the receptive role in anal intercourse, could result in their increased vulnerability to HIV.

Limitations of this study included an approach that allowed some themes to emerge naturally and, therefore, all participants did not address all issues. Consequently, a direct comparison of the experiences of the Brazilian, Colombian, and Dominican men on all topics was not possible. The men in this study discussed numerous cultural characteristics and stereotypes that they believed were relevant to their choosing or being chosen as a sexual partner. Although such stereotypes clearly affected the experience of our participants and their partners, we recognize that such stereotypes do not accurately describe entire ethnic or national groups. Moreover, the questions that opened the topic in the in-depth interviews first addressed sexual attraction and then queried preferences for partners based on ethnicity, style, age, and gender roles. Such questions could have influenced respondents to focus their attention on short-term sexual encounters, in which partner characteristics such as appearance would be more salient than qualities such as intelligence or honesty. Because short-term encounters are common among MSM, however, it is important to explore partner selection in this context as well.

Further research is needed to investigate the prevalence of the perceptions and experiences reported here and the extent to which these gender- and culture-related factors influence sexual behavior and sexual risk. In addition, future research could explore mechanisms by which sexual stereotyping might contribute to self- and other-objectification, as well as the consequences of objectification among MSM and their relationships with sexual partners. It is also important to further understand the difficulties that MSM encounter in trying to conform to gendered expectations of themselves and the strategies used to deal with these difficulties especially for men how are perceived as effeminate and stigmatized. Gaining greater insight into the context of Latino gay men's

sexual encounters and relationships within the U.S. is crucial in order to design effective HIV/AIDS prevention programs and meet the mental health needs of this population.

Acknowledgement The preparation of this paper was supported by Award Number R01 HD046258 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development (NICHD). The content is solely the responsibility of the authors and does not necessarily represent the official views of NICHD or the NIH.

References

Arciniega, G. M., Anderson, T. C., Tovar-Blank, Z. G., & Tracey, T. J. G. (2008). Toward a fuller conception of machismo: Development of a traditional machismo and caballerismo scale. *Journal of Counseling Psychology, 55*, 19–33.

Bailey, J. M., Gaulin, S., Agyei, Y., & Gladue, B. A. (1994). Effects of gender and sexual orientation on evolutionarily relevant aspects of human mating psychology. *Journal of Personality and Social Psychology, 66*, 1081–1093.

Bailey, J. M., Kim, P. Y., Hills, A., & Linsenmeier, J. A. W. (1997). Butch, femme, or straight-acting? Partner preferences of gay men and lesbians. *Journal of Personality and Social Psychology, 73*, 960–973.

Bergling, T. (2001). *Sissyphobia: Gay men and effeminate behavior*. Binghamton, NY: Haworth.

Buss, D. M. (1985). Human mate selection. *American Scientist, 73*, 47–51.

Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review, 100*, 204–232.

Buss, D. M., Shackelford, T. K., Kirkpatrick, L. A., & Larsen, R. J. (2001). A half century of mate preferences: The cultural evolution of values. *Journal of Marriage and Family, 63*, 491–503.

Centers for Disease Control (CDC). (2008). HIV/AIDS surveillance in men who have sex with men (MSM). http://www.cdc.gov/hiv/topics/surveillance/resources/slides/msm/slides/msm9.pdf. Retrieved 6 Oct 2008.

Connell, R. W. (2005). *Masculinities* (2nd ed.). Berkeley, CA: University of California Press.

D'Augelli, A. R., Grossman, A. H., & Starks, M. T. (2006). Childhood gender atypicality, victimization, and PTSD among lesbian, gay, and bisexual youth. *Journal of Interpersonal Violence, 21*, 1462–1482.

Díaz, R. M., & Ayala, G. (1999). Love, passion and rebellion: Ideologies of HIV risk among Latino gay men in the USA. *Culture, Health, & Sexuality, 1*, 277–293.

Drummond, M. J. N. (2005). Asian gay men's bodies. *Journal of Men's Studies, 13*, 291–300.

Durik, A. M., Hyde, J. S., Marks, A. C., Roy, A. L., Anaya, D., & Schultz, G. (2006). Ethnicity and gender stereotypes of emotion. *Sex Roles, 54*, 429–445.

Fredrickson, B. L., Roberts, T. A., Noll, S. M., Quinn, D. M., & Twenge, J. M. (1998). That swimsuit becomes you: Sex differences in self-objectification, restrained eating, and math performance. *Journal of Personality and Social Psychology, 75*, 269–284.

Frieze, I. H. (2008). Publishing qualitative research in *Sex Roles* [Editorial]. *Sex Roles, 58*, 1–2.

Glaser, B. G. (1992). *Emergence vs. forcing: Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.

Gobrogge, K. L., Perkins, P. S., Baker, J. H., Balcer, K. D., Breedlove, S. M., & Klump, K. L. (2007). Homosexual mating preferences from an evolutionary perspective: Sexual selection theory revisited. *Archives of Sexual Behavior, 36*, 717–723.

Gómez, C. A., Mason, B., & Alvarado, N. J. (2005). Culture matters: The role of race and ethnicity in the sexual lives of HIV-positive gay and bisexual men. HIV+ sex. In P. N. Halkitis, C. A. Gomez, & R. J. Wolitski (Eds.), *The psychological and interpersonal dynamics of HIV-seropositive gay and bisexual men's relationships* (pp. 87–100). Washington, DC: American Psychological Association.

Gutmann, M. C. (1996). *The meanings of macho: Being a man in Mexico City*. Berkeley, CA: University of California Press.

Halkitis, P. N. (2001). An exploration of perceptions of masculinity among gay men living with HIV. *Journal of Men's Studies, 9*, 413–429.

Herek, G. M. (2007). Confronting sexual stigma and prejudice: Theory and practice. *Journal of Social Issues, 63*, 905–925.

Lippa, R. A. (2007). The preferred traits of mates in a cross-national study of heterosexual and homosexual men and women: An examination of biological and cultural influences. *Archives of Sexual Behavior, 36*, 193–208.

Logan, J. R. (September 10, 2001). The new Latinos: Who they are, where they are. http://mumford.albany.edu/census/HispanicPop/HspReport/page5.html. Retrieved 8 Jan 2007.

Luiz, É. (May 29, 2005). Brazilians in America: 1.3 million and growing fast. http://www.brazzil.com/index2.php?option=com_content&do_pdf=1&id=9294. Retrieved 8 Jan 2007.

Martins, Y., Tiggemann, M., & Kirkbride, A. (2007). Those Speedos become them: The role of self-objectification in gay and heterosexual men's body image. *Personality and Social Psychology Bulletin, 33*, 634–647.

McKinley, N. M. (1999). Women and objectified body consciousness: Mothers' and daughters' body experience in cultural, developmental, and familial context. *Developmental Psychology, 35*, 760–769.

Mirandé, A. (1997). *Hombres y machos: Masculinity and Latino culture*. Boulder, CO: Westview Press.

Sandfort, T. G. M., Melendez, R. M., & Diaz, R. M. (2007). Gender nonconformity, homophobia, and mental distress in Latino gay and bisexual men. *Journal of Sex Research, 44*, 181–189.

Silverthorne, Z. A., & Quinsey, V. L. (2000). Sexual partner age preferences of homosexual and heterosexual men and women. *Archives of Sexual Behavior, 29*, 67–76.

Taywaditep, K. J. (2001). Marginalization among the marginalized: Gay men's anti-effeminacy attitudes. *Journal of Homosexuality, 42*, 1–28.

Teunis, N. (2007). Sexual objectification and the construction of whiteness in the gay male community. *Culture, Health & Sexuality, 9*, 263–275.

Torres, J. B., Solberg, V. S. H., & Carlstrom, A. H. (2002). The myth of sameness among Latino men and their machismo. *American Journal of Orthopsychiatry, 72*, 163–181.

Zea, M. C., Reisen, C. A., Poppen, P. J., & Bianchi, F. T. (in press). Unprotected anal intercourse among Latino MSM: The role of characteristics of the person and the sexual encounter. *AIDS and Behavior*.

Arch Sex Behav (2010) 39:1331–1341
DOI 10.1007/s10508-009-9539-x

ORIGINAL PAPER

Sexual Behaviors among Club Drug Users: Prevalence and Reliability

Enbal Shacham · Linda B. Cottler

Received: 16 October 2008 / Revised: 23 July 2009 / Accepted: 24 July 2009 / Published online: 9 September 2009

Abstract HIV prevention efforts require a focus on reducing high risk sexual behavior. Because these are self-reported, assessments that reduce memory bias and improve elicitation of data are needed. As part of a multi-site psychometric study of club drug use, abuse, and dependence, data were collected with a test-retest design that measured the reliability of the Washington University Risk Behavior Assessment for Club Drugs (WU-RBA-CD). Reliability was assessed separately by sex via kappa coefficients and intraclass correlation coefficients (ICC); z tests compared coefficients by sex. A total of 603 participants were interviewed by independent assessors with 5 days in between interviews. Reliability for all 51 items of the sexual activity section of the WU-RBA-CD ranged from .23 to 1.00; 71% ($n = 36$) of items resulted in moderate to high reliability (.55–1.00). Number of lifetime sex partners was consistently reported for same-sex partners for both men and women and opposite-sex partners. Items with high reliability included reporting ever being under the influence of ecstasy (.87) or GHB (.87) while having sex. Items with lower reliability included those that queried the determinants of condom use (.45–.82) and about behaviors and attitudes experienced while using drugs (.23–.87). Very few sex differences were revealed in the reliability of reported sexual activities. Overall, the WU-RBA-CD performed with fairly high reliability rates. Assessing situations of when, how, and why individuals use condoms may offer the clearest evaluation of determinants of sexual behaviors, yet those items are not as reliable.

Keywords Sexual behaviors · Measurement · Reliability · Club drug use

E. Shacham (✉)
George Warren Brown School of Social Work,
Washington University, 700 Rosedale Avenue,
Campus Box 1009, St. Louis, MO 63112, USA
e-mail: eshacham@wustl.edu

L. B. Cottler
Department of Psychiatry, Washington University School
of Medicine, St. Louis, MO, USA

Introduction

As part of the response to HIV prevention efforts, public health professionals have developed methods to assess high risk behaviors for HIV infection. To measure both the risk for and effectiveness of HIV prevention efforts, reliable and valid assessments are needed. The development of appropriate measures has evolved throughout the HIV epidemic as risk behaviors have changed and different populations have become infected. Primarily, measurement error can be attributed to memory error, poor item content, flawed administration of instruments, as well as many other factors (Fenton, Johnson, McManus, & Erens, 2001; Johnson et al., 2000; Schroder, Carey, & Vanable, 2003a, 2003b).

Drug-using populations are frequently thought to be unreliable concerning their self-reported measures of risk behaviors, yet data do not support this claim (Goldstein et al., 1995; Klinkenberg et al., 2002; McElrath, Chitwood, Griffin, & Comerford, 1994; Needle et al., 1995; Petry, 2001; Twitchell, Huber, Reback, & Shoptaw, 2002). Studies among drug using populations report moderate to high reliability in reports of sex and drug-related behaviors, which used test-retest and inter-rater reliability methods (Compton, Cottler, Dorsey, Spitznagel, & Mager, 1996; Cottler et al., 1997, 1998b; Goldstein et al., 1995; Horton, Compton, & Cottler, 2000).

Memory error, recall bias, and poor question construction are consistently documented as factors of measurement error

and, in turn, they contribute to the limitations of study findings. Specific to sexual behaviors, memory error is found to be dependent on length of reference intervals, behavior measures, and frequencies of the targeted behavior (Schroder et al., 2003a). Assessments must balance the length of time an individual is able to recall accurately and the length of time that is representative of the individuals' practices. Two to three months reference intervals are often justified as an intermediate length of time that assesses risk behaviors more appropriately (Kauth, Lawrence, & Kelly, 1991; Schroder et al., 2003a). Alternatively, longer reference intervals offer opportunities to assess behaviors that would influence overall levels of sexual risk, if reliably reported. Self-reported sexual behaviors have been found to be more reliable when response options are dichotomous. The limitation of dichotomous items is that they only capture whether or not the behavior occurred and not the trend of the behavior over time, which may limit a thorough assessment of current level of risk (Tourangeau & Smith, 1996). Research has suggested that indicators, such as age at first sex, are strong predictors of risk and are reliable (Catania, Gibson, Chitwood, & Coates, 1990).

Assessing risk by using frequencies are scrutinized due to research suggesting that as frequency of behaviors increase, the accuracy of reporting that target behavior decreases (Schroder et al., 2003b). With higher frequency behaviors, rough estimates tend to be provided. For example, number of sex partners is "rounded" to increments of five or ten with higher frequencies (Tourangeau & Smith, 1996). Although the actual count will offer more data for risk analysis, a valid count may be difficult to obtain both due to the reference interval and to challenges in recalling the actual frequency (Schroder et al., 2003b). High frequency behaviors may encourage adopting a shorter recall period. Low frequency behaviors seem to be more memorable and are more likely to be reported even if that behavior did not occur during the specified timeframe. Incidence data are reported to be more reliable, except when referring to behaviors that may not be engaged in regularly. Research suggests directly querying the frequency of these low frequency behaviors, thereby learning how often one is engaging in the activity. This method offers an opportunity for individuals to perceive this behavior as more common and increase comfort to respond honestly, unlike a dichotomous measure regarding their engagement in behaviors in which they may not typically engage (Durant & Carey, 2000; Schroder et al., 2003b; Tourangeau & Smith, 1996).

Social desirability has also been documented as a significant contributor to recall bias as individuals may respond in a manner that assumes a more socially conservative response when assessing sexual behaviors, which incorrectly minimizes the level of risk (Fenton et al., 2001; Schroder et al., 2003b). Research has found that reporting of sexual behaviors differs by sex and may influence biases in responses. This has most commonly been found in questions concerning number of opposite-sex partners with men reporting higher frequency than women. This has been documented in situations where communities would converge with number of sex partners due to a finite group of sex partner possibilities (Catania et al., 1990, 1995; Tourangeau & Smith, 1996). Other items which seem to reveal sex-based biases include frequency of self-stimulation, oral sex with an opposite sex partner, and frequency of unprotected sex (Tourangeau & Smith, 1996). It is unclear that sex differences exist consistently based on behavior or which sex gives more accurate reports (Catania et al., 1990, 1995; Durant & Carey, 2000).

With these challenges to reliability and validity, it is necessary that the field continues to develop assessments that are reliable for the most sensitive behaviors, so that risk can be accurately assessed. One of the most widely used assessments of sexual behaviors, the NIDA Risk Behavior Assessment (RBA), was validated and originally used in a multi-city study of a drug-using population (Needle et al., 1995). Versions of the RBA continue to be used to assess levels of risk behaviors associated with sex, drug use, attitudes, and beliefs about associated behaviors varying populations (Cottler et al., 1998a; Dowling-Guyer, 1994; Inciardi, Surratt, Kurtz, & Weaver, 2005; Needle et al., 1995; Twitchell et al., 2002). In the present study, revisions to the RBA were made to enrich diagnostic data collected for a club drug (ecstasy, rohypnol, ketamine, and/or GHB) use, abuse, and dependence study. As part of that study, reliability data were generated and the current analyses assessed reliability of reports of sexual behaviors from the revised version of the RBA by sex.

Method

Participants

The data come from a tri-city study of club drug use, abuse, and dependence conducted in St. Louis, Miami, and Sydney. The study was a nosological assessment of the constructs and relevance of DSM-IV criteria for ecstasy, ketamine, rohypnol, and GHB. In addition, behaviors and attitudes associated with club drug use were evaluated. The analyses presented test the reliability of the sexual behavior measures of the Washington University Risk Behavior Assessment (WU-RBA) between sexes.

Each study site utilized identical research protocols. Interested participants contacted researchers after they were recruited through flyer distributions in high volume areas such as community centers, dance clubs, and in school newspapers. The flyers were designed to recruit current club drug users to participate in a research study. Eligible individuals were 15 years of age and older who had used ecstasy more than five times lifetime and at least once in the past 12 months, who spoke English. A total of 927 club drug users were screened

and eligible at all three sites. Approximately 300 (35%) of these participants did not return for retest interviews. Two participants identified as transgendered and were not included for the sex-based comparisons. The study design was developed to make eligible participants who reported recent ecstasy use, at the time a commonly used club drug. This created an opportunity to examine the role of multi-drug use among an ecstasy-using sample.

A total of 603 individuals from St. Louis ($n = 278$), Miami ($n = 182$), and Sydney ($n = 143$) completed test-retest interviews. Over one-third of the sample was non-White (36%) and female (40%); the mean age was 23.3 years ($SD = 5.15$; range, 16-59). Specific demographic characteristics by study site are described in Table 1. The interviews and protocol procedures were approved by the Washington University Human Research Protection Office, University of Delaware (for the Miami site, where co-investigators were affiliated), and the University of New South Wales Human Research Ethics Committee.

Measures

Demographic characteristics were collected during the initial interview. These included age, sex, and race/ethnicity. The Washington University Risk Behavior Assessment (WU-RBA-CD) was modified from the NIDA's Risk Behavior Assessment, which was originally developed for use in cooperative agreement studies of HIV risk behaviors (Coyle, 1998; Needle et al., 1995). This assessment was originally designed to collect self-reported sexual behaviors, HIV risk perceptions, patterns of drug use, and influence of drug use on sexual behaviors among an adult population. The revised RBA included questions related to situations where club drugs are used and the impact they may have on sex partner selection, types of sexual acts, condom use while using club drugs, and sexual pleasure while using club drugs. In addition, attitudes regarding condom use, age of first sex, ever oral/vaginal/anal sex, forced sex, number of sex partners in the past 3 months, trading sex for club drugs/ housing/ food/money, and condom use consistency were measured. Specific items are included in Tables 3 and 4.

In addition, the Club Drug-Substance Abuse Module (CD-SAM) was modified from the CIDI-Substance Abuse Module (SAM) version developed by Washington University for the DSM-IV field trials (Cottler & Helzer, 1989; Cottler et al., 1995; Horton et al., 2000). Diagnostic items were added to the CD-SAM specific to club drugs that mimicked question patterns of other drug use, abuse, and dependence diagnoses. This enhanced the existing comprehensive assessment of use and physical, psychological, and withdrawal symptoms listed for illicit drugs covered in the SAM. These data were included to describe the lifetime drug-using patterns of the sample.

Procedure

Data were collected between 2002 and 2005. Intensive training and certification on administering the instruments and the overall interview protocol were required for all interviewers. Interviewer training typically consisted of 3–5 days of training on the instruments and testing with volunteer participants. Informed consent was given for participation in the study. For individuals under 18 years of age, parental informed consent was provided. Initial and retest interviews were conducted with independent raters and completed between 5 and 7 days after the initial interview. This protocol aimed to limit recall of their previous responses and not encounter many true score changes. Participants were remunerated for each of their interviews.

Data analyses were conducted with SAS (version 9.1) and STATA (version 8.0). Descriptive and bivariate analyses were conducted to provide the prevalence rates of lifetime substance use disorders, as assessed by the CD-SAM. Test-retest reliability analyses were conducted in two ways: kappa coefficients were conducted for the dichotomous variables and intraclass correlation coefficients (ICC) were conducted

Table 1 Demographic characteristics by study site ($N = 603$)

	St. Louis ($n = 278$)		Miami ($n = 182$)		Sydney ($n = 143$)		p
	M	*SD*	M	*SD*	M	*SD*	
Age (in years)	23.3	5.5	23.5	4.9	22.9	4.8	ns
Number of ecstasy pills (lifetime)	186.4	541.4	224.9	528.7	234	348.4	ns
Female	43.5%		39.0%		40.7%		ns
Race/ethnicity							ns
African American	15.1%		2.2%		0.0%		
Asian	1.8%		1.6%		12.0%		
Caucasian	74.9%		32.8%		82.4%		
Hispanic or Latino	2.5%		56.3%		0.0%		
Other	5.7%		7.1%		5.6%		

for the continuous variables. The kappa statistic ranges from −1.00, which signifies total disagreement, to 1.00, signifying total agreement. Typically, kappa values less than 0.40 indicate marginal agreement, 0.40–0.60 indicate moderate agreement, 0.60–0.75 indicate good agreement, and values greater than 0.75 indicate excellent agreement (Bishop, Fienberg, & Holland, 1975). The ICC was calculated by obtaining the difference in the error for each of the mean squares and divided by the sum of those errors. Similarly, ICC range from −1.00 to 1.00 with values less than 0.40 indicating poor agreement, 0.40–0.75 indicating moderate to good agreement, and values greater than 0.75 indicating excellent agreement (Fleiss, 1986). *Z* statistical tests were conducted to compare the reliability coefficients by sex.

Results

Drug Use Patterns

High proportions of the sample reported lifetime use of opiates (30%), hallucinogens (26%), ecstasy (25%), ketamine (22%), cannabis (23%), cocaine (21%), alcohol (17%), GHB (10%), rohypnol (6%), and amphetamines (5%). Although not shown, sex comparisons by drug use were conducted by χ^2 tests revealed that more women reported using cannabis than men (33.1% vs. 17.6%; $p < .001$). Other drug use was not significantly different by sex.

Sexual Behavior Patterns by Sex

Sex differences in reported sexual behaviors were measured by the WU-RBA-CD. ANOVAs were conducted for categorical and continuous variables; χ^2 analyses were conducted for dichotomous variables. As shown in Table 2, women more often than men reported ever having vaginal sex ($p = .02$), although 94.9% of men in sample also endorsed having vaginal sex. Most of the sample reported having had oral sex (98.2%) and significantly more men reported having anal sex than women (49.0% vs. 40.8%; $p = .05$). Men reported their first sexual encounter at a younger age than women (15.0 vs. 15.5 years; $p = .01$). The mean number of lifetime male sex partners for men was 7.0 ($SD = 57.1$) and for women 12.7 ($SD = 15.7$). The number of lifetime female sex partners differed significantly ($p < .001$) between sex, with men reporting a mean of 24.8 ($SD = 76.4$) lifetime female sex partners and women reporting a mean of 1.7 ($SD = 3.9$) female sex partners. Men reported having three or more sex partners within the past three months more often than women ($p < .001$).

There were no significant sex differences in reports of condom use for oral sex, 3.8% reporting almost always or always using condoms during oral sex. Men reported more consistent condom use during anal sex ($p = .02$) and vaginal sex ($p < .03$).

Most of the sample (92.5%) reported consenting to sex the first time, of those who had sex; yet, 9.7% of women reported more often than men that they had not consented to sex the first time ($p = 0.01$). Further, 16.4% of the sample reported having experienced forced sexual contact with a dating partner; women more often than men reported this behavior (23.0% vs. 8.6%; $p < .001$).

There were no significant differences between sex for those who had traded sex for club drugs (4.8%) or club drugs traded for sex (3.8%). Men reported more often than women having exchanged club drugs for money, food, clothing or housing (32.1% vs. 13.3%; $p < .001$). In response to questions about sexual repertoire and perceptions of pleasure from drugs, the majority of the sample that had reported using ecstasy (76.0%), ketamine (41.5%), GHB (61.5%), and/or rohypnol (36.1%) reported that it made sex more pleasant. Most participants reported that the club drug they had used during sex had no effect on their partner selection (ecstasy [67.5%], ketamine [63.1%], GHB [57.7%] or rohypnol [55.6%]). Participants reported similar sexual experiences across sites and by sex when using ecstasy (62.8%), ketamine (60.0%), GHB (56.4%) or rohypnol (63.9%). Participants responded similarly when asked about the likelihood of using condoms when using the specific drugs ecstasy (69.2%), ketamine (80.0%), GHB (64.1%) or rohypnol (72.2%).

Reliability of Self-Reported Sexual Activities

We found that the standard kappa and ICC reliability categories were less relevant in this study as many of our items were found to be very reliable; therefore, we utilized more stringent reliability categories where kappas and ICCs between .80 and 1.00 were identified as highly reliable, scores of 0.55–0.79 were moderately reliable, and items <0.55 had low reliability. Shown in Tables 3 and 4 are the 51 items and the reliability coefficients assessed. The test-retest reliability analyses ranged from a low of 0.36 to a high of 1.00. Overall, 16 (31.4%) items resulted in high reliability, 22 (43.1%) with moderate reliability, and 12 (23.5%) items resulted in low reliability. Rates of reliability were compared between sites with no significant differences.

Reliability of Items Related to Sexual Initiation

Rates of reliability for most items related to the initiation of sexual activities ranged from 0.80 to 1.00. Of the 11 items related to sexual initiation, there was one item with significant sex differences. These items included: ever having vaginal sex (0.91), ever having anal sex (0.92), having been pregnant (0.96), consenting the first time to have sex (0.80),

Table 2 Differences in sexual behaviors among club drug users by sex as conducted by ANOVAS

	Male ($N = 353$)		Female ($N = 248$)		p
	M	*SD*	M	*SD*	
Lifetime number of sex partners					
Male partners	7.0	57.1	12.7	15.7	.001
Female partners	24.8	76.4	1.7	3.9	.001
Age at first sex (in years)	15.0	2.8	15.5	2.1	.010
	n	%	*n*	%	*p*
Ever had vaginal sex	335	94.9	246	98.4	.020
Ever had oral sex	348	98.6	245	98.0	
Ever had anal sex	173	49.0	102	40.8	.050
Ever have been pregnant	–	–	69	27.8	
Did not consent to sex first time	16	4.6	24	9.7	.010
Ever experienced forced sexual contact	30	8.6	57	23.0	.001
Gave club drugs for sex	9	2.5	4	1.6	
Been given club drugs for sex	13	3.7	8	3.2	
Traded club drugs for money, food, clothing or housing	113	32.1	33	13.3	.001
Never use condom for vaginal sex	51	15.2	56	22.8	.030
Never use condom for oral sex	282	81.0	214	87.4	
Never use condom for anal sex	72	40.9	50	48.1	.020
Sexual behaviors while using club drugs					
Had sex on ecstasy ($n = 450$)	266	75.6	196	79.0	
On ketamine ($n = 47$)	39	11.1	26	10.5	
On GHB ($n = 68$)	52	14.8	24	9.7	
On rohypnol ($n = 24$)	19	5.4	14	5.7	
Sex is more pleasant when using ecstasy ($n = 450$)	200	75.2	131	66.8	
Using ketamine ($n = 47$)	15	38.5	8	30.8	
Using GHB ($n = 68$)	33	66.0	15	62.5	
Using rohypnol ($n = 24$)	8	50.0	4	28.6	
More selective about sex partners when using ecstasy ($n = 450$)	25	9.4	17	8.7	
Using ketamine ($n = 47$)	2	5.1	4	15.4	
Using GHB ($n = 68$)	3	5.8	1	4.2	
Using rohypnol ($n = 24$)	1	5.3	0	0	
More varied sex acts when using ecstasy ($n = 450$)	81	30.5	55	28.1	
Using ketamine ($n = 47$)	8	20.5	5	19.2	
Using GHB ($n = 68$)	13	26.5	8	33.3	
Using rohypnol ($n = 24$)	4	25.0	2	14.3	
More likely to use condoms when using ecstasy ($n = 450$)	17	6.4	12	6.1	
Using ketamine ($n = 47$)	1	2.6	3	11.5	
Using GHB ($n = 68$)	0	0	4	16.7	.020
Using rohypnol ($n = 24$)	1	5.3	2	14.3	

and attraction to men (0.99) and women (0.97). There were no significant sex differences among these items. However, men, compared to women, were less reliable in their reports of their lifetime female sex partners (0.79 vs. 0.99) and women were less reliable than men in their reports of number of lifetime male sex partners (0.90 vs. 1.00). Both men (1.00)

Table 3 Kappa test-retest reliability results of sexual activity items by sex

Items	Male ($N = 353$)					Female ($N = 248$)					
	Responses			Reliability coefficient	95% confidence Interval	Responses			Reliability coefficient	95% confidence Interval	M–F Z Test
	Time 1	Time 2				Time 1	Time 2				
		No	Yes				No	Yes			
Have you ever had vaginal sex?	No	18	2	0.94	(0.87–1.00)	No	2	0	0.66	(0.23–1.00)	1.24
	Yes	0	333			Yes	2	246			
Have you ever had anal sex?	No	172	9	0.90	(0.86–0.95)	No	143	2	0.94	(0.90–0.98)	1.23
	Yes	8	164			Yes	5	100			
Have you ever had oral sex?	No	4	4	0.61	(0.29–0.92)	No	2	1	0.49	(0.06–0.92)	0.43
	Yes	1	344			Yes	3	244			
Have you ever been sexually attracted to men?	No	283	0	0.97	(0.94–1.00)	No	3	0	0.86	(0.57–1.0)	0.82
	Yes	3	66			Yes	1	246			
Have you ever been sexually attracted to women?	No	19	1	0.97	(0.92–1.00)	No	118	4	0.96	(0.93–0.99)	0.40
	Yes	0	333			Yes	1	127			
Have you ever been pregnant?	No	n/a	n/a	n/a	n/a	No	176	2	0.96	(0.92–1.00)	
	Yes	n/a	n/a	n/a	n/a	Yes	2	67			
In the past 3 months have you had 3 or more sex partners?	No	286	10	0.78	(0.69–0.88)	No	226	7	0.71	(0.53–0.89)	0.71
	Yes	8	39			Yes	2	12			
The first time you had sex, did you consent to it?	No	10	3	0.68	(0.48–0.88)	No	23	5	0.87	(0.77–0.97)	1.71
	Yes	6	331			Yes	1	219			
Have you ever experienced forced sexual content with a dating partner?	No	307	8	0.64	(0.50–0.78)	No	175	10	0.71	(0.61–0.82)	0.80
	Yes	13	22			Yes	16	47			
Have you ever traded club drugs to get money, food clothes, or a place to stay?	No	203	20	0.64	(0.56–0.73)	No	194	7	0.59	(0.45–.072)	0.72
	Yes	36	91			Yes	21	26			
Have you ever been under the influence of Ecstasy while you were having any kind of sex?	No	75	8	0.85	(0.79–0.92)	No	46	2	0.90	(0.83–0.97)	1.00
	Yes	11	256			Yes	6	194			
Have you ever been under the influence of GHB while you were having any kind of sex?	No	292	7	0.84	(0.76–0.92)	No	221	0	0.93	(0.86–1.00)	1.70
	Yes	7	44			Yes	3	24			
Have you ever been under the influence of ketamine while you were having any kind of sex?	No	300	11	0.66	(0.54–0.79)	No	216	6	0.74	(0.60–0.88)	0.76
	Yes	12	27			Yes	6	20			
Have you ever been under the influence of rohypnol while you were having any kind of sex?	No	323	6	0.63	(0.45–0.81)	No	230	3	0.74	(0.56–0.93)	0.84
	Yes	8	13			Yes	4	11			
You didn't use a condom because condoms or other barrier protection caused less sensation	No	141	27	0.62	(0.53–0.70)	No	125	16	0.61	(0.50–0.71)	0.14
	Yes	36	126			Yes	28	68			

Table 3 continued

Items	Male ($N = 353$)					Female ($N = 248$)					
	Responses			Reliability coefficient	95% confidence Interval	Responses			Reliability coefficient	95% confidence Interval	M–F Z Test
	Time 1	Time 2				Time 1	Time 2				
		No	Yes				No	Yes			
You didn't use a condom because your religion did not allow you to	No	328	2	1.00		No	0	238	1.00		
	Yes	0	0			Yes	0	0			
You didn't use a condom because condoms or other barrier protection cost too much	No	314	2	0.65	(0.43–0.88)	No	227	2	0.70	(0.44–0.95)	0.24
	Yes	6	8			Yes	3	6			
You didn't use a condom because you were too drunk or high	No	221	23	0.67	(0.58–0.76)	No	158	11	0.72	(0.62–0.82)	0.64
	Yes	19	67			Yes	16	53			
You didn't use a condom because you didn't have one available	No	168	28	0.61	(0.53–0.70)	No	136	16	0.72	(0.62–0.81)	1.61
	Yes	33	101			Yes	15	71			
You didn't use a condom because a partner objected	No	277	17	0.57	(0.43–0.71)	No	206	6	0.57	(0.39–0.74)	0.05
	Yes	12	24			Yes	12	14			
You didn't use a condom because you didn't think about it	No	175	44	0.51	(0.42–0.61)	No	136	27	0.60	(0.49–0.70)	1.15
	Yes	30	81			Yes	16	59			
You didn't use a condom because you think you or your partner could be infected with HIV or an STD	No	133	37	0.53	(0.43–0.62)	No	71	26	0.42	(0.30–0.53)	1.48
	Yes	41	119			Yes	43	98			
You didn't use a condom because you were afraid to ask your partner	No	322	4	0.39	(0.00–0.78)	No	225	3	0.61	(0.35–0.88)	0.94
	Yes	2	2			Yes	4	6			
You didn't use a condom because condoms or other barrier protection didn't taste good during oral sex	No	236	35	0.50	(0.38–0.61)	No	165	17	0.55	(0.42–0.67)	0.56
	Yes	19	40			Yes	21	35			
You didn't use a condom because you or your partners didn't want to take the time to put one on	No	258	18	0.47	(0.33–0.60)	No	188	14	0.43	(0.27–0.59)	0.07
	Yes	27	27			Yes	19	17			
You didn't use a condom because of another reason	No	224	16	0.52	(0.42–0.63)	No	140	20	0.59	(0.48–0.69)	0.46
	Yes	41	49			Yes	23	55			

Table 4 Intraclass correlation coefficient test-retest reliability results of sexual activity items by sex

Items	Male ($N = 353$)			Female ($N = 248$)			
	n	Reliability coefficient	95% Confidence interval	*n*	Reliability coefficient	95% confidence interval	M–F *Z* Test
How old were you the first time you had sex of any kind?	350	0.80	(0.63–0.97)	248	0.83	(0.67–0.99)	0.25
In your lifetime, how many of your sex partners were male?	349	1.00		248	0.90	(0.83–0.96)	3.07*
In your lifetime, how many of your sex partners were female?	341	0.79	(0.69–0.89)	248	0.99	(0.99–1.00)	3.93*
In the last 12 months, how often have you used a condom or other barrier protection when having anal sex?	164	0.62	(0.24–0.99)	100	0.45	(0.02–0.88)	0.57
In the last 12 months, how often have you used a condom or other barrier protection when having oral sex?	344	0.85	(0.64–1.07)	244	0.77	(0.45–1.09)	0.43
In the last 12 months, how often have you used a condom or other barrier protection when having vaginal sex?	333	0.73	(0.47–0.99)	246	0.69	(0.40–0.99)	0.18
How many times have you been given a club drug like ecstasy, ketamine, GHB, or rohypnol in exchange for having sex?	350	0.99	(0.96–1.01)	248	0.94	(0.82–1.06)	0.75
How many times have you given a club drug like ecstasy, ketamine, GHB, or rohypnol in exchange for sex?	350	0.97	(0.91–1.02)	248	0.95	(0.83–1.06)	0.28
How many times have you traded club drugs in exchange for money, food, a place to stay, or clothes?	88	0.83	(0.72–0.95)	26	0.96	(0.92–1.01)	2.15*
When using ecstasy was sex more pleasant?	256	0.68	(0.20–1.15)	194	0.66	(0.17–1.15)	0.05
When using ecstasy have you been more likely to use a condom?	256	0.40	(0.00–0.95)	194	0.70	(0.22–1.18)	0.81
When using ecstasy were you more selective about whom you had sex with?	256	0.49	(0.00–1.03)	194	0.62	(0.11–1.13)	0.34
When using ecstasy have you engaged in more varied sex acts?	256	0.58	(0.00–1.20)	194	0.57	(0.00–1.17)	0.03
When using ketamine was sex more pleasant?	27	0.50	(0.00–1.08)	20	0.45	(0.00–1.17)	0.10
When using ketamine have you been more likely to use a condom?	27	0.00	(0.00–0.35)	20	0.61	(0.01–1.19)	1.73
When using ketamine were you more selective about who you had sex with?	27	0.26	(0.00–0.88)	20	0.94	(0.77–1.11)	2.07*
When using ketamine have you engaged in more varied sex acts?	27	0.33	(0.00–0.95)	20	0.84	(0.45–1.23)	1.37
When using GHB was sex more pleasant?	44	0.67	(0.22–1.13)	24	0.04	(0.00–0.42)	2.07*
When using GHB were you more selective about who you had sex with?	44	0.63	(0.05–1.20)	24	0.23	(0.00–0.90)	0.88
When using GHB have you been more likely to use a condom?	44	0.26	(0.00–0.86)	24	0.44	(0.00–1.08)	0.08
When using GHB have you engaged in more varied sex acts?	44	0.59	(0.02–1.15)	24	0.87	(0.57–1.16)	0.86
When using rohypnol was sex more pleasant?	13	0.14	(0.00–0.85)	11	1.00		
When using rohypnol have you been more likely to use a condom?	13	0.45	(0.00–1.28)	11	0.78	(0.32–1.23)	0.04
When using rohypnol were you more selective about who you had sex with?	13	0.20	(0.00–0.94)	11	0.73	(0.11–1.35)	1.09
When using rohypnol have you engaged in more varied sex acts?	13	0.61	(0.00–1.23)	11	0.79	(0.26–1.32)	0.43

* *Z* test statistically significant

and women (0.99) were highly reliable in their reports of lifetime number of same-sex partners.

Reliability of Condom-Related Items

Many of the items associated with the determinants of condom use resulted in reliability coefficients ranging between 0.55 and 0.79. There were no sex differences in reliability among of the 16 items measuring condom-related behaviors. These determinants included items about condoms causing less sensation (0.62), costing too much (0.67), being too drunk or high to use condoms (0.69), not having a condom available (0.66), having a partner object (0.57), and not thinking about it (0.55). Not using condoms for religious reasons was never misreported by any participant, (no participant endorsed this item). Participants were unreliable regarding other reasons for not using a condom such as: being afraid to ask their partner (0.54) to use a condom, thinking they or their partners could be infected with an STI or HIV (0.48), believing that barrier methods did not taste good during oral sex (0.52), or not wanting to take the time to put a condom on (0.45). Consistent reports, regardless of sex, in the frequency of use of condoms during sex varied, but all were reliable: vaginal sex had an ICC of 0.71, anal sex had an ICC of 0.55, and oral sex had an ICC of 0.82.

Reliability of Sex-Related Substance Using Behaviors

Items related to sexual behaviors and attitudes during concomitant club drug use resulted in a range of reliability. Of the 24 items pertaining to sex-related substance-using behaviors, there were three items that were significantly different in reliability by sex. Numbers of times someone had either been given club drugs (0.99) or had given club drugs (0.97) in exchange for sex were highly consistent and reliable. The number of times someone had traded club drugs for exchange of food, money, place to stay, or clothes resulted in an ICC of 0.83. Rates of reliability when queried about having sex while using ecstasy or GHB were also high (kappa = 0.87), and for rohypnol and ketamine, moderately reliable (0.68 and 0.70). Men were less reliable than women (0.83 vs. 0.96) when reporting the number of times they exchanged club drugs for food, housing or clothing.

Participants were asked if their partner selectivity changed with each drug use, which tended to have low reliability: being more selective with a sex partner under the influence of ecstasy had a reliability coefficient of 0.56, rohypnol 0.51, ketamine 0.60, and GHB 0.42. Women were more reliable in their responses to whether ketamine made them more selective in their sex partners (0.94). The level of pleasure related to sex when using club drugs ranged from 0.39–0.67. Specifically, women were less reliable than men regarding GHB's effect on sexual pleasure (0.04 vs. 0.67). Participants were moderately reliable when they were queried if they had more varied sexual acts when on ketamine (0.57), ecstasy (0.62), GHB (0.70), and rohypnol (0.72). The reliability for whether participants were more likely to use condoms while using each club drug resulted with lower reliability with kappa coefficients for ketamine (0.23), ecstasy (0.52), GHB (0.41), and rohypnol (0.63). Other significant differences did not occur between men and women for items related to sex- and drug- using patterns.

Discussion

This study aimed to assess the reliability of a self-reported HIV risk behavior instrument that was modeled after one used in numerous studies with drug abusers (RBA). However, for this study, the RBA was revised to assess contextual factors related to club drug use, abuse, and dependence in three different communities: the Miami club scene, the St. Louis rave scene, and the Sydney rave and club scene. For the first time, reliability of club drug use patterns and sexual behaviors reported by club drug users have been collected using identical methods across multiple sites.

The findings of this geographically diverse study population revealed that these club drug users, regardless of location, were highly reliable reporters of their behaviors. There were also very few differences in reliability of responses by sex. Similar to previous research (Catania et al., 1990; Fenton et al., 2001; Schroder et al., 2003a; Tourangeau & Smith, 1996), the rates of reliability regarding sexual debut (ever having oral, vaginal or anal sex and age of sexual initiation) were highly reliable. Items that queried participants on their condom use attitudes and beliefs, such as whether they would take the time to put one on, were high or drunk, or had a condom to use had moderate levels of reliability and no differences by sex in reliability. Items where further development and research needs to be conducted include challenges with measuring condom use negotiation and perceived susceptibility of STD/HIV. There was perfect agreement for both men and women regarding condom use where attitudes were well-established, such as not using a condom because one's religion doesn't allow it.

Sex differences in reliability did occur when reporting number of lifetime sex partners. There was near perfect agreement for number of lifetime same-sex partners, yet significant differences in reliability between sexes in number of lifetime opposite-sex partners. The study did not assess sexual orientation, just number of male and female sex partners for all participants. This is a limitation of the study findings, as reported behaviors differ from self-identified sexual orientation. Other studies have found that same-sex partners tend to be a low frequency behavior, and therefore individuals would know exactly how many same-sex partners they have had (Tourangeau & Smith, 1996). This may also be an effect of

"rounding," memory error, or being cognitively unable to accurately report for those with a higher number of sex partners (Bogart et al., 2007; Schroder et al., 2003a; Tourangeau & Smith, 1996).

The assessments of specific club drugs and their effect on sexual behavior during use were moderately reliable and require further examination. Social desirability and its role on low frequency behaviors may have affected the reliability of responses to these items (Durant & Carey, 2000; Fenton et al., 2001; Schroder et al., 2003a; Tourangeau & Smith, 1996). Additionally, these items had lower prevalence rates, as expected with lower frequency behaviors. Yet, it is a well-documented limitation of the kappa statistic that when there are low prevalence rates, rates of reliability are lower as well (Byrt, Bishop, & Carlin, 1993; Spitznagel & Helzer, 1985; Viera & Garrett, 2005). Specifically, women had markedly higher rates of reliability on items regarding selectivity of sex partners when using ketamine and number of times they reported exchanging club drugs, while men had significantly higher rates of reliability in responding to whether GHB enhanced sexual pleasure. All of these items had less than 4% of the overall sample endorsing using GHB or ketamine.

While these findings revealed high reliability among most of the items by sex and across study sites, some of the confidence intervals associated with the ICC or kappa statistics of each item were very wide. These wide confidence intervals more often occurred with the items that had smaller sample sizes, an expected effect similar to that discussed previously. Some items had unexpectedly low kappa scores, which tends to occur when there is low variability in responses. When near perfect agreement with little or no variance occurs, kappa scores are lower than anticipated. Therefore, the lower kappa scores were superficially low, such as ever having oral sex (0.61), consented to sex (0.68), and experienced forced sex (0.64).

Future areas of methodological research should focus on improving questions with low reliability and developing additional risk assessments. Examining beliefs that influence condom use offers insight to determining level of risk, yet those beliefs are assumed to be static. These findings suggest that beliefs that influence condom use are not static, are to be determined at the time of sexual activity, and often, dependent on partners. Items querying type of sex partner, specific sex partner, and shorter recall time frames may yield more insight than broader assessments of risk (Kauth et al., 1991; Schroder et al., 2003a, 2003b; Tourangeau & Smith, 1996).

This study was conducted to assess the reliability of club drug abuse and dependence criteria; the investigators also developed an opportunity to assess sex- and drug-using behaviors in these populations. Studies such as these are important as they serve as opportunities to promote the quality of assessments in the field that measure risk behaviors and evaluate change in behaviors. Based on these results, it is clear that few sex differences in reliability exist, and that self-reported measures are reliable measures of risk with carefully implemented assessments and protocols. Through these test-retest reliability analyses, the WU-RBA-CD was found to be a highly reliable measure of sexual behaviors that may be used by others.

Acknowledgements This investigation was supported by the National Institutes of Health from the National Institute on Drug Abuse (DA 14854-02S1) and partially supported by National Research Service Award (T32 DA07313) from the National Institute of Drug Abuse. Additionally, this publication was made possible from the National Center for Research Resources (UL1 RR024992), specifically, the KL2RR024994 component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research. Its contents are solely the responsibility of the authors and do not necessarily represent the official view of NCRR or NIH.

References

Bishop, Y. M., Fienberg, S., & Holland, P. (1975). *Discrete multivariate analyses*. Cambridge: MIT Press.

Bogart, L. M., Walt, L. C., Pavlovic, J. D., Ober, A. J., Brown, N., & Kalichman, S. C. (2007). Cognitive strategies affecting recall of sexual behavior among high-risk men and women. *Health Psychology, 26*, 787–793.

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology, 46*, 423–429.

Catania, J. A., Binson, D., Dolcini, M. M., Stall, R., Choi, K. H., Pollack, L. M., et al. (1995). Risk factors for HIV and other sexually transmitted diseases and prevention practices among US heterosexual adults: changes from 1990 to 1992. *American Journal of Public Health, 85*, 1492–1499.

Catania, J. A., Gibson, D. R., Chitwood, D. D., & Coates, T. J. (1990). Methodological problems in AIDS behavioral research: Influences on measurement error and participation biases in studies of sexual behavior. *Psychological Bulletin, 108*, 339–362.

Compton, W. M., Cottler, L. B., Dorsey, K. B., Spitznagel, E. L., & Mager, D. E. (1996). Comparing assessments of DSM-IV substance dependence disorders using CIDI-SAM and SCAN. *Drug & Alcohol Dependence, 41*, 179–187.

Cottler, L. B., Compton, W. M., Ben Abdallah, A., Cunningham-Williams, R., Abram, F., Fichtenbaum, C., et al. (1998a). Peer-delivered interventions reduce HIV risk behaviors among out-of-treatment drug abusers. *Public Health Reports, 113*(Suppl. 1), 31–41.

Cottler, L. B., Grant, B. F., Blaine, J., Mavreas, V., Pull, C., Hasin, D., et al. (1997). Concordance of DSM-IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDI and SCAN. *Drug & Alcohol Dependence, 47*, 195–205.

Cottler, L. B., & Helzer, J. E. (1989). Diagnostic agreement between DSM-III and DSM-III-R dependence disorders. *NIDA Research Monograph, 95*, 380–381.

Cottler, L. B., Leukefeld, C., Hoffman, J., Desmond, D., Wechsberg, W., Inciardi, J. A., et al. (1998b). Effectiveness of HIV risk reduction initiatives among out-of-treatment non-injection drug users. *Journal of Psychoactive Drugs, 30*, 279–290.

Cottler, L. B., Schuckit, M. A., Helzer, J. E., Crowley, T., Woody, G., Nathan, P., et al. (1995). The DSM-IV field trial for substance use disorders: Major results. *Drug & Alcohol Dependence, 38*, 59–69.

Coyle, S. L. (1998). Women's drug use and HIV risk: Findings from NIDA's cooperative agreement for community-based outreach/intervention research program. *Women & Health, 27*, 1–18.

Dowling-Guyer, S. (1994). Reliability of drug users' self-reported HIV risk behaviors and validity of self-reported recent drug use. *Assessment, 1*, 383–392.

Durant, L. E., & Carey, M. P. (2000). Self-administered questionnaires versus face-to-face interviews in assessing sexual behavior in young women. *Archives of Sexual Behavior, 29*, 309–322.

Fenton, K. A., Johnson, A. M., McManus, S., & Erens, B. (2001). Measuring sexual behavior: Methodological challenges in survey research. *Sexually Transmitted Infections, 77*, 84–92.

Fleiss, J. (1986). *The design and analysis of clinical experiments*. New York: Wiley.

Goldstein, M., Friedman, S., Neaigus, A., Jose, B., Ildefonso, G., & Curtis, R. (1995). Self-reports of HIV risk behavior by injecting drug users: Are they reliable? *Addiction, 90*, 1097–1104.

Horton, J., Compton, W., & Cottler, L. B. (2000). Reliability of substance use disorder diagnoses among African-Americans and Caucasians. *Drug & Alcohol Dependence, 57*, 203–209.

Inciardi, J. A., Surratt, H. L., Kurtz, S. P., & Weaver, J. C. (2005). The effect of serostatus on HIV risk behaviour change among women sex workers in Miami, Florida. *AIDS Care, 17*, 88–104.

Johnson, M., Fisher, D. G., Montoya, I., Booth, R., Rhoades, F., Andersen, M., et al. (2000). Reliability and validity of not-in-treatment drug users' follow-up self-reports. *AIDS and Behavior, 4*, 373–380.

Kauth, M. R., Lawrence, J. S., & Kelly, J. A. (1991). Reliability of retrospective assessments of sexual HIV risk behavior: A comparison of biweekly, three-month, and twelve-month self-reports. *AIDS Education & Prevention, 3*, 207–214.

Klinkenberg, D., Calsyn, R., Morse, G., McCudden, S., Richmond, T., & Burger, G. (2002). Consistency of recall of sexual and drugs-using behaviors for homeless persons with dual diagnosis. *AIDS and Behavior, 6*, 295–302.

McElrath, K., Chitwood, D., Griffin, D., & Comerford, M. (1994). The consistency of self-reported HIV risk behavior among injection drug users. *American Journal of Public Health, 84*, 1965–1970.

Needle, R., Weatherby, N., Chitwood, D., Brown, B., Cesari, H., Booth, R., et al. (1995). Reliability of self-reported HIV risk behaviors of drug users. *Psychology of Addictive Behaviors, 9*, 242–250.

Petry, N. (2001). Reliability of drug users' self-reported HIV risk behaviors using a brief, 11-item scale. *Substance Use & Misuse, 36*, 1731–1747.

Schroder, K. E., Carey, M. P., & Vanable, P. A. (2003a). Methodological challenges in research on sexual risk behavior: II. Accuracy of self-reports. *Annals of Behavioral Medicine, 26*, 104–133.

Schroder, K. E., Carey, M. P., & Vanable, P. A. (2003b). Methodological challenges in research on sexual risk behavior: I. Item content, scaling, and data analytical options. *Annals of Behavioral Medicine, 26*, 76–103.

Spitznagel, E. L., & Helzer, J. (1985). A proposed solution to the base rate problem in the kappa statistic. *Archives of General Psychiatry, 42*, 725–728.

Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly, 60*, 275–304.

Twitchell, G., Huber, A., Reback, C., & Shoptaw, S. (2002). Comparison of general and detailed HIV risk assessments among methamphetamine abusers. *AIDS and Behavior, 6*, 153–162.

Viera, A., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine, 37*, 360–363.

Arch Sex Behav (2010) 39:1343–1352
DOI 10.1007/s10508-009-9522-6

ORIGINAL PAPER

Test–Retest Reliability and Predictors of Unreliable Reporting for a Sexual Behavior Questionnaire for U.S. Men

Alan G. Nyitray · Robin B. Harris · Andrew T. Abalos · Carrie M. Nielson · Mary Papenfuss · Anna R. Giuliano

Received: 25 November 2008 / Revised: 29 May 2009 / Accepted: 29 May 2009 / Published online: 25 August 2009

Abstract Accurate knowledge about human sexual behaviors is important for increasing our understanding of human sexuality; however, there have been few studies assessing the reliability of sexual behavior questionnaires designed for community samples of adult men. A test–retest reliability study was conducted on a questionnaire completed by 334 men who had been recruited in Tucson, Arizona. Reliability coefficients and refusal rates were calculated for 39 non-sexual and sexual behavior questionnaire items. Predictors of unreliable reporting for lifetime number of female sexual partners were also assessed. Refusal rates were generally low, with slightly higher refusal rates for questions related to immigration, income, the frequency of sexual intercourse with women, lifetime number of female sexual partners, and the lifetime number of male anal sex partners. Kappa and intraclass correlation coefficients were substantial or almost perfect for all non-sexual and sexual behavior items. Reliability dropped somewhat, but was still substantial, for items that asked about household income and the men's knowledge of their sexual partners' health, including abnormal Pap tests and prior sexually transmitted diseases (STD). Age and lifetime number of female sexual partners were independent predictors of unreliable reporting while years of education was inversely associated with unreliable reporting. These findings among a community sample of adult men are consistent with other test–retest reliability studies with populations of women and adolescents.

Keywords Reliability · Test–retest · Sexual behavior · Men

A. G. Nyitray (✉) · M. Papenfuss · A. R. Giuliano
H. Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Dr., MRC-CANCONT, Tampa, FL 33612, USA
e-mail: alan.nyitray@moffitt.org

A. G. Nyitray · R. B. Harris · A. T. Abalos
Mel and Enid Zuckerman College of Public Health, Tucson, AZ, USA

C. M. Nielson
Public Health and Preventive Medicine Department, Oregon Health and Science University, Portland, OR, USA

Introduction

Our knowledge of human sexual behavior informs our understanding of human sexuality. Data describing sexual behaviors are usually collected through self-report; however, there is concern that study participants' self-reported sexual behaviors may not be valid (Catania, Gibson, Chitwood, & Coates, 1990; Franco, 1997; Potterat, Phillips, & Muth, 1987; Stoneburner, Chiasson, Solomon, & Rosenthal, 1986; Weinhardt, Forsyth, Carey, Jaworski, & Durant, 1998) due to several sources of measurement error, including the demands of the recall task and the sociological context of the survey (Schroder, Carey, & Vanable, 2003). For example, self-presentation bias—the result of a participant's effort to manage information in order to shape his or her image (Catania, 1999)—is a threat to validity if the participant's presentation diverges from reality.

Unfortunately, the validity of sexual behavior measurements is difficult or impossible to assess due to cultural taboos that disapprove of the direct observation of human sexual behavior. Lacking a direct assessment of validity, studies often collect data through repeated self-reported measures of a sexual behavior to identify unreliable or inconsistent responses that may point to measurement error and, therefore, a lack of validity (Carballo-Dieguez, Remien, Dolezal, & Wagner, 1999; Müller, 2007; Rohan, McLaughlin, &

Harnish, 1994; Schlecht et al., 2001; Schrimshaw, Rosario, Meyer-Bahlburg, & Scharf-Matlick, 2006; Van Duynhoven, Nagelkerke, & Van de Laar, 1999).

Test–retest studies have measured the reliability of sexual behavior items in a variety of populations (Brener et al., 2002; Kalichman, Kelly, & Stevenson, 1997; Schrimshaw et al., 2006; Sohler, Colson, Meyer-Bahlburg, & Susser, 2000; Taylor, Rosen, & Leiblum, 1994). Some studies find that self-administered instruments, either of the paper and pencil variety or computer assisted, reduce measurement error and tend to produce reports of higher frequencies of sensitive sexual and drug using behaviors than do face-to-face interview methods (Boekeloo et al., 1994; Des Jarlais et al., 1999; Le, Blum, Magnani, Hewett, & Do, 2006; Romer et al., 1997; Simoes, Bastos, Moreira, Lynch, & Metzger, 2006a; Turner et al., 1998). It is thought that higher frequencies of sexual behaviors, particularly stigmatized behaviors, reflect greater validity since adults may seek to manage their public image by underreporting such behaviors (Catania et al., 1990).

Eight test–retest studies of self-administered sexual behavior questionnaires (SAQ) have been published (Brener et al., 2002; Durant & Carey, 2002; Kalichman et al., 1997; Müller, 2007; Schlecht et al., 2001; Sieving et al., 2005; St. Lawrence et al., 1999; Taylor et al., 1994). With the exception of one study from Germany that included only 12 men (Müller, 2007), these studies have assessed reliability for questionnaires designed for either adolescents or women. However, the expression of measurement error may differ by age, sex, or other population characteristics. For example, the novelty of sexual behavior for adolescents may lead to better recall (Tourangeau, 2000) than would occur among older, more experienced males. Also, a generally higher cumulative number of sexual partners for adults, as opposed to adolescents, may make recall of these larger numbers more difficult and thus lead to less reliable measures (Downey, Ryan, Roffman, & Kulich, 1995; Durant & Carey, 2002; Van Duynhoven et al., 1999).

When low reliability items are found in sexual behavior questionnaires, it may also be useful to understand characteristics, demographics or otherwise, that are associated with less reliable items; however, predictors of lower test–retest reliability for specific sexual behavior items have only occasionally been reported in the literature (Morrison-Beedy, Carey, & Tu, 2006; Needle et al., 1995; Schlecht et al., 2001; Williams et al., 2000). Likewise, non-response to questionnaire items, or refusal to answer specific questions, also increases measurement error (Catania, McDermott, & Pollack, 1986).

Finally, studies often limit their reporting of reliability to sexual behaviors and do not report the reliability of demographic characteristics and non-sexual behaviors collected by the same instrument. A comparison of the reliability of responses to questions about sexual behaviors and non-sexual behaviors (some of which may be less associated with image management) may provide further insight into presentation bias.

The objective of the present study was to assess the test–retest reliability of a wide range of sexual and nonsexual items on a self-administered questionnaire delivered to a community sample of 334 men in Tucson, Arizona. The study also sought to assess predictors of unreliable reporting for lifetime number of female sexual partners—a sexual behavior that is important for better understanding the transmission of STDs.

Method

Participants

Men were recruited between September 2003 and May 2005 to a longitudinal study of the natural history of genital human papillomavirus (HPV). Men were enrolled if they met eligibility criteria that included an age of 18–44 years, a residence in southern Arizona, no prior penile or anal cancers or genital warts, a willingness to comply with a total of four visits over 18 months, and no plans to relocate within the next two years.

Recruitment was conducted at a variety of community locales through flyers targeting a large university campus, advertisements in local newspapers and on radio, direct mail advertisements, and face-to-face recruitment at a public STD clinic. Participants were offered a total of $110 for completing all visits of the longitudinal study. Men consented to participation in the study using forms and protocols approved by the Institutional Review Board of the University of Arizona. Additional details of recruitment and study design have been described elsewhere (Giuliano et al., 2008). Men were included in the reliability study if they attended both the run-in visit (an initial visit to establish the participant's suitability for and interest in the study) and the baseline visit approximately two weeks later.

Participants attending both visits had a mean age of 30.2 years (median 28.7) at baseline. A majority identified as white (75.9%) while 19.5% reported a Hispanic ethnicity. Table 1 includes additional characteristics of the study population.

Procedure

Men expressing interest in participating in the study received an appointment to come to the clinic for an initial visit. At this run-in visit, those who consented to the study experienced all participant activities to help them determine whether they wanted to remain in the longitudinal study which would require their participation for an additional 18 months. Participants completed a paper and pencil 86-item self-administered questionnaire, and a study clinician collected exfoliated skin cells at external genital sites for HPV testing. Men were then scheduled for the baseline visit approximately two weeks later

Table 1 Characteristics of study participants at baseline ($n = 334$)

Variable	*N*	%
Age (in years)		
$M = 30.2$; range, 18–44	331	–
Refuse[a]	3	0.9
Race[b]		
White	224	75.9
Black	12	4.1
American Indian/Alaskan Native	8	2.7
Asian	6	2.0
Other	36	12.2
Don't know	1	0.3
Refuse	8	2.7
Ethnicity		
Hispanic	65	19.5
Non-Hispanic	263	78.7
Refuse	6	1.7
Marital status		
Single, never married	208	62.3
Married	46	13.8
Cohabitating	33	9.9
Divorced/widowed	39	11.7
Refuse	8	2.4
Education (in years)		
Less than 12	20	6.0
12–15	174	52.1
More than 15	136	40.7
Refuse	4	1.2
Household income, monthly		
Less than $1500	124	37.1
$1500–$3999	121	36.2
More than $3999	35	10.5
Don't know	26	7.8
Refuse	28	8.4
Age at first sexual intercourse (years)[b]		
Less than 15	45	14.0
16–17	136	42.2
18–19	86	26.7
20–29	45	14.0
Refuse	10	3.1
Lifetime number of female sexual partners[b]		
$M = 15.0 \pm 19.6$ (1–99)	291	–
Refuse	22	7.0
Ever had sex with a man		
Yes	70	21.7
No	249	77.1
Refuse	4	1.2

[a] A refusal was defined as leaving the question blank or choosing the refuse option for questions that were not otherwise precluded by a skip pattern

[b] Numbers do not add to 334 due to skip patterns

($M = 21.5$ days, $SD = 23.8$, median = 16 days, range 9–292 days), at which time they completed the same self-administered questionnaire again. Men received their first HPV test results at their follow-up visit, approximately six months later. No participants were excluded from the reliability study based on their test–retest intervals.

Of 379 men who completed a run-in questionnaire, 334 returned to complete a baseline questionnaire and were included in the reliability study. A comparison of study participants and those who did not return for the baseline visit found no statistically significant differences with regard to any variable.

Measures

The questionnaire contained 86 items with 16 demographic questions, seven alcohol use questions, eight tobacco use questions, and 55 sexual history/sexual behavior questions. The questionnaire had previously been used to assess sexual behavior in women (Giuliano et al., 2002) and had been adapted for and pre-tested with men.

The sexual history and sexual behavior items assessed history of STDs; circumcision; incidence and frequency of penetrative sexual behaviors (vaginal, anal, and oral sex) with women and men; age at first intercourse; number of female and male sexual partners; frequency of condom use with vaginal and anal sex; incidence and frequency of sex with "steady" and other partners; time since last vaginal sex and anal sex; and history of paying for sex. Participants were asked to recall their frequency of substance use and sexual behaviors for varying periods of time, including the last month, the last three months, and lifetime.

A subset of items from the questionnaire was selected for assessment of reliability with preference given to items where reliability coefficients would not be biased by the test–retest interval. For example, while the questionnaire required recall of behaviors over several time periods, items with a one-month recall period were not assessed because the time interval between test and retest ($M = 21.5$ days) would have led the participant to recall sexual behavior for largely different time periods. A total of 39 variables were assessed, including 28 categorical, three ordinal, and eight discrete variables. Demographic, tobacco use, sexual history, and sexual behavior variables were included in the assessment. A discrete variable for number of days between the test and retest was created to estimate how a lengthening interval might affect reliability. In addition, a variable for number of years since first sexual intercourse was created.

Data Analysis

Refusal rates for each questionnaire item were calculated. A refusal was recorded if a participant chose the refusal option

on any particular item or if the participant declined to answer an item that was not precluded by a skip pattern. If a participant responded to a question that should have been skipped because of skip instructions, the response was not used in calculating refusal rates or reliability coefficients.

For categorical variables, the kappa statistic (κ) was calculated. Unlike percent agreement, κ identifies the level of agreement after correcting for chance agreement (Cohen, 1960). Benchmarks for interpreting κ values followed Landis and Koch (1977). Since the κ statistic can be unstable in situations where there are a small number of cases (Maclure & Willett, 1987), reliability was not assessed for binary variables with a case total of less than five (Schroder et al., 2003).

For ordinal variables, a weighted κ statistic was calculated (Cicchetti & Allison, 1971) in order to distinguish among levels of disagreement. For instance, in responding to a question about the frequency of condom use with five possible responses, if a subject chose the adjacent categories of "more than half the time" and "half the time" on run-in and baseline, respectively, simple κ will identify zero agreement between the two responses; however, weighted κ will give credit for the partial agreement.

Discrete variables were assessed using the intraclass correlation coefficient (ICC) (McGraw & Wong, 1996). All ICCs created from skewed variables were transformed using Fisher's z transformation before calculating confidence intervals (Rosner, 2000). Confidence intervals were then transformed back to the original scale. Outliers were assessed using scatter plots and none were considered highly influential.

Logistic regression was performed to identify predictors of unreliable reporting for lifetime number of female sexual partners. This variable was chosen because studies have identified lifetime number of sexual partners as particularly susceptible to measurement error (Schlecht et al., 2001; Van Duynhoven et al., 1999) and because it has been strongly associated with an increased risk for HPV and other STDs (Nielson et al., 2007; Nyitray et al., 2008). Also, compared with other candidate variables in the questionnaire, lifetime number of female sexual partners had more discordant responses which led to more stable regression coefficients. A dichotomous variable was created to delineate unreliable vs. reliable reporting of number of partners. Unreliable reporting was defined as a difference between test and retest of five or more partners. The cut point identifying unreliable reporting was intended to highlight discrepancies that could more seriously bias estimates of effect. For example, because persons who have fewer than approximately five lifetime female sexual partners may be at lower risk for genital HPV infection (Nielson et al., 2007; Vaccarella et al., 2006), a discrepancy of five or more partners on test and retest could bias estimates of the effect of lifetime number of sexual partners on HPV infection. Conversely, the cut point of five partners difference between test and retest allowed for true variation in number of lifetime female partners that may have occurred between test and the retest conducted an average of 21.5 days later. Bivariate associations were computed between this dependent variable and ten potential predictors of unreliable reporting (age, race, ethnicity, marital status, education, income, country of birth, number of years since first intercourse, lifetime number of female sexual partners, and length of test–retest interval). A likelihood ratio test with a p value of less than .20 identified candidate variables for multivariate regression. Independent predictors of unreliable reporting collected at the baseline visit were identified using a backwards selection process. Potential confounders were identified by their association with both the dependent and independent variable and by their ability to substantially modify the independent variable odds ratio. Finally, goodness-of-fit for the final model was assessed using the Hosmer–Lemeshow test.

Results

Item refusal rates were generally below 5% with the exception of income, country of residence, length of residence in the U.S., and three sexual behavior questions. The sexual behavior questions asked about the lifetime number of female sexual partners (refusal rate 3.7% on test and 7.0% on retest), the frequency of sexual intercourse with women in the past three months (refusal rate 6.5% on test and 7.8% on retest), and the lifetime number of male anal sex partners (refusal rate 5.5% on test and 14.7% on retest).

Test–retest reliability for 37 items in the sexual behavior questionnaire is shown in Table 2. Reliability was substantial or almost perfect for demographic variables. Of these, household income per month was the least reliable (weighted $\kappa = .74$). Most of the misclassification on income involved participants switching between adjacent income categories. κ was .90 or greater for all tobacco-use items.

Most sexual behavior items also had almost perfect reliability. Less reliable (but still substantially reliable) were questions asking the participant to report the health conditions of their sexual partners. For example, κ was .67 for an item asking the participant if he had ever had a sex partner with an STD. Two additional variables (ever diagnosed with genital warts and ever diagnosed with syphilis) are not included in Table 2 because too few positive responses for each item precluded calculation of a stable κ statistic.

In bivariate analyses (Table 3), variables associated with unreliable reporting of lifetime number of female sexual partners were older age (OR, 1.13; 95% CI, 1.07–1.20), race (OR, 4.34; 95% CI, 1.03–18.23 for black race compared with white race), an increasing number of years since first intercourse (OR, 1.12; 95% CI, 1.07–1.18), and an increasing lifetime number of female sex partners (OR, 1.03; 95% CI, 1.02–1.05).

Table 2 Kappa and intraclass correlation coefficients by questionnaire item

Variable	N[a]	κ[b] or ICC[c]	95% CI[d]
Demographics			
Race	239	.83	.75–.91
Ethnicity	322	.98	.95– 1.00
Country of birth	330	.97	.92– 1.00
Years lived in the U.S. if not born in the U.S.	20	.99	.97– 1.00
Country of residence for most of life	302	1.00	–
Birth date	330	.95	.93–.96
Marital status	320	.93	.89–.97
Years of education[e]	329	.92	.89–.96
Household income per month[e]	301	.74	.68–.81
Tobacco use			
Ever used tobacco	322	.90	.85–.95
Smoked at least 100 cigarettes	204	.99	.97– 1.00
Age started smoking cigarettes	130	.91	.87–.93
Years of smoking cigarettes	126	.93	.91–.95
Sexual health history			
Ever diagnosed with a sexually transmitted disease	330	.86	.80–.92
Ever diagnosed with genital herpes	324	.88	.74–1.00
Ever diagnosed with Chlamydia	324	.88	.80–.97
Ever diagnosed with gonorrhea	323	.81	.63–.99
Ever diagnosed with non-gonococcal urethritis	324	.82	.74–.91
Ever diagnosed with hepatitis B	322	.80	.60–.99
Ever diagnosed with hepatitis C	321	.80	.57–1.00
Ever diagnosed with HIV	324	.92	.77–1.00
Ever had sex partner with a sexually transmitted disease	331	.67	.61–.74
Ever had sex partner with genital warts	332	.71	.63–.78
Ever had sex partner with an abnormal Pap smear	330	.71	.65–.77
Circumcision	330	.93	.88–.99
Sexual behavior (lifetime)			
Ever had vaginal sex	331	.98	.93–1.00
Ever had anal sex	327	.95	.92–.98
Ever had oral sex	329	.76	.59–.93
Age at first sexual intercourse	311	.91	.89–.93
Lifetime number of female sexual partners	289	.89	.86–.91
Use of a condom at first sex with steady partner	188	.76	.67–.85
Ever paid a woman for sex	311	.96	.92–1.00
Ever had sex with a man	317	.94	.90–.99
Lifetime number of male anal sex partners	49	1.00	–
Ever paid a man for sex	63	.78	.49–1.00
Sexual behavior (past three months)			
Frequency of condom use for vaginal sex[e]	238	.83	.78–.89
Frequency of intercourse with women	219	.85	.80–.88

[a] N excludes refusals and missing observations

[b] κ: kappa

[c] ICC: intraclass correlation coefficient

[d] CI: Confidence interval. For ICC interval estimation, ICCs were z transformed and then converted back to the original scale

[e] Weighted κ after Cicchetti and Allison (1971)

In multivariate analyses (Table 4), older age (OR, 1.12; 95% CI, 1.06–1.20) and lifetime number of female sexual partners (OR, 1.02; 95% CI, 1.01–1.04) were independent predictors of unreliable reporting. Also, more than 15 years of education, compared with 12–15 years of education, was inversely associated with unreliable reporting (OR, .33; 95% CI, .13–.88). Age and lifetime number of female sexual partners were moderately correlated (Spearman $r = .39$;

Table 3 Factors predicting unreliable reporting between run-in and baseline visits for self-reported lifetime number of female sexual partners: bivariate analyses

Variable[a]	N[b]	OR[c]	95% CI[d]
Age	286	1.13	1.07–1.20
Race			
White	199	Reference	
Black	10	4.34	1.03–18.23
Other	39	1.49	.52–4.28
Ethnicity			
Non-Hispanic	230	Reference	
Hispanic	55	1.23	.47–3.20
Marital status			
Single, never married	174	Reference	
Married	45	1.44	.49–4.23
Cohabitating	29	1.33	.36–4.94
Divorced/widowed	36	2.30	.82–6.46
Education (years)			
Less than 12	13	2.19	.55–8.69
12–15	149	Reference	
More than 15	126	.43	.17–1.06
Household income, monthly			
Less than $1500	105	Reference	
$1500–3999	110	1.07	.42–2.74
More than $3999	31	1.14	.29–4.51
Country of birth			
U.S.	269	Reference	
Other	19	.50	.06–3.88
Years since first sexual intercourse	285	1.12	1.07–1.18
Lifetime number of female sexual partners	289	1.03	1.02–1.05

[a] At baseline

[b] Numbers do not add to 334 due to refusals and missing observations

[c] OR: odds ratio

[d] CI: confidence interval

Table 4 Factors predicting unreliable reporting between run-in and baseline visits for self-reported lifetime number of female sexual partners: multivariate analyses

Variable[a]	Adjusted OR[b]	95% CI[c]
Age	1.12	1.06–1.20
Education (years)		
Less than 12	.90	.18–4.57
12–15	Reference	
More than 15	.33	.13–.88
Lifetime number of female sexual partners	1.02	1.01–1.04

[a] At baseline

[b] Odds ratios adjusted for other variables in the model

[c] CI: confidence interval

$p < .0001$) (data not shown). Holding lifetime number of female partners constant at the sample mean of 15.0, there was a 2% probability that a 25-year-old with more than 15 years of education would be unreliable on test and retest for lifetime number of female partners. In contrast, there was a 10% probability that a 40-year-old with more than 15 years of education would provide unreliable answers on this item. There would be a 25% probability that a 40-year-old man with a high school education would provide unreliable answers on his lifetime number of female sexual partners.

The final model fit the data well (Hosmer–Lemeshow $p = .43$) and no confounders of the association between the independent and dependent variables were identified. Furthermore, the test–retest interval was not statistically associated with the reliability of reporting lifetime number of female sexual partners ($p = .31$, data not shown).

Discussion

The 334 men in this study, recruited from the general community, provided consistent answers to most demographic, tobacco use, and sexual behavior questions on both test and retest. The sexual behavior items were generally as reliable as the demographic items. Reliability dropped somewhat for items that asked about household income and the health of the men's sexual partners (e.g., partners' histories of STDs or abnormal Pap smears). The study's overall findings were consistent with other test–retest reliability studies with populations of women and adolescents that also found generally high reliability for sexual behavior questionnaires (Brener et al., 2002; Durant & Carey, 2000; Sieving et al., 2005).

While men in this study were explicitly given permission to refuse to answer items on the questionnaire, with a few exceptions, most men completed all the questions. This is consistent with observations from other studies (Peterson & Catania, 1997). Items with higher non-response in the current study included questions related to immigration. For example, persons not born in the U.S. were asked how long they had lived in the U.S. A total of 8.3% (2/24) on test and 12.5% (3/24) on retest refused to answer the question. The refusal rates on immigration-related questions may be due, in part, to participants who perceive a more threatening climate for immigrants in the U.S. An item asking about income was refused by 5.4 and 8.4% of participants on the test and retest, respectively. The perceived threat associated with divulging income information has been documented previously (Peterson & Catania, 1997).

Regarding higher refusal rates for sexual behavior items, 6.5 and 7.8% of participants on test and retest refused to report the frequency with which they had sexual intercourse with women in the past three months. Also, on test and retest, 5.5% (4/73) and 14.7% (10/68) of men acknowledging sex

with other men did not report the lifetime number of male anal sex partners. Since this item came near the end of the 86-item questionnaire, it is possible that participant fatigue led to the higher refusal rates. Refusal rates should be considered only a subset of the total amount of refused questionnaire items. Some men may have felt more comfortable with answering a question falsely and less comfortable with skipping a question or checking the refuse option (Catania et al., 1990).

Increasing age and a greater lifetime number of female sexual partners were independent predictors of unreliable reporting while more than 15 years of education was associated with reliable reporting for lifetime number of female partners. Increasing measurement error on questionnaires in association with fewer years of education and increasing age has been seen in prior research (Schlecht et al., 2001; Van Duynhoven et al., 1999). It is also possible that the effect of increasing age is simply requiring the older person to remember more years of sexual activity, which likely limits reliability due to memory failure (Fenton, Johnson, McManus, & Erens, 2001; Kauth, St. Lawrence, & Kelly, 1991; Lagarde, Enel, & Pison, 1995; Tourangeau, 2000). The present study's finding among mostly heterosexual men that a higher lifetime number of female sexual partners was associated with more unreliable reporting of those partners echoes prior research among women and gay men where higher frequency behaviors were also associated with increased measurement error (Downey et al., 1995; Durant & Carey, 2002; Morrison-Beedy et al., 2006). Our definition of unreliable reporting, a difference between test and retest of five or more partners, may be appropriate given the participants' mean of 15.0 lifetime female sexual partners. A lower cut point (e.g., a difference of only two or more partners) would have labeled more participants as inconsistent even though the actual difference between test and retest would be small. Also, for men who added to their lifetime number of sexual partners between test and retest, a lower cutpoint would label more of these men as unreliable even though they may have reported their number of partners accurately. Conversely, allowing for a greater amount of inconsistency in number of partners reported may have been too low a standard for reliability.

We chose to assess the predictors of unreliable reporting for lifetime number of female sexual partners because number of sexual partners is a risk factor that is strongly associated with HPV infection in men and therefore plays a crucial role in determining risk for anogenital HPV in men. Initially, we also planned to assess predictors of unreliable reporting for a history of sex with men and for age at first sexual intercourse—two variables that are also important in determining risk for anogenital HPV; however, only lifetime number of female sexual partners generated a large enough number of discordant answers between test and retest to allow creation of a binary variable with adequate cell sizes for men who reported their number of partners reliably and men who did not.

The two variables of age and lifetime number of sexual partners were moderately correlated as one would expect given that the lifetime number of sexual partners cannot decrease as age increases. Given this correlation, one practical use may be to support memory recall of lifetime number of sexual partners in men over a certain age, for example, using Timeline Follow-back methods (Carey, Carey, Maisto, Gordon, & Weinhardt, 2001; Weinhardt et al., 1998).

Due to the time interval between test and retest (mean 21.5 days, median 16 days), participants were asked to recall their sexual behavior for slightly different time periods. For example, the reliability of the item, "In the past 3 months, how many times did you have sexual intercourse with a woman?" may be expected to decline if the frequency of the men's sexual behavior changed during the test–retest interval in comparison to the three months prior to the test. To minimize potential bias caused by the test–retest interval, items measuring recall for time periods of one month were not assessed. Also, a test–retest interval variable was created and no statistically significant association was found between interval length and the reliability of reporting lifetime number of female sexual partners. Other researchers have also found the number of partners reported on test and retest was unaffected by the recall interval (McElrath, Chitwood, Griffin, & Comerford, 1994).

The interval between test and retest also could have promoted bias if the period was short enough to allow the recall of prior answers (Nunnally, 1978); however, such explanations have been disputed as a significant source of bias (McKelvie, 1992). In addition, this study's test–retest interval was a common span of time in test–retest studies that may balance the competing goals of limiting recall of prior answers while providing a largely overlapping reference period for both test and retest questions (Weinhardt et al., 1998; Wiederman, 2002). Even so, it seems plausible, for example, that a man intent on concealing a history of sex with men may have invoked a strategy to remember his answers to such questions, a situation that will bias the reliability coefficient upwards. It also reminds us that highly reliable data cannot be assumed to be valid. Alternatively, a participant may have decided to conceal a behavior on the test but then decided to answer truthfully on retest, which would have attenuated the reliability coefficient.

The reliability coefficients we calculated may be inflated because, compared with non-volunteers, men who volunteered may be less likely to be evasive about reporting their sexual behaviors; thus, the study may not be generalizable to a general population of non-volunteers. Also, the Landis and Koch (1977) categories that label levels of reliability as "almost perfect," "substantial," etc. are subjective;

however, they are useful for discussions that compare reliability coefficients.

This study could have been improved upon by the use of a retest that explicitly referenced the same time period as the original test. Such a measure may have reduced bias introduced by the temporal offset of the test–retest interval. However, such a measure would have required construction of a different baseline questionnaire which then would have been inappropriate for the longitudinal nature of the larger study.

As computer technology becomes more commonplace, it seems likely that computer-assisted self-interview technology will increasingly become the survey method of choice for collecting accurate human sexual behavior data; however, it is not clear that CASI survey methods increase the reliability of sexual behavior survey methods in all situations (Jaya, Hindin, & Ahmed, 2008; Jennings, Lucenko, Malow, & Devieux, 2002; Jobe, Pratt, Tourangeau, Baldwin, & Rasinski, 1997; Johnson et al., 2001; Le et al., 2006; Metzger et al., 2000; Morrison-Beedy et al., 2006; Simoes, Bastos, Moreira, Lynch, & Metzger, 2006b; Turner et al., 1998; Van den Brakel, Vis-Visschers, & Schmeets, 2006). While there are studies comparing CASI (or audio-CASI) with other survey methods like self-administered questionnaires or face-to-face interviews, we are aware of few studies reporting the test–retest reliability of this method for a sexual behavior questionnaire (Krawczyk et al., 2003; Romer et al., 1997; Williams et al., 2000; Wolford et al., 2008). Since none of these studies have assessed reliability in a general community sample of adults, future research should assess the level of test–retest reliability obtainable by CASI for sexual behavior data collected from general populations.

Acknowledgments The authors are especially thankful to the men who gave of their time in this study. We also wish to thank Eleanor Pomaski, Bonnie Weible, and Martha Abrahamsen for their support in collecting these data. This article was completed as part of dissertation work for the Epidemiology Graduate Program at the Mel and Enid Zuckerman College of Public Health. This research was supported by the Arizona Disease Control Research Commission (Contract 7016; funding period, 2003–2006).

References

Boekeloo, B. O., Schiavo, L., Rabin, D. L., Conlon, R. T., Jordan, C. S., & Mundt, D. J. (1994). Self-reports of HIV risk factors by patients at a sexually transmitted disease clinic: Audio vs. written questionnaires. *American Journal of Public Health, 84*, 754–760.

Brener, N. D., Kann, L., McManus, T., Kinchen, S. A., Sundberg, E. C., & Ross, J. G. (2002). Reliability of the 1999 Youth Risk Behavior Survey questionnaire. *Journal of Adolescent Health, 31*, 336–342.

Carballo-Dieguez, A., Remien, R. H., Dolezal, C., & Wagner, G. (1999). Reliability of sexual behavior self-reports in male couples of discordant HIV status. *Journal of Sex Research, 36*, 152–158.

Carey, M. P., Carey, K. B., Maisto, S. A., Gordon, C. M., & Weinhardt, L. S. (2001). Assessing sexual risk behaviour with the Timeline Followback (TLFB) approach: Continued development and psychometric evaluation with psychiatric outpatients. *International Journal of STD and AIDS, 12*, 365–375.

Catania, J. A. (1999). A framework for conceptualizing reporting bias and its antecedents in interviews assessing human sexuality. *Journal of Sex Research, 36*, 25–38.

Catania, J. A., Gibson, D. R., Chitwood, D. D., & Coates, T. J. (1990). Methodological problems in AIDS behavioral research: Influences on measurement error and participation bias in studies of sexual behavior. *Psychological Bulletin, 108*, 339–362.

Catania, J. A., McDermott, L. J., & Pollack, L. M. (1986). Questionnaire response bias and face-to-face interview sample bias in sexuality research. *Journal of Sex Research, 22*, 52–72.

Cicchetti, D. V., & Allison, T. (1971). A new procedure for assessing reliability of scoring EEG sleep recordings. *American Journal of Electroneurodiagnostic Technology, 11*, 101–109.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37–46.

Des Jarlais, D. C., Paone, D., Milliken, J., Turner, C. F., Miller, H., Gribble, J., et al. (1999). Audio-computer interviewing to measure risk behaviour for HIV among injecting drug users: A quasi-randomised trial. *Lancet, 353*, 1657–1661.

Downey, L., Ryan, R., Roffman, R., & Kulich, M. (1995). How could I forget—inaccurate memories of sexually intimate moments. *Journal of Sex Research, 32*, 177–191.

Durant, L. E., & Carey, M. P. (2000). Self-administered questionnaires versus face-to-face interviews in assessing sexual behavior in young women. *Archives of Sexual Behavior, 29*, 309–322.

Durant, L. E., & Carey, M. P. (2002). Reliability of retrospective self-reports of sexual and nonsexual health behaviors among women. *Journal of Sex and Marital Therapy, 28*, 331–338.

Fenton, K. A., Johnson, A. M., McManus, S., & Erens, B. (2001). Measuring sexual behaviour: Methodological challenges in survey research. *Sexually Transmitted Infections, 77*, 84–92.

Franco, E. L. (1997). Statistical issues in studies of human papillomavirus infection and cervical cancer. In E. Franco & J. Monsonego (Eds.), *New developments in cervical cancer screening* (pp. 39–50). Oxford: Blackwell Science.

Giuliano, A. R., Harris, R., Sedjo, R. L., Baldwin, S., Roe, D., Papenfuss, M. R., et al. (2002). Incidence, prevalence, and clearance of type-specific human papillomavirus infections: The Young Women's Health Study. *Journal of Infectious Diseases, 186*, 462–469.

Giuliano, A. R., Lu, B., Nielson, C. M., Flores, R., Papenfuss, M. R., Lee, J. H., et al. (2008). Age-specific prevalence, incidence, and duration of human papillomavirus infections in a cohort of 290 US men. *Journal of Infectious Diseases, 198*, 827–835.

Jaya, Hindin, M. J., & Ahmed, S. (2008). Differences in young people's reports of sexual behaviors according to interview methodology: A randomized trial in India. *American Journal of Public Health, 98*, 169–174.

Jennings, T. E., Lucenko, B. A., Malow, R. M., & Devieux, J. G. (2002). Audio-CASI vs interview method of administration of an HIV/STD risk of exposure screening instrument for teenagers. *International Journal of STD and AIDS, 13*, 781–784.

Jobe, J. B., Pratt, W. F., Tourangeau, R., Baldwin, A. K., & Rasinski, K. (1997). Effects of interview mode on sensitive questions in a fertility survey. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, et al. (Eds.), *Survey measurement and process quality* (pp. 311–329). New York: John Wiley.

Johnson, A. M., Copas, A. J., Erens, B., Mandalia, S., Fenton, K., Korovessis, C., et al. (2001). Effect of computer-assisted self-interviews on reporting of sexual HIV risk behaviours in a general population sample: A methodological experiment. *AIDS, 15*, 111–115.

Kalichman, S. C., Kelly, J. A., & Stevenson, L. Y. (1997). Priming effects of HIV risk assessments on related perceptions and behavior: An experimental field study. *AIDS and Behavior, 1*, 3–8.

Kauth, M. R., St. Lawrence, J. S., & Kelly, J. A. (1991). Reliability of retrospective assessments of sexual HIV risk behavior: A comparison of biweekly, three-month, and twelve-month self-reports. *AIDS Education and Prevention, 3*, 207–214.

Krawczyk, C. S., Gardner, L. I., Wang, J. C., Sadek, R., Loughlin, A. M., Anderson-Mahoney, P., et al. (2003). Test–retest reliability of a complex human immunodeficiency virus research questionnaire administered by an audio computer-assisted self-interviewing system. *Medical Care, 41*, 853–858.

Lagarde, E., Enel, C., & Pison, G. (1995). Reliability of reports of sexual behavior: A study of married couples in rural west Africa. *American Journal of Epidemiology, 141*, 1194–1200.

Landis, J. R., & Koch, G. G. (1977). Measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

Le, L. C., Blum, R. W., Magnani, R., Hewett, P. C., & Do, H. M. (2006). A pilot of audio computer-assisted self-interview for youth reproductive health research in Vietnam. *Journal of Adolescent Health, 38*, 740–747.

Maclure, M., & Willett, W. C. (1987). Misinterpretation and misuse of the kappa statistic. *American Journal of Epidemiology, 126*, 161–169.

McElrath, K., Chitwood, D. D., Griffin, D. K., & Comerford, M. (1994). The consistency of self-reported HIV risk behavior among injection drug users. *American Journal of Public Health, 84*, 1965–1970.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*, 30–46.

McKelvie, S. J. (1992). Does memory contaminate test–retest reliability? *Journal of General Psychology, 119*, 59–72.

Metzger, D. S., Koblin, B., Turner, C., Navaline, H., Valenti, F., Holte, S., et al. (2000). Randomized controlled trial of audio computer-assisted self-interviewing: Utility and acceptability in longitudinal studies. *American Journal of Epidemiology, 152*, 99–106.

Morrison-Beedy, D., Carey, M. P., & Tu, X. (2006). Accuracy of audio computer-assisted self-interviewing (ACASI) and self-administered questionnaires for the assessment of sexual behavior. *AIDS and Behavior, 10*, 541–552.

Müller, M. J. (2007). Development and retest reliability of a German version of the Sexual Behaviour Questionnaire (SBQ-G). *Archives of Andrology, 53*, 67–69.

Needle, R., Fisher, D. G., Weatherby, N., Chitwood, D., Brown, B., Cesari, H., et al. (1995). Reliability of self-reported HIV risk behaviours of drug users. *Psychology of Addictive Behaviors, 9*, 242–250.

Nielson, C. M., Harris, R. B., Dunne, E. F., Abrahamsen, M., Papenfuss, M. R., Flores, R., et al. (2007). Risk factors for anogenital human papillomavirus infection in men. *Journal of Infectious Diseases, 196*, 1137–1145.

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

Nyitray, A., Nielson, C. M., Harris, R. B., Flores, R., Abrahamsen, M., Dunne, E. F., et al. (2008). Prevalence of and risk factors for anal human papillomavirus infection in heterosexual men. *Journal of Infectious Diseases, 197*, 1676–1684.

Peterson, J., & Catania, J. (1997). Item nonresponse in the national AIDS behavioral surveys among African American and white respondents. In J. Bancroft (Ed.), *Researching sexual behavior: Methodological issues* (pp. 106–109). Bloomington: Indiana University Press.

Potterat, J. J., Phillips, L., & Muth, J. B. (1987). Lying to military physicians about risk factors for HIV infections. *Journal of the American Medical Association, 257*, 1727.

Rohan, T. E., McLaughlin, J. R., & Harnish, D. G. (1994). Repeatability of interview-derived information on sexual history: A study in women. *Epidemiology, 5*, 360–363.

Romer, D., Hornik, R., Stanton, B., Black, M., Li, X. M., Ricardo, I., et al. (1997). ''Talking'' computers: A reliable and private method to conduct interviews on sensitive topics with children. *Journal of Sex Research, 34*, 3–9.

Rosner, B. (2000). *Fundamentals of biostatistics* (5th ed.). Pacific Grove, CA: Duxbury.

Schlecht, N. F., Franco, E. L., Rohan, T. E., Kjaer, S. K., Schiffman, M. H., Moscicki, A. B., et al. (2001). Repeatability of sexual history in longitudinal studies on HPV infection and cervical neoplasia: Determinants of reporting error at follow-up interviews. *Journal of Epidemiology and Biostatistics, 6*, 393–407.

Schrimshaw, E. W., Rosario, M., Meyer-Bahlburg, H. F., & Scharf-Matlick, A. A. (2006). Test–retest reliability of self-reported sexual behavior, sexual orientation, and psychosexual milestones among gay, lesbian, and bisexual youths. *Archives of Sexual Behavior, 35*, 225–234.

Schroder, K. E., Carey, M. P., & Vanable, P. A. (2003). Methodological challenges in research on sexual risk behavior: II. Accuracy of self-reports. *Annals of Behavioral Medicine, 26*, 104–123.

Sieving, R., Hellerstedt, W., McNeely, C., Fee, R., Snyder, J., & Resnick, M. (2005). Reliability of self-reported contraceptive use and sexual behaviors among adolescent girls. *Journal of Sex Research, 42*, 159–166.

Simoes, A. A., Bastos, F. I., Moreira, R. I., Lynch, K. G., & Metzger, D. S. (2006a). Acceptability of audio computer-assisted self-interview (ACASI) among substance abusers seeking treatment in Rio de Janeiro, Brazil. *Drug and Alcohol Dependence, 82*(Suppl. 1), S103–S107.

Simoes, A. A., Bastos, F. I., Moreira, R. I., Lynch, K. G., & Metzger, D. S. (2006b). A randomized trial of audio computer and in-person interview to assess HIV risk among drug and alcohol users in Rio De Janeiro, Brazil. *Journal of Substance Abuse Treatment, 30*, 237–243.

Sohler, N., Colson, P. W., Meyer-Bahlburg, H. F. L., & Susser, E. (2000). Reliability of self-reports about sexual risk behavior for HIV among homeless men with severe mental illness. *Psychiatric Services, 51*, 814–816.

St. Lawrence, J. S., Chapdelaine, A. P., Devieux, J. G., O'Bannon, R. E., Brasfield, T. L., & Eldridge, G. D. (1999). Measuring perceived barriers to condom use: Psychometric evaluation of the Condom Barriers Scale. *Assessment, 6*, 391–404.

Stoneburner, R. L., Chiasson, M. A., Solomon, K., & Rosenthal, S. (1986). Risk factors in military recruits positive for HIV antibody. *New England Journal of Medicine, 315*, 1355.

Taylor, J. F., Rosen, R. C., & Leiblum, S. R. (1994). Self-report assessment of female sexual function: Psychometric evaluation of the brief index of sexual functioning for women. *Archives of Sexual Behavior, 23*, 627–643.

Tourangeau, R. (2000). Remembering what happened: Memory errors and survey reports. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The science of self-report: Implications for research and practice* (pp. 29–47). Mahwah, NJ: Lawrence Erlbaum Associates.

Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Pleck, J. H., & Sonenstein, F. L. (1998). Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science, 280*, 867–873.

Vaccarella, S., Lazcano-Ponce, E., Castro-Garduno, J. A., Cruz-Valdez, A., Diaz, V., Schiavon, R., et al. (2006). Prevalence and determinants of human papillomavirus infection in men attending vasectomy clinics in Mexico. *International Journal of Cancer, 119*, 1934–1939.

Van den Brakel, J. A., Vis-Visschers, R., & Schmeets, J. J. G. (2006). An experiment with data collection modes and incentives in the Dutch family and fertility survey for young Moroccans and Turks. *Field Methods, 18*, 321–334.

Van Duynhoven, Y., Nagelkerke, N. J. D., & Van de Laar, M. J. W. (1999). Reliability of self-reported sexual histories: Test–retest and interpartner comparison in a sexually transmitted diseases clinic. *Sexually Transmitted Diseases, 26*, 33–42.

Weinhardt, L. S., Forsyth, A. D., Carey, M. P., Jaworski, B. C., & Durant, L. E. (1998). Reliability and validity of self-report measures of HIV-related sexual behavior: Progress since 1990 and recommendations for research and practice. *Archives of Sexual Behavior, 27*, 155–180.

Wiederman, M. W. (2002). Reliability and validity of measurement. In M. W. Wiederman & B. E. Whitley (Eds.), *Handbook for conducting research on human sexuality* (pp. 25–50). Mahwah, NJ: Erlbaum.

Williams, M. L., Freeman, R. C., Bowen, A. M., Zhao, Z., Elwood, W. N., Gordon, C., et al. (2000). A comparison of the reliability of self-reported drug use and sexual behaviors using computer-assisted versus face-to-face interviewing. *AIDS Education and Prevention, 12*, 199–213.

Wolford, G., Rosenberg, S. D., Rosenberg, H. J., Swartz, M. S., Butterfield, M. I., Swanson, J. W., et al. (2008). A clinical trial comparing interviewer and computer-assisted assessment among clients with severe mental illness. *Psychiatric Services, 59*, 769–775.

Arch Sex Behav (2010) 39:1353–1358
DOI 10.1007/s10508-010-9653-9

ORIGINAL PAPER

Sexual Activity in HIV-Positive African American Crack Cocaine Smokers

Sandra C. Timpson · Mark L. Williams · Anne M. Bowen · John S. Atkinson · Michael W. Ross

Received: 14 April 2008 / Revised: 26 February 2010 / Accepted: 26 February 2010 / Published online: 14 July 2010

Abstract The AIDS epidemic in the United States continues to disproportionately affect minorities of color, especially African Americans. The purpose of this study was to explore the sexual behaviors of a sample of African American HIV positive crack smokers aware of their serostatus. Participants (100 men, 37 women) were included in this study based on the following criteria: a minimum age of 18 years, HIV positive serostatus, treatment with HIV antiretroviral medications for a minimum of 3 months prior to interview, crack cocaine use at least once in the 7 days prior to being interviewed, willingness to provide a urine sample to confirm recent drug use, and vaginal or anal sex at least once in the past 7 days. The questionnaire was a compilation of other reliable surveys and was designed to collect sociodemographic data, drug use, sexual behavior, condom use intentions and motivators, STD and HIV infection history, HIV medications, and adherence requirements. Participants reported having 1,266 different partners in the 30 days prior to the interview and had traded sex for money or drugs with 68%. A total of 79 participants had multiple partners and accounted for 1,247 partnerships. Rates of consistent condom use across partnerships were low, indicating that more interventions in this at-risk population are needed.

Keywords African-American · Crack cocaine · HIV+ · Sexual activity

S. C. Timpson · M. L. Williams (✉) · J. S. Atkinson · M. W. Ross
Center for Health Promotion and Prevention Research, University of Texas School of Public Health, 7000 Fannin, Suite 2516, Houston, TX 77030, USA
e-mail: Mark.L.Williams@uth.tmc.edu

A. M. Bowen
Department of Psychology, University of Wyoming, Laramie, WY, USA

Introduction

The AIDS epidemic in the United States continues to disproportionately affect minorities of color, especially African Americans. African Americans now comprise a majority of individuals diagnosed with AIDS in the United States and a majority of deaths due to the disease. Infection with HIV among African Americans is rising at a much faster rate than for other racial/ethnic groups in the country (CDC, 2009). Once diagnosed with AIDS, African Americans succumb sooner than Whites (CDC, 2009). A meta-analysis by Smith et al. (2000) found that AIDS is now the leading cause of death among African Americans aged 25–44. The co-epidemics of crack cocaine and other sexually transmitted diseases among African Americans residing in urban areas may partially explain their higher levels of HIV infection (Timpson, Williams, Ross, & Keel, 2005; Williams et al., 2000). Factors related to cocaine use itself may increase the risk for HIV among crack smokers.

In an analysis of the HIV infection rate in New York City, Des Jarlais et al. (1999, 2000) reported that the epidemic was declining among injection drug users. However, these investigators also speculated that city-wide the epidemic may not be in decline because of increasing rates of infection among crack cocaine smokers. Ross, Risser, Peters, and Johnson (2006) noted the strong association between crack use and syphilis in African Americans in the early 2000s, suggesting that, in this population, sexually transmissible infections are closely associated with crack smoking. The behavioral effects that result from crack smoking also increase the chances for HIV infection (Timpson, Williams, Bowen, & Keel, 2003). Smoking crack induces a strong sense of euphoria and other feelings of mood elevation that may lead to an increased willingness to engage in sexual behaviors that are highly risky (McCoy & Inciardi, 1995). For some new users, the disinhibiting effect of crack results in more extreme sexual orgasms and hypersexuality, creating a self- reinforcing behavioral pattern. The effects of crack smoking also can induce cycles

of binging and crashing that increase risk of sexual risk. Users smoke crack in binging cycles, smoking as much as resources allow for up to 3 or 4 days, then crashing for similar lengths of time (Ross, Timpson, Williams, & Bowen, 2003; Williams et al., 2001). Binge cycles may include frequent and prolonged sexual activity with unknown partners (McCoy & Inciardi, 1995). Sexual risk of HIV infection may be particularly pronounced if an individual's psychological functioning is impaired (Riehman, Wechsberg, Zule, Lam, & Levine, 2008; Risser, Timpson, McCurdy, Ross, & Williams, 2006; Williams & Latkin, 2004).

Research studies show that the sexual behaviors of crack smokers may be amenable to brief interventions that focus on reducing the number of sexual partners and, to a lesser extent, condom use. Kotranski et al. (1998) found that 684 not-in-treatment drug users decreased the number of partners with whom they had sex after participating in a brief risk reduction intervention, and that 37% initiated or increased condom use during vaginal sex. Cottler et al. (1998) found similar results. Of 715 crack smokers participating in brief intervention, 75% reduced the number of sexual partners and 5% eliminated their sexual risk, defined as having no sex or consistently using condoms during vaginal sex. Kwiatkowski, Stober, Booth, and Zhang (1999) reported similar results in a sample of 3,357 heterosexual, sexually active drug users participating in a multi-site risk reduction study. In this sample, there was a 7% increase in condom use 6 months after participating in a brief intervention. However, 71% of the sample reported no change or less condom use after 6 months. Factors predicting increased condom use included HIV positive status, being single with multiple sex partners, or engaging in money for drugs or sex. Even so, Kwiatkowski et al. reported that only half of HIV positive drug users increased their condom use. In a study of 1,362 participants, who were both injectors and crack cocaine smokers, interventions were compared using a nine session enhanced cognitive-behavioral intervention and a two session standard counseling and testing intervention developed by the National Institute on Drug Abuse (Hershberger, Wood, & Fisher, 2003). The enhanced intervention was rarely found to be different from NIDA's standard counseling and testing intervention in "reducing both drug and sex-related risks" (Hershberger et al., 2003). Both interventions resulted in significant reduction of most HIV risk behaviors. In another study (Sterk, Theall, & Elifson, 2003), 265 African American women who only used crack cocaine participated in an intervention that used a randomized block design that assigned them to one of two enhanced interventions or to the two-session NIDA standard intervention. The two enhanced designs included a four-session motivation intervention and a four-session negotiation intervention. Participants were interviewed at baseline, post-intervention, and 6-month follow-up. Women in all three intervention groups improved in sexual and drug using behaviors. The entire sample reported reductions in crack use, trading sex, and having sex while high. In a more recent study of a sample of 514 crack users, Ross, Timpson, Williams, and Bowen (2007) found that 29% had participated in a previous HIV prevention intervention in the past 12 months. Those who had reported using condoms scored significantly higher on all measures of condom use, had positive condom use expectations, and had higher condom related self-efficacy beliefs.

The great majority of studies investigating behaviors of drug users at risk for HIV infection have examined seroprevalence rates and risk behaviors as these are related to primary infection. However, it is assumed that once individuals are aware of being HIV positive, they will be more likely to use condoms during sexual activity (Friedman et al., 1999; Ross et al., 2007). In a relatively early study of serodiscordant couples where one of the partners was HIV seropositive and a drug injector and the other not a drug injector, Friedman et al. (1994) reported a high level of consistent condom use. Although the study did not squarely address the issue of secondary prevention, it is significant because it can be assumed that the seropositive partner was aware that he or she was HIV positive and was behaving in light of this knowledge.

While primary prevention is a major concern, secondary prevention is also important in preventing the HIV epidemic. However, little is known about HIV positive drug users who are aware that they are seropositive. In a preliminary investigation of the risk behaviors of 137 African American crack smokers who were aware of their seropositive HIV status, we found that the majority of the sample smoked crack daily, had sex frequently, and used condom inconsistently with all partner types (Timpson et al., 2003). We also found that knowledge of HIV serostatus was not related to condom use and that, except for primary partners, most sexual encounters for HIV positive crack smokers were with partners whose status was unknown or who were HIV negative. The purpose of the present study was to further explore the sexual behaviors of a sample HIV positive crack smokers aware of their serostatus. Awareness of HIV positive serostatus was confirmed by treatment for HIV disease.

Method

Participants

Participants were included in this study based on the following criteria: African-American, a minimum age of 18 years, HIV positive serostatus, treatment with HIV antiretroviral medications for a minimum of 3 months prior to interview, crack cocaine use at least once in the 7 days prior to being interviewed, willingness to provide a urine sample to confirm recent drug use, and vaginal or anal sex at least once in the past 7 days. Individuals were recruited from neighborhoods in which they lived and from agencies offering services to people who are HIV positive. Telephone screening was preceded by being informed of the intent of the screening and verbal consent. Those who appeared to meet eligibility requirements as a result of the telephone screening were asked to come to a neighborhood data collection center, provided informed consent, and completed a more in depth screening

instrument. Participants were also asked to bring with them bottles for all or some of the medications they were currently receiving for HIV infection. After receiving informed consent, participants were screened by a research assistant and asked to show the medication bottles they had brought with them. Of the 225 people who were screened for the study, 137 were eligible and were interviewed. None who were told that they were eligible refused to participate.

If participants met eligibility requirements, they were asked to respond to a risk behavior questionnaire. A research assistant administered the 2-h survey. Participants were paid $20 for their time. Study procedures were reviewed by the committee for the protection of human subjects at the University of Texas, and all participants signed informed consents.

The African American sample of 137 participants was mostly male ($n = 100$, 73%) and single ($n = 100$, 73%). The age of participants ranged from 21 to 53 years. Fifty-one percent of the sample was 40 years of age or older and both the mean and median age was 40 years. Individuals were almost equally proportioned in sexual orientation with 49 (36%) who self-identified as heterosexual, 49 (36%) as bisexual, and 44 (32%) as homosexual. Most ($n = 125$, 91%) had completed a high school education. Slightly more than a quarter of the sample ($n = 36$, 26%) considered themselves to be homeless while 17 (12%) were living in a shelter at the time of the interview. Fifty-eight (42%) participants had ever injected drugs, and 11 (8%) had injected at least once in the previous 30 days.

Measures

Data for this cross-sectional study were collected between March 1999 and July 2000, in Houston, Texas using the Elicitation of Compliance and Adherence Behaviors Questionnaire (ECAB). The ECAB, a compilation of other reliable surveys, was designed to collect data regarding sociodemographic characteristics, drug use and sexual behavior, condom use intentions and behaviors, STD and HIV infection history, and HIV medication and adherence history.

Sociodemographic data items on the ECAB were adapted from the Texas Christian University Short Assessment Form (Simpson, Knight, & Ray, 1993). This instrument has been shown to be reliable and valid (Simpson et al., 1994). Drug use information, both lifetime and current drug use, was collected on the use of marijuana, methamphetamines, crack cocaine, cocaine, heroin, and speedball (a mix of cocaine and heroin). Thirty-day recall of current drug use has been found to be reliable and to have a high level of validity in other studies using the same or similar measures (Watters et al., 1992; Weatherby et al., 1994).

Sexual behaviors were measured for 6 months and 30 days prior to the interview. Data were collected by partner type, gender of the sexual partner, and type of sexual behavior. Partner types were defined as main, casual, drug trading (in which the participant was given drugs in exchange for sex), money trading (in which the participant was given money in exchange for sex), and new. Sexual behaviors asked about were oral, anal, and vaginal sex. Condom use with a main partner was assessed by asking participants if they always used condoms with that partner. The total number of partners was computed as the sum of casual, trading, and new partners. Main partners were not included in the total. Condom use with casual, sex for drugs, sex for money, and new partners was assessed by asking if condoms were always used with the most recent partner of the respective type.

Two-tailed independent-samples *t*-tests (with equal variances assumed) were used to assess differences in the mean number of partnerships reported by men and women by each partnership type and in the total number of partners reported. Fisher's exact test, two-tailed, was used to assess differences in the proportion of men and women reporting partnerships of a given type and to assess differences in the proportions of total partnerships accounted for by each partnership type. All analyses were conducted with SPSS 17.02.

Results

Seventy-seven (77%) men and 21 (57%) women reported having at least one casual, trading, or new partner ($p < .05$). Sixty-seven (67%) men and 11 (30%) women had at least one casual partner ($p < .001$). Forty-two (42%) men and 12 (32%) women had a money for drugs partner. Twenty-eight (28%) men and 12 (32%) women had a money for sex partner. Nineteen (19%) men and 5 (14%) women reported a new partner. Fifty-four (54%) men had a main partner as did 27 (73%) of women.

A total of 1,266 sex partners in the 30 days prior to the interview was reported (range, 0–200). The most frequent type of partner was those with whom the participant had traded sex for drugs ($n = 440$, 35%), followed by sex for money partners ($n = 416$, 33%), casual partners ($n = 252$, 20%), and new sex partners ($n = 158$, 12%). Individuals in the sample had an average of 9.2 (SD = 24.9) sex partners in the 30 days prior to the survey compared to 17.4 (SD = 87.6) in the prior 6 months. Thirty-nine participants (29%) reported no partners, 19 (14%) reported one partner, and the remaining 79 (58%) reported two or more partners. Forty-two participants (31%) reported having only one type of partner. The 79 individuals who had multiple partners accounted for 1,247 of the 1,266 total partners (98%) or an average of 15.8 (SD = 31.3) partners in the 30 days prior to the interview.

As shown in Table 1, men accounted for 60% of the reported partnerships. Twenty-nine percent of the men's partners had traded drugs with the participant for sex, compared to 44% of women's partners. Similarly, 28% of the men's partners had paid them money for sex compared to 40% of the women's partners. Men had a casual relationship with 27% of their total partners compared to 9% for the women. Sixteen percent and 7% of the

Table 1 Total, mean, and range of respondent's sex partners by type in the last 30 days

	Male respondent's partners				Female respondent's partners			
	Total (%)	*M*	*SD*	Range	Total (%)	*M*	*SD*	Range
Total sex partners	755	7.55	18.17	0–169	511	13.81	37.44	0–200
Men	539	5.39	17.93	0–169	468	12.65	35.64	0–200
Women	216	2.16	4.36	0–22	43	1.16	4.69	0–28
Total drug trade partners	218 (29)*	2.18**	3.83	0–20	222 (44)	6.00	17.50	0–100
Men	123	1.23**	3.01	0–20	207	5.59	17.08	0–100
Women	95	.95	2.58	0–20	15	.41	1.57	0–9
Total money trade partners	211 (28)*	2.11	8.30	0–75	205 (40)	5.54	17.41	0–100
Men	191	1.91	8.24	0–75	192	5.19	17.03	0–100
Women	20	.20	.85	0–7	13	.35	1.51	0–9
Total casual sex partners	205 (27)*	2.05	2.89	0–20	47 (9)	1.27	5.09	0–31
Men	119	1.19	1.75	0–10	35	.95	3.63	0–22
Women	86	.86	2.45	0–20	12	.32	1.51	0–9
Total new sex partners	121 (16)*	1.21	7.18	0–71	37 (7)	1.00	4.94	0–30
Men	106	1.06	7.17	0–71	34	.92	4.61	0–28
Women	15	.15	.69	0–5	3	.08	.36	0–2

Note. Male participants ($n = 100$); female participants ($n = 137$)

* $p < .0001$ for percentage comparisons between males and females, Fisher's exact test

** $p < .05$ for mean comparisons between males and females, independent-samples *t*-test (two-tailed, equal variances assumed)

partners were new for male and female participants, respectively. For each of these percentage comparisons, $p < .0001$ using Fisher's exact test. As also shown in the table, women had a significantly higher mean number of total drug trade partners and a higher mean number of male drug trade partners. All other differences in the mean number of partners were non-significant.

With regard to condom use, 18 (22%) of the 81 participants with a main partner had always used a condom with that partner. Of the 78 participants with at least one casual partner, 26 (33%) used a condom every time they had sex. Of those with a new partner, 3 (13%) of 24 had always used condoms. Eight (16%) of 51 participants who traded sex for drugs consistently used condoms as had 10 (26%) of the 38 of participants who traded sex for money.

Discussion

A sample of 100 male and 37 female African-American crack smokers being treated for HIV reported having 1,266 partners in the 30 days prior to the interview and had traded sex for drugs or money with 67% of them. The 79 participants with multiple partners were responsible for 1,247 (98%) of the partnerships. A greater proportion of men than women reported any sex partner, and a greater proportion of men than women reported casual partners. Women had a higher mean number of total sex for drugs partners and a greater mean number of male drugs for sex partners than men. A greater proportion of female participants' partners were trading partners compared to males. Males had a higher proportion of casual and new partners than did females. Men had approximately the same proportion of casual partners as trading partners and a lower proportion of new partners. Women had approximately equal proportions of sex for drugs and sex for money partners and approximately equal, and lower, proportions of casual and new partners. The 79 individuals with multiple partners represent a core group who have the potential of infecting large numbers of other individuals in their social and sexual networks. Each time a person with HIV has an unprotected sexual encounter they place that individual at greater risk for contracting HIV or contracting a new strain of HIV.

The continued need for research and the development of effective interventions to increase condom use in this population is evident. Sexual risk reduction for individuals who are HIV+ include abstinence or lifetime condom use or they risk re-infection with a more virulent strain of HIV and/or transmitting the virus to their sex partners. Previous studies have indicated that less than half of HIV positive drug users consistently use condoms (Avants, Warburton, Hawkins, & Margolin, 2000; Friedman et al., 1999; Higgenbotham et al., 2000; Kwiatkowski & Booth, 1998; Nilsson Schonnesson et al., 2008; Timpson et al., 2003). The results of this study of African American crack users support these findings, with few of the participants reporting consistent condom use. These HIV+ crack smokers continued to use drugs, even while on antiretroviral medication, thus increasing their likelihood of unsafe sex. As greater numbers of heterosexual minority drug users are diagnosed with HIV disease, secondary prevention of HIV infection will become increasingly important.

Limitations

The cross-sectional design, small sample size, and use of a convenience sample limits the generalizability and direction of the relationship between condom use and its predictors. However, it does give us an indication of factors that may be related to inconsistent condom use in this difficult to reach population. One of the more alarming findings of this study concerns the large number of individuals who know they are HIV+ yet continue to have unprotected sex. This suggests a strong need to continue research and the development of effective and long lasting sexual risk reduction interventions with sexually active HIV+ crack cocaine smokers. The shifting demographics of the HIV pandemic to low-income heterosexual minority populations make it imperative to reduce the risk of HIV transmission to sexual partners and to reduce the risk of re-infection in HIV+ persons.

Acknowledgements Support for this research was provided by a grant from the National Institute on Drug Abuse, RO3 DA12328, Dr. Mark Williams, Principal Investigator.

References

Avants, S., Warburton, L., Hawkins, K., & Margolin, A. (2000). Continuation of high-risk behavior by HIV-positive drug users: Treatment implications. *Journal of Substance Abuse Treatment, 19*, 15–22.

Centers for Disease Control and Prevention. (2009). *HIV/AIDS surveillance report, 2007* (Vol. 19). Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. http://www.cdc.gov/hiv/topics/surveillance/resources/reports/.

Cottler, L., Compton, W., Ben Abdallah, A., Cunningham-Williams, R., Abram, F., Fichtenbaum, C., et al. (1998). Peer-delivered interventions to reduce HIV risk behaviors among out-of-treatment drug abusers. *Public Health Reports, 113*, 31–41.

Des Jarlais, D., Friedman, S., Perlis, T., Chapman, T., Sotheran, J., Paone, D., et al. (1999). Risk behavior and HIV infection among drug injectors in the era of AIDS in New York City. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology, 20*, 67–72.

Des Jarlais, D., Marmor, M., Friedmann, P., Titus, S., Aviles, E., Deren, S., et al. (2000). HIV incidence among injection drug users in New York City, 1992–1997: Evidence for a declining epidemic. *American Journal of Public Health, 90*, 352–359.

Friedman, S., Chapman, T., Perlis, T., Rockwell, R., Paone, D., Sotheran, J., et al. (1999). Similarities and differences by race/ethnicity in changes of HIV seroprevalence and related behaviors among drug injectors in New York City, 1991–1996. *Journal of Acquired Immune Deficiency Syndromes, 22*, 83–91.

Friedman, S., Jose, B., Neaigus, A., Goldstein, M., Curtis, R., Ildefonso, G., et al. (1994). Consistent condom use in relationships between seropositive injecting drug users and sex partners who do not inject drugs. *AIDS, 8*, 357–361.

Hershberger, S., Wood, M., & Fisher, D. (2003). A cognitive-behavioral intervention to reduce HIV risk behaviors in crack and injection drug users. *AIDS and Behavior, 7*, 229–243.

Higgenbotham, S., Holmes, R., Stone, H., Beil, J., Datu, G., Costa, S., et al. (2000). Adoption of protective behaviors among persons with recent HIV infection and diagnosis—Alabama, New Jersey, and Tennessee, 1997–1998. *MMWR Morbidity and Mortality Weekly Report, 49*, 512–515.

Kotranski, L., Semaan, S., Collier, K., Lauby, J., Halbert, J., & Feighan, K. (1998). Effectiveness of an HIV risk reduction counseling intervention for out-of-treatment drug users. *AIDS Education and Prevention, 10*, 19–33.

Kwiatkowski, C., & Booth, R. (1998). HIV-seropositive drug users and unprotected sex. *AIDS and Behavior, 2*, 151–159.

Kwiatkowski, D., Stober, D., Booth, R., & Zhang, Y. (1999). Predictors of increased condom use following HIV intervention with heterosexually active drug users. *Drug and Alcohol Dependence, 54*, 57–62.

McCoy, C., & Inciardi, J. (1995). *Sex, drugs, and the continuing spread of AIDS*. Los Angeles: Roxbury Publishing Company.

Nilsson Schonnesson, L., Atkinson, J., Williams, M., Bowen, A., Ross, M., & Timpson, S. (2008). A cluster analysis of drug use and sexual HIV risks and their correlates in a sample of African American crack cocaine smokers with HIV infection. *Drug and Alcohol Dependence, 97*, 44–53.

Riehman, F., Wechsberg, W., Zule, W., Lam, W., & Levine, B. (2008). Gender differences in the impact of social support on crack use among African Americans. *Substance Use and Misuse, 43*, 85–104.

Risser, J., Timpson, S., McCurdy, S., Ross, M., & Williams, M. (2006). Psychological correlates of trading sex for money among African American crack cocaine smokers. *American Journal of Drug and Alcohol Abuse, 32*, 645–653.

Ross, M. W., Risser, J., Peters, R. J., & Johnson, R. J. (2006). Cocaine use and syphilis trends: Findings from the Arrestee Drug Abuse Monitoring (ADAM) program and syphilis epidemiology in Houston. *American Journal on Addictions, 15*, 473–477.

Ross, M., Timpson, S., Williams, M., & Bowen, A. (2003). Situational correlates of condom use in a sample of African-American drug users who are primarily crack cocaine users. *AIDS and Behavior, 7*, 55–60.

Ross, M., Timpson, S., Williams, M., & Bowen, A. (2007). The impact of HIV-related interventions on HIV risk behavior in a community sample of African American crack cocaine users. *AIDS Care, 19*, 608–616.

Simpson, D., Camacho, L., Vogtsberger, K., Williams, M., Stephens, R., Jones, A., et al. (1994). Reducing AIDS risks through community outreach for drug injectors. *Psychology of Addictive Behavior, 8*, 86–101.

Simpson, D., Knight, K., & Ray, S. (1993). Psychosocial correlates of AIDS-risk drug use and sexual behaviors. *AIDS Education and Prevention, 5*, 121–130.

Smith, D., Gwinn, M., Selik, R., Miller, K., Dean-Gaitor, H., Ma'at, P., et al. (2000). HIV/AIDS among African Americans: Progress or progression. *AIDS, 14*, 1237–1248.

Sterk, C., Theall, K., & Elifson, K. (2003). Effectiveness of a risk reduction intervention among African American women who use crack cocaine. *AIDS Education and Prevention, 15*, 15–32.

Timpson, S., Williams, M., Bowen, A., & Keel, B. (2003). Condom use behaviors in HIV-infected African Americans crack cocaine users. *Substance Abuse, 24*, 211–220.

Timpson, S., Williams, M., Ross, M., & Keel, B. (2005). African-American crack users' attitudes and beliefs about male and female condoms. *Journal of Substance Use, 10*, 207–213.

Watters, J., Needle, R., Brown, B., Weatherby, N., Booth, R., & Williams, M. (1992). The self-reporting of cocaine use. *Journal of the American Medical Association, 268*, 2374–2377.

Weatherby, N., Needle, R., Cesari, H., Booth, R., McCoy, C., Watters, J., et al. (1994). Validity of self-reported drug use among injection drug users and crack cocaine users recruited through street outreach. *Evaluation and Program Planning, 17*, 347–355.

Williams, M., Bowen, A., Elwood, W., McCoy, C., McCoy, V., Freeman, R., et al. (2000). An elicitation of the determinants of

condom use among African Americans who smoke crack cocaine. *Culture, Health, & Sexuality, 2*, 15–32.

Williams, C., & Latkin, C. (2004). The role of depressive symptoms in predicting sex with multiple and high-risk partners. *Journal of Acquired Immune Deficiency Syndrome, 38*, 69–73.

Williams, M., Ross, M., Bowen, A., Timpson, S., McCoy, H., Keel, K., et al. (2001). An investigation of condom use by frequency of sex. *Sexually Transmitted Infections, 77*, 433–435.

Arch Sex Behav (2010) 39:1359–1374
DOI 10.1007/s10508-010-9600-9

ORIGINAL PAPER

Sexual Self-Schemas of Female Child Sexual Abuse Survivors: Relationships with Risky Sexual Behavior and Sexual Assault in Adolescence

Ashley F. Niehaus · Joan Jackson · Stephanie Davies

Received: 12 December 2008 / Revised: 18 January 2010 / Accepted: 21 January 2010 / Published online: 14 March 2010

Abstract Childhood sexual trauma has been demonstrated to increase survivors' risk for engaging in unrestricted sexual behaviors and experiencing adolescent sexual assault. The current study used the sexual self-schema construct to examine cognitive representations of sexuality that might drive these behavioral patterns. In Study 1 ($N = 774$), we attempted to improve the content validity of the Sexual Self Schema Scale for child sexual abuse (CSA) survivors, introducing a fourth sexual self-schema factor titled the "immoral/irresponsible" factor. In Study 2 ($N = 1150$), the potential differences in sexual self-views, as assessed by the four sexual self-schema factors, between CSA survivors and non-victims were explored. In addition, Study 2 evaluated how these sexual self-schema differences may contribute to participation in unrestricted sexual behaviors and risk for sexual assault in adolescence. Results indicated that a history of CSA impacted the way women viewed themselves as a sexual person on each of the four factors. CSA survivors were found to view themselves as more open and possessing more immoral/irresponsible cognitions about sexuality as compared to women who did not have a CSA history. In addition, the CSA survivors endorsed less embarrassment and passionate/romantic views of their sexual selves. The interaction of CSA severity and the sexual self-schemas explained variance in adolescent sexual assault experiences above and beyond the severity of CSA history and participation in risky sexual behaviors. The findings suggest that sexual self-views may serve to moderate the relationship between CSA and adolescent sexual assault. Implications of these findings and directions for future research are discussed.

Keywords Child sexual abuse · Sexual self-schema · Sexual risk · Adolescent sexual assault

A. F. Niehaus (✉)
Psychology Service (116B), Boston VA Healthcare System,
150 S. Huntington Avenue, Boston, MA 02130, USA
e-mail: ashley.niehaus@va.gov

J. Jackson
Department of Psychology, University of Georgia,
Athens, GA, USA

S. Davies
Ottawa, ON, Canada

Introduction

A growing body of research demonstrates that child sexual abuse (CSA) interferes with the development of sexuality (Cole & Putnam, 1992; Noll, Trickett, & Putnam, 2003; Polusny & Follette, 1995). A CSA history is a risk factor for difficulties with sexual functioning and engaging in unrestricted sexual behavior (Schloredt & Heiman, 2003). One way researchers have attempted to explain the sexual behavior differences observed between CSA survivors and nonvictims is by developing a better understanding of cognitive representations of sexuality. While many sexuality researchers have studied sexuality using attitudinal, behavioral, or physiological methodologies, cognitive representations of sexuality have only been explored in a limited fashion (Andersen & Cyranowski, 1994). Although there are multiple ways to understand cognitive sexuality, we have chosen to use the sexual self-schema framework because of the apparent usefulness of schema theory for understanding the impact of CSA.

Sexual self-schemas have been defined as generalized cognitive views about oneself as a sexual person (Andersen & Cyranowski, 1994). One's sexual self-schema is posited to be a representation of one's past experiences with sexuality and relationships, contributed to by the thoughts, emotions, and behaviors of the past events. As the cumulative record of one's sexual history, one's sexual self-schema is the underlying mechanism

that drives one's current and future decisions, judgments, and behaviors in the sexual domain (Andersen & Cyranowski, 1994). Due to the largely interpersonal nature of sexuality, these sexual self views would be expected to impact one's sexual behaviors as well as interpersonal processes. In their original research on the sexual self-schema construct in women, Andersen and Cyranowski (1994) identified three sexual self-schema factors: open/direct, romantic/passionate, and embarrassed/conservative.

The theoretical basis of sexual self-schemas posits that these schemas are based upon the individual's sexual history. Consequently, it would seem that traumatic childhood experiences involving sexuality would significantly impact sexual self-schemas, and the imprint of these experiences on CSA survivors' cognitive representations of sexuality might explain sexual behavior differences observed between CSA survivors and nonvictims. However, past studies examining these ideas have yielded variable results. Despite their predictions that CSA experiences would predict more negative sexual self-schemas, Andersen and Cyranowski (1994) failed to find a significant relationship between CSA history and sexual self-schemas in adulthood when CSA was defined in a more restricted manner, encompassing only exposure and sexual touching. More intrusive activities, such as childhood rape, were excluded from the CSA definition. When more intrusive CSA experiences were included in the definition in a subsequent study, a significant association was found between presence of a CSA history and less positive sexual self-schemas; that is, lower scores on the open/direct and romantic/passionate factors (Reissing, Binik, Khalife, Cohen, & Amsel, 2003).

Meston, Rellini, and Heiman (2006) further explored the association between sexual self-schemas and CSA history, including a broader range of abuse events in their assessment of CSA. Meston et al. found that a CSA history negatively impacted the romantic/passionate factor, and then investigated how this relationship might explain sexual dysfunction experienced by survivors of sexual abuse. Their results indicated lower levels on the romantic/passionate schema factor predicted higher negative sexual affect during sexual arousal. This relationship was independent of the effects of several other variables which were also included in the regression model, including anxiety and depressive symptoms.

As the theory predicts that one's sexual self-schema would shape current and future sexual behaviors, it is notable that these previous studies examining the sexual self-schemas of CSA survivors have only found CSA history to impact the positive schema factors but have found no significant effect on the negative sexual schema factor—the embarrassed/conservative factor (Andersen & Cyranowski, 1994). It seems that some of the negatively valenced views of sexuality endorsed by CSA survivors in the literature more broadly were not captured by the sexual self-schema scale. Prior studies have found support for negative sexual self-perceptions among CSA survivors that encompass qualities such as guilt regarding sexual activities and feelings (Walser & Kern, 1996), perceptions of self as promiscuous (Fromuth, 1986), views of sexuality in hostile terms (Schloredt & Heiman, 2003), and feeling as if sex is bad and dirty due to the abuse events and messages received from the perpetrator (Maltz, 1991). In addition, negative sexual self-perceptions have been associated with physical and psychological abuse as well as sexual coercion in adult intimate relationships (Offman & Matheson, 2004). The discrepancies between these findings and those of previous studies examining the sexual self-schemas of CSA survivors call into question the content validity of the Sexual Self-Schema Scale for CSA survivors.

Moreover, the details of the cognitive processes that may be driving the unrestricted sexual behavior consistently observed, particularly in young adult CSA survivors, are not well understood. Young women with a history of CSA have been shown to have significantly higher rates of younger age at first consensual intercourse (Miller, Monsoon, & Norton, 1995; Noll et al., 2003; Siegel & Williams, 2003), a higher number of consensual sexual partners (Fergusson, Horwood, & Lynskey, 1997), less frequent use of contraception (Noll et al., 2003), greater likelihood of engaging in HIV-risk behaviors (Elze, Auslander, McMillen, Edmond, & Thompson, 2001), more teenage pregnancies (Fergusson et al., 1997), a higher endorsement of engaging in risky sexual behaviors including extramarital affairs, contracting sexually transmitted infections, engaging in sexual intercourse on the first date (Walser & Kern, 1996), and finally, adolescent sexual assault (Messman-Moore & Long, 2003). Furthermore, CSA history has been found to be associated with increased alcohol misuse (Wilsnack, Vogeltanz, Klassen, & Harris, 1997), a behavioral pattern that has been consistently related to adolescent and adult sexual assault risk as well as increased engagement in unrestricted sexual behavior (Siegel & Williams, 2003). Overall, CSA survivors have been found to be at two to three times higher risk to experience adult sexual assault than women without a CSA history (Marx, Heidt, & Gold, 2005). Developing a more comprehensive understanding of the cognitive representations of sexuality for CSA survivors is an important area for further investigation as it has the potential to identify modifiable targets for prevention.

The current study attempted to improve on the content validity of the sexual self-schema construct for CSA survivors by introducing a sexual self-schema factor that captured the negatively valenced self-assessments described in the literature reviewed above. We then examined whether sexual self-schemas as assessed by the Modified Sexual Self-Schema Scale (SSSS) differed between women with and without a CSA history. In addition, the ability of sexual self-schemas to explain the processes by which early sexual abuse experiences become linked to adolescent sexual trauma experiences was tested. Specifically, we hypothesized that women with a history of CSA would endorse more negative and less positive sexual self-schemas than women who did not endorse a CSA history. In addition, we hypothesized that these overall more negative sex-

ual self-schemas would interact with the severity of participant's abuse history to increase adolescent sexual assault. Role of sexual self-schemas in explaining adolescent risky behavior was also addressed. This study extends past research of this nature by investigating whether differences in sexual self-schemas describe the experience of adolescent sexual assault of women with a history of CSA above and beyond the early sexual trauma as well as participation in risky sexual behaviors, including number of consensual sexual partners and alcohol usage.

Study 1

Method

Participants

A total of 774 college women were recruited from introductory-level undergraduate psychology courses and participated in the various stages of Study 1. Participants were recruited for three separate phases of Study 1. The majority identified themselves as Caucasian (87.3%), Protestant (61.5%), and never married (97.1%). The mean age of the participants was 18.9 years. They met with the investigator in an auditorium in group sessions. Participants were administered packets of questionnaires to complete anonymously following the completion of informed consent forms. Approval for this study was obtained by the university's Institutional Review Board.

Measures

Background Information Basic demographic information was assessed, including age, marital status, and ethnicity. Information regarding participant's child sexual abuse histories was also gathered using an abbreviated version of the Life Experiences Questionnaires (Jackson, Calhoun, Amick, & Maddever, 1990). Events assessed ranged from exposure and fondling to completed vaginal or anal intercourse, and participants responded to the items dichotomously (yes = 1 or no = 0). For the purposes of Study 1, CSA was defined as endorsement of at least one form of sexual contact that occurred prior to age 18 by someone at least five years older than the participant, or of any age if the contact was not consented to or involved coercion.

Sexual Self-Schema Scale (Andersen & Cyranowski, 1994) This scale consists of 50 trait adjectives (26 scored and 24 fillers) for which participants were asked to rate the degree to which each word described themselves on a 7-point scale ranging from 0 (not at all descriptive of me) to 6 (very much descriptive of me). A major strength of the scale is that participants do not report awareness that a sexual construct is being measured. The scale's total score was obtained by subtracting the negative factor score (the embarrassed/conservative factor) from the sum of the two positive factors (the open/direct and romantic/passionate factors). The embarrassed/conservative factor consisted of items such as "self-conscious" and "cautious." Example items from the open/direct and romantic/passionate factors include "frank" and "revealing" and "loving" and "arousable," respectively. The scale demonstrated high test–retest reliabilities at two and nine weeks ($r = .89$ and .88, respectively) and good internal consistency (alpha = .82). Analysis of response bias showed that the correlations of the scale with measures of social desirability and with positive and negative affectivity were non-significant, $r = .11$ with social desirability and .26 with positive affect and −.13 with negative affect. Convergent validity was demonstrated by the significant and positive correlations of the total score with measures of attitudes towards sex ($r = .26$), of sexual behaviors ($r = .30$), of sexual arousability ($r = .25$), and of romantic relationships ($r = .32$), and significant and negative correlations with sex guilt ($r = -.16$). In addition, the expected correlations between each of these measures and the three factors were found. The scale was not found to be associated with self-esteem or extroversion and thus was found to demonstrate good discriminant validity (Andersen & Cyranowski, 1994).

Positive and Negative Affectivity Scale (Watson, Clark, & Tellegen, 1988) This scale consists of 10 words that measure pure markers of positive affect or mood and 10 that measure negative affect. Participants were asked to rate the extent they were experiencing the different emotions at the current moment on a 5-point Likert scale ranging from 1 (not at all) to 5 (extremely), and the scale demonstrated significant stability over every time frame. The internal consistency was .89 for positive affectivity and .85 for negative affectivity.

Rosenberg Self-Esteem Scale (Rosenberg, 1965) This 10-item self-report measure assesses self-esteem as a positive or negative attitude toward the self using a five point Likert scale with item scores added. The scale items show good face validity, and self-reported self-esteem was found to be correlated with depression scores assigned by others, indicating good predictive validity.

Sexual Arousability Inventory (Hoon, Hoon, & Wincze, 1976) This 14-item self-report questionnaire assesses the self-perceived ability of a woman to become sexually aroused across a variety of experiences. Participants indicated on a scale of −1 (adversely affects arousal) to 5 (always causes sexual arousal) whether or not particular activities were sexually arousing. Activities include "when a loved one kisses you passionately," "when a loved one undresses you," and "when you make love in a new or unusual place." Alpha coefficients of .91 and .92 for two samples and Spearman-Brown split half coefficients of .92 for both samples indicated excellent internal consistency. The test–retest

reliability coefficient over eight weeks was .69. Data from female college students illustrated the scales concurrent *validity* with respect to *sexual* experience, activity, and satisfaction, and discriminates between clinical and normal populations (Hoon et al., 1976). More recently, women with congenital adrenal hyperplasia were found to have lower sexual arousability on the SAI when compared to healthy women (Zucker et al., 2004).

Sexual Desire Inventory (Spector, Carey, & Steinberg, 1996) This 13-item self-report instrument is designed to measure level of sexual desire or yearning for sexual activity. It evaluates two dimensions of sexual desire: solitary and dyadic desire. Participants responded to items using a 7 or 8 point Likert scale, indicating their strength of sexual desire (when you have sexual thoughts, how strong is your desire to engage in sexual behavior with a partner) or the frequency they have experienced a wish for sexual activity (during the past month, how often would you have like to behave sexually by yourself?). Reliability of the measure was assessed by determining internal consistency which ranged from .86 to .96, and test–retest reliability at one month was $r = .76$.

Index of Sexual Satisfaction (Hudson, Harrison, & Crosscup, 1981) On this 25-item scale, participants were asked to rate their degree of sexual satisfaction in their sexual relationship with their partner by applying a 5-point Likert scale, ranging from 1 (rarely or none of the time) to 5 (most or all of the time). Example items include, "I feel that my partner enjoys our sex life" and "I feel that my sex life is boring." The higher the score, the more sexual dissatisfaction the participant is experiencing. Prior to responding to the scale, participants were asked to choose from five items to indicate the relationship status perspective from which they were responding to the scale items. Specifically, they indicated if they were basing their answers on a current sexual relationship, basing their answers on a previous sexual relationship, on a previous intimate relationship where intercourse did not occur, on a current intimate relationship where intercourse did not occur, or finally, these questions were not applicable to them. The measure has shown good reliability, with an alpha coefficient of .92 and a test–retest coefficient of .93. The scale correlated significantly with a measure of marital adjustment and of marital satisfaction. The scale was also found to distinguish people with known problems in sexual satisfaction from those known not to have these problems.

Passionate Love Scale (Hatfield & Sprecher, 1986) The 15-item short form of this questionnaire was used in this study. This scale measures "longing for union with another" by asking participants to circle a number between 1 (not at all true) and 9 (definitely true). The short form of the scale has an alpha comparable to the full scale (.91 vs. .94). The scale is uncontaminated by social desirability bias, but is highly correlated with other measures of relationship intimacy and love.

Sexual Experience Scale from the Derogatis Sexual Functioning Inventory (Derogatis & Melisaratos, 1979) This subscale evaluates the variety of participant's sexual experiences repertoire by asking whether they have ever had any of the 24 listed sexual experiences (yes/no) and which ones they have had in the past 30 days. A broad range of sexual experiences were assessed including "kissing on the lips," "masturbating alone," "breast petting," and multiple positions of coitus (e.g., intercourse, side by side). This subtest of the DSFI shows very high internal consistency (Cronbach's alpha = .97) and two-week test–retest reliability ($r = .92$).

Marlowe–Crowne Social Desirability Scale (Crowne & Marlowe, 1960) This 33-item scale is designed to measure participants' response style and need for approval by answering items in socially appropriate ways. Example scale items include, "I have never intensely disliked anyone," and "I'm always willing to admit it when I make a mistake." Participants respond to the scale items using "true" or "false." The internal consistency coefficient for this scale was .88.

Procedure

Scale Construction of the Modified Sexual Self-Schema Scale An additional sexual self-schema factor was developed to improve the content validity of the original Sexual Self-Schema Scale (Andersen & Cyranowski, 1994) by broadening the domain of adjectives that might capture the impact of childhood sexual trauma. An expert panel of 17 clinical psychology faculty and graduate students with expertise in child sexual abuse was consulted and assisted in the generation of adjectives to describe a "sexual woman." This process resulted in a 60-item list of descriptors. This measure was administered to 142 participants who were asked to rate how much each word was descriptive of a "sexual woman" on a 7-point Likert scale. Participants were also asked to suggest words they believed were missing from the list in addition to completing a brief self-report assessment of their child sexual abuse history. This step of the procedure resulted in eight new adjectives being suggested by participants with CSA histories as well as five descriptors with severely restricted ranges being eliminated from the measure. In addition, the item "prudent" from the original Sexual Self-Schema Scale was replaced by its synonym "sensible" due to frequent reports it was misunderstood by participants.

To ensure that the adjectives included in the final version of the measure were distinct from self-esteem, social desirability, and positive and negative affectivity, a new group of participants comprised of 256 women completed the Modified Sexual Self-Schema Scale, the Rosenberg Self-Esteem Scale (Rosenberg, 1965), the Positive and Negative Affectivity Schedule (Watson et al., 1988), and the Marlowe–Crowne Social Desirability Scale (Crowne & Marlowe, 1960). The items on the

revised scale were eliminated if their correlations with these measures exceeded .30. This procedure yielded a list of 36 words.

A common factor analysis with an oblique rotation of the 36 words was conducted to assess the internal factor structure of the scale. Four factors were identified by examining the eigenvalue scree plot and factor interpretability, including the original sexual self-schema factors delineated by Andersen and Cyranowski (1994), in addition to a new 9-item factor labeled, "Immoral–Irresponsible." The scale also included 23 filler adjectives. The rotated factor pattern with loadings for each of the 36 items is shown in Table 1. The open/direct factor explained 7% of the total variance in the scale, while the passionate–romantic, immoral–irresponsible, and embarrassed–conservative factors explained 8%, 6%, and 7%, respectively. The possible range for each sexual self-schema scale factor was as follows: 0–48 for the open–direct factor, 0–66 for the embarrassed–conservative factor, 0–48 for the passionate–romantic factor, and 0–54 for the immoral–irresponsible factor.

Reliability and Validity of the Modified SSSS A total of 376 women completed the Modified SSSS in addition to the Rosenberg Self-Esteem Scale (Rosenberg, 1965), the Positive and Negative Affectivity Schedule (Watson et al., 1988), the Marlowe–Crowne Social Desirability Scale (Crowne & Marlowe, 1960), the Sexual Arousability Inventory (Hoon et al., 1976), the Sexual Desire Inventory (Spector et al., 1996), the Index of Sexual Satisfaction (Hudson et al., 1981), the Passionate Love Scale (Hatfield & Sprecher, 1986), the Sexual Experiences Subscale for the Derogatis Sexual Functioning Inventory (Derogatis & Melisaratos, 1979), and an instrument which solicited demographics and briefly assessed childhood and adolescent sexual abuse experiences. To evaluate the test–retest reliability of the measure, 101 participants completed the Modified SSSS after a two-week period.

Results

In regards to reliability, the Modified Sexual Self-Schema Scale demonstrated good internal consistency (Cronbach's alphas ranged from .54 to .83) as well as good two-week test–retest reliability (*r*s ranged from .76 to .86). Convergent validity was shown by correlating each sexual self-schema factor with sexuality and relationship measures. The new factor, immoral–irresponsible, showed significant positive correlations with lifetime sexual experiences and current sexual desire level. Consistent with the findings of Andersen and Cyranowski (1994), the original subscales correlated with all the measures of sexuality (see Table 2). The same procedure Andersen and Cyranowski utilized to demonstrate the discriminant validity of the original scale was used to determine the discriminant validity of the modified scale. The modified scale's ability to predict incremental variance in sexual arousability and lifetime sexual experiences beyond that explained by self-esteem, social desirability, and positive and negative affectivity was assessed, and results indicated that all four factors accounted for significant increments in explained variance in the prediction of lifetime sexual experiences (see Table 3). Finally, the incremental validity of the immoral–irresponsible factor was tested with a series of hierarchical regression analyses to determine if it added to the ability of the scale to predict sexual and relationship measures beyond the original three factors. The new factor accounted for a significant amount of variance in predicting dyadic, solitary, and total sexual desire, in addition to current and lifetime sexual behaviors, demonstrating the unique added contribution of the new factor. These results are shown in Table 4.

Discussion

The new sexual self-schema factor, labeled immoral–irresponsible, developed in the current study assessed if women viewed their sexual self as "bad," "loose," "immoral," or "irresponsible." These adjectives are consistent with the negatively valenced views CSA survivors have used to describe their sexual selves, and it is posited that this factor improves the content validity of the Sexual Self-Schema Scale for CSA survivors. This fourth sexual self-schema factor did not detract from the validity of the original scale, and offered incremental validity in predicting aspects of sexual experience and desire. Consequently, it seems that the additional factor augmented the scale's comprehensiveness in measuring components of women's sexual self-schemas.

Study 2 was designed with several aims in mind. First, we hypothesized that women who endorsed a CSA history would exhibit higher levels of the negative sexual self-schema factors (immoral/irresponsible and embarrassed/conservative) and lower levels of the positive sexual self-schema factors (open/direct and passionate/romantic). In addition, we proposed that sexual self-schemas would explain subsequent sexual revictimization in adolescence, both directly and via engagement in risky sexual behaviors (number of consensual partners and alcohol usage).

Study 2

Method

Participants

Participants in the study were 1,150 female undergraduate students recruited through the research pool at a large southeastern university. The mean age for participants was 19.4 years ($SD = 1.64$, range = 18–41). Most of the women (96.1%) had never been married. However, 55.4% reported they were currently involved in a romantic relationship. Racial composition of the

Table 1 Factor loadings of the Modified Sexual Self-Schema Scale

Item	Factor			
	III Embarrassed–conservative	I Passionate–romantic	IV Immoral–irresponsible	II Open–direct
Timid	**−0.66**	0.07	0.04	−0.16
Embarrassed	**−0.65**	0.00	0.25	−0.16
Self-conscious	**−0.59**	0.15	0.17	0.05
Cautious	**−0.51**	0.06	−0.19	0.04
Stimulating	**0.49**	0.19	0.11	0.31
Conservative	**−0.45**	0.06	−0.06	−0.08
Risk-taking	**0.44**	0.05	0.23	0.20
Uninhibited	**0.41**	0.18	0.17	0.11
Vulnerable	**−0.37**	0.25	0.33	−0.06
Naïve	**−0.30**	0.04	0.14	−0.16
Inexperienced	**−0.26**	−0.11	−0.14	−0.14
Romantic	0.08	**0.73**	−0.05	−0.10
Unromantic	−0.23	**−0.66**	0.20	0.14
Passionate	0.15	**0.65**	0.01	0.11
Warm	−0.07	**0.61**	−0.19	−0.04
Sympathetic	−0.23	**0.60**	0.03	0.12
Loving	−0.06	**0.57**	0.04	0.02
Feeling	−0.13	**0.49**	−0.11	0.20
Arousable	0.11	**0.33**	0.15	0.17
Bad	−0.19	−0.04	**0.63**	0.07
Irresponsible	0.06	−0.09	**0.57**	0.03
Immoral	−0.11	0.02	**0.53**	0.10
Sensible	−0.25	0.02	**−0.46**	0.39
Dependent	−0.22	0.31	**0.44**	−0.15
Self controlled	−0.16	0.11	**−0.44**	0.23
Dirty	−0.01	−0.12	**0.42**	0.09
Loose	0.10	−0.08	**0.39**	0.06
Experienced	0.31	0.23	**0.34**	0.24
Straightforward	0.06	−0.01	−0.01	**0.80**
Frank	0.03	−0.19	0.07	**0.79**
Direct	0.16	0.02	0.08	**0.73**
Broad minded	0.22	0.12	0.02	**0.45**
Casual	0.01	0.05	−0.01	**0.45**
Open minded	0.14	0.13	0.03	**0.40**
Outspoken	0.33	0.11	−0.03	**0.39**
Revealing	0.14	0.11	0.19	**0.24**

Factor identification numbers are based on Andersen and Cyranowski's (1994) factor numbers with Factor IV indicating the new factor identified in the present study

participants was 74.5% Caucasian, 6.8% African American, 3.7% Asian, and the remaining 11.2% endorsed Native American, Pacific Islander, Hawaiian, or Hispanic racial backgrounds. In regards to religious affiliation, 52.5% of the participants endorsed being Protestant, 14.1% Catholic, and 13.7% nonaffiliated. Of the participants, 97.7% reported being exclusively or mostly heterosexual, and 2.3% reported being mostly or exclusively homosexual. The study was approved by the university's Institutional Review Board.

Procedure

Participants received partial fulfillment of a departmental research requirement when they volunteered for a study about beliefs and attitudes regarding relationships. The study involved a 2-h group session that was conducted in an auditorium in a university setting. Groups of 50–200 participants completed an anonymous and confidential questionnaire that gathered information about demographics, early sexual trauma experiences, cognitions about sexuality, risky sexual behaviors, and adolescent sexual assault experiences. After obtaining informed consent, guaranteeing anonymity, and emphasizing the voluntary nature of the research participation, the surveys were administered in fixed order. When each participant finished completing the survey packet, she received a written debriefing form and was encouraged to contact the researcher with any questions that arose after survey completion. No participants chose to withdraw or hand in a blank questionnaire.

Measures

Child Sexual Abuse History A modified version of the Life Experiences Questionnaire (LEQ) (Jackson et al., 1990) was used to assess CSA history. The questions solicited information regarding demographics, family relationships, and 12 unwanted sexual experiences. Detailed information about childhood sexual abuse characteristics included the frequency, duration, and nature of the abuse activities, age of the victim and perpetrator at the time of abuse, relationship between victim and perpetrator, whether the victim disclosed the abuse event to anyone, and finally, emotional reactions to the disclosure experience.

For the purposes of the present study, childhood sexual abuse was defined as either contact or non-contact abuse occurring prior to the age 14 and was experienced as coercive and forced. This definition was based on the definition of CSA established by Wyatt, Lawrence, Vodounon, and Mickey (1992). However, the current study's definition of abuse included individuals who experienced abuse events perpetrated by someone more than five years older than the victim as well as events perpetrated by someone of similar age. This decision was based on the recommendation of Roodman and Clum (2001) who found that limiting the definition of CSA to events perpetrated by someone more than five years older than the victim was unnecessarily restrictive, particularly when attempting to identify those CSA survivors at risk for revictimization. Twelve abuse events were assessed by the LEQ and, for each event, participants circled a

Table 2 Convergent validity of the Modified Sexual Self-Schema Scale: Pearson product-moment correlations of factor and total sexual self-schema scores with sexuality and relationship measures

Scales	Factor 1 Passionate–romantic	Factor 2 Open–direct	Factor 3 Embarrassed–conservative	Factor 4 Immoral–irresponsible	Total scale
ISS[a]	0.40***	0.19**	−0.26***	−0.09	0.40***
PLS[b]	0.38***	0.11**	−0.12**	0.01	0.25***
SAI[c]	0.40***	0.19***	−0.11**	0.07	0.27***
SDI[d]	0.28***	0.15**	−0.16**	0.18***	0.19***
SES[e] (current)	0.23***	0.11*	−0.14**	0.11*	0.17***
SES (lifetime)	0.26***	0.22***	−0.33***	0.20***	0.30***

* $p < .05$; ** $p < .01$; *** $p < .001$

[a] Index of Sexual Satisfaction

[b] Passionate Love Scale

[c] Sexual Arousability Inventory

[d] Sexual Desire Inventory

[e] Sexual Experiences Subscale of the Derogatis Sexual Functioning Inventory

Table 3 Discriminant validity of the Modified Sexual Self-Schema Scale: percent incremental variance accounted for by factors and total scale, above self-esteem, social desirability, and positive and negative affectivity, in predicting lifetime sexual experiences and sexual arousability

	Lifetime sexual experiences % Incremental variance	Sexual arousability % Incremental variance
Passionate–romantic	3.74***	10.74***
Open–direct	1.99***	1.05*
Embarrassed–conservative	7.42***	0.29
Immoral–irresponsible	3.97***	0.35
Total	4.85***	3.45***

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 4 Percent incremental variance accounted for by immoral–irresponsible factor, above and beyond original factors, in the predictions of related criterion variables

	% Incremental variance
Sexual dissatisfaction	0.62
passionate love	0.04
Sexual arousability	0.65
Solitary sexual desire	1.35*
Dyadic sexual desire	2.76**
Total sexual desire	2.97***
Lifetime sexual experiences	2.47**
Current sexual experiences	1.10**

* $p < .05$; ** $p < .01$; *** $p < .001$

number between 0 (never) and 5 (5 or more times), indicating the number of times they had experienced the particular event. For the analyses in which an abuse severity index was used, the abuse severity score was calculated based on the nature and frequency of the abuse activities participants endorsed. More intrusive CSA experiences were assigned more points, and the point value of a participant's most intrusive experience was multiplied by the frequency the event occurred to arrive at her severity score. Higher numbers indicated more frequent and intrusive abuse activities.

In the total sample of 1150 women, 238 (20.7%) participants were identified as child sexual abuse survivors and 912 (79.3%) were classified as non-victims. Of the participants identified as CSA survivors, 144 (12.5%) reported they had experienced multiple types of abusive activities. The most common CSA activity was exposure of the perpetrator's genitals and 157 (13.7%) women endorsed this experience. The second most commonly reported CSA experience was genital fondling, reported by 131 (11.4%) participants; 33 women (2.9%) reported having experienced forced vaginal intercourse; 94 women (39.5%) were abused by a family member, while the remaining women were abused by acquaintances (12.2%), peers (44.5%), or strangers (3.8%). A total of 71 women (29.8% of the CSA survivors in the sample) reported CSA experiences between the ages of three and five years. The distribution of age of onset of CSA was: 49 women (20.6%) 6–8, 34 women (14.3%) 9–11, and 84 women (35.3%) 12–13. CSA survivors did not differ significantly from non-victims on race, sexual orientation, marital status, family's religious affiliation, current religious affiliation, or current involvement in a romantic relationship.

Sexual Self-Schemas The Modified Sexual Self-Schema Scale was used to assess participants' sexual self-schemas. This scale includes the original sexual self-schema factors (passionate–romantic, open–direct, and embarrassed–conservative) developed by Andersen and Cyranowski in

addition to the immoral–irresponsible factor developed in Study 1 described above.

Unrestricted Sexual Behaviors The Heterosexual Behavior Scale-Short form (HBI) (Bentler, 1968) is a 15-item questionnaire that assesses non-coercive heterosexual experiences. Fifteen sexual behaviors are listed in ascending order, from kissing to vaginal and oral intercourse. Participants were instructed to report the number of partners she had for each of the listed behaviors on a 5-point Likert scale (0, 1–2, 3–5, 6–9, 10 or more). The primary variables of interest in the current study were the number of individuals with whom the participant had engaged in intercourse as well as the variety of consensual sexual activities in which the woman had participated. The number used for the analysis in this study was the total number of consensual sexual partners reported for the vaginal intercourse item.

The Rutgers Alcohol Problem Index (RAPI; White & Labouvie, 1989) is a 23-item self-report measure that assesses adolescent problem drinking. Participants indicated on a 5-point Likert scale, ranging from 0 (none) to 4 (10 times) "How many times has this happened to you while you were drinking or because of your drinking during the last year?" The scale has been demonstrated to have a reliability of .92 and a 3-year stability coefficient of .40.

Adolescent Sexual Assault Experiences The Sexual Experiences Survey (SES) (Koss, 1985) is a 10-item self-report instrument that assesses sexual victimization at or after the age of 14. Adolescent sexual victimization was defined as endorsement of one or more of the unwanted sexual activities. The SES assesses a variety of sexual assault experience ranging from forced sex play to attempted and completed intercourse. Participants answered each item reporting the occurrence of a particular activity since age 14 (yes/no). Participants also indicated the number of times the experience had occurred since age 14 on a 5-point scale ranging from 1 (1 occurrence) to 5 (5 or more occurrences) in addition to the number of occurrences specifically during the last year on the same 5-point scale. Finally, the participants indicated if the event had occurred during the past 3 months (yes/no). Scores were computed as the sum of the responses to the 10 abuse event items, with high scores signifying extensive trauma. The SES has shown one week test–retest reliability of .93 and an internal consistency of .74.

Adolescent sexual assault experiences were reported by 330 (27.9%) women in the sample and 320 (27.8%) women endorsed having experienced multiple types of adolescent sexual assault experiences. In this sample, the most common adolescent unwanted sexual experience was participation in sex play activities due to being overwhelmed by a man's continual arguments and pressure, reported by 248 (21.6%) participants. In addition, 145 (12.6%) participants endorsed having experienced forced sexual intercourse.

Results

The first step in the data analytic plan for Study 2 was to conduct bivariate correlations and independent samples *t* tests to examine the inter-relationships among variables and determine whether participants with a CSA history reported different levels of the sexual self-schema factors than the women who did not endorse a history of CSA. Then, the multiple regression analytic strategy outlined by Baron and Kenny (1986) was used to test the sexual self-schema factors as a mediator in the relationship between severity of CSA and risky sexual behaviors (consensual sexual partners and alcohol usage). According to the guidelines of Baron and Kenny, three criteria must be met to establish a variable as a mediator, and these conditions can be tested using a series of linear regression analyses. First, a significant relationship must be established between the predictor variable (CSA severity) and the potential mediator (sexual self-schema factor). Next, the mediator and the dependent variable (consensual sexual behavior and alcohol usage) must exhibit a significant relationship. Finally, the relationship between the predictor and the dependent variable must be significantly reduced when the mediator is introduced into the regression equation. In order to reduce the probability of committing a Type I error, a Bonferroni-correction procedure was used in both the linear and hierarchical regression analyses. In the linear regression analyses, an alpha level of .0125 was used to determine if a value was significant. Lastly, a set of hierarchical regression analysis were conducted to determine the additive contribution of the sexual self-schema factors, beyond the effects of abuse severity and risky sexual behaviors, in predicting adolescent sexual revictimization. A Bonferroni-correction was also used for the hierarchical regression analyses, and the alpha level was set to .00625.

Bivariate Correlations and t Tests

The bivariate correlations among the variables are shown in Table 5. The relationship of sexual abuse history to sexual self-schema factors was evaluated using planned pairwise comparisons. Separate *t* tests for each factor were conducted to investigate differences between women with and without a history of CSA. Results of *t* tests indicated that there were significant differences among CSA survivors on each of the sexual self-schema factors. Descriptive statistics (*M* and *SD*s) for survivors and non-victims on each of the scales are shown in Table 6.

Relationship Among CSA, Sexual Self-Schemas, and Risky Sexual Behaviors

To this point, the results have indicated that there were group differences between women with and without a CSA history on the sexual self schema factors. Next, the relationships between

Table 5 Pearson correlation coefficients among Study 2 variables

Variable	1	2	3	4	5	6	7	8
1. CSA severity	–	.34**	.28**	−.25**	−.31**	.32**	.26**	.57**
2. Immoral/irresponsible schema		–	.35**	−.04	.10**	.38**	.35**	.33**
3. Open/direct schema			–	.01	−.15**	.27**	.19**	.28**
4. Passionate/romantic schema				–	.27**	−.04	−.41**	−.27**
5. Embarrassed/conservative schema					–	−.26**	−.10**	−.22**
6. Consensual partners						–	.47**	.45**
7. Alcohol usage							–	.36**
8. Adolescent sexual assault								–

** $p < .001$

Table 6 Group differences in sexual self-schemas, risky sexual behaviors, and adolescent sexual assault experiences

Measure	Nonabused women ($n = 766$)		Sexually abused women ($n = 239$)		$t(1003)$	Cohen's d
	M	SD	M	SD		
Sexual self-schemas						
Open/direct	28.77	6.29	31.24	7.08	−5.11***	0.37
Immoral/irresponsible	21.65	4.76	23.36	6.52	−4.41***	0.30
Embarrassment/conservatism	34.63	6.27	32.64	7.29	4.12***	0.29
Romantic/passionate	36.84	5.84	35.85	8.20	2.04**	0.14
Risky sexual behaviors						
Consensual sexual partners	2.02	2.52	3.54	3.62	−7.28***	0.49
Alcohol usage	9.69	12.87	11.90	13.71	−2.22**	0.17
Adolescent sexual assault experiences	0.38	0.78	1.55	2.35	−11.54***	0.67

** $p < .01$; *** $p < .001$. Cohen's $d = \frac{\bar{x}_1 - \bar{x}_2}{s}$

CSA severity and sexual self-schema factors were explored. According to Baron and Kenny (1986), these pairwise relationships between constructs are required as a first step for evaluating a model in which sexual self schema factors mediate the relationship between CSA severity and risky sexual behaviors. Four separate linear regressions were conducted on severity of CSA and each of the sexual self-schema factors, with CSA severity being the predictor and each of the schema factors being the outcome variables. CSA severity significantly predicted 9.2% of the variance in the immoral/irresponsible factor, $F(1, 1097) = 112.72$, $p < .0125$, 7.6% of the variance in the open/direct factor, $F(1, 1097) = 91.20$, $p < .0125$, 9.4% of the variance in the passionate/romantic factor, $F(1, 1097) = 113.38$, $p < .0125$, and 5.9% of the variance in the embarrassed/conservative factor, $F(1, 1097) = 70.34$, $p < .0125$.

Eight hierarchical linear regressions were conducted to predict risky sexual behaviors with each of the sexual self-schema factors serving as a potential mediator of CSA severity. Each set of hierarchical regressions was composed of two steps. In the first step, CSA was the only predictor and, in the second step, the respective sexual self-schema factor was added to the list of predictors (see Table 7). This method allowed assessment of whether the sexual self-schema factor acted as a mediator in the relationship between CSA and number of consensual sexual partners and CSA and alcohol usage. Results from the models assessing the consensual partners outcome variable will be reviewed first.

Mediational Models Predicting Consensual Sexual Partners

The results of the first step in each model indicated that CSA severity was a significant predictor of the variance in number of consensual sexual partners, $t(1, 1091) = 11.44$, $p < .0125$. This step accounted for 10.6% of the variance in consensual sexual partners. In the first model, adding the immoral/irresponsible factor (Step 2) increased the percentage of consensual sexual partners explained to 19.7%. A review of the beta coefficients in the first model revealed that CSA continued to provide a unique contribution to the explanation of this risky sexual behavior after the sexual self-schema factor was added ($B = .23$, $p < .0125$). Thus, although mediation was not supported in the first model, the immoral/irresponsible factor provided a unique explanation of number of consensual sexual partners (immoral/irresponsible factor: $B = .32$, $p < .0125$), suggesting that this sexual self-schema factor may have explained this sexual behavior independently from CSA severity. Results of the second model indicated that adding the open/direct factor to the model (Step 2) increased the percentage of variance in consensual sexual partners explained to 14.4%. A review of the beta coefficients in the

Table 7 Results for the hierarchical regression of the risky sexual behaviors

Outcome	R^2	ΔR^2	B	$SE\,B$	$t(1091)$
Consensual sexual partners					
Step I	0.17*				
CSA severity			0.33	0.05	11.44*
Step II	0.20*	0.09*			
CSA severity			0.23	0.05	7.94*
Immoral/irresponsible schema			0.32	0.07	11.12*
Step II	0.14*	0.04*			
CSA severity			0.27	0.05	9.24*
Open/direct schema			0.20	0.06	6.81*
Step II	0.19	0.01			
CSA severity			0.34	0.05	11.40*
Passionate/romantic schema			0.05	0.06	1.62*
Step II	0.150*	0.04*			
CSA severity			0.28	0.05	9.61*
Embarrassed/conservative schema			−0.21	0.06	−7.43*
Alcohol usage					
Step I	0.07*				
CSA severity			0.26	0.05	8.63*
Step II	0.15*	0.08*			
CSA severity			0.16	0.05	5.30*
Immoral/irresponsible schema			0.31	0.07	10.18*
Step II	0.08*	0.01*			
CSA severity			0.22	0.05	7.16*
Open/direct schema			0.13	0.06	4.02*
Step II	0.07	0.01			
CSA severity			0.24	0.05	7.66*
Passionate/romantic schema			−0.06	0.06	−1.79
Step II	0.07	0.01			
CSA severity			0.25	0.05	7.98*
Embarrassed/conservative schema			−0.05	0.06	−1.76

* $p < .0125$

second model revealed that CSA continued to provide a unique contribution to the explanation of the risky sexual behavior ($B = .27$, $p < .0125$), but the open/direct factor provided unique explanation as well ($B = .20$, $p < .0125$). Thus, the final criterion for mediation was not met.

In the third model in which the romantic/passionate factor was the proposed mediator, CSA continued to provide a unique contribution to the explanation of number of consensual sexual partners ($B = .34$, $p < .0125$), even after the romantic/passionate factor was added to the model (Step 2). The romantic/passionate factor was not a significant predictor ($B = .05$, $p = .105$). In this model, the addition of the sexual self-schema factor did not result in a significant increase in the percentage of the variance in consensual sexual partners explained. In the fourth model predicting consensual sexual partners in adolescence, adding the embarrassed–conservative factor to the model increased the percentage of variance in consensual sexual partners explained to 14.9%. Results of the final model indicated that adding the embarrassed/conservative factor provided a unique explanation of consensual sexual partners when it was added to the model in Step 2 ($B = -.21$, $p < .0125$). A review of the beta coefficients in the second step revealed that CSA continued to provide a unique contribution to the explanation of consensual sexual partners ($B = .28$, $p < .0125$), and consequently, the criteria for mediation were not fully met.

Mediational Models Predicting Alcohol Usage

The results of the first step in each model predicting alcohol usage indicated that CSA severity was a significant predictor of the variance, $t(1, 1091) = 8.63$, $p < .0125$. This step accounted for 6.6% of the variance in alcohol usage. In the first model, adding the immoral/irresponsible factor (Step 2) increased the percentage of alcohol usage explained to 15%. A review of the beta coefficients in the first model revealed that CSA continued to provide a unique contribution to the explanation of alcohol usage after the sexual self-schema factor was added ($B = .16$, $p < .0125$). Although mediation was not met, the immoral/irresponsible factor provided a unique explanation of alcohol usage (immoral/irresponsible factor: $B = .31$, $p < .0125$), suggesting that this sexual self-schema factor may have explained this risky behavior independently from CSA severity. Results of the second model indicated that adding the open/direct factor to the model (Step 2) increased the percentage of alcohol usage explained to 7.9%. A review of the beta coefficients in the second model revealed that CSA continued to provide a unique contribution to the explanation of the risky sexual behavior ($B = .22$, $p < .0125$), but the open/direct factor provided unique explanation as well ($B = .14$, $p < .0125$).

In the third model in which the romantic/passionate factor was the proposed mediator, CSA continued to provide a unique contribution to the explanation of alcohol usage ($B = .24$, $p < .0125$), even after the romantic/passionate factor was added to the model (Step 2). In this model, the romantic/passionate factor was not a significant predictor ($B = -.06$, $p = .074$). Results of the fourth model indicated that adding the embarrassed/conservative factor did not provide a unique explanation of alcohol usage when it was added to model in Step 2 ($B = -.05$, $p = .079$). A review of the beta coefficients in the second step revealed that CSA continued to provide a unique contribution to the explanation of alcohol usage ($B = .25$, $p < .0125$). The addition of the sexual self-schema factors to the models did not increase the percentage of the alcohol usage variance explained in either of the final two models.

Relationships Among CSA, Risky Behaviors, Sexual Self-Schemas, and Adolescent Sexual Assault Experiences

Given that the relationship between CSA, sexual self-schemas, and risky sexual behaviors was established in the previous analyses, we investigated whether the sexual self-schemas of CSA survivors predicted adolescent sexual assault experiences in this population. A four-step hierarchical linear regression was computed to predict adolescent sexual assault experiences. In the first step, the predictor was CSA severity; in the second step, risky sexual behaviors were added to the list of predictors; and in the third step, the predictors included CSA severity, risky sexual behaviors, the immoral/irresponsible schema factor, the open/direct schema factor, the romantic/passionate schema factor, and the embarrassed/conservative schema factor. Finally, in the fourth step, the interaction between each of the schema factors and CSA were added to the predictor list. These terms included the interaction between CSA and the immoral/irresponsible factor (CSA × Immoral), CSA and the open/direct factor (CSA × Open), CSA and the romantic/passionate factor (CSA × Romantic), and CSA and the embarrassed/conservative factor (CSA × Embarrass). The terms were added in the order listed here. Risky sexual behaviors were included in the second step to test whether the variables in subsequent steps added to the explanation of variance in adolescent sexual assault experiences, above and beyond this well-established risk factor. If schema and severity of CSA were linked to adolescent sexual assault experiences independently from risky sexual behaviors, then adding schemas to the model that already contained risky sexual behaviors would show a significant increase in R^2. The interaction term variables were developed to assess the role of sexual self-schemas for women with varying levels of CSA severity.

The first step comprised of CSA severity explained 47% of the variance in adolescent sexual assault experiences, $F(1, 999) = 871.1, p < .00625$, suggesting that CSA severity accounted for a significant amount of the variance in adolescent sexual assault experiences. The second step, which used CSA severity and adolescent risky sexual behaviors as predictors of adolescent sexual assault experiences, explained 52.3% of the variance in adult sexual assault experiences, showing a statistically significant increase in R^2. This step suggests that risky sexual behaviors contributed significantly to the explanation of adolescent sexual assault experiences related to CSA, $F(1, 999) = 363.90$, $p < .00625$. Both increased engagement in consensual sexual activity and higher alcohol usage were positively related to number of adolescent sexual assault experiences.

Next, the third step, which included CSA severity, risky sexual behaviors, and the four sexual self-schema factors explained 53% of the variance in adolescent sexual assault experiences. The immoral–irresponsible and romantic/passionate factors emerged as the most noteworthy schema factors in predicting adolescent sexual assault $F(1,999) = 159.483, p < .001$ in this step, although they were not significant at $p < .00625$. These results indicate that the relationships between these sexual self-schema factors and adolescent sexual assault did not add significant explanation to adolescent sexual assault experiences, above and beyond the more well-established sexual trauma history and participation in risky sexual behaviors.

Finally, the fourth step, which included CSA severity, number of consensual sexual partners, alcohol usage, all four sexual self-schema factors, CSA × Immoral, CSA × Open, CSA × Romantic, and CSA × Embarrass explained 58% of the variance in adolescent sexual assault experiences, which corresponded to a statistically significant increase in R^2 from the model tested in Step 3, $F(1, 999) = 121.56, p < .00625$. Each of the interaction variables provided unique contributions to the model, indicating that there were unique relationships between CSA severity and adolescent sexual assault at varying levels of the sexual self-schema factors. Post hoc analysis of these interactions following the procedures of Aiken and West (1991) revealed that at higher levels of the immoral–irresponsible self schema factor the expected positive relationship between CSA severity and adolescent sexual assault was observed. However, at the lower level of the immoral-responsible self-schema, this relationship was not found. Conversely, more positive relationships between CSA severity and adolescent sexual assault were observed at *lower* levels of the remaining sexual self-schema factors—open–direct, romantic–passionate, and embarrassed–conservative than at the higher levels. Further, at the higher levels of these self-schema factors the relationship between CSA severity and adolescent sexual assault was actually negative (Table 8).

General Discussion

On the basis of the idea that sexual self-schemas are shaped by one's history with sexual experiences, it was hypothesized that significant differences would be observed between the sexual self-schemas of women who had a history of childhood sexual trauma and those who did not. It was also hypothesized that sexual self-schemas impacted by a history of CSA may partially explain the increased engagement in unrestricted sexual behaviors and heightened risk for adolescent sexual assault observed in populations of CSA survivors. While the results of the current study supported the idea that CSA experiences affect cognitive representations of sexuality in significant ways, it seemed that the interaction between these sexual self-schemas and victimization experiences was most useful in explaining why CSA survivors were at heightened risk for adolescent sexual assault experiences.

The current study found support for the validity of a new sexual self-schema factor, the immoral–irresponsible factor, developed in the current study to assess if women viewed their sexual self as "bad," "loose," "immoral," or "irresponsible." A major finding of the current study was that there were significant

Table 8 Results for the hierarchical regression of adolescent sexual assault

Step	R^2	ΔR^2	*B*	*SE B*	*t*(996)
I	0.47				
CSA severity			0.68	0.00	29.51*
II	0.52	0.06*			
CSA severity			0.59	0.00	25.43*
Consensual partners			0.17	0.00	6.55*
Alcohol usage			0.13	0.00	5.13*
III	0.53	0.08			
CSA severity			0.57	0.01	22.06*
Consensual partners			0.16	0.00	7.11*
Alcohol usage			0.12	0.00	0.55
Immoral/irresponsible schema			0.06	0.01	2.16
Open/direct schema			−0.03	0.01	−1.17
Passionate/romantic schema			−0.06	0.01	−2.50
Embarrassed/conservative schema			−0.03	0.01	−1.20
IV	0.58	0.05*			
CSA severity			0.42	0.01	10.39*
Consensual partners			0.14	0.00	5.31*
Alcohol usage			0.13	0.00	5.15*
Immoral/irresponsible schema			0.02	0.01	0.64
Open/direct schema			−0.02	0.01	−0.88
Passionate/romantic schema			−0.01	0.01	−0.37
Embarrassed/conservative schema			−0.02	0.01	−0.59
CSA × Immoral			0.31	0.00	2.84*
CSA × Open			−0.48	0.00	−4.54*
CSA × Romantic			−0.25	0.00	−2.58*
CSA × Embarrass			−0.18	0.00	−2.14*

* $p < .00625$

differences between CSA survivors and nonvictims on each of the sexual self-schema factors, including the immoral–irresponsible factor. This was the first study to demonstrate that CSA survivors endorse the immoral/irresponsible factor that explores cognitive representations of sexuality as immoral, irresponsible, and bad significantly more than nonvictims. CSA survivors also endorsed higher levels of the open/direct factor, indicating CSA survivors were more open, direct, revealing, and straightforward about sexuality. Finally, CSA survivors endorsed the embarrassed/conservative and passionate/romantic factors significantly less than nonvictims, indicating less self-consciousness, caution, love, and passion in thinking about sexuality. Overall, CSA survivors appeared less inhibited and more likely to approach rather than avoid sexuality and sexual situations when compared to nonvictims, but they also appeared to experience a sense that they were bad and dirty rather than loving and passionate when experiencing sexual feelings and engaging in sexual behaviors.

These findings were consistent with the literature showing that CSA survivors report more negative attitudes about sex (Orr & Downs, 1985), including experiencing more sex guilt (Walser & Kerns, 1996), considering sex as a hostile behavior (Scholerdt & Heiman, 2003), and labeling themselves as promiscuous even when engaging in similar levels of sexual behavior as their peers who do not attach labels with negative connotations to their sexual behavior patterns (Fromuth, 1986). Prior research has also found that CSA survivors report feeling as if sex is something that is bad, dirty, and even controlling, possibly due to the events of the abuse and messages received from the perpetrator (Maltz & Holman 1987; Maltz, 1991). However, it should be noted, despite these overall more negative views regarding their sexual selves, CSA survivors endorsed less embarrassment regarding their sexuality as assessed by the sexual self-schema factors, so this quality may be viewed as a potential protective factor for sexual functioning.

In light of the finding that CSA survivors endorsed ideas about sexuality being immoral and irresponsible, it makes sense that these women were less able to view themselves as passionate and romantic. Viewing one's sexual self as immoral and irresponsible would likely be inconsistent with thinking about the sexual self as passionate and romantic. The finding that CSA survivors endorse less passionate and romantic views about sexuality is consistent with two prior studies (Meston et al., 2006; Reissing et al., 2003).

Interestingly, this is the first study to find CSA survivors reporting more open/direct sexual self-schemas despite the extensive empirical literature supporting the higher levels of consensual sexual activity of CSA survivors when compared to non-victims and the evidence that sexual self-schemas reflect past sexual behaviors (Andersen & Cyranowski, 1994). However, much of this body of research that reports CSA survivors engage in increased levels of sexual behavior is drawn from samples of young adult women such as college students (Alexander & Lupfer, 1987; Meston, Heiman, & Trapnell, 1999). There is another body of studies documenting that CSA survivors have an increased likelihood of experiencing sexual avoidance and dysfunctions which primarily comes from studies with clinical samples (Becker et al., 1984; Jehu, 1988; Sarwer & Durlak, 1996; Westerlund, 1992). CSA survivors in the clinical samples were usually older than the women in studies examining college-age women. A pattern emerges in which some CSA survivors engaged in high levels of sexual behaviors with many partners while in their dating years; however, when they become involved in more serious, committed relationships as they mature, they experience difficulties with sexual desire and arousal.

This study's sample was primarily comprised of young and relatively high functioning women and, consequently, it makes sense that the results were consistent with the pathway toward an increase in sexual behavior and interest. As these patterns of sexual behavior have been linked to risk for revictimization in

addition to other negative outcomes (Loeb et al., 2002; Mayall & Gold, 1995), this new finding regarding higher levels of the open/direct factor among CSA survivors raises questions about Anderson and Cyranowski's classification of the open/direct factor as qualitatively positive. It seems it is positive to be open with some moderation, but when openness becomes extremely high or low, it can negatively impact sexuality and risk for victimization. High levels of the open/direct factor may be more likely to be problematic when women also exhibit co-occurring low levels of the passionate/romantic factor, as was the case with the CSA survivors in the current study. Such co-existing sexual self-schemas may result in engaging in sexual behavior of a more casual nature, involving more sexual partners but with low levels of passion and caring involved.

Given the significant differences between CSA survivors and non-victims in regards to sexual self-schemas, it seemed important to continue investigating these cognitions and their impact on engagement in risky sexual behaviors and risk for adolescent sexual assault. Although the sexual self-schema factors did not mediate the relationship between CSA severity and adolescent risky sexual behavior, in most instances, the sexual self-schema factors explained additional variance in the risky behavior of interest, beyond severity of CSA history. While the immoral/irresponsible and open/direct factors explained additional variance in both consensual sexual partners and alcohol usage, the passionate/romantic factor did not explain additional variance in either outcome's variance. It may be that because sexual behavior often occurs more casually among college women, outside the context of a long-term relationship or the desire for a romantic relationship, the passionate/romantic factor was less predictive of the risky sexual behaviors. In addition, the embarrassed/conservative factor only added significantly to the prediction of consensual sexual partners, not alcohol usage. College women who are less embarrassed/conservative regarding sexuality are more likely to engage in sexual behavior with more partners, above and beyond severity of their CSA history.

An additive model was tested using hierarchical regression to determine if the sexual self-schema factors explained unique variance in adolescent sexual assault, above and beyond abuse characteristics and risky sexual behaviors. Interaction terms were added to test for the moderating effects of the sexual self-schema factors. The interaction between CSA and the immoral/irresponsible factor, CSA and the open/direct factor, CSA and the embarrassed/conservative factor, and CSA and the romantic/passionate factor added significant contributions to the explanation of adolescent sexual assault in CSA survivors. For individuals high on the immoral/irresponsible factor, the relationship between CSA severity and adolescent sexual assault experiences was significantly strengthened in a positive direction. However, for individuals low on the immoral–irresponsible factor, there was a decrease in adolescent sexual assault experiences as severity of CSA experiences decreased. A possible explanation for this finding when women are high on the immoral/irresponsible factor they may have more difficulty engaging in assertive and self-protective behaviors if they encounter a risky sexual situation, particularly when they have severe CSA history. The negative feelings (i.e., dirty, loose) evoked by sexual situations may impede self-protective behaviors. This finding is consistent with previous research which has found that women with CSA histories were lower in sexual refusal assertiveness and reported they would be less likely to use active forms of resistance when faced with hypothetical date rape scenarios (VanZile-Tamsen, Testa, & Livingston, 2005).

The interaction between CSA severity and the romantic/passionate factor was seemingly in direct contrast to the interaction explained above with the immoral/irresponsible factor. At high levels of the passionate/romantic factor, the relationship between CSA severity and adolescent sexual assault was negative. At low levels of the passionate/romantic factor, the relationship between CSA severity and adolescent sexual assault was strongly positively. It makes theoretical sense that young women who are choosing potential partners, but making these decisions without viewing sexuality in a romantic way, may be associating with individuals and attracting partners who are not interested in forming relationships. These potential partners may be more interested in interactions based on short-term, self-serving motivations. Although much sexual activity, particularly in college populations, occurs outside the context of any established relationship, for CSA survivors who may also view intimacy and sexuality as immoral and dirty, a less passionate caring interaction may decrease the CSA survivor's feelings of comfort expressing her needs and desires.

Examination of the interaction between CSA and the embarrassed/conservative factor revealed that at high levels of the factor, CSA severity's relationship with adolescent sexual assault was negative. At low levels of the embarrassed/conservative factor, the relationship between CSA severity and adolescent sexual assault was significantly positive. A potential explanation is that embarrassment and conservatism regarding sexuality among CSA survivors is protective against subsequent victimization as these cognitions are driving behaviors that are avoidant of sexual situations (Andersen & Cyranowski, 1994). Conversely, when CSA survivors are low in embarrassment regarding sexuality, they are likely going to expose themselves to more sexual situations, and therefore, from a simple probabilistic standpoint, heighten their risk for adolescent sexual assault.

Finally, the pattern of the interaction between CSA and the open/direct factor was inconsistent with the positive correlations between the open/direct factor and both CSA severity as well as adolescent sexual assault experiences. In addition, the *t* tests revealed the mean for the open/direct factor was higher for the CSA survivors than for non-survivors. Based on these data it was expected that at higher levels of the open–direct factor, a positive relationship would be observed between CSA severity and adolescent sexual assault. However, this was not found. At high levels of the open–direct factor, there was a negative

relationship between CSA severity and adolescent sexual assault. At low levels of the open–direct factor, there was a positive relationship between CSA severity and adolescent sexual assault. A coherent picture between the interaction pattern and the earlier results emerged by focusing on the overall pattern of these findings as well as the two adjectives that are the cornerstones of this factor, openness and directness. The factor seems to have a positive relationship with victimization experiences when the factor was considered in its entirety. When entered into the regression equation, it seemed likely that the unrestricted sexual behavior variables partialled out much of the variance associated with the openness component, leaving behind the directness component of the factor. It makes theoretical sense that, for women with low levels of directness, adolescent sexual assault experiences would increase as CSA severity increases, as it seems this construct may be associated with sexual assertiveness.

Due to the stringent control variables entered into the regression equation prior to each of these interaction terms, it appears the impact of the interaction of the immoral/irresponsible, open/direct, romantic/passionate, and embarrassed/conservative factors with childhood sexual trauma experiences is quite robust. These findings underscore the comprehensive effect of CSA on cognitive representations of sexuality as well as the importance of attending to these specific types of cognitions about sexuality among CSA survivors in a clinical setting. They also highlight the unique relationship between the sexual self-schema factors and CSA severity. It will be important for future studies that examine these relationships to take abuse severity into account in addition to examining a broad range of abuse experiences.

The current study had several limitations that should be acknowledged. A major limitation of the current study was its retrospective nature, and consequently, memory bias and other unknown variables could have influenced the variables of interest being assessed in this study. Due to its correlational design, it was not possible to determine from the current study which variables preceded or followed others, and no causal relationship can be inferred from the results. Longitudinal data are needed to determine the causal effect of sexual trauma experiences on cognitive representations of sexuality as well as cognitive representations of sexuality on sexual revictimization. In addition, the current study used a convenience sample comprised of mostly white, heterosexual, well-educated women who are not representative of the whole population of CSA survivors. Additional research is needed to examine the immoral–irresponsible factor in a clinical and more ethnically diverse sample of women with a history of CSA to speak to the generalizability of the findings of this line of research. The current body of research on sexual self-schema factors does not address the impact of other types of traumatic events on sexual self-schemas such as physical assault and emotional abuse, and future research is needed to address the effects of cumulative trauma on sexual self-schemas.

As this was the first study to introduce the immoral–irresponsible factor, future research should address the role of this factor in explaining sexual dysfunction among childhood sexual trauma survivors. Future research geared at incorporating these findings into therapy programs for survivors of CSA is also needed. Subsequent revictimization among CSA survivors is prevalent, and these cognitions about sexuality would be modifiable targets for intervention and prevention endeavors. Also, CSA survivors who present for treatment suffering from sexual dysfunctions may be assisted by addressing these cognitions, especially helping the CSA survivors increase their identity as women who are passionate or romantic as opposed to dirty and immoral. Finally, as CSA survivors are at heightened risk for PTSD, future research should investigate the impact of PTSD on sexual self-schemas as well as how psychological symptomatology and sexual self-schema factors may interact to increase risk for revictimization.

This study contributes several important empirical findings to the literature. The immoral–irresponsible factor is a new sexual self-schema factor that augments the original scale and improves its content validity for CSA survivors. This study demonstrated that sexual trauma in childhood impacts multiple facets of survivors' sexual self-perceptions, and these cognitions provide important information about when CSA survivors are at heightened risk for adolescent sexual assault experiences. In addition, the interaction between CSA severity and each of the sexual self-schema factors explained unique variance in adolescent sexual assault, above and beyond well-established contributors to adolescent sexual assault. Despite the general assumption of mental health professionals that CSA impacts cognitions about sexuality, very few studies have tested this idea empirically. This is the first study to examine the impact these changes in cognitions about sexuality might have on risk for adolescent revictimization.

In summary, the primary purposes of this study were to investigate differences between CSA survivors and non-victims in cognitions about sexuality as assessed by the sexual self-schema factors and to determine how these cognitions might impact the risk of adolescent sexual assault. Perhaps the most important result to highlight from this study is that CSA seems to have a comprehensive effect on the sexual self-view of survivors. Survivors were more likely to view themselves as more sexually open and less conservative than non-victims, and survivors were more likely to view their sexuality in a negatively valenced manner, associating their sexuality with descriptors such as immoral and dirty as opposed to passionate and loving.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage Publications.

Alexander, P. C., & Lupfer, S. L. (1987). Family characteristics and long-term consequences associated with sexual abuse. *Archives of Sexual Behavior, 16*, 235–245.

Andersen, B. L., & Cyranowski, J. M. (1994). Women's sexual self-schema. *Journal of Personality and Social Psychology, 67*, 1079–1100.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Becker, J. V., Skinner, L. J., Abel, G. G., Axelrod, R., & Cichon, J. (1984). Sexual problems of sexual assault survivors. *Women and Health, 9*, 5–20.

Bentler, P. M. (1968). Heterosexual behavior assessment: II. Females. *Behaviour Research and Therapy, 6*, 27–30.

Cole, P. M., & Putnam, F. W. (1992). Effect of incest on self and social functioning: A developmental psychopathology perspective. *Journal of Consulting and Clinical Psychology, 60*, 174–184.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349–354.

Derogatis, L. R., & Melisaratos, N. (1979). The DSFI: A multidimensional measure of sexual functioning. *Journal of Sex and Marital Therapy, 5*, 244–281.

Elze, D. E., Auslander, W., McMillen, C., Edmond, T., & Thompson, R. (2001). Untangling the impact of sexual abuse on HIV risk behaviors among youth in foster care. *AIDS Education and Prevention, 13*, 377–389.

Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. (1997). Childhood sexual abuse, adolescent sexual behaviors and sexual revictimization. *Child Abuse and Neglect, 21*, 789–803.

Fromuth, M. E. (1986). The relationship of childhood sexual abuse with later psychological and sexual adjustment in a sample of college women. *Child Abuse and Neglect, 10*, 5–15.

Hatfield, E., & Sprecher, S. (1986). Measuring passionate love in intimate relationships. *Journal of Adolescence, 9*, 383–410.

Hoon, E. F., Hoon, P. W., & Wincze, J. P. (1976). An inventory for the measurement of female sexual arousability: The SAI. *Archives of Sexual Behavior, 5*, 291–300.

Hudson, W. W., Harrison, D. F., & Crosscup, P. C. (1981). A short-form scale to measure sexual discord in dyadic relationships. *Journal of Sex Research, 17*, 157–174.

Jackson, J. L., Calhoun, K. S., Amick, A. E., & Maddever, H. M. (1990). Young adult women who report childhood intrafamilial sexual abuse: Subsequent adjustment. *Archives of Sexual Behavior, 19*, 211–221.

Jehu, D. (1988). *Beyond sexual abuse: Therapy with women who were victims in childhood*. Chichester: Wiley.

Koss, M. P. (1985). The hidden rape victim: Personality, attitudinal, and situational characteristics. *Psychology of Women Quarterly, 9*, 193–212.

Loeb, T. B., Williams, J. K., Carmona, J. V., Rivkin, I., Wyatt, G. E., Chin, D., … Asuan-O'Brian, A. (2002). Child sexual abuse: Associations with the sexual functioning of adolescents and adults. *Annual Review of Sex Research, 13*, 307–345.

Maltz, W. (1991). *The sexual healing journey: A guide for survivors of sexual abuse*. New York, NY: Harper Collins.

Maltz, W., & Holman, B. (1987). *Incest and sexuality: A guide to understanding and healing*. Lexington, MA: Lexington Books.

Marx, B. P., Heidt, J. M., & Gold, S. D. (2005). Perceived uncontrollability and unpredictability, self-regulation, and sexual revictimization. *Review of General Psychology, 9*, 67–90.

Mayall, A., & Gold, S. R. (1995). Definitional issues and mediating variables in the sexual revictimization of women sexually abused as children. *Journal of Interpersonal Violence, 10*, 26–42.

Messman-Moore, T. L., & Long, P. J. (2003). The role of childhood sexual abuse sequelae in the sexual revictimization of women: An empirical review and theoretical reformulation. *Clinical Psychology Review, 23*, 537–571.

Meston, C. M., Heiman, J. R., & Trapnell, P. D. (1999). The relation between early abuse and adult sexuality. *Journal of Sex Research, 36*, 385–395.

Meston, C. M., Rellini, A. H., & Heiman, J. R. (2006). Women's history of sexual abuse, their sexuality, and sexual self-schemas. *Journal of Consulting and Clinical Psychology, 74*, 229–236.

Miller, B. C., Monson, B. H., & Norton, M. C. (1995). The effects of forced sexual intercourse on white female adolescents. *Child Abuse and Neglect, 19*, 1289–1303.

Noll, J. G., Trickett, P. K., & Putnam, F. W. (2003). A prospective investigation of the impact of childhood sexual abuse on the development of sexuality. *Journal of Consulting and Clinical Psychology, 71*, 575–586.

Offman, A., & Matheson, K. (2004). The sexual self-perceptions of young women experiencing abuse in dating relationships. *Sex Roles, 51*, 551–560.

Orr, D. P., & Downs, M. C. (1985). Self-concept of adolescent sexual abuse victims. *Journal of Youth and Adolescence, 14*, 401–410.

Polusny, M. A., & Follette, V. M. (1995). Long-term correlates of child sexual abuse: Theory and review of the empirical literature. *Applied & Preventive Psychology, 4*, 143–166.

Reissing, E. D., Binik, Y. M., Khalife, S., Cohen, D., & Amsel, R. (2003). Etiological correlates of vaginismus: Sexual and physical abuse, sexual knowledge sexual self-schema and relationship adjustment. *Journal of Sex and Marital Therapy, 29*, 47–59.

Roodman, A. A., & Clum, G. A. (2001). Revictimization rates and method variance: A meta-analysis. *Clinical Psychology Review, 21*, 183–204.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Sarwer, D. B., & Durlak, J. A. (1996). Childhood sexual abuse as a predictor of adult female sexual dysfunction: A study of couples seeking sex therapy. *Child Abuse & Neglect, 20*, 963–972.

Schloredt, K. A., & Heiman, J. R. (2003). Perceptions of sexuality as related to sexual functioning and sexual risk in women with different types of childhood abuse histories. *Journal of Traumatic Stress, 16*, 275–284.

Siegel, J. A., & Williams, L. M. (2003). Risk factors for sexual victimization of women: Results from a prospective study. *Violence Against Women, 9*, 902–930.

Spector, I. P., Carey, M. P., & Steinberg, L. (1996). The Sexual Desire Inventory: Development, factor structure, and evidence of reliability. *Journal of Sex and Marital Therapy, 22*, 175–190.

VanZile-Tamsen, C., Testa, M., & Livingston, J. A. (2005). The impact of sexual assault history and relationship context on appraisal of and responses to acquaintance sexual assault risk. *Journal of Interpersonal Violence, 20*, 813–832.

Walser, R. D., & Kern, J. M. (1996). Relationships among childhood sexual abuse, sex guilt, and sexual behavior in adult clinical samples. *Journal of Sex Research, 33*, 321–326.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070.

Westerlund, E. (1992). *Women's sexuality after childhood incest*. New York: W.W. Norton.

White, H. R., & Labouvie, E. W. (1989). Towards the assessment of adolescent problem drinking. *Journal of Studies on Alcohol, 50*, 30–37.

Wilsnack, S. C., Vogeltanz, N. D., Klassen, A. D., & Harris, T. R. (1997). Childhood sexual abuse and women's substance abuse: National survey findings. *Journal of Studies on Alcohol, 58*, 264–271.

Wyatt, G. E., Lawrence, J., Vodounon, A., & Mickey, M. R. (1992). The Wyatt Sex History Questionnaire: A structured interview for female sexual history taking. *Journal of Child Sexual Abuse, 1*, 51–68.

Zucker, K. J., Bradley, S. J., Oliver, G., Blake, J., Fleming, S., & Hood, J. (2004). Self-reported sexual arousability in women with congenital adrenal hyperplasia. *Journal of Sex and Marital Therapy, 30*, 343–355.

Arch Sex Behav (2010) 39:1375–1379
DOI 10.1007/s10508-009-9555-x

ORIGINAL PAPER

Family Structure as a Risk Factor for Women's Sexual Victimization: A Study Using the Danish Registry System

Ask Elklit · Mark Shevlin

Received: 24 September 2008 / Revised: 21 April 2009 / Accepted: 17 August 2009 / Published online: 9 October 2009

Abstract This study estimated the risk of sexual victimization associated with different family structures. Based on the Danish Civil Registration System, all female visitors to the Centre for Rape Victims (CRV) at the University Hospital in Aarhus, during a two-year period (January 2005 to December 2006) were identified ($N = 214$) along with a control group ($N = 4{,}343$) that was matched by age and residential location. The family structure in the preceding year was used as a predictor variable in a logistic regression model. Results indicted that, compared to those who were married with children at home, being single with children at home significantly increased the likelihood of having visited the CRV. This is consistent with the research literature that has shown that single women with children are at risk for disadvantage on a range of socioeconomic and psychological factors as well as at risk for interpersonal violence.

Keywords Rape · Sexual victimization · Family structure · Single mothers · Marriage

Introduction

Prevalence estimates of 6–14% for sexual assault of women have been reported in large population based surveys in the U.S. (Golding, 1999; Sorenson, Stein, Siegal, Golding, & Burnam, 1987). There is limited information on self-reported experiences of sexual victimization in Denmark. A national youth probability study, based on participants aged between 13 and 15 years, found that 2.6% of the females reported having been raped (Elklit, 2002). Sundaram, Laursen, and Helweg-Larsen (2008) reported a higher estimate for adolescent females (14–16 years) of 4% based on an analysis of the Danish Youth Survey (Helweg-Larsen & Bøving Larsen, 2003). Estimates from adult samples in Denmark and other Nordic countries are consistently higher. Sundaram et al. (2008) reported that 14% of females from the Danish National Health and Morbidity Survey claimed at least once experience of actual, or attempted, forced sexual activity. The estimate for forced sexual activity based on the Prevalence and Health Sequels of Violence project (Helweg-Larsen, Sundaram, Heiskanen, & Piispa, 2003) was 4.6%, which was lower than the same estimate from Finland (10%). In a large study of females across five Nordic countries, 11.7% of the Danish sample reported having experienced severe sexual abuse, which was defined as actual or attempted penetration (Wijma et al., 2003); this was lower than the estimate for the Finish sample (15%) and higher than that for Sweden (5.8%). The use of inconsistent definitions of sexual assault and rape and the under-reporting of such experiences (Jewkes & Abrahams, 2002) leads to caution being required in the interpretation of such estimates and may explain the variation in the prevalence rates.

The research evidence also suggests particular groups have increased risk of sexual victimization. Increased rates of sexual assault have been reported for females from racial minorities (Elliott, Mok, & Briere, 2004), victims of child sexual abuse (Campbell, Greeson, Bybee, & Raja, 2008), women with severe disability impairments (Casteel, Martin, Smith, Gurka, & Kupper, 2008), and women who report their general health as poor (Cloutier, Martin, & Poole, 2002).

A. Elklit
Department of Psychology, Aarhus University, Aarhus, Denmark

M. Shevlin (✉)
School of Psychology, Faculty of Life and Health Sciences, University of Ulster, Magee Campus, Londonderry BT48 7JL, Northern Ireland, UK
e-mail: m.shevlin@ulster.ac.uk

Experience of sexual victimization and rape has been shown to have negative consequences in terms of mental health, physical health, and interpersonal relationships. Research has provided evidence that sexual victimization is associated with a range of psychological and behavioral problems (Resick, 1993), including posttraumatic stress disorder (Bromet, Sonnega, & Kessler, 1998), recurrent depression, self-harming behavior, and suicidal attempts (Floen & Elklit, 2007), alcohol abuse (Hankin et al., 1999), disordered eating (Ackard & Neumark-Sztainer, 2002), symptoms of psychosis (Shevlin, Adamson, & Dorahy, 2007), and the abuse of prescription and recreational drugs (Kilpatrick, Acierno, Resnick, Saunders, & Best, 1997; Sturza & Campbell, 2005).

The dissolution of romantic relationships has also been identified as a consequence of sexual assault or rape. Several studies have noted that partnership break-ups are quite common (35–59%) following sexual victimization (Becker, Skinner, Abel, & Treacy, 1982; Burgess & Holmstrom, 1979; Norris & Feldman-Summers, 1981). This may be, in part, attributable to the social and interpersonal costs associated with the aftermath of sexual victimization. Zweig, Crockett, Sayer, and Vicary (1999) reported lower levels of relationship quality and intimacy competence in victims of violent sexual coercion and decreased interest and satisfaction in sex (Siegel, Golding, Stein, Burnam, & Sorenson, 1990).

These findings indicate a significant methodological problem when estimating the risk associated with different family structures and the risks of sexual victimization. The assessment of family structure is generally concurrent with the self-reported accounts of sexual victimization. For example, a population based survey study of 281 female registered voters in one North Carolina city found that single status, divorce status, and having children were associated with increased risks of sexual assault (odds ratios 3.19, 9.56, and 2.09, respectively) (Smith, Thornton, DeVellis, Earp, & Coker, 2002). Using data from the National Survey of Mental Health in Australia, Butterworth (2004) compared incidences of physical and sexual violence reports of single mothers ($N = 662$) and partnered mothers ($N = 1{,}602$). Single mothers were found to be over three times more likely (OR = 3.41) than partnered mothers to report at least one form of physical or sexual violence. These findings suggest that living without a partner and having children increased the risk of sexual victimization. However, it could also be argued that current family structure was a result of prior sexual victimization; previously partnered women may have become unpartnered due to their experience of sexual victimization. Furthermore, family structure may be a proxy indicator of other variables that predict sexual victimization, such as low socioeconomic status (McLanahan & Percheski, 2008).

In order to establish the effects of different family types on risk of sexual victimization, it would be preferable to conduct a prospective study where family structure and previous occurrences of sexual victimization were known and subsequent occurrences of sexual victimization are then observed. The use of such a cohort design has the desired property of allowing the temporal ordering of events to be established. However, prospective designs are rarely employed as they are costly in terms of time to implement. Our study used the Danish Civil Registration System (CRS), which stores information on all Danish residents, to implement a quasi-prospective study. Females who experienced sexual victimization were identified and their pre-victimization family structure and previous experiences of sexual victimization were taken from the CRS. This meant that the temporal ordering of family structure and sexual victimization was established. In addition, the use of information on annual income allowed for the statistical control of this potentially confounding variable.

This study aimed to estimate the risk of sexual victimization associated with different family structures. This study will add to the existing research by (1) examining a broad range of family structures, (2) addressing the temporal ordering problem by using family structure information before the sexual victimization (and excluding cases of previous victimization), (3) using a matched control group to assess differences between those exposed and non-exposed to sexual victimization, and (4) statistically controlling for differences in financial status across the family structures.

Method

Participants

This study was based on the Danish Civil Registration System (CRS). A detailed description of the structure of CRS was provided by Pedersen, Gøtzsche, Møller, and Mortensen (2006). All females ($N = 214$), or cases, who visited the Centre for Rape Victims (CRV) at the University Hospital in Aarhus, Denmark during a two year period (January 2005 to December 2006) were identified and demographic information was added based on the individual civil registry number (CPR). The females who visited the CRV were aged from 13 to 87 years ($M = 26$ years, $SD = 13.41$). The CPR data were used to extract a control group of 20 women for each case with the same age and living in the same municipality; this equates to a prevalence rate of sexual victimization of 5%, which is between the lower official statistics and the higher self-report estimates. Any women who had visited the CRV since 1999 were not included in the control group. For the time period January 2005 to December 2006, data from Denmark Statbank, the central government agency for statistics, reported that there were 54 rapes officially recorded in the Aarhus municipality. The average female population in this municipality was 150,660 for the 2 year time period. Therefore, the officially

recorded rapes and CRV attendance represented .04 and 1.4% of the population, respectively.

Family structure was a composite variable created from register information on marital status, co-habiting status, and number of children. The variable had six categories: (1) married, children at home; (2) married, no children at home; (3) single, children at home; (4) single, no children at home; (5) cohabiting, children at home; (6) cohabiting, no children at home. For the cases and controls, their family structure during the year 2004 was recorded. The participant's annual income, in Danish krone (DKK), for the year 2004 was also recorded.

Results

Data were missing on family structure for 151 controls (3.52%) so the effective sample size for cases and controls was 4,343. The cross-tabulated frequencies of cases and controls by family structure are reported in Table 1. There were more than expected cases in two family types (single, children at home and cohabiting, children at home) although the difference between the observed and expected counts was very small for the group who were cohabiting with children at home. The overall association between family structure and CRV attendance was statistically significant ($\chi^2 = 58.86, df = 5, p < .01$).

Data on income were missing for 421 controls (9.83%) and 25 cases (11.68%). The cases and controls differed significantly on annual income. The mean income for controls ($M =$ 115,438 DKK, $SD =$ 109,126) was higher than for cases ($M =$ 70,451 DKK, $SD =$ 74,935) and this difference was statistically significant, $t(4046) = 5.60, p < .01$. To identify the associations between specific family types and sexual victimization, two logistic regression models were estimated. For both models, the dependent variable was attendance at the CRV in the 24-month period (2005–2006). For the first model, the predictor variable was the family structure 2004, and in the second model annual income added as a covariate. For both analyses, the "married, children at home" group was the reference category.

The results are shown in Table 2. For the first model, the likelihood ratio test for the family structure variable was significant ($\chi^2 = 54.14, df = 5, p < .01$) and the odds ratios indicated that being single with children at home significantly increased the likelihood of having visited the CRV, and being married with no children at home reduced the likelihood compared to those who were married with children at home. For the second model, the likelihood ratio tests for the family structure variable ($\chi^2 = 40.33, df = 5, p < .01$) and annual income ($\chi^2 = 359.10, df = 5, p < .01$) were statistically significant. The odds ratios indicated that being single with children at home remained statistically significant, and the effect for being married with no children at home became non-significant.

Discussion

This study aimed to assess the relationship between family structure and sexual victimization measured by attendance at a CRV. The cross-tabulation results showed that higher than expected rates of CRV attendance were found for single and cohabiting women with children at home. Lower than expected rates of CRV attendance were found for married women both with and without children, and single and cohabiting women with no children. The logistic regression indicated that single women with children at home were 2.77 times more likely to have attended the CRV and married women without children at home were 84% less likely to have attended the CRV than the married women with children. The effect for married women without children at home became non-significant when annual income was statistically controlled for and the effect for single women with children at home increased (OR = 3.27). Therefore, this study suggested that family structures differed in terms of the risk associated with sexual victimization while controlling for different levels of income.

Table 1 Cross-tabulation of family structure and sexual victimization

	Family structure 2004					
	Married, children at home	Married, no children at home	Single, children at home	Single, no children at home	Cohabiting, children at home	Cohabiting, no children at home
Cases						
Count	65	2	57	59	18	13
Expected count	73.2	13.6	24.9	62.5	17.7	22.0
%	30.4	0.9	26.6	27.6	8.4	6.1
Controls						
Count	1420	275	449	1210	342	433
Expected count	1411.8	263.4	481.1	1206.5	342.3	424.0
%	34.4	6.7	10.9	29.3	8.3	10.5

Table 2 Unadjusted and adjusted logistic regression estimates for family structure and sexual victimization analysis

Family structure	*B*	*SE*	*p*	OR (95% CI) Unadjusted	OR (95% CI) Adjusted[a]
Married, children at home	Reference category				
Married, no children at home	−1.840	.72	.01	0.16 (0.04–0.65)	0.32 (0.08–1.35)
Single, children at home	1.020	.19	<.01	2.77 (1.91–4.02)	3.27 (2.14–5.00)
Single, no children at home	0.063	.18	ns	1.07 (0.74–1.53)	1.40 (0.94–2.11)
Cohabiting, children at home	0.140	.27	ns	1.15 (0.67–1.96)	0.61 (0.89–2.91)
Cohabiting, no children at home	−0.422	.31	ns	0.66 (0.36–1.20)	0.99 (0.52–1.88)

[a] Estimates adjusted for the effects of annual income

Single women with children at home have a greater risk of CRV attendance compared to their married counterparts, independent of income. This is consistent with the research literature that has shown that single women with children are at risk of disadvantage on a range of socioeconomic and psychological factors as well as risk of interpersonal violence. Single mothers have been found to have lower income compared to married mothers (Brown & Moran, 1997; Thomas & Sawhill, 2005), lower levels of educational achievement (Butterworth, 2004), and higher rates of unemployment (Macran, Clarke, & Joshi, 1996). Lower levels of psychological well being for single mothers have been consistently reported in the literature (Butterworth, 2004; Cooper et al., 2008; Franz, Lensche, & Schmitz, 2003; Hope, Power, & Rodgers, 1999) along with elevated rates of substance and alcohol abuse (Barrett & Turner, 2006). Single mothers, compared to partnered mothers, were reported to be significantly more likely to have been victims of rape, sexual molestation, and severe physical assault or threatened with a weapon (Butterworth, 2004).

The mechanism by which this particular family type increases the risks of a broad range of negative outcomes is not clear. However, alternative theoretical models have been proposed to explain the association between family type and substance abuse (Barrett & Turner, 2006) and psychological distress (Hope et al., 1999) and these may offer some guidance in understanding the family structure-sexual victimization link. These models share some explanatory variables, such as the direct effect of economic deprivation and low levels of social support. However, none of the theoretical models seem to accommodate an explanation of sexual victimization as such a model would need to accommodate explanations of the motives of the assailant as well as the victim. As such, any explanatory model based exclusively on characteristics of the victim will be limited. An additional problem is that any explanatory model will have to accommodate increasingly large and heterogeneous family types (Andersson, 2002).

This study had some limitations. First, there was no independent verification that every woman who attended the CRV had actually been subjected to an attempted or completed rape nor were there details of the nature of the assaults. Importantly, the degree of physical force used in the assaults is unknown. However, attendance at the CRV is an objective behavioral indicator that is likely to be indicative of some type of sexual aggression having been experienced. Second, the control group was matched only on the basis of age and living in the same area. It is not known if the groups differed systematically on potential confounding variables that are related to both family structure and sexual victimization, such as socioeconomic status. Third, not all women who experience sexual victimization will have attended the CRV. It may be that married or cohabiting women can find support from their husbands or partners and are therefore less likely to attend the CRV. However, Ullman and Filipas (2001) found that marital status was not a significant predictor of formal help-seeking in victims of sexual assault. Fourth, it is possible that the family structure of some participants may have changed prior to experiencing sexual victimization. Finally, the percentage of the municipality population that attended the CRV is much smaller than would be expected given the prevalence estimates of sexual victimization reported in the Introduction, and so the sample may not be completely representative of the population. However, the percentage is much larger than the official government statistics, and this is consistent with previous research that has reported higher rates of women attending CRV's than reporting to the police in Nordic countries (Nesvold, Worm, Vala, & Agnarsdóttir, 2005).

References

Ackard, D. M., & Neumark-Sztainer, D. (2002). Date violence and date rape among adolescents: Associations with disordered eating behaviors and psychological health. *Child Abuse and Neglect, 26*, 455–473.

Andersson, G. (2002). Children's experience of family disruption and family formation: Evidence from 16 FFS countries. *Demographic Research, 7*, 343–364.

Barrett, A. E., & Turner, J. (2006). Family structure and substance use problems in adolescence and early adulthood: Examining explanations for the relationship. *Addiction, 101*, 109–120.

Becker, J. V., Skinner, L. J., Abel, G. G., & Treacy, E. (1982). Incidence and types of sexual dysfunctions in rape and incest victims. *Journal of Sex and Marital Therapy, 8*, 65–74.

Bromet, E., Sonnega, A., & Kessler, R. C. (1998). Risk factors for DSM-III-R posttraumatic stress disorder: Findings from the National Comorbidity Survey. *American Journal of Epidemiology, 147*, 353–361.

Brown, G. W., & Moran, P. M. (1997). Single mothers, poverty and depression. *Psychological Medicine, 27*, 21–33.

Burgess, A. W., & Holmstrom, L. L. (1979). Adaptive strategies and recovery from rape. *American Journal of Psychiatry, 136*, 1278–1282.

Butterworth, P. (2004). The prevalence of psychiatric disorders among lone mothers: Association with physical and sexual violence. *British Journal of Psychiatry, 184*, 21–27.

Campbell, R., Greeson, M. R., Bybee, D., & Raja, S. (2008). The co-occurrence of childhood sexual abuse, adult sexual assault, intimate partner violence, and sexual harassment: A mediational model of posttraumatic stress disorder and physical health outcomes. *Journal of Consulting and Clinical Psychology, 76*, 194–207.

Casteel, C., Martin, S. L., Smith, J. B., Gurka, K. K., & Kupper, L. L. (2008). National study of physical and sexual assault among women with disabilities. *Injury Prevention, 14*, 87–90.

Cloutier, S., Martin, S. L., & Poole, C. (2002). Sexual assault among North Carolina women: Prevalence and health risk factors. *Journal of Epidemiology and Community Health, 56*, 265–271.

Cooper, C., Bebbington, P. E., Meltzer, H., Bhugra, D., Brugha, T., Jenkins, R., et al. (2008). Depression and common mental disorders in lone parents: Results of the 2000 National Psychiatric Morbidity Survey. *Psychological Medicine, 38*, 335–342.

Elklit, A. (2002). Victimization and PTSD in a Danish national youth probability sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*, 174–181.

Elliott, D. M., Mok, D. S., & Briere, J. (2004). Adult sexual assault: Prevalence, symptomatology, and sex gender differences in the general population. *Journal of Traumatic Stress, 17*, 203–211.

Floen, S. K., & Elklit, A. (2007). Psychiatric diagnoses, trauma, and suicidality. *Annals of General Psychiatry, 6*, 12. Retrieved April 21, 2009, from http://www.annals-general-psychiatry.com/content/6/1/12.

Franz, M., Lensche, H., & Schmitz, N. (2003). Psychological distress and socioeconomic status in single mothers and their children in a German city. *Social Psychiatry and Psychiatric Epidemiology, 38*, 59–68.

Golding, J. M. (1999). Sexual assault history and medical care seeking: The roles of symptom prevalence and illness behavior. *Psychology and Health, 14*, 949–957.

Hankin, C. S., Skinner, K. M., Sullivan, L. M., Millar, D. R., Frayne, S., & Tripp, T. J. (1999). Prevalence of depressive and alcohol abuse symptoms among women VA outpatients who report experiencing sexual assault while in the military. *Journal of Traumatic Stress, 12*, 601–612.

Helweg-Larsen, K., & Bøving-Larsen, H. (2003). Ethical issues in youth surveys: Potentials for conducting a national questionnaire study on adolescent schoolchildren's sexual experiences with adults. *American Journal of Public Health, 93*, 1878–1882.

Helweg-Larsen, K., Sundaram, V., Heiskanen, M., & Piispa, M. (2003). *The development and implementation of questions on violence in national health surveys in Denmark and Finland: A comparative study under the DAPHNE program 2000–2003*. Copenhagen: National Institute of Public Health, Statistics Finland.

Hope, S., Power, C., & Rodgers, B. (1999). Does financial hardship account for elevated psychological distress in lone mothers? *Social Science and Medicine, 49*, 1637–1649.

Jewkes, R., & Abrahams, N. (2002). The epidemiology of rape and sexual coercion in South Africa: An overview. *Social Science and Medicine, 55*, 1231–1244.

Kilpatrick, D. G., Acierno, R., Resnick, H. S., Saunders, B., & Best, C. L. (1997). The relationship between assault and alcohol, marijuana, and hard drug use in women. *Journal of Consulting and Clinical Psychology, 65*, 834–847.

Macran, S., Clarke, L., & Joshi, H. (1996). Women's health: Dimensions and differentials. *Social Science and Medicine, 42*, 1203–1216.

McLanahan, S., & Percheski, C. (2008). Family structure and the reproduction of inequalities. *Annual Review of Sociology, 34*, 257–276.

Nesvold, H., Worm, A. M., Vala, U., & Agnarsdóttir, G. (2005). Different Nordic facilities for victims of sexual assault: A comparative study. *Acta Obstetricia et Gynecologica Scandinavica, 84*, 177–183.

Norris, J., & Feldman-Summers, S. (1981). Factors related to the psychological impacts of rape on the victim. *Journal of Abnormal Psychology, 90*, 562–567.

Pedersen, C. B., Gøtzsche, H., Møller, J. O., & Mortensen, P. B. (2006). The Danish Civil Registration System: A cohort of eight million persons. *Danish Medical Bulletin, 53*, 441–449.

Resick, P. A. (1993). The psychological impact of rape. *Journal of Interpersonal Violence, 8*, 223–255.

Shevlin, M., Adamson, G., & Dorahy, M. (2007). Childhood traumas and hallucinations: An analysis of the National Comorbidity Survey. *Journal of Psychiatric Research, 41*, 222–228.

Siegel, J. M., Golding, J. M., Stein, J. A., Burnam, M. A., & Sorenson, S. B. (1990). Reactions to sexual assault: A community study. *Journal of Interpersonal Violence, 5*, 229–246.

Smith, P. H., Thornton, G. E., DeVellis, R., Earp, J., & Coker, A. L. (2002). A population-based study of the prevalence and distinctiveness of battering, physical assault, and sexual assault in intimate relationships. *Violence Against Women, 8*, 1208–1232.

Sorenson, S. B., Stein, J. A., Siegel, J. M., Golding, J. M., & Burnam, M. A. (1987). The prevalence of adult sexual assault: The Los Angeles Epidemiological Catchment Area Project. *American Journal of Epidemiology, 126*, 1154–1164.

Sturza, M. L., & Campbell, R. (2005). An exploratory study of rape survivors' prescription drug use as a means of coping with sexual assault. *Psychology of Women Quarterly, 29*, 353–363.

Sundaram, V., Laursen, B., & Helweg-Larsen, K. (2008). Is sexual victimisation gender-specific? The prevalence of forced sexual activity among men and women in Denmark, and self-reported well-being among survivors. *Journal of Interpersonal Violence, 23*, 1414–1450.

Thomas, A., & Sawhill, I. (2005). For love and money? The impact of family structure on family income. *The Future of Children, 15*, 57–74.

Ullman, S. E., & Filipas, H. H. (2001). Correlates of formal and informal support seeking in sexual assault victims. *Journal of Interpersonal Violence, 16*, 1028–1047.

Wijma, B., Schei, B., Swahnberg, K., Hilden, M., Offerdal, K., Pikarinen, U., et al. (2003). Emotional, physical, and sexual abuse in patients visiting gynaecological clinics: A Nordic cross-sectional study. *Lancet, 361*, 2107–2111.

Zweig, J. M., Crockett, L. J., Sayer, A., & Vicary, J. R. (1999). A longitudinal examination of the consequences of sexual victimization for rural young adult women. *Journal of Sex Research, 36*, 396–409.

Arch Sex Behav (2010) 39:1381–1388
DOI 10.1007/s10508-009-9520-8

ORIGINAL PAPER

The Relationship Between Rejection Sensitivity and Compliant Condom Use

Gaynor L. Edwards · Bonnie L. Barber

Received: 26 November 2008 / Revised: 25 June 2009 / Accepted: 28 June 2009 / Published online: 4 September 2009

Abstract Those who are rejection sensitive anxiously expect and readily perceive rejection. Rejection sensitivity is hypothesized to predict behavior; however, this link may be more evident in some contexts than others. The current study examined the link between rejection sensitivity and condom use. Australian young adults in romantic ($n = 649$, 70% female) and casual ($n = 144$, 76.2% female) relationship contexts completed measures on rejection sensitivity, condom use preferences, and condom use. Regression analysis showed that rejection sensitivity predicted condom use when participants' condom use preferences were at odds with those they thought their partner held. Specifically, highly rejection-sensitive individuals who preferred more frequent condom use reported using condoms less often, if that was what they thought their partners wanted. The results lend support to the model of rejection sensitivity in that individuals comply more with their perceived partner's preferences if they are more rejection-sensitive. The results also highlight the need to take the situational context into consideration when examining links between personality dispositions and behavior.

Keywords Condom use · Rejection sensitivity · Condom use preferences · Relationship context

Introduction

In an effort to reduce rates of HIV and STIs, many researchers have examined factors related to condom use. Although health implications may be a primary motivating factor for conducting such research, health implications may not be the only motivating factor for condom use. Many scholars have suggested that interpersonal factors may be more important than health concerns to people making risky sex decisions and that a greater emphasis is needed on the interpersonal dimensions of sexual behavior (Amaro, 1995; Bromnick & Swinburn, 2003; Crockett, Raffaelli, & Moilanen, 2003; Rosenthal, Hall, & Moore, 1992; Sheeran, Abraham, & Orbell, 1999). Thus, the current study aimed to examine condom use as a behavior motivated by interpersonal concerns. In particular, we were interested in how the romantic views held by individuals related to condom use. Romantic views refer to cognitive representations of relationships that people hold (Furman & Simon, 1999) and may include constructs such as rejection sensitivity and attachment style.

Rejection Sensitivity and Sexual Behavior

Those who are rejection sensitive readily perceive and anxiously expect rejection from others (Downey, Bonica, & Rincon, 1999). In an effort to prevent rejection, individuals may comply with what they perceive their partners to want. Downey et al. suggested that one area in which compliance may manifest is sexual behavior. We argue that those who fear rejection may make condom use decisions based on what their partner wants, rather that what they themselves want. In effect, sexual compliance may manifest in condom use behavior.

In examining the link between rejection sensitivity and sexual behavior, Young and Furman (2008) found that rejection sensitivity predicted increased risk of sexual victimization among adolescents. However, the relationship context under which sexual victimization was taking place was unknown. Additionally, the measure of sexual victimization included experiences of the use of physical force to obtain sex. As the construct of compliance involves voluntarily engaging in a

G. L. Edwards (✉) · B. L. Barber
School of Psychology, Murdoch University, South Street, Murdoch, WA 6150, Australia
e-mail: g.edwards@murdoch.edu.au

behavior that one would rather not do, it seems inappropriate to consider forced sex as sexual compliance. Condom use may be a more appropriate measure of sexual compliance.

The link between condom use and romantic views has been investigated by attachment researchers. The construct of adult attachment style shares similarities with rejection sensitivity. In particular, attachment anxiety is characterized by fears of abandonment and fears of being unloved (Collins & Read, 1990). Research suggests a link between attachment anxiety and less frequent condom use (Feeney, Kelly, Gallois, Peterson, & Terry, 1999; Feeney, Peterson, Gallois, & Terry, 2000; Kershaw et al., 2007), but also to more frequent condom use (Bogaert & Sadava, 2002). One reason for these mixed results may be that relationship status hasn't always been considered. More information is needed regarding how fears of rejection may predict condom use in casual and romantic relationship contexts. Research has shown that individuals expect to use condoms more in casual versus romantic relationship contexts (Ellen, Cahn, Eyre, & Boyer, 1996). In addition, those who have casual sexual encounters use condoms more than those who have romantic sexual encounters (Gebhardt, Kuyper, & Greunsven, 2003; Juarez & Martín, 2006; Sheeran et al., 1999). Therefore, it is apparent that relationship context is an important variable to consider when predicting condom use.

When Should Rejection Sensitivity Matter?

In both romantic and casual relationships, it is also important to consider whether each partner wants to use condoms. Mischel and Shoda (1995) argued that in order to understand how a personality disposition relates to behavior, it is not only important to examine how behavior varies as a function of the personality construct, but also the contexts in which this link is strongest. If rejection sensitivity affects sexual behavior at some times but not others, this may be a reflection of the true nature of rejection sensitivity, rather than an inconsistent link between rejection sensitivity and behavior.

The link between rejection sensitivity and sexual compliance should be greatest when there is a chance of conflict or rejection. When negotiating condom use, it is unlikely that rejection sensitivity will affect behavior if both partners agree to use a condom, or agree not to use a condom. In contrast, when one partner wishes to use a condom but the other doesn't, the chance of conflict should be higher. These situations are more likely to trigger rejection sensitivity. It is also these situations in which we should expect those with higher levels of rejection sensitivity to be more compliant and, therefore, it is important to consider the condom use preferences of both partners.

Condom Use Preference

Many studies have examined attitudes towards condoms and intentions to use condoms. Surprisingly, few studies have asked participants whether they *wanted* to use condoms. At first glance, these concepts appear similar, but there are important distinctions to be made. Attitudes towards a behavior refer to whether the individual's evaluations of that behavior are favorable or unfavorable (Ajzen, 1991). However, some individuals with negative attitudes towards condoms may still want to use condoms to protect themselves against unwanted pregnancy and/or STIs. Intentions to perform a behavior refer to the extent to which a person will try to perform a behavior (Ajzen, 1991). However, an individual may not intend to use condoms but still want to use them. For example, some individuals may not intend to use condoms due to their belief that they could not convince their partners to use condoms, but they may still want to use condoms. In order to accurately measure sexual compliance, research should examine not whether individuals like condoms, or whether they intend to use condoms, as such constructs do not consider the overall condom use preference. Instead, researchers should examine condom use preferences when examining condom use as a form of sexual compliance.

In summary, the current study aimed to examine condom use as a function of rejection sensitivity and condom use preference discrepancies. It was expected that rejection sensitivity would predict behavior in situations where individuals' condom use preferences were at odds with their partner's perceived condom use preferences. The current study grouped participants into four groups based on such preferences. Two were congruent groups (both partners did or did not want to use condoms) and two were incongruent groups (participants either did or did not want to use condoms, and thought their partners wanted the opposite to themselves). In the two incongruent groups, it was expected that highly rejection sensitive individuals would be compliant to what they perceived their partner's condom use preferences to be, whether that be by using condoms more or less frequently. In the two congruent groups, it was not expected that rejection sensitivity would be related to condom use. As condom use often differs according to other birth control use, this variable was controlled in all regression analyses (Sheeran et al., 1999). Additionally, since condom use and expectations about condom use differ for romantic and casual relationships, these effects were examined across both relationship contexts.

As an additional aim, the current study also sought to describe the differences between condom use preference groups. First, whether levels of condom use differed between condom use preference groups, and second, whether men and women were distributed significantly differently across groups. As women expect their male partners to want use condoms less than themselves (Edwards & Barber, 2009), it was hypothesised that women would be overrepresented in the group that wanted to use condoms more than they thought their partners did.

Method

Participants

The current study included participants with romantic ($n = 649$), and casual ($n = 144$) relationship experience. All participants were from one Western Australian university, were 18–25 years of age, unmarried, and reported having sex with an opposite sex partner during the past year. Seventy-one percent of participants were born in Australia. The mean age of participants reporting romantic relationship experience was 20.9 years ($SD = 2.06$). Seventy percent were female. Those reporting casual relationships were of similar age ($M = 21.9$; $SD = 2.09$) and gender composition (76.2% female).

Measures

An online survey included sections on demographic information, rejection sensitivity, relationship type, condom use preferences, and sexual behavior.

Rejection Sensitivity

The Rejection Sensitivity Questionnaire (RSQ; Downey & Feldman, 1996) presents participants with a number of situations where something is asked of another (e.g., you ask your boyfriend/girlfriend if he/she really loves you). Participants rated how concerned or anxious they would be over the person's response (1 = very unconcerned to 6 = very concerned) and whether they expected the person to respond in an accepting manner (1 = very unlikely to 6 = very likely). The RSQ consists of 18 items using situations involving romantic partners, friends, parents, and significant others ($\alpha = .83$); however, the current study focuses on the six items that were specific to a romantic relationship or a potential romantic relationship situation ($\alpha = .75$).

A total rejection sensitivity score was calculated by following the instructions outlined by Downey and Feldman (1996). This involved reversing the participants' scores on expectancy of acceptance to reflect expectancy of rejection, and then multiplying the reversed score by their degree of anxiety. Each of the multiplied scores was then averaged across all six items, to provide a final rejection sensitivity score. Higher scores reflect greater rejection sensitivity.

Relationship Type

Participants were asked about the different relationships they had experienced during the last year. The definitions of relationships are shown in the Appendix. All "other relationships" will be referred to as casual relationships in analyses that follow. Participants who had experienced a romantic relationship within the last year were asked to think of their current or most recent romantic partner when responding to the survey items on condom use preferences, condom use and birth control use. The same method was used for casual relationship experiences. Participants could respond to either section on romantic or casual relationships, or both.

Condom Use Preferences

To assess participants' own preference to use condoms, participants were asked "How often do you want to use a condom?" (1 = never to 5 = every time). To assess their perception of their partner's condom use preference, participants were asked "How often do you think your partner wants to use a condom?" (1 = never to 5 = every time). Some participants in casual relationships indicated that they had sex with their partner only once. Accordingly, such participants were presented with items that were more relevant to their experiences as it did not make sense to present a never-always frequency scale when they had sex with their partners only once. Specifically, they were asked: "Did you want to use a condom" and "Do you think your partner wanted to use a condom"? (1 = No, not at all to 5 = Yes, very much). To bring the responses in line with those having sex more than once, participants responding at the lower end of the scale (with a 1 or 2) were recoded as 1 (never) and those responding with a 4 or 5 were coded as 5 (every time). A small number of participants responded with a 3 and were excluded from analysis as it was unclear which end of the scale would best represent their condom use preferences ($n = 2$ for own preferences and $n = 1$ for partner's perceived preferences).

To examine sexually compliant condom use behavior, participants were divided into four groups based on their own condom use preferences (more or less frequently) and congruence with their partner's perceived condom use preferences (congruent or incongruent). The "Both Want" and "Both Don't Want" groups had condom use preferences that were congruent with their partner. Specifically, when reporting condom use preferences, these participants had preferences that were no more than one point away from their partner's perceived preferences. In the "Both Don't Want" group, both partners wanted to use condoms less frequently. In the "Both Want" group, both partners wanted to use condoms more frequently. Participants were placed in one of two incongruent groups (the "Participant Wants" and "Partner Wants" groups) when participants had condom use preferences that were at least two points away from their partner's perceived condom use preference. In the "Partner Wants" group, participants wanted to use condoms less frequently than they thought their partners did. In the "Participant Wants" group, participants wanted to use condoms more frequently than they thought their partners did.

If the participants rated their own or their partner's condom use preference as a 3, they were included in a congruent group if they were only one point away from their partner, but in an incongruent group if they were two points away from

their partner. For example, if a participant rated their condom use preference as a 3, and perceived their partner's condom use preference as a 2, then they were included in the "Both Don't Want" group. In contrast, if they perceived their partner's condom use preference to be a 1, then they were included in the "Participant Wants" group. This ensured that all participants in the incongruent groups had condom use preferences that were at least two points away from their partner's perceived condom use preferences. If both the participant and their partner's perceived level of condom use preference were rated as 3, they were excluded from analysis ($n = 22$ for romantic relationships; $n = 3$ for casual relationships).

Birth Control Use

Participants indicated which method of birth control they generally used with that particular partner from a list. These included not using protection or birth control, condoms, the pill, and other forms of birth control (e.g., implanon, a diaphragm, having a vasectomy). Participants could endorse more than one option if applicable (e.g., condoms and the pill).

Condom Use

Participants were asked how often they used condoms with their partner (1 = never to 5 = always). Participants who had been in a casual relationship in which they had sex with their partner only once indicated which method/s of birth control they used that time. This resulted in a dichotomous measure of condom use (yes/no). Similar to the condom use preference measure, their responses were recoded to allow participants to be included in the analyses. Those who used a condom with their partner were recoded with a 5 as they technically used a condom every time with their partner. Those who did not use a condom were recoded with a 1, meaning they never used protection with that partner.

Procedure

Participants were recruited to complete an online survey via one of two ways: the School of Psychology homepage or an email sent to university guild members. Posters were displayed on notice boards at various buildings around the university directing students to the School of Psychology homepage. All participants were required to enter their student number and password to access the survey. This ensured that only the university's students participated and that responses were entered only once. Participants were assured that although their student number was required to enter the survey, identifying information would be kept separate from their responses. The survey took approximately 15 min to complete. To compensate participants for their time, they were able to enter a lottery for a MP3 Player upon completion of the survey.

Results

Descriptive information for rejection sensitivity, condom use preference, partners' perceived condom use preferences, and condom use is shown in Table 1 for men and women in romantic and casual relationship contexts. Analyses were performed to examine whether the mean differences in condom use frequency differed among condom use preference groups. The mean levels of condom use for each group are shown in Table 2. In analyses focused on casual relationships, the "Partner Wants" group was excluded due to low sample size ($n = 9$). A one-way between-groups ANOVA was used with post-hoc contrasts using Tukey's tests to detect significant differences in condom use frequency among condom use preference groups. In romantic relationships, the ANOVA revealed significant differences between groups, $F(3, 616) = 276.50$; $p < .01$, with "Both Don't Want" participants using condoms less than all other groups ($p < .01$) and "Both Want" participants using condoms more than all other groups ($p < .01$).

Table 1 Mean levels of rejection sensitivity, condom use preference, and condom use, by relationship type and gender

Variables	Romantic relationships						Casual relationships					
	Men $n = 188$		Women $n = 458$		Total $n = 649$		Men $n = 33$		Women $n = 110$		Total $n = 144$	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Rejection sensitivity[a]	11.06	4.11	11.86	4.89	11.62	4.69	11.27	5.13	12.29	5.28	12.09	5.26
Condom use preference[b]	2.51	1.63	2.62	1.69	2.60	1.68	3.46	1.88	3.64	1.75	3.61	1.77
Partner's perceived condom use preference[b]	2.71	1.68	2.02	1.46	2.23	1.57	3.34	1.77	2.61	1.74	2.80	1.77
Condom use[b]	3.01	1.62	2.68	1.60	2.78	1.61	3.41	1.71	3.28	1.77	3.32	1.75

[a] Participants scores ranged from 1.50 to 29.50

[b] Absolute range 1–5

Table 2 Mean levels of condom use frequency according to relationship context and condom use preference group

Condom use frequency	Romantic relationships $N = 628$			Casual relationships $N = 132$		
	M	*SD*	*n*	*M*	*SD*	*n*
Both don't want	1.67_a	1.07	342	1.80_a	1.24	36
Partner wants	3.74_b	1.17	35			
Participant wants	3.40_b	1.38	100	2.58_b	1.78	43
Both want	4.56_c	0.70	140	4.75_c	0.71	52

Values with differing subscripts in each column are significantly different from one another

In casual relationships, the between-groups ANOVA showed a significant difference in condom use frequency between groups, $F(2, 130) = 63.82$; $p < .01$. Tukey's post hoc tests revealed that all groups significantly differed from one another with "Both Don't Want" participants using condoms the least ($p < .05$) and "Both Want" participants using condoms the most ($p < .01$).

Chi-square analyses were conducted to determine whether men and women were distributed significantly differently across condom use preference groups in romantic and casual relationships. Seventy percent of participants in romantic relationships were female, but results showed that men and women were distributed differently across condom use preference groups, $\chi^2(3, 656) = 49.04$, $p < .01$. Further analysis suggested that the significant chi-square result was due to the large residuals found in the "Partner Wants," "Participant Wants" and "Both Want" groups. Specifically, fewer men (15%) and more women (85%) were found in the "Participant Wants" group than expected. Additionally, the "Partner Wants" group contained more men (69%) and fewer women (31%) than expected. Finally, more men (41%) than expected were found in the "Both Want" group. The chi-square analysis was not significant for casual relationships, $\chi^2(2, 144) = 3.44$.

Multivariate Analyses

Multiple linear regression analyses were used to test the main and interactive effects of rejection sensitivity and condom use preference group on condom use frequency. Dummy variables were created to allow the condom use preference groups to be entered into the regression analyses. The "Both Don't Want" group was the referent group. Analyses were run separately for romantic and casual relationship contexts. At the first step, the main effects of rejection sensitivity and the three condom use preference group dummy variables were entered, along with gender and other birth control use as control variables. At the second step, the interaction terms were entered. The interaction terms were created by multiplying dummy variable codes with rejection sensitivity *z* scores.

Regression Analysis for Romantic Relationship Contexts

Statistics for the regression analysis are shown in Table 3. At the first step, all three dummy variables and not being on another form of birth control significantly predicted more condom use. Neither rejection sensitivity nor gender significantly predicted condom use. At the second step, the dummy variables and other contraceptive use remained significant predictors of condom use. The dummy variable for the "Participant Wants" group significantly interacted with rejection sensitivity.

To test the nature of the interaction, condom use was regressed onto rejection sensitivity separately for each condom use preference group with gender and birth control retained as control variables. Among "Participant Wants" participants, rejection sensitivity significantly predicted less condom use ($B = -.05$; $SE\ B = .03$; $\beta = -.20$; $p < .05$). Among

Table 3 Summary of hierarchical regression analysis for variables predicting condom use among sexually experienced individuals in romantic relationships ($N = 617$)

Variable	Model 1			Model 2		
	B	*SE B*	*β*	*B*	*SE B*	*β*
Gender[a]	−.06	.09	−.02	−.06	.09	−.02
Birth control use[b]	−.73	.10	−.20**	−.72	.10	−.20**
Rejection sensitivity	−.01	.01	−.02	−.01	.01	−.01
"Partner Wants" group	1.79	.19	.25**	1.91	.20	.27**
"Participant Wants" group	1.61	.12	.36**	1.63	.12	.37**
"Both Want" group	2.66	.11	.68**	2.65	.11	.68**
RS × "Partner Wants"				.32	.18	.05
RS × "Participant Wants"				−.26	.12	−.06*
RS × "Both Want"				−.03	.12	−.01
Adjusted R^2	.62			.63		
F for change in R^2	164.59**			3.15*		

[a] Females = 0, males = 1

[b] Non-use = 0, use = 1

* $p < .05$; ** $p < .01$

Table 4 Summary of hierarchical regression analysis for variables predicting condom use among sexually experienced individuals in casual relationships ($N = 131$)

Variable	Model 1			Model 2		
	B	*SE B*	*β*	*B*	*SE B*	*β*
Gender[a]	.11	.27	.02	.16	.26	.04
Birth control use[b]	−.21	.23	−.06	−.17	.22	−.05
Rejection sensitivity	−.08	.02	−.22**	−.03	.04	−.07
"Participant Wants" group	.87	.29	.23**	.86	.28	.22**
"Both Want" group	2.90	.28	.79**	2.89	.28	.79**
RS × "Participant Wants"				−.69	.28	−.23*
RS × "Both Want"				.02	.28	.01
Adjusted R^2	.53			.56		
F for change in R^2	30.57**			4.98**		

[a] Females = 0, males = 1

[b] Non-use = 0, use = 1

* $p < .05$; ** $p < .01$

"Partner Wants" participants, rejection sensitivity did not significantly predict condom use, but the coefficient was in the expected direction with higher levels of rejection sensitivity relating to more condom use, which would be the compliant response for this group ($B = .06$; $SE\ B = .04$; $\beta = .26$; $p > .05$). For the congruent groups, rejection sensitivity did not significantly predict condom use in the "Both Don't Want" group ($B = .01$; $SE\ B = .01$; $\beta = .01$; $p > .05$) or the "Both Want" group ($B = -.02$; $SE\ B = .01$; $\beta = -.11$; $p > .05$).

Regression Analysis for Casual Relationship Contexts

Statistics for the regression analysis are shown in Table 4. At the first step, rejection sensitivity significantly predicted less condom use and the two dummy variables significantly predicted more condom use. Neither gender nor birth control use was significant. At the second step, the interaction between the dummy variable, reflecting "Participant Wants" group membership, and rejection sensitivity was significant. This indicated that rejection sensitivity interacted with being in the group in which the participants wanted to use condoms more often than they thought their partners did. Rejection sensitivity was not a significant predictor of condom use at the second step, once the interaction terms were taken into account.

Regression analyses were run separately for the three condom use preference groups to investigate the direction of the interaction with rejection sensitivity. Condom use was regressed onto rejection sensitivity in each group with gender and other birth control use retained as control variables. Rejection sensitivity significantly predicted less condom use in the "Participant Wants" group with a standardized beta of −.57 ($B = -.18$; $SE\ B = .04$; $p < .01$). Rejection sensitivity did not predict condom use among participants in the "Both Don't Want" group ($B = -.02$; $SE\ B = .04$; $\beta = -.08$; $p > .05$) and the "Both Want" group ($B = -.03$; $SE\ B = .02$; $\beta = -.17$; $p > .05$).

Discussion

The current study examined the link between rejection sensitivity and condom use, while taking into consideration condom use preferences, to determine in what circumstances rejection sensitivity predicted condom use behavior. In both romantic and casual relationship contexts, the link between rejection sensitivity and condom use was moderated by condom use preference group membership. Specifically, higher levels of rejection sensitivity predicted less condom use in the "Participant Wants" groups. In groups where the participants' condom use preferences were congruent with their perceptions of their partners' preferences, the link between rejection sensitivity and condom use was not significant. In the remaining "Partner Wants" group, rejection sensitivity did not significantly predict condom use but the relationship was in the predicted direction with higher levels of rejection sensitivity being associated with more condom use in romantic relationships. It is likely that the small sample size of the "Partner Wants" group resulted in insufficient statistical power to detect a significant relationship. In summary, rejection sensitivity predicted sexually compliant condom use behavior when there was a discrepancy between the participants' condom use preferences and their partner's condom use preferences.

The results of the current study lend support to the argument presented by Mischel and Shoda (1995) that a personality disposition should relate to behavior particularly in situations which act as a trigger to the personality disposition. In this case, rejection sensitivity related to behavior in situations that were ripe for conflict and rejection, that is, that the participant perceived they had different preferences for condom use from their partner. The results also suggest that the model of rejection sensitivity (Downey et al., 1999) can be suitably applied to compliant sexual behavior in both romantic and casual relationship contexts among young adults.

Young and Furman (2008) found that rejection sensitivity predicted sexual victimization and suggested that youths who were rejection sensitive may comply with their partner's pressures due to fear of rejection. The current results were consistent with this suggestion and extend research on rejection sensitivity to condom use in both casual and romantic relationship contexts. The current results also add to the literature on attachment anxiety and condom use. Past research findings on attachment anxiety and condom use have been mixed with one study finding attachment anxiety to be associated with more condom use (Bogaert & Sadava, 2002), but most studies have found attachment anxiety to be associated with less condom use (Feeney et al., 1999, 2000; Kershaw et al., 2007). The current results are consistent with previous research finding attachment anxiety to be associated with less condom use, at least when participants wanted to use condoms. Young and Furman found that rejection sensitivity, but not preoccupied attachment predicted sexual victimization, even though the two constructs reflecting romantic views overlapped. They suggest that rejection sensitivity may be more relevant to situations where setting boundaries with a partner is of concern as the measure of rejection sensitivity focuses on acceptance from a partner, whereas preoccupied attachment focuses on the emotional dependence upon a partner. The current results suggest that there is a link between fears of rejection and condom use that is dependent upon the context of the sexual encounter, including the perceived condom use preference discrepancy of the dyad. Thus, the results support the idea that rejection sensitivity as a construct is relevant to boundary setting with a sexual partner.

We also note that the results for condom use varied among condom use preference groups. There were clear differences in the frequency of condom use among the condom use preference groups. Specifically, when both partners wanted to use condoms (the "Both Want" group), condom use was close to every time. In contrast, in the "Both Don't Want" group, condom use was close to never. Condom use fell around the mid-range point on the frequency of use scale in the groups where one partner wanted to use a condom, but the other did not (the "Participant Wants" and "Partner Wants" groups). Such results may help inform health promotion efforts. Perhaps in couples where neither partner wishes to use condoms, a more realistic goal than increasing condom use may be to ensure that both partners have STI and HIV tests before making the decision to have unprotected sex. For couples with both partners wanting to use condoms, attention could be paid to the maintenance of condom use. Finally, if an individual belongs to an incongruent group, health promotion efforts could be directed towards respecting one's partner's preferences to use condoms or in working towards building communication skills about negotiating safer sex with one's partner.

One caveat to this health promotion approach is that participants were coded as belonging to a condom use preference group based on their *perceptions* of their partner's condom use preferences. We did not have the data to test the accuracy of their perceptions. Therefore, it is possible that some individuals who were assigned to the "Participant Wants" group had partners who wanted to use condoms more frequently than they imagined. In such cases, it may be counterproductive to educate individuals about skills to convince their partner to use a condom (or to accept that they wear a condom). To a rejection-sensitive person, the notion that their partner may indeed be unwilling to use a condom, and that they must convince them, would likely be an anxiety provoking task that may be too much for them to attempt. Perhaps it is simply their perceptions of their partner's condom use preferences that are inaccurate. If so, educating individuals that others most likely want to use condoms as often as they do themselves (Edwards & Barber, 2009) may prove a more effective strategy for promoting safer sex.

Correcting misperceptions about partner condom use preferences may be particularly important for women. The current study showed that women were overrepresented in the "Participant Wants" group. If women underestimate the frequency at which their partners want to use condoms, they may actually belong to the "Both Want" group, rather than the "Participant Wants" group. In such cases intervention could be directed towards correcting the misperception that many men need to be convinced to wear condoms by educating women that generally men want to use condoms as often as women do (Edwards & Barber, 2009).

Several limitations of the current study should be noted. The number of participants recruited in the current study was satisfactory, but few men with experience in casual relationships participated. As such, the current results may not generalize well to young men's casual relationship experience. There was also a lack of participants in the incongruent group where participants wanted to use condoms less than they thought their partner did ("Partner Wants" group). This may have contributed to the inability to detect a significant relationship between rejection sensitivity and condom use in this group. Finally, the measure of condom use was a single-item measure, although alternatives are rare in this field of research (Noar, Cole, & Carlyle, 2006). The measure of condom use preferences also may be limited by the fact that participants were retrospectively reporting their condom use preferences. It may be that participants adjusted their condom use preferences to align with their behavior, consistent with cognitive dissonance theory (Festinger, 1957).

Other recommendations for investigating condom use behavior, however, have been followed (Noar et al., 2006). This included controlling for birth control use and measuring condom use specific to one particular partner which is not often taken into account when predicting condom use, according to the review by Noar et al. The other main strength of the current study was that condom use was examined as an interpersonal behavior, rather than just as health behavior.

In summary, the current results extend the literature on rejection sensitivity by applying the model to condom use behavior. Rejection sensitivity was found to predict condom use only when condom use preferences were perceived to be discrepant from one's partner. Thus, the results highlight the importance of taking into account the condom use preferences of individuals and their perceived partners' preferences when predicting condom use. When using an individual difference variable to predict behavior, it is important to take the situation into account and this claim has been supported in determining the link between rejection sensitivity and condom use.

Appendix

Definitions of relationship types presented to participants

Relationship type	Relationship sub-type	Presented definition
Romantic relationships	No sub-types	This is when you have had a boyfriend or girlfriend. You are both aware that you are in an actual relationship together. You may or may have not have had sex
Other relationships		This will include a variety of casual relationships and casual sex encounters including:
	A one night stand	Sex with someone you have just met
	Sex with a friend	You are primarily friends but also have had sex
	Sex with an ex-partner	Having sex after you have broken up
	Casual relationships	You have been seeing someone casually, but the relationship may not be "official". You may be still seeing other people or may be unsure about where you stand with each other. You may or may not have had sex

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179–211.

Amaro, H. (1995). Love, sex, and power: Considering women's realities in HIV prevention. *American Psychologist, 50*, 437–447.

Bogaert, A. F., & Sadava, S. (2002). Adult attachment and sexual behavior. *Personal Relationships, 9*, 191–204.

Bromnick, R., & Swinburn, P. (2003). Young people's social representations of a sexual experience. *Journal of Adolescence, 26*, 375–379.

Collins, N. L., & Read, S. J. (1990). Adult attachment, working models, and relationship quality in dating couples. *Journal of Personality and Social Psychology, 58*, 644–663.

Crockett, L. J., Raffaelli, M., & Moilanen, K. L. (2003). Adolescent sexuality: Behavior and meaning. In G. R. Adams & M. D. Berzonsky (Eds.), *Blackwell handbook of adolescence* (pp. 371–392). Malden, MA: Blackwell Publishing.

Downey, G., Bonica, C., & Rincon, C. (1999). Rejection sensitivity and adolescent romantic relationships. In W. Furman, B. B. Brown, & C. Feiring (Eds.), *The development of romantic relationships in adolescence* (pp. 148–174). New York: Cambridge University Press.

Downey, G., & Feldman, S. I. (1996). Implications of rejection sensitivity for intimate relationships. *Journal of Personality and Social Psychology, 70*, 1327–1343.

Edwards, G. L., & Barber, B. L. (2009). Women may underestimate their partners' desires to use condoms: Possible implications for behaviour. *Journal of Sex Research*. doi:10.1080/00224490902867863.

Ellen, J. M., Cahn, S., Eyre, S. L., & Boyer, C. B. (1996). Types of adolescent sexual relationships and associated perceptions about condom use. *Journal of Adolescent Health, 18*, 417–421.

Feeney, J. A., Kelly, L., Gallois, C., Peterson, C., & Terry, D. J. (1999). Attachment style, assertive communication, and safer-sex behavior. *Journal of Applied Social Psychology, 29*, 1964–1983.

Feeney, J. A., Peterson, C., Gallois, C., & Terry, D. J. (2000). Attachment style as a predictor of sexual attitudes and behavior in late adolescence. *Psychology & Health, 14*, 1105–1122.

Festinger, L. (1957). *A theory of cognitive dissonance*. Oxford: Row, Peterson.

Furman, W., & Simon, V. A. (1999). Cognitive representations of adolescent romantic relationships. In W. Furman, B. B. Brown, & C. Feiring (Eds.), *The development of romantic relationships in adolescence* (pp. 75–98). New York: Cambridge University Press.

Gebhardt, W. A., Kuyper, L., & Greunsven, G. (2003). Need for intimacy in relationships and motives for sex as determinants of adolescent condom use. *Journal of Adolescent Health, 33*, 154–164.

Juarez, F., & Martín, T. C. (2006). Safe sex versus safe love? Relationship context and condom use among male adolescents in the Favelas of Recife, Brazil. *Archives of Sexual Behavior, 35*, 25–35.

Kershaw, T. S., Milan, S., Westdahl, C., Lewis, J., Rising, S. S., Fletcher, R., et al. (2007). Avoidance, anxiety and sex: The influence of romantic attachment on HIV-risk among pregnant women. *AIDS and Behavior, 11*, 299–311.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*, 246–268.

Noar, S. M., Cole, C., & Carlyle, K. (2006). Condom use measurement in 56 studies of sexual risk behavior: Review and recommendations. *Archives of Sexual Behavior, 35*, 327–345.

Rosenthal, D. A., Hall, C., & Moore, S. M. (1992). AIDS, adolescents, and sexual risk taking: A test of the Health Belief Model. *Australian Psychologist, 27*, 166–171.

Sheeran, P., Abraham, C., & Orbell, S. (1999). Psychosocial correlates of heterosexual condom use: A meta-analysis. *Psychological Bulletin, 125*, 90–132.

Young, B. J., & Furman, W. (2008). Interpersonal factors in the risk for sexual victimization and its recurrence during adolescence. *Journal of Youth and Adolescence, 37*, 297–309.

Arch Sex Behav (2010) 39:1389–1401
DOI 10.1007/s10508-009-9592-5

ORIGINAL PAPER

Pornography, Normalization, and Empowerment

Martin S. Weinberg · Colin J. Williams · Sibyl Kleiner · Yasmiyn Irizarry

Received: 25 November 2008 / Revised: 13 November 2009 / Accepted: 13 November 2009 / Published online: 3 February 2010

Abstract Opponents and proponents of erotic representations (referred to hereafter as "pornography") have described the effects of pornography from their perspective. Little, however, has been done in the way of research to investigate these claims from the consumer's point of view. This especially has been so regarding the positive impact of such consumption on a person's sex life. Using a study group of 245 college students, we examined this question in a framework of scripting theory. We wanted to see whether viewing pornography appeared to expand sexual horizons through normalization and facilitate a willingness to explore new sexual behaviors and sexual relationships through empowerment. The data supported this viewpoint and further showed the effects to be mediated by gender and sexual preference identity. They suggested, however, that established scripts were extended rather than abandoned. We conclude with connections between our findings and the widespread viewing of pornography in contemporary society.

Keywords Pornography · Sexual scripts · Gender · Sexual preference · Normalization · Empowerment

M. S. Weinberg (✉) · S. Kleiner · Y. Irizarry
Department of Sociology, Indiana University, 1020 E. Kirkwood Avenue, Bloomington, IN 47405, USA
e-mail: weinberm@indiana.edu

C. J. Williams
Department of Sociology, Indiana University-Purdue University at Indianapolis, Indianapolis, IN, USA

Introduction

Pornography is intertwined in the sexualities of many people. Although the consumption of pornography in the United States is difficult to measure, Rich (2001), has estimated that $10-$14 billion annually is earned from "video pornography…porn networks and pay-per-view movies on cable and satellite, Internet Web sites, in-room hotel movies, phone sex, sex toys and…magazines….[P]ornography is a bigger business than professional football, basketball and baseball put together." Rich goes on to point out that: "People pay more money for pornography in America in a year than they do on movie tickets, more than they do on all the performing arts combined" (p. 462).

Despite pornography's central place in American entertainment, it still is the focus of great social controversy. On the one hand, pornographic images have been attacked as a source of social and moral evil and, on the other, have been defended as contributing to sexual liberation. Attacks on pornography come from both religious conservatives, who see pornography as encouraging impersonal, casual sex, and from feminists who claim that pornography makes the plight of women worse by encouraging misogynist, sexist, and patriarchal attitudes. This is seen as leading to the objectification of women, a loss of respect for them, and to rape and violence against them (Dworkin, 1981; Paul, 2005; Porn, Feminism, and the Meese Report, 1987).

Although the attacks on pornography have centered on the fear of its supposed negative effects, there has been little empirical investigation of whether and how such effects may occur. Moreover, when this has been attempted, results that do not support pornography's negative image are often ignored or argued away. For example, the in-depth research of the 1970 Commission on Obscenity and Pornography showed no causal link between pornography and violence (e.g., Goldstein, Kant, Judd,

Rice, & Green, 1971). Failure to find negative effects in the 1970 research was displeasing to those with an anti-pornography agenda (Rubin, 1984), and for whom being anti-porn was part of a wider conservative agenda. This shows how anti-pornography sentiments do not necessarily rest on empirical evidence. Rather, a negative effect is often believed to be the case whether or not supporting evidence exists—especially if this fits a particular world view.

This can also be the case for those who take a liberal point of view toward pornography. Defenders of pornography have been firm in asserting the benign or positive effects of pornography without strong empirical evidence to support their claims. Other arguments that are common by defenders are directed at the issue of censorship and the fact that the anti-pornography focus is primarily "violent porn" and "child porn" rather than what the majority of adults view (Klein, 2006). One interesting change in the liberal view is the increasing support it receives from feminists. Earlier (second wave) feminists were characterized by their anti-pornography platform, but a generation change has give rise to a "third wave feminism" (Bailey, 1997) in which female sexuality is re-interpreted as a means of self-definition and expression (see Alfonso & Triglio, 1997). Thus, Beggan and Allison (2002) characterize the movement as "…expand(ing) the boundaries in which women are free to express their sexuality" (p. 104)—which includes their consumption of pornography (Chapkis, 1997; McElroy, 1995). The change is seen as "a method of empowerment for women" to control their own sexuality (Beggan & Allison, 2002, pp. 106–107). These issues of boundaries and empowerment will be considered later in this article. At this point, we note the emergence of the claim that pornography is educational, that it allows individuals to explore in a safe way new forms of sex (Barbour, 1995; Duncan & Donnelly, 1991; Duncan & Nicholson, 1991), especially for women, whose sexual expression historically has been discouraged (Palac, 1995; Tanenbaum, 2004; Tiefer, 2004).

It is difficult to discern a "winner" in these battles. The technological revolution, particularly the Internet, has increased the availability of pornographic images to more people than has existed at any other time. At the same time, anti-pornography forces have worked hard to pass laws to restrict the viewing of pornography (especially child pornography). What remains true is that there is insufficient empirical research on many of these issues. The work done on mainstream pornography has usually looked at its possible negative effects (for a review of this literature, see Davis & Bauserman, 1993). Little work has been done on the supposed "educational" effect as claimed by some of the above commentators. Research on this topic would fit neatly into one of the major schemas in sex research, "sexual scripts," offered by Gagnon and Simon (1973), which centers on the acquisition of cognitive schemes that connect to sexual desires and behaviors. Extant research illustrates the plausibility of such an approach and provides a direction to follow. Some studies suggest that while pornography use does not directly increase sexual satisfaction (Štulhofer, Buško, & Landripet, 2008), both men and women attribute positive influences to pornography (Ciclitira, 2004; Hald & Malamuth, 2008; Loftus, 2002; Shaw, 1999). These tend to be effects that represent a qualitative broadening of sexual horizons, such as learning new forms of sexual behavior or finding new resources for fantasy construction (Tiefer, 2004). Several recent studies also replicate older findings that pornography use is associated with quantitatively greater levels of activities, such as masturbation, as well as overall numbers of partners (Janghorbani, Lam, & The Youth Sexuality Study Task Force, 2003; Laumann, Gagnon, Michael, & Michaels, 1994; Lewin, 1997; Traeen, Sorheim-Nielsen, & Stigum, 2006), suggesting an expansion of sexual horizons. Given pornography's potential as a resource for exploring sexuality (Attwood, 2005), greater attention to the correlates of pornography use is warranted.

Of course, we do not want to propose a simplistic model. The effects of viewing pornography can be seen as wholly absent, wholly negative, wholly positive, or a mixture of positive and negative. Examples of these different points of view can be found in Shaw (1999) for the putative effects on women and Hardy (1998) for the putative effects on men. The movement from exposure to pornography toward actual sexual experimentation may be complex. For example, people who try out pornographic scenarios can end up feeling ridiculous (Morrison & Tallack, 2005). The unrealistic portrayal of sexuality was also noted by Paul (2005). In this article, we are concerned only with examining any positive effects while also recognizing that a person may feel it has had negative effects or no effects at all (cf. Parvez, 2006).

Nor do we expect the effects of pornography viewing to be uniform, but rather to be mediated by other variables (for one list of such variables, see Hald & Malamuth, 2008). In the present research, we will consider gender and sexual preference identity as mediating variables.

Gender has been shown to be a consistent predictor of both the consumption of pornography (Janghorbani et al., 2003) and its effects (Hald & Malamuth, 2008). Gender has also been shown to relate to the appeal of pornographic themes—with men on average preferring more hard-core and women, more soft-core, themes (Hald, 2006; cf. Janssen, Carpenter, & Graham, 2003). Compared to women, men have also been shown to find pornography more sexually exciting and enhancing (Traeen, Spitznogle, & Beverfjord, 2004).

The second mediating variable, sexual preference identity, has been neglected in mainstream pornography research (Morrison & Tallack, 2005). Yet, it is important as it allows for a consideration of the effects of gender nonconformity on sexuality. Thus, non-heterosexual people are freer from traditional gender roles; for example, non-heterosexual women are less subject than heterosexual women to the expectations of "ideal femininity" (cf. Gordon, 2002) and any double standard

attached to it. In this instance, it could mitigate the traditional expectation that "good girls" do not show too much interest in pornography or sexual variety (cf. Tanenbaum, 2004). Moreover, Stein (1997) described how younger lesbians in the 1990s created a more sexualized lesbian subculture in which they did not model themselves after the more sexually restricted lesbians of the past. This led to more sexual experimentation, including the increased use of pornography, which can help to explain the adoption of more expansive sexual scripts (heralded in the lesbian publication, *On Our Backs*) with a wider variety of sexual acts and partners seen as appropriate (cf. Califia, 1980, 1994).

Theoretical Considerations

As previously mentioned, the influence of pornography on shaping sexuality can be conceptualized in terms of the development of a person's "sexual scripts" (Gagnon & Simon, 1973). From this perspective, persons are engaged in defining situations as sexual through developing "organized cognitive [sexual] schemas" (Gagnon, 1990, p. 6), which they have learned and elaborated. Pornography is an important source through which individuals can acquire or reinforce sexual scripts (for one approach using script theory, see Štulhofer et al., 2008). It can present appealing views of a variety of sexual behaviors as well as portraying the pleasure possible from activities like oral sex, anal sex, a variety of coital positions, etc. Nothing about this process, however, is automatic. As Simon and Gagnon (1987, p. 365) point out, "Scenarios have to be tried on for confirmation and possibly modified where stress or discomfort on either interpersonal or intrapsychic levels is experienced." In addition, the scripting schema does not deal with the processes that are involved in making a script "work."

We believe it is unassailable that the central function of pornography is the creation or enhancement of sexual fantasy and/or arousal. That is, it presents bodies, behaviors, and situations in a way that is intended to sexually inspire or excite the viewer, regardless of whether such bodies, behaviors, and situations would be available or even desirable for the viewer to experience in real life. Thus, the discussion of the social utility of pornography centers on the acceptability of sexual arousal and what it leads to. The question becomes: How does pornography come to have a socializing effect on its viewers through the development of sexual scripts?

We offer the following theoretical considerations on how sexually explicit images could have a positive effect on a person's sexuality. Pornography can promote a "sensual slide" from everyday life into an erotic reality through the erotic construction of time, space, persons, and situations in sexual depictions (Davis, 1983). It provides both the presentation of idealized bodies and the opportunities to visualize them in a variety of sexual situations performing an abundance of sexual acts—many of which could be considered to transgress mainstream culture. Collectively, this expansive set of sexual possibilities presented in pornography has been called a "pornotopia" (Marcus, 1966; Peckham, 1969; Williams, 1999). One key to porn's popularity is a "validation of the viewer's vision of erotic abundance" (Klein, 2006, p. 138).

We consider two major processes as operative with regard to the effects of viewing pornography. The first we refer to as "normalization." The more frequently a person enters the world of pornotopia, the more s/he will view a variety of sexual behaviors as being normal. That is, what may have once been seen as odd becomes viewed as a variation of normal behavior—in this instance, "normal" sexual behavior (cf. Rubington & Weinberg, 1996), which has a decreasing capacity to shock or offend (for a different conceptualization of this in terms of "satiation," see Zillmann & Bryant, 1984).

The second process that the consumption of pornography can promote is a sense of erotic empowerment, the ability not only to create or alter sexual scripts but also a desire to act on them. Thus, in the words of Palac (1995), "Once I figured out how to use porn and come…my life was irrevocably changed…. For the first time in my life, I felt sexually autonomous" (pp. 34–35). According to Klein (2006), the "paradigm of pornography's truths is what sex therapists try to get couples to understand…[that] the keys to satisfying sexual relationships are self-acceptance and self-empowerment…" (p. 137). Often this is done by projecting oneself into the scene as a participant (Hardy, 1998; Loftus, 2002); hence, the strong association between pornography and self-masturbation. Persons become empowered by the fantasy of "easy sex," free from the real life concerns such as anxieties over attractiveness, performance and, especially, entangling relationships (see a description of the parameters of "easy sex" in Weinberg & Williams, 1975). A feeling of such erotic empowerment also can transfer to a person's actual sex life. What is learned from the sexual scenes experienced in pornography can give a person both the interest in, and the confidence to experiment with, sexual behaviors s/he had previously never tried. While we will examine such effects in a positive framework, they can also be considered as negative consequences (e.g., for adolescents and young adults, see Zillmann, 2000).

We do not, however, see people as cultural dopes (Garfinkel, 1967). We do not believe that people indiscriminately and automatically take on the behaviors they see in erotic representations, but rather select those that fit in most easily with other aspects of their social and sexual socialization (on the domestication of pornography, see Juffer, 1998). For example, the social location of our study group may well serve to limit the degree to which pornography modifies their sexual lives. Thus, we suggest that they may increase the variety of sexual behaviors they engage in by integrating them into their existing sexual scripts without much damage to core aspects of such scripts.

Research Questions

To summarize the above, we present the following research questions:

Normalization question. Is the greater frequency of pornography consumption associated with expanding the boundaries of what is considered acceptable sexual behavior?
Empowerment question. Is the greater frequency of pornography consumption associated with engaging in a greater variety of sexual behaviors and is it also associated with an expansion of the boundaries of who would be considered an acceptable sexual partner?
Gender question. Are the answers to the above questions mediated by the gender of the consumer so that such effects are more likely to appear for one gender rather than another?
Sexual preference identity question. Are the answers to the above questions mediated by the sexual preference identity of the consumer so that such effects are more likely to appear for one sexual preference identity than another?

Method

Participants

The data used for the research come from a study of students at a midwestern state university. Although a study group of college students limits the social class background and the age range of participants, pornography use is more common among younger adults and among those with greater education (Buzzell, 2005). As the purpose of this study was to examine how the use of sexually explicit materials was related to sexual attitudes and behaviors, for our purposes it was desirable to analyze the data from a study group that contained a higher concentration of pornography users than would be found in the general population.

Two data collections were conducted: In the first (labeled "the quantitative research"), we used closed-ended questions in a self-administered questionnaire to discover whether associations existed between the frequency of viewing pornography and the variables reflecting normalization and empowerment. The study group was comprised of students who were recruited from sociology courses containing 50 or more students. In the initial phase of data collection, gathering a student study group that contained at least as many women as men and represented a variety of sexual identities was of great concern because the relationship between pornography and women's sexuality, and more so gay/lesbian/bisexual pornography use, have been largely ignored in prior empirical research (Plante, 2006). To aid in the recruitment of gay, lesbian, and bisexual identified participants, we described the research at meetings of such students and sent announcements to university gay, lesbian, and bisexual email distribution lists. All students were offered the choice of either a $10 phone card or a $10 Starbucks gift certificate for their participation in the study.

Of the 172 participants in the study group obtained in these ways, 101 were women and 71 were men. Sixty-nine women identified themselves as heterosexual and 32 as non-heterosexual (14 identified as lesbian and 18 as bisexual). Among the men, 52 identified themselves as heterosexual and 19 as non-heterosexual (16 identified as gay and 3 as bisexual). The ages of study participants ranged from 18 to 34 years with a mean of 21.3. Two-thirds of the persons in the study group were between the ages of 18 and 22. Ethnically, 70% described themselves as White, 6% Latino/Hispanic, 15% African American, and 10% in some other way. Finally, using parental education as an indicator of social class background, over half of both mothers and fathers had a college degree. Although we found a greater proportion of African Americans among the heterosexual women and a slightly older average age for the non-heterosexual men, these differences in social characteristics by gender and sexual preference identity did not prove to be significantly related to the responses given to the questions.

The second data collection (labeled "the qualitative research") was conducted several semesters after the first one. We believe that any study of the effects of pornography should include a strong qualitative aspect in which the research participants themselves provide their own interpretations of the experience (see arguments for this by Attwood, 2005). Thus, we wanted to see, in the participants own words, how viewing pornography can relate to normalization and empowerment which, in turn, can relate to one's sexual life. We provided an open-ended questionnaire to students in their classes to voluntarily and anonymously complete. This additional phase of data collection was intended to obtain a more in-depth view of the putative effects of viewing pornography from firsthand accounts of viewers. A total of 73 students completed and returned these questionnaires. In order to keep the questionnaire short, gender and sexual preference identity were the only demographics requested. Of the 73 participants, 74% were women and 26% were men (similar to the gender distribution in the classes). Ninety-three percent of the women identified as heterosexual, 4% as lesbian, and 6% as bisexual. Sixty-three percent of the men defined as heterosexual, 21% as gay, 11% as bisexual, and one as queer.

Procedure

In the quantitative study, participants were interviewed by six well trained undergraduate interviewers (three women and three men) in a larger study covering a number of topics in addition to the topic of this article (Weinberg & Williams, 2005). Participants were matched to an interviewer by gender, and were asked whether they would rather be interviewed by a heterosexual or non-heterosexual person. Throughout the training,

pre-testing, and data collection stages, all interviews were carefully reviewed and any problems discussed with the interviewers. We also talked to study participants after each interview; all rated their interviewer as competent and the experience as enjoyable. To gather data for the part of the study on pornography viewing and its relationship to sexual variables, we used a closed-ended self-administered questionnaire that was handed to the study participant by the interviewer, answered privately by the participant, put in and sealed in a large envelope we gave to them, and then handed back to the interviewer. This questionnaire contained items meant to measure attitudes toward the appeal of a variety of sexual practices, the number and type of sexual experiences engaged in during a set time period, and the social relationship between the study participant and his/her sex partner(s).

In the qualitative study, students who wished to participate in the study took a questionnaire that was available at the end of class time, completed it outside of class, and returned it anonymously to a collection box in the main office of the Sociology Department.

Measures

The Quantitative Study

For the quantitative part of the research, principal components factor analyses (varimax rotation) were carried out to aid in the construction of composite measures. One composite was related to pornography viewing and was comprised of two items: In the last 12 months, how many times have you viewed an X-rated movie or video? In the last 12 months, how many times have you viewed soft or hard-core porn in magazines or on-line (alpha = .75). There were 11 response categories for each item ranging from 0 to over 100. Means and *SD* for the each of the four gender-sexual preference identity subgroups are shown in Table 1. The only significant difference in pornography viewing was across gender ($p \leq .05$)—with men viewing pornography more frequently than women.

Another set of variables was constructed from a list of items that were rated in terms of their appeal. The study participants were informed that these ratings need not necessarily correspond to actual behaviors or plans to engage in such behaviors. The response categories were: 1 = very appealing, 2 = appealing, 3 = neither appealing nor unappealing, 4 = unappealing, 5 = very unappealing. Scores were reversed so that a higher value indicated greater appeal. These composite measures based on the factor analyses noted above (factor loadings, Cronbach's alpha, and the items contained in the composites are provided in the Appendix) were as follows: Composite #1: Appeal of Using a Vibrator/Sex Toy; Composite #2: Appeal of Oral-Genital Activity; Composite #3: Appeal of Anal Activity; Composite #4: Interest in Third Parties (e.g., watching others, engaging in group sex).

In regards to reporting actual behavior, we asked about the number of times they had engaged in particular acts—with 11 response categories ranging from 0 to over 100. For acts of partnered sex, in general, we also asked the number of persons they had engaged in the act within the last 12 months—with 9 response categories ranging from 0 to over 30. For anal sex, we asked this with regard to their lifetime. Based on principal components factor analyses, the following composites (which combined number of partners and frequency) were constructed for sexual behaviors: Composite #1: Self-Masturbation; Composite #2: Manual Sex; Composite #3: Oral Sex; Composite #4: Coitus; Composite #5: Anal Sex (see Appendix for factor loadings, Cronbach's alpha, and detailed item listings).

The final measures were about the social relationship participants had with their sex partner(s). The categories of sex partners presented to study participants included: didn't know or had met that day (e.g., Spring Break); knew as a friend or acquaintance but hadn't "dated"; were on a first "date" with; had previously "dated" (not the first date or was an ex-partner or boyfriend or girlfriend); were significantly involved with at the time (partner/boyfriend/girlfriend). After questions on the number of sex partners for various sexual activities (in the last year, except for anal sex where we used "lifetime"), we then asked how many

Table 1 Means and *SD* for pornography use (past 12 months) by gender and sexual preference identity

Frequency of viewing pornography	Women				Men				Combined	
	Heterosexual		Non-heterosexual		Heterosexual		Non-heterosexual			
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
No. times viewed X-rated film or video[a]	2.06	1.85	2.59	1.66	5.38	3.05	6.11	3.03	3.61	2.90
No. times viewed magazines or online[a]	2.20	1.41	3.44	2.78	6.15	3.51	7.37	2.75	4.20	3.49
Pornography composite	2.13	1.79	3.02	1.93	5.77	2.74	6.74	2.55	3.90	2.87
N	69		32		52		19		172	

All means statistically significant across gender at $p \leq .05$

[a] Individual Items: 0, 1, 2, 3 = (3–5), 4 = (6–10), 5 = (11–15), 6 = (16–20), 7 = (21–30), 8 = (31–50), 9 = (51–100), 10 ≥ 100

of these persons fell into each of the above social relationship categories. Thoits (1995) found that recall error increased the most after a 12-month period. Thus, we restricted our time period to "the last 12 months" with only one exception—the measure for anal sex. We extended the time period for this event to "lifetime" because of its relatively low incidence and the assumption that the young age of the study participants would make it fairly easy for them to recall the approximate number of such experiences.

The Qualitative Study

The qualitative study contained the following open-ended questions:

1. How has pornography affected your attitudes toward specific sexual acts (e.g., to find them more or less appealing)? Describe what this change in attitude/appeal has been and how pornography came to change it.
2. How has viewing pornography affected your actual sexual behaviors (e.g., the particular sexual acts that you engage in)? Describe what this change in sexual behavior has been and how pornography came to change it.

Analysis

For the quantitative part of the research, ordinal logistic regressions were run to explore the associations between frequency of pornography viewing (composite) and the other variables. Statistically significant beta coefficients between the frequency of pornography viewing and any of the other variables do not necessarily indicate causality in a particular direction. Our interpretation of the direction of these relationships is based on results from the qualitative data collection, which described the perceived effects from viewing pornography. This does not mean that in some cases the direction could not be a different or a more interactive one.

In order to account for the possibility of other variables entering the relationship between pornography viewing and the other variables, in the quantitative part of the research, we controlled in all of our models for demographic factors other than gender and sexual preference identity that could be confounding variables—viz., year in college, age, and religiosity (and although there was a fairly strong correlation between year in college and age, no collinearity was found in any of the models using variance inflation factors). For this study group, controlling for these variables did not change the results.

For the qualitative research, answers to questions 1 and 2 were read over a number of times and the data coded in terms of the themes of normalization and empowerment. We coded statements as illustrating "normalization" when they described how viewing pornography had made the study participant see certain behaviors as more "normal" or "less strange" or "less perverted" then they had before such viewing. We coded statements as illustrating "empowerment" when the participant noted how viewing pornography increased her/his confidence or courage in trying a particular sexual behavior. There was almost complete agreement between the two authors who did the coding; in the few cases where there was not, after discussion, we decided that these statements were too vague to be coded as one or the other. The qualitative material was used to interpret the casual direction underlying the associations found in the quantitative study.

Results

Does Pornography Consumption Increase the Appeal of a Variety of Sexual Behaviors?

As shown in Table 2, for all four gender-sexual preference identity groups, we found positive associations between the greater viewing of pornography and a more expansive sexuality in terms of what were considered appealing acts. For all four groups, the frequency of viewing pornography was related to the appeal of the sexual presence of a third party (i.e., watching other people engage in sexual activity both in porn videos and watching them in-person, and in having sex with more than one person at a time). For two of the four groups (the exception being non-heterosexual women and men), there was also a positive relationship between the frequency of viewing pornography and the appeal of using a vibrator/sex toy (i.e., using one on themselves or having a sex partner use a vibrator or other sex toy on them). For the two groups of men as well as the heterosexual women, there was also an association between the frequency of viewing pornography and the appeal of anal sex (i.e., giving and receiving manual anal stimulation and engaging in anal intercourse). Finally, for the heterosexual women, the frequency of viewing pornography was related to the greater appeal of oral sex (performing and receiving).

The qualitative data provided many comments on how the exposure to pornography altered study participants' views of what had been considered odd, unusual, or deviant. For example, one woman described how anal sex became normalized: "I think porn has helped to open my eyes….Anal sex no longer seems mysterious, [it] just seems normal." Another woman reported how seeing many acts of fellatio led her to see it in a new way: "Blow jobs don't seem as much of a gross thing as they used to seem (probably because there is oral sex in just about all porn)." A third woman also noted how seeing sexual acts visually portrayed made them more real and less unusual: "Viewing pornography has made more sex acts (fellatio, cunnilingus) more appealing. Before I didn't know what it [oral sex] looked like…watching made it real."

Particularly supportive of our theoretical ideas were many comments that compared what they had learned from their

Table 2 Ordinal logistic regression beta coefficients and odds ratios: appeal of various activities by frequency of viewing sexually explicit materials

Appeal composites	Women				Men			
	Heterosexual		Non-heterosexual		Heterosexual		Non-heterosexual	
	β	OR	β	OR	β	OR	β	OR
Use of a vibrator/sex toy[a]	0.40*	1.50	0.36	1.44	0.23*	1.25	0.24	1.27
Oral-genital activity[a]	0.36**	1.44	−0.24	0.79	0.15	1.16	0.19	1.20
Anal activity[a]	0.28*	1.33	0.27	1.31	0.37**	1.45	0.63**	1.87
Interest in third parties[a]	0.53***	1.70	0.70**	2.02	0.29**	1.34	0.51*	4.00

No statistically significant differences across gender or sexual preference

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

[a] Composite range: 1 = very unappealing – 5 = very appealing

earlier sexual socialization and their exposure to pornography. Thus, a man described how he came to see the sexual perspective of his parents differently: "Since I am a very sexual being, it [pornography] helps me to know that what I see is not an abnormal act like my parents told me." In the words of a woman: "Coming from a home where talking about sex was discouraged, I was taught that any sex, let alone 'kinky' sex, was bad. Pornography allowed me to see that these that these variations are ok, more normal, and often very enjoyable. I have become more open, accepting, and interested in most sexual acts because of porn."

Does Pornography Consumption Increase the Types of Sexual Behaviors Engaged In?

If the frequency of viewing pornography increases the appeal of various behaviors, does this translate into a similar relationship in promoting expansiveness in engaging in the sexual behaviors themselves? We first examined this relationship with solo sex and then partnered sex.

Solo sex, i.e., self-masturbation, is the most common sexual act associated with viewing pornography (Polsky, 1998) and the most direct in its connection to pornography's presentation of a pornotopia. As shown in Table 3, for three of the four gender-sexual preference identity groups (the exception being the non-heterosexual women), there was a significant relationship between the frequency of viewing pornography and the frequency of self-masturbation.

We also found a significant relationship between pornography viewing and the frequency of heterosexual coitus for three of the four groups (in this case, the non-heterosexual men being the exception). The strength of this relationship was significantly different across the gender-sexual preference identity groups with the non-heterosexual women having the

Table 3 Ordinal logistic regression beta coefficients and odds ratios: sexual behaviors by frequency of viewing sexually explicit materials

Sexual behavior composites	Women				Men			
	Heterosexual		Non-heterosexual		Heterosexual		Non-heterosexual	
	β	OR	β	OR	β	OR	β	OR
Self-masturbation[a,b]	0.55***	1.73	0.25	1.28	0.92***	2.51	0.66*	1.93
Manual sex[a,c]	0.21	1.24	0.28	1.33	0.12	1.13	0.28	1.32
Oral sex[a,b]	0.26**	1.29	0.17	1.18	0.20*	1.22	−0.01	0.99
Coitus[a,d,e,f,g]	0.25*	1.29	0.86**	2.37	0.21*	1.24	−0.22	0.81
Anal sex[b,e,h]	0.31*	1.36	0.74**	2.09	0.39**	1.48	0.03	1.03

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

[a] Time period is past year

[b] Composite range: 0–10

[c] Composite range: 0–8

[d] Beta coefficients statistically significant across gender at $p \leq .05$

[e] Beta coefficients statistically significant across sexual preference at $p \leq .05$

[f] Beta coefficients for gender/sexual preference interaction statistically significant at $p \leq .05$

[g] Composite range: 0–7

[h] Time period is lifetime

strongest relationship followed by the heterosexual women. Significant relationships between the frequency of anal sex and the frequency of viewing pornography were also found for three of the four groups (the non-heterosexual men again being the exception). In addition, we found that for the heterosexual women and men, there was a significant association between the frequency of viewing pornography and the frequency of oral-genital activity.

A relationship between the frequency of viewing pornography and an expansion of one's sexuality, then, is not confined to sexual fantasies or interests. In general, both the frequency of solo and partnered sex were significantly related to the frequency of pornography viewing, viz., more self-masturbation, oral sex, heterosexual coitus, and anal sex.

We previously showed with the qualitative data that pornography can widen a person's sexual horizon through the normalization of various sexual acts. When it comes to translating these new perspectives into behavior, however, we argued that pornography seems to function not only as a learning experience, but a source of erotic empowerment for the viewer—building up the confidence to try new things. These theoretical interpretations also received support from the data of the qualitative study.

Many study participants made comments such as the following. From a woman: "Porn has served as a source of ideas for me to try." From another woman: "Watching porn…caused me to be more experimental." And from a man: "It helped me broaden what I do sexually."

Learning about oral-genital activity was most frequently cited in this regard, as illustrated in the response of one woman: "Everything I know about oral sex I learned from porn." Other new behaviors referred to included learning new sexual positions. In the words of one man, "I like to try a lot of positions that I've seen on some of these videos." And, from another man, "There were things and positions I would never have thought of that after seeing it I thought I have to try this out." An interest in trying anal sex was also mentioned by some of the study participants. In the words of one woman: "Anal sex in porn made me curious so I have tried it several times and I enjoy it."

We were especially struck by the number of women who voiced a sense of empowerment that was attributed to their pornography viewing. One woman said, "It allowed me to be more open to trying new things. I am more willing to try something to see if I like it, instead of being scared of trying and never doing it." And another woman: "I am now more open to different sex acts. I like to be adventurous and try new things." A third woman stated it this way: "Watching porn… [led me to be] less afraid to be loud, made me feel less guilty about wanting sex, wanting pleasure, and directly asking for it." Such confidence was not confined to women. One man had this to say:

> Because I watch so much porn, I have confidence when I give oral sex to my female partner. If anything, it's made me more confident in my decision making in a sexual situation….Without porn, I probably never would have been confident enough to perform cunnilingus on a female.

Thus, as well as increasing the range of their sexual behaviors, feelings of empowerment also were associated with the more frequent viewing of pornography.

Does Pornography Consumption Expand the Boundary for Acceptable Sex Partners?

As a consequence of increased normalization and empowerment, does more frequent viewing of pornography expand

Table 4 Ordinal logistic regression beta coefficients and odds ratios: relationship to the number of partners they performed oral sex on (past 12 months) by frequency of viewing sexually explicit materials

Social relationship	Women				Men			
	Heterosexual		Non-heterosexual		Heterosexual		Non-heterosexual	
	β	OR	β	OR	β	OR	β	OR
Didn't know or met that day[a]	0.17	1.19	0.66	1.93	0.15	1.16	0.26	1.30
Knew as a friend or acquaintance but never dated[b]	0.26	1.30	0.49	1.63	0.16	1.17	0.09	1.10
Were on a first date with[c,d]	0.51	1.66	1.74	5.70	0.05	1.05	0.01	1.01
Had previously dated (was not the first date or was an ex-partner/boyfriend/girlfriend)[e]	0.11	1.11	0.55*	1.73	0.04	1.04	0.51	1.67
Were significantly involved with at the time[d]	0.31*	1.37	−0.21	0.81	0.20	1.23	0.18	1.20

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

[a] Range: 0–15

[b] Range: 0–8

[c] Beta coefficients statistically significant across gender at $p \leq .05$

[d] Range: 0–5

[e] Range: 0–3

the boundaries of with whom a person will engage in sex (e.g., to include people who are strangers or acquaintances)? As presented in Tables 4, 5, and 6, we found that the more frequent consumption of pornography was related to a greater number of partners heterosexual women had oral and coital sex with. The greater number of partners, however, was restricted to "significant others" and did not reflect experiences with a greater number of "casual sexual partners."

We obtained a similar result for the non-heterosexual women with regard to the number of "significant others" as coital partners. Additionally, however, we also found relationships between the extent of pornography viewing and oral and coital sex with partners with whom they were not in a significant relationship. This relationship with heterosexual intercourse reflected the bisexual interests of many of the non-heterosexual women.

For both groups of men, however, there was no significant relationship between frequency of pornography viewing and their number of sex partners or the social relationship with their sexual partners. Thus, the results on partners illustrate that the

Table 5 Ordinal logistic regression beta coefficients and odds ratios: relationship to the number of partners they received oral sex from (past 12 months) by frequency of viewing sexually explicit materials

Social relationship	Women				Men			
	Heterosexual		Non-heterosexual		Heterosexual		Non-heterosexual	
	β	OR	β	OR	β	OR	β	OR
Didn't know or met that day[a]	0.13	1.14	0.35	1.42	0.10	1.11	0.23	1.27
Knew as a friend or acquaintance but never dated[b]	0.05	1.05	0.48	1.62	0.20	1.22	0.01	1.01
Were on a first date with[c]	0.03	1.03	–	–	−0.06	0.94	0.03	1.03
Had previously dated (was not the first date or was an ex-partner/ boyfriend/girlfriend)[d]	0.04	1.04	0.55*	1.74	0.01	1.02	0.22	1.25
Were significantly involved with at the time[e]	0.41*	1.51	−0.14	0.87	0.21	1.23	0.10	1.11

Note: Models that did not converge are noted with dashes. No statistically significant differences across gender or sexual preference

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

[a] Range: 0–12

[b] Range: 0–8

[c] Range: 0–10

[d] Range: 0–3

[e] Range: 0–5

Table 6 Ordinal logistic regression beta coefficients and odds ratios: relationship to the number of partners they engaged in coitus with (past 12 months) by frequency of viewing sexually explicit materials

Social relationship	Women				Men			
	Heterosexual		Non-heterosexual		Heterosexual		Non-heterosexual	
	β	OR	β	OR	β	OR	β	OR
Didn't know or met that day[a]	0.21	1.23	0.64	1.89	0.05	1.05	0.10	1.11
Knew as a friend or acquaintance but never dated[b]	−0.09	0.91	0.35	1.42	0.17	1.19	–	–
Were on a first date with[c]	0.04	1.04	–	–	0.00	1.00	–	–
Had previous dated (was not the first date or was an ex-partner/ boyfriend/girlfriend)[d]	−0.07	0.93	0.92*	2.50	−0.05	0.95	–	–
Were significantly involved with at the time[a,e,f]	0.39**	1.48	0.54*	1.72	0.10	1.10	−6.14	0.00

Note: Models that did not converge are noted with dashes

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

[a] Range: 0–5

[b] Range: 0–8

[c] Range: 0–1

[d] Range: 0–2

[e] Beta coefficients statistically significant across gender at $p \leq .05$

[f] Beta coefficients statistically significant across sexual preference at $p \leq .05$

association with pornography consumption was mediated by gender and sexual preference identity—having more of an effect on the women than the men and, among the women, with differences between the heterosexuals and the non-heterosexuals.

Discussion

The quantitative data showed a relationship between the viewing of pornography and the variables of sexual appeal and behavior. The qualitative data helped support our contention that the viewing of pornography can broaden the appeal, and practice, of a variety of sex acts. In other words, pornography's profusion and dissemination of sexual scripts seems to have, for some, a liberalizing effect. For example, for heterosexual men, there was an association between the frequency of viewing pornography and the appeal of using a vibrator, which parallels the common depiction of the use of vibrators in pornography. Generally, these findings were consistent with our theoretical assumptions that pornography can shape a person's sex life through the normalization of sexual behaviors and a feeling of empowerment that makes certain types of sexual experimentation more probable. The normalization of pornography did not, however, necessarily lead to the desire and empowerment to engage in sex with partners outside of a relationship. The relationship between the frequency of viewing pornography and the greater appeal of a variety of sexual activities did not differ by gender or sexual preference identity. Women did, however, show more behavioral associations with the greater viewing of pornography than did the men. Sexual preference identity also affected the relationship with sexual behaviors.

The above findings suggest a number of things. First, the relationship between pornography consumption and number of coital partners for the heterosexual women was restricted to their significant relationships. We interpret this as an extension of their sexual script rather than a leap to a radically different one; namely, they still follow a norm of "relationship sex." The finding of an association between pornography consumption and the greater likelihood of performing coitus on a first date or with someone they previously dated on the part of non-heterosexual women can be seen as reflecting a less conventional gender script. Finally, the interaction between gender and sexual preference identity showed that non-heterosexual women and men differed more in associations between pornography consumption and frequency of heterosexual coitus than did heterosexual women and men. This finding for non-heterosexuals has also been found in other research (Bell & Weinberg, 1978). It has been interpreted as involving women's history of accommodation to the other sex and their greater conformity to societal expectations (the traditional script of women servicing men).

Before discussing the wider implications of these results, a number of caveats should be made about this study. First, our study was limited by the nature of the study group, college men and women, who lived in a milieu where erotic images were widely available and sexual norms relatively permissive (Armstrong, Hamilton, & Sweeny, 2006). This will affect the generalizability of the findings. For example, Parvez (2006) has shown that social class background differentiates women's enjoyment of pornography. Second, the data are open to different interpretations based on one's sexual values. Individuals with conservative values may differ in their perspective from those who are more liberal (such as our participants) in their interpretation of the expansion of sexual repertoires as positive. Rather, this could be viewed as reflecting an unhealthy focus on the sexual. So too may greater acceptance of pornography be interpreted from a conservative point of view as desensitizing and contributing to undesired effects (see Hald & Malamuth, 2008), whereas those with more liberal sexual values see pornography as an important part of their sexual learning experiences. A third caveat is a cross-sectional study such as this one does not allow us to definitively know the causal direction of relationships. For example, it could be that some people who have developed their sexuality and sexual repertoire to a greater degree are also likely to view pornography more. This could be the interpretation given to the finding in a Norwegian study by Traeen et al. (2006) that experience with group sex predicted the amount of exposure to pornography. Or it could be that some persons with a stronger interest in sex have both a greater interest in viewing pornography as well as finding more appeal in and engaging more frequently in a variety of sexual activities. Haavio-Mannila and Kontula (2003) found the viewing of pornography to be frequent among highly sexually-active individuals in a Finnish study group. There is no reason, then, to believe that there is a monolithic casual relationship. Our interpretation, one that receives support from the qualitative data, is that, for some people, viewing pornography plays a role in making a variety of sexual acts more appealing as well as creating a greater desire to experiment with them.

On a broader level, we can say that the expansion of pornography in the U.S. has not abated despite the attempts of its opponents to restrict its spread (Klein, 2006). According to Paul (2005), for large segments of the population, viewing pornography is so common that it has become a routine part of sexual socialization. We agree. It is the rare student, for example, who has not seen pornography or does not know how to access it (e.g., on her/his laptop). As pornography becomes more a part of mainstream culture (and mainstream culture has become more pornographic), permissive sexual scripts have become widely disseminated. Scripts such as "sex is fun," "sex need not be saved for marriage," and "sex should be available and imaginative" are increasingly dispersed, which can affect, as suggested by our findings, a greater appeal and more frequent practice of a variety of sex acts. This seems to be especially the case for women. Research has shown an increase in the viewing of pornography by women (Ciclitira, 2004) and the incorporation of pornography into their sex lives with their

partners as well as in solo sex. This is again in line with our findings and supports Tiefer's (2004) view that pornography and self-masturbation can play an important part in women's sexual learning.

The growth of free and amateur pornography that is now available also provides new ways of consumption undreamt of by those involved in the pornography debates of the 1970s. Thus, "pornotopia" has expanded its boundaries to intersect more with the outside world. In this way, pornography has played an important role in the widespread normalization of sexuality in postmodern societies (referred to by McNair, 1996, as "pornographication"). Whether this is followed in the larger society (i.e., beyond college campuses and in older age groups) by more varied sexual practices is not easy to estimate. Certainly this is possible. As noted in the Introduction, though, we do not see people as cultural dopes (Garfinkel, 1967): they seem to consume pornography like any other cultural product in the postmodern marketplace. It may be important and influential at different times in their lives (e.g., while in college) and in different situations (e.g., after divorce), and may induce a person to experiment a few times with unconventional sexualities (e.g., swinging, SM), but it will always be just one element, and not necessarily the most important one, in their sexual socialization. Gagnon (2004) noted that in postmodern societies there is an increasing "disconnection between the amount of erotica and the absolute conventionality of…people" (p. 320). Moreover, survey research results in the U.S. adult population do not show a large number of people experiencing a personal pornotopia, even though their sexual horizons have been expanded (Laumann et al., 1994). Pornography will continue to retain its symbolic uses as a signifier of wider social concerns but, as its mostly benign effects are increasingly recognized, it will, we expect, cease to be demonized. Gagnon and Simon (1970) have referred to written pornography as a "paper tiger." Perhaps they could speak similarly today, this time referring to a "pixel tiger."

Acknowledgments An early version of this paper was presented at the Society for the Scientific Study of Sexuality meetings, Las Vegas, November 2006. We are indebted to The Kinsey Institute for the Study of Sex, Gender, and Reproduction for its graduate student grant to Sybil Kleiner for her work on the study, and to the Editor and the anonymous reviewers for their helpful suggestions for improving the paper.

Appendix: Composite Measures, Cronbach's Alpha, and Factor Scores

Appeal Measures

Composite #1: *Appeal of Use of a Vibrator/Sex Toy (alpha = .67)*: Two items—the appeal: of using a vibrator/sex toy on yourself (.86); of having a partner using a vibrator/sex toy on you (.86).

(Composite #2) *Appeal of Oral-Genital Activity (alpha = .46)*: Two items—the appeal: of receiving oral-genital activity (.81); of performing oral-genital activity (.81).

(Composite #3) *Appeal of Anal Activity (alpha = .83)*: Three items—the appeal: of receiving manual-anal stimulation (.85); of performing manual-anal stimulation (.91); of performing anal intercourse [if a female, with a dildo] (.84).

(Composite #4) *Third Party Interest (alpha = .72)*: Three items—the appeal: of watching others have sex on a video (.79); of, in person, watching others have sex (.83); of having sex with more than one person at a time (.81).

Behavior Measures

(Composite #1) *Self-Masturbation (alpha = .79)*: Two items—number of times masturbated over last 12 months (.93); number of times masturbated over last 12 months while viewing sexually explicit material (.93).

(Composite #2) *Manual Sex (alpha = .88)*: Two items—number of people you masturbated in last 12 months (.97); number of people who masturbated you in the last 12 months (.97).

(Composite #3) *Oral Sex (alpha = .89)*: Two items—number of times you performed oral sex in last 12 months (.95); number of times you received oral sex in the last 12 months (.95).

(Composite #4) *Coitus (alpha = .53)*: Two items—number of times you performed coitus in the last year (.83); number of people you engaged in coitus within the last 12 months (.83).

(Composite #5) *Anal Sex (alpha = .90)*: Three items—in your lifetime, number of people you performed finger anal stimulation on (.80); in your lifetime, number of people who performed finger anal stimulation on you (.89); in your lifetime, number of people you have engaged in anal intercourse with (.88).

References

Alfonso, R., & Triglio, J. (1997). Surfing the third wave: A dialogue between two third wave feminists. *Hypatia, 12*, 7–16.

Armstrong, E. A., Hamilton, L., & Sweeny, B. (2006). Sexual assault on campus: A multilevel, integrative approach to party rape. *Social Problems, 53*, 483–499.

Attwood, F. (2005). What do people do with porn? Qualitative research into the consumption, use and experience of pornography and other sexually explicit media. *Sexuality and Culture, 19*, 65–86.

Bailey, C. (1997). Making waves and drawing lines: The politics of defining the vicissitudes of feminism. *Hypatia, 12*, 17–28.

Barbour, C. (1995). Looking at pictures. In A. M. Stan (Ed.), *Debating sexual correctness* (pp. 190–192). New York: Dell Publishing.

Beggan, J. K., & Allison, S. T. (2002). The *Playboy* playmate paradox: The case against the objectification of women. In P. Gagné &

R. Tewskbury (Eds.), *Gendered sexualities* (pp. 103–156). Oxford: Elsevier.
Bell, A. P., & Weinberg, M. S. (1978). *Homosexualities: A study of diversity among men and women*. New York: Simon and Schuster.
Buzzell, T. (2005). Demographic characteristics of persons using pornography in three technical contexts. *Sexuality and Culture, 9*, 28–48.
Califia, P. (1980). *Sophistry: The book of lesbian sexuality*. Tallahassee, FL: Naiad Press.
Califia, P. (1994). *Public sex: The culture of radical sex*. San Francisco: Cleis.
Chapkis, W. (1997). *Live sex acts: Women performing erotic labor*. New York: Routledge.
Ciclitira, K. (2004). Women and feminism: Between pleasure and politics. *Sexualities, 7*, 281–301.
Davis, C. M., & Bauserman, R. (1993). Exposure to sexually explicit materials; An attitude change perspective. *Annual Review of Sex Research, 4*, 121–209.
Davis, M. (1983). *Smut: Erotic reality/obscene ideology*. Chicago: University of Chicago Press.
Duncan, D. F., & Donnelly, J. W. (1991). Pornography as a source of sex information for students at a private northeastern university. *Psychological Reports, 68*, 782.
Duncan, D. F., & Nicholson, T. (1991). Pornography as a source of sex information for students at the Southeastern State University. *Psychological Reports, 68*, 802.
Dworkin, A. (1981). *Pornography: Men possessing women*. London: Penguin.
Gagnon, J. H. (1990). The explicit and implicit use of the scripting perspective in sex research. *Annual Review of Sex Research, 1*, 1–45.
Gagnon, J. H. (2004). *An interpretation of desire: Essays in the study of sexuality*. Chicago: University of Chicago Press.
Gagnon, J. H., & Simon, W. (1970). Pornography: Raging menace or paper tiger? In J. H. Gagnon & W. Simon (Eds.), *The sexual scene* (pp. 137–150). New York: Aldine.
Gagnon, J. H., & Simon, W. (1973). *Sexual conduct: The social sources of human sexuality*. Chicago: Aldine.
Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.
Goldstein, M., Kant, H., Judd, L., Rice, C., & Green, R. (1971). Experience with pornography: Rapists, pedophiles, homosexuals, transsexuals, and controls. *Archives of Sexual Behavior, 1*, 1–15.
Gordon, L. (2002). Lesbians' resistance to culturally defined attractiveness. In E. Rubington & M. S. Weinberg (Eds.), *Deviance: The interactionist perspective* (pp. 310–315). Boston: Allyn and Bacon.
Haavio-Mannila, E., & Kontula, O. (2003). *Sexual trends in the Baltic Sea area*. Helsinki: Publication of the Population Research Institute.
Hald, G. M. (2006). Gender differences in pornography consumption among young heterosexual Danish adults. *Archives of Sexual Behavior, 35*, 577–585.
Hald, G., & Malamuth, N. (2008). Self-perceived effects of pornography consumption. *Archives of Sexual Behavior, 37*, 614–625.
Hardy, S. (1998). *The reader, the author, his woman and her lover: Soft core pornography and heterosexual men*. London: Cassell.
Janghorbani, M., Lam, T. H., & The Youth Sexuality Study Task Force. (2003). Sexual media use by young adults in Hong Kong: Prevalence and associated factors. *Archives of Sexual Behavior, 32*, 545–553.
Janssen, E., Carpenter, D., & Graham, C. A. (2003). Selecting films for sex research: Gender differences in erotic film preference. *Archives of Sexual Behavior, 32*, 243–251.
Juffer, J. (1998). *At home with pornography: Women sex, and everyday life*. New York: New York University Press.
Klein, M. (2006). *America's war on sex: The attack on law, lust and liberty*. Westport, CT: Praeger.
Laumann, E. O., Gagnon, J. H., Michael, R. T., & Michaels, S. (1994). *The social organization of sexuality: Sexual practices in the United States*. Chicago: University of Chicago Press.
Lewin, B. (1997). Part 2: Pornography—Attitudes and use. In B. Lewin (Ed.), *Sex in Sweden: On the Swedish sexual life* (pp. 252–263). Stockholm: The National Institute of Public Health.
Loftus, D. (2002). *Watching sex: How men really respond to pornography*. New York: Thunder's Mouth Press.
Marcus, S. (1966). *The other victorians*. New York: Basic Books.
McElroy, W. (1995). *XXX: A woman's right to pornography*. New York: St. Martin's Press.
McNair, B. (1996). *Mediated sex: Pornography and postmodern culture*. London: Hodder Arnold.
Morrison, T., & Tallack, D. (2005). Lesbian and bisexual women's interpretations of lesbian and ersatz lesbian pornography. *Sexuality and Culture, 9*, 3–30.
Palac, L. (1995). How dirty pictures changed my life. In A. M. Stan (Ed.), *Debating sexual correctness* (pp. 236–252). New York: Dell Publishing.
Parvez, Z. F. (2006). The labor of pleasure: How perceptions of emotional labor impact women's enjoyment of pornography. *Gender and Society, 2*, 605–631.
Paul, P. (2005). *Pornified: How pornography is transforming our lives, our relationships, and our families*. New York: Holt.
Peckham, M. (1969). *Art and pornography: An experiment in explanation*. New York: Basic Books.
Plante, R. F. (2006). *Sexualities in context: A social perspective*. Ithaca, NY: Westview Press.
Polsky, N. (1998). *Hustlers, beats, and others*. New York: The Lyons Press.
Porn, Feminism, and the Meese Report. (1987). *Proletarian Revolution, 27*(1), 7–14.
Rich, F. (2001). Naked capitalists. In M. Stombler, E. O. Burgess, D. Donnelly, & W. Simonds (Eds.), *Sex matters: The sexuality and society reader* (pp. 462–466). Boston: Pearson Education Inc.
Rubin, G. (1984). Thinking sex: Notes for a radical theory of the politics of sexuality. In C. Vance (Ed.), *Pleasure and danger: Exploring female sexuality* (pp. 267–319). Boston: Routledge and Kegan Paul.
Rubington, E., & Weinberg, M. S. (1996). *Deviance: The interactionist perspective* (6th ed.). Boston: Allyn and Bacon.
Shaw, S. M. (1999). Men's leisure and women's lives: The impact of pornography on women. *Leisure Studies, 18*, 197–212.
Simon, W., & Gagnon, J. H. (1987). A sexual scripts approach. In J. H. Geer & W. T. O'Donohue (Eds.), *Theories of human sexuality* (pp. 363–384). New York: Plenum Press.
Stein, A. (1997). *Sex and sensibility: Stories of a lesbian generation*. Berkeley: University of California Press.
Štulhofer, A., Buško, V., & Landripet, I. (2008). Pornography, sexual socialization, and satisfaction among young men. *Archives of Sexual Behavior*. doi:10.1007/s10508-008-9387-0.
Tanenbaum, L. (2004). Slut! Insult of insults. In M. Stombler, D. M. Baunach, E. O. Burgess, D. Donnelly, & W. Simonds (Eds.), *Sex matters: The sexuality and society reader* (pp. 403–411). Boston: Pearson Education Inc.
Thoits, P. A. (1995). Identity-relevant events and psychological symptoms: A cautionary tale. *Journal of Health and Social Behavior, 36*, 72–82.
Tiefer, L. (2004). *Sex is not a natural act and other essays* (2nd ed.). Boulder, CO: Westview Press.
Traeen, B., Sorheim-Nielsen, T., & Stigum, H. (2006). Use of pornography in traditional media and on the internet in Norway. *Journal of Sex Research, 41*, 193–200.
Traeen, B., Spitznogle, K., & Beverfjord, A. (2004). Attitudes and use of pornography in the Norwegian population 2002. *Journal of Sex Research, 41*, 193–200.
Weinberg, M. S., & Williams, C. J. (1975). Gay baths and the social organization of impersonal sex. *Social Problems, 23*, 124–136.
Weinberg, M. S., & Williams, C. J. (2005). Fecal matters: Habitus, embodiments, and deviance. *Social Problems, 52*, 315–336.

Williams, L. (1999). Film bodies: Gender, genre and excess. In S. Thornham (Ed.), *Feminist film theory* (pp. 267–281). New York: University Press.

Zillmann, D. (2000). Influence of unrestrained access to erotica on adolescents' and young adults' dispositions toward sexuality. *Journal of Adolescent Health, 27*, 41–44.

Zillmann, D., & Bryant, J. (1984). Effects of massive exposure to pornography. In N. Malamuth (Ed.), *Pornography and sexual aggression* (pp. 115–138). New York: Academic Press.

Arch Sex Behav (2010) 39:1403–1414
DOI 10.1007/s10508-009-9597-0

ORIGINAL PAPER

The Relation Between Sexual Orientation and Suicide Attempts in Austria

Martin Plöderl · Karl Kralovec · Reinhold Fartacek

Received: 25 November 2008 / Revised: 15 December 2009 / Accepted: 15 December 2009 / Published online: 4 February 2010

Abstract Previous studies indicate that homosexual or bisexual individuals are at a higher risk of attempting suicide compared to heterosexuals. To overcome biases in these studies, more rigorous definitions of "suicide attempts" and the assessment of multiple dimensions of sexual orientation are needed. In addition, studies from the German speaking region are sparse, especially those not recruiting participants from the gay or lesbian communities. We solicited self-reported suicide attempts among 1,382 Austrian adults recruited through structured snowball sampling from students' social networks. Suicide attempts were more frequently reported by those participants with homosexual or bisexual fantasies, partner preference, behavior, and self-identification, compared to their heterosexually classified counterparts. This was true for any dimension of sexual orientation and for suicide attempts with intent to die or suicide attempts that required medical treatment. Our Austrian study confirmed existing evidence that homosexual and bisexual individuals are at an increased risk for attempting suicide. This should be considered in suicide preventive efforts.

Keywords Suicide attempts · Sexual orientation · Homosexuality

M. Plöderl (✉)
University Institute for Clinical Psychology, Paracelsus Private Medical University and Christian-Doppler-Klinik, Salzburg, Austria
e-mail: m.ploederl@salk.at

M. Plöderl · K. Kralovec · R. Fartacek
Suicide Prevention Research Program, Institute of Public Health, Paracelsus Private Medical University, Salzburg, Austria

M. Plöderl · K. Kralovec · R. Fartacek
Department of Suicide Prevention, University Clinic of Psychiatry and Psychotherapy I, Christian-Doppler-Klinik, Salzburg, Austria

Introduction

Reviews of numerous studies conclude that lesbian, gay, and bisexual (LGB) individuals report higher rates of suicide attempts than heterosexuals (Bagley & Tremblay, 2000; King et al., 2008; McDaniel, Purcell, & D'Augelli, 2001; Meyer, 2003; Plöderl, Sauer, & Fartacek, 2006; Russell, 2003). These reviews also indicate that the increased suicide risk of LGB individuals continued to appear in recent studies with methodologically improved methods such as using birth cohorts, twin registries, prospective designs, representative and large-scaled adolescent surveys, or matching LGB individuals with their heterosexual siblings.

It has, however, been argued that the increased suicide attempt risk among sexual minorities may be a product of methodological biases in these studies. For example, biased results may stem from vague definitions of "suicide attempt." Most studies only used a single item on suicide attempts, thus leaving it open to the study participants to decide what counts as a suicide attempt. When suicide attempts were more rigorously assessed, the difference in sexual orientation lost its significance in one study (Savin-Williams, 2001, Study 2), but not in another study, where the risk difference actually increased (Bagley & Tremblay, 1997). In the Savin-Williams (2001) study, participants were given follow-up questions to classify suicide attempts as "false attempts," i.e., attempts that were not carried out and were only suicide plans or "true attempts," i.e., attempts that were carried out. No significant sexual orientation differences occurred after restricting the analysis to true suicide attempts. Bagley and Tremblay's (1997) study is, to our knowledge, the only study that differentiated between suicidal gestures (with no clear intention to die and without a lethal method) and serious suicide attempts with a clear intention to die and/or the use of a potentially lethal method (Bagley, Wood, & Young, 1994). Notably, the intent to die is crucial for distinguishing suicide attempts from other suicide related behav-

iors (Silverman, Berman, Sanddal, O'Carroll, & Joiner, 2007). Using the "serious suicide attempt" variable resulted in a substantially greater sexual orientation difference than using the "suicidal gesture" variable. Given the mixed results of the two studies and the lack of related studies assessing intent to die, more research with improved assessments of suicide attempts is clearly needed.

Another methodological problem is the measurement of sexual orientation. Most studies used only one dimension, such as sexual attraction, sexual behavior, or self-identification (King et al., 2008; Plöderl et al., 2006; Russell, 2003). It is still an open question what dimension of sexual orientation is most appropriate (Herek & Garnets, 2007; Savin-Williams, 2006, 2008). Relying on the behavioral dimensions makes it impossible to classify individuals who did not have sexual contacts but are certain about their sexual identification. Classification problems also arise if there is a discrepancy between dimensions of sexual orientation. For example, a clear majority of men who had sex with men identified as heterosexual (Pathela et al., 2006). Similarly, the majority of bisexually attracted men and women considered themselves as heterosexual (McCabe, Hughes, Bostwick, West, & Boyd, 2009). Moreover, some individuals are not sure about their sexual orientation and this was found to be associated with increased distress (Espelage, Aragon, Birkett, & Birkett, 2008). Such individuals often do not find appropriate categories in questionnaires, and it is not justified to place them simply in a homosexual, bisexual, or heterosexual category. Suicide risk may vary with the dimension of sexual orientation. Thus, studies with multiple dimensions of sexual orientation would therefore be beneficial but are largely lacking (Russell, 2003). Meyer (2003) discussed that a person's homosexual identity may lead to more proneness toward minority stress, but also is a source for affiliation and support from other minority members. Indeed, a review of studies concluded that gay identified males have the highest risk, and those with same-sex behavior and same-sex orientation have a somewhat lower, but still increased, risk compared to their heterosexual counterparts (Russell, 2003). Similarly, another study reported that women and men who identified themselves as lesbian or gay generally had greater odds of substance abuse and substance dependence than when minority sexual orientation was defined by same-sex attraction or behavior (McCabe et al., 2009). However, in another of the few available studies with multiple dimensions of sexual orientation, same-sex sexual behavior was a stronger predictor of suicide attempts than homosexual identity or attraction, perhaps because same-sex behavior is a "social fact that makes the stress of non-heterosexuality more acute" (Wichstrøm & Hegna, 2003, p. 149). Given these mixed findings, more research with multiple metrics of sexual orientation is needed.

Most of the reviewed studies are from English-speaking countries. There are only few studies from German-speaking European countries and these studies found high rates of suicide attempts among sexual minorities (Berliner Senatsverwaltung für Bildung, Jugend und Sport, 1998; Biechele, 2001; Dannecker & Reiche, 1974; Plöderl & Fartacek, 2005). Participants in these studies were recruited from gay and lesbian communities (LGB support groups, bars, festivals, youth summer camps, LGB-press). It is not clear if such a sampling procedure leads to overestimation or underestimation of mental health problems, because participating in a sexual minority community may bear both risk factors (e.g., alcohol consumption in gay bars) or protective factors (social support) for mental health (Herek & Garnets, 2007), or because such organizations may be selectively visited by LGB individuals with more mental health problems (Muehrer, 1995; Savin-Williams, 2001, 2008). Empirical studies found that attachment to the LGB community was associated with lower levels of depression (Lewis, Derlega, Berndt, Morris, & Rose, 2001; Mao et al., 2009; Mills et al., 2004), but also with more drug and alcohol use (Knox, Kippax, Crawford, Prestage, & Van De Ven, 1999; Plöderl, 2005). No effects of gay community involvement were found for suicide attempts and suicide ideation in an Austrian study (Plöderl, 2005). However, gay youth sampled from support groups had higher rates of suicide attempts than those sampled from non-support groups (Savin-Williams & Ream, 2003). "LGB community" seems to be an umbrella term for a variety of organizations, with different effects on mental health, and this may explain the above mixed findings. Sampling from the general population instead of from LGB organizations may overcome possible biases.

The present study attempted to contribute to the existing literature about the association of suicide attempts and sexual orientation by overcoming important caveats of previous studies. This was realized by applying multiple dimensions of sexual orientation, more rigorous classifications of suicide attempts, and by using a sample that was not recruited from LGB organizations or venues.

Method

Participants

Study participants were recruited with a structured snowball sampling procedure via a social network of psychology students (see procedure section for details). Participants were classified as heterosexual, bisexual, or homosexual with different metrics of sexual orientation (see measures section). About half ($n = 726$, 53%) of the 1,382 study participants were female. The mean age was 37.77 ($SD = 14.45$) and ranged from 18 to 84 years. Nearly half (47%) had a degree of education lower than Matura (equivalent to German Abitur or the British A-level), 38% had Matura, and 14% were academics. Compared to the Austrian population in 2005 (Statistik Austria, 2006, 2009), our participants were younger (Austrian mean

age in the range of 18 to 84 years: 46.64 years), and they had a higher degree of education (Austria: 77% lower than Matura, 14% Matura, 10% academics).

Measures

Sexual Orientation

The behavioral dimension of sexual orientation was assessed with the items "Sex partner(s) since the age of 18" and "Sex partner(s) in the past 12 months" with the response options: women/men/women and men/no sex. Sexual fantasies and sexual attraction were assessed with "In your sexual fantasies there are …" and "You would like to have sex with …" and five response options: only women/mostly women/men and women/mostly men/only men). Self-identification was assessed with "How do you describe yourself?" with the response options: heterosexual/mostly heterosexual/bisexual/mostly homosexual/homosexual/transsexual/not sure. Participants were classified as bisexual if they ranged somewhere between the heterosexual end and the homosexual end of the response options. For some analyses, a stricter cut-off for bisexuality was used, where mostly heterosexually identified, mostly opposite sex attracted, and those with mostly opposite sex fantasies were categorized as heterosexual.

Suicide Attempts

The gate question on suicide attempts was "Have you ever attempted suicide?" with "yes/no" options separately for the past year and for previous years. To separate suicide attempts from other suicide-related behavior, follow-up items had to be completed by participants who reported a suicide attempt in the gate question. In case of multiple self-reported suicide attempts, participants had to refer to the most serious attempt. The follow-up items were inspired by studies that used more rigorous assessments of suicide attempts among sexual minorities (Bagley & Tremblay, 1997; Savin-Williams, 2001).

Open-ended questions were used to assess the motive behind the suicide attempt, the suicide method, resulting injuries, and resulting treatment of the injuries. Injuries resulting from suicide attempts were categorized into those treated in a hospital, treated by a doctor (outside a hospital), self-treated injuries, and no injuries.

To determine if the attempt was actually carried out, the multiple choice item "What describes the suicide attempt best?" was used, having six response options: (1) I only thought seriously about hurting/harming myself; (2) I had everything prepared but did not try to attempt suicide; (3) I stopped hurting/harming myself in the last second. I knew that it would not have been lethal; (4) I stopped hurting/harming myself in the last second. I knew that it would have been lethal; (5) I hurt/harmed myself, but I knew that I would not have died from this; (6) I hurt/harmed myself and I knew that I would die from this.

Finally, intent to die was quantified with the multiple-choice item "How strong was the wish to die?" with the response options: very strong/strong/likely not strong/I definitely did not want to die. The first three response options were collapsed into a "some intent" category. A "strong intent" category was created by collapsing those who responded with "very strong" or "strong" on the item.

We also tried to uncover false negative (unreported) suicide attempts among participants who did not report a suicide attempt in the gate question. Therefore, we used the expressions "incident" or "harm/injury" instead of "suicide attempt." Such incidents were solicited with the item: "There was an incident where I hurt or harmed myself, or where I intended to do so, but I do not think this was a suicide attempt (e.g., overdose of alcohol or medication, standing in front of an abyss, jumping in front of a car, cutting wrists, and the like)." Follow-up items identical to those described above were given for a further description of the incident.

Categorization of Suicide Attempts

We created five different types of suicide attempt categories. First, as in most previous studies, a participant was classified as a suicide attempter if he or she reported a suicide attempt in the past year or in previous years according to the simple gate question "Have you ever attempted suicide?"

Second, for a more rigorous classification of a suicide attempt, participants were classified as attempters only if (1) they reported a suicide attempt in the gate question or an incident of self harm in the follow-up item (for those who did not initially reported a suicide attempt), (2) there was some intent to die, and (3) the attempt/incident was either carried out or aborted but with resulting injuries, and not just ideated, planned, or aborted without injuries (see also Plöderl, Kralovec, Yazdi, & Fartacek, in press).

A third category was similar to the second one, except that only attempts with strong intent to die were considered.

The fourth category was suicide attempts that were carried out with a clear expectation to die, i.e., participants who chose the follow-up item: "I hurt/harmed myself and I knew that I would die from this."

The fifth and last type of suicide attempts was the kind that resulted in medically treated injuries. The fourth and fifth categories also included the second criterion, i.e., only attempts with some intent to die and attempts that were actually carried out.

Procedure

Each of 85 psychology students who already took part in survey-methodology courses distributed 20 questionnaires

within his or her social network of friends, colleagues, family members, or relatives, making up a total of 1,700 distributed questionnaires. In this structured snowball sampling procedure, a sampling scheme was followed as close as possible for the student, so that each subsample of 20 participants reflected the Austrian adult population with respect to sex, age, and education. The students were instructed to guarantee the anonymity of the participants and to achieve a high response rate. Guarantee of anonymity was attempted by leaving it open to the participant if he or she wanted to hand back the questionnaire in a sealed envelope to the student or, alternatively, send it back in a prepaid envelope by regular mail directly to the first author. The students were also instructed to tell the participants how important it was for the quality of the data to give honest answers; to achieve a high response rate (by explaining that non-responding may bias the results); that the results were necessary to answer important research questions; and that study results may be obtainable upon request. Sampling issues were discussed in regular meetings.

Of the 1,435 returned questionnaires (84% response rate), some were excluded because they were blank or contained obviously joking answers ($n = 7$), because gender was not reported ($n = 1$) or because the age was lower than 18 years ($n = 34$) or because the gate question about suicide attempts was left blank ($n = 8$). Furthermore, three subjects were excluded because they described themselves as transsexuals. Finally, 1,382 (81%) questionnaires were entered into the data set for the analysis.

Data Analysis

Data were analyzed with R 2.7.0 (R Development Core Team, 2008). Odds ratios and 95% median-unbiased confidence intervals were calculated with the "epitools" package (Aragon, 2007). The associations of suicide attempts with sexual orientation were calculated for each metric of sexual orientation and separately for men and women. For the latter two types of suicide attempts, however, we did not report the results separated by gender because of low counts.

Results

Association of Sexual Orientation with Age and Education

The association of age with sexual orientation was inspected via regression analysis with age as the dependent variable and sexual orientation as the predictor, with heterosexuals as reference category. Bisexually classified participants were significantly ($p < .05$) younger than heterosexuals on the sexual orientation dimensions sexual fantasies ($\beta = -3.78, SE = 0.93, t = -4.06$), partner preference ($\beta = -4.41, SE = 1.16, t = -3.81$), and self-identification ($\beta = -3.30$, $SE = 1.45$, $t = -2.77$). Participants ranging on the homosexual end of the sexual orientation dimensions tended to be younger than heterosexuals, but the differences did not achieve statistical significance (for descriptive results, see Table 1). Participants who had no sex in the past year were significantly older than their heterosexually active counterparts ($\beta = 3.38, SE = 1.46, t = -2.33, p < .05$), whereas this age difference was reversed for sex in previous years ($\beta = -15.62, SE = 3.01, t = -5.19$).

Table 1 Age and education and relation to dimensions of sexual orientation

Sexual orientation	Age		Education (%)[a]		
	M	*SD*	Low	Medium	High
Fantasies					
Heterosexual	38.63	14.87	49	38	13
Bisexual	34.85	12.66	41[†]	42	18[†]
Homosexual	35.85	10.08	42	32	26
Preferred partner					
Heterosexual	38.37	14.70	49	38	13
Bisexual	33.96	12.27	39	43	18
Homosexual	34.42	9.63	42	37	21
Self-Identification					
Heterosexual	37.91	14.41	48	39	13
Bisexual	34.61	13.81	37	42	21
Homosexual	33.69	8.40	44	31	25
Not sure	32.89	16.28	33	56	11
Behavior (12 mos)					
Heterosexual	37.53	13.84	47	38	14
Bisexual	31.40	12.62	27	53	20
Homosexual	35.88	12.18	38	46	17
No sex	40.91	19.29	54	39	7
Behavior (prev.)					
Heterosexual	38.23	14.47	48	38	14
Bisexual	35.23	12.40	38	38	25*
Homosexual	33.33	9.90	33	42	25
No sex	22.61	9.50	35	65*	0[†]

[†] Standardized residual $p < .10$; * standardized residual $p < .05$

[a] Percentages add up to 100 row-wise

Associations of education with sexual orientation were explored with χ^2-tests and analyses of standardized residuals. There were significant effects ($p < .05$) of education for sexual fantasies, $\chi^2(4) = 11.40$, and sex since the age of 18, $\chi^2(6) = 16.34$ (Table 1). The significant result for sexual fantasies was produced by the disproportionately high percentage of highly educated bisexual participants. The significant effect for sex since the age of 18 resulted from the disproportionately high fraction of participants who were highly educated and bisexual and those with medium education who had no sex.

Table 2 Correlations of sexual orientation measures (women)

Sexual orientation	Fantasies	Preferred partner	Identification	Behavior (12 mos)
Preferred partner	.56	–		
Self-identification	.32	.49	–	
Behavior (12 mos)	.28	.37	.42	–
Behavior (prev.)	.34	.44	.53	.64

Note: The original items with five response options were used for sexual fantasies, partner preference, and self-identification. Those who did not report sexual behavior were treated as missing. All correlations were nonparametric Spearman rank correlations and were statistically significant ($p < .0001$)

Table 3 Correlations of sexual orientation measures (men)

Sexual orientation	Fantasies	Preferred partner	Identification	Behavior (12 mos)
Preferred partner	.77	–		
Self-identification	.66	.80	–	
Behavior (12 mos)	.42	.53	.50	–
Behavior (prev.)	.45	.54	.54	.69

Note: The original items with five response options were used for sexual fantasies, partner preference, and self-identification. Those who did not report sexual behavior were treated as missing. All correlations were nonparametric Spearman rank correlations and were statistically significant ($p < .0001$)

Intercorrelations of Sexual Orientation Dimensions

The intercorrelations of the sexual orientation metrics were all statistically significant for both men and women and tended to be smaller among women than men (see Tables 2 and 3).

Suicide Attempts and Sexual Orientation

Suicide attempts as a function of different measures of sexual orientation are shown in Tables 4–8. Based on the gate question, suicide attempts were more frequently reported by participants with homosexual or bisexual fantasies, partner preference, behavior, and self-identification, compared to their heterosexually classified counterparts (see Tables 4 and 5). Most of the differences achieved statistical significance.

Analysis of suicide attempts that were actually carried out and involved at least some intent to die resulted in comparable or even increased sexual orientation differences compared to results based on the simple gate question (Tables 4 and 5). Again, a clear majority of the differences were statistically significant. When considering only suicide attempts with strong intent to die, men and women with homosexual or bisexual fantasies/partner preference/behavior/identification still reported higher incidences of suicide attempts than their heterosexual counterparts (see Tables 4 and 5). Most of the sexual orientation differences remained statistically significant among men; among women, the results lost significance for sexual fantasies, partner preference, and self-identification.

The results for suicide attempts with a clear expectation of death and with medically treated injuries are shown in Table 6. Because the incidences of such attempts were low, results were not separated by gender. The analysis for the gate-question was given as a baseline. For suicide attempts with clear expectation of death ("I hurt/harmed myself and I knew that I would die from this"), the differences were significant for participants who ranged on the homosexual end of the sexual orientation dimensions, but not for bisexuals, and some differences could not be calculated because no suicide attempter was in the bisexual group. For attempts with medical treatment, the majority of differences were statistically significant and tended to be larger or comparable than for the gate question.

Participants classified as bisexual had comparable or smaller risk differences than those classified as homosexual in most cases when they were compared with heterosexuals (see odds ratios in Tables 4–6). If homosexual and bisexual participants were collapsed into one group, they nearly always reported significantly more suicide attempts compared to heterosexuals.

Using a stricter cut-off for bisexuality for the sexual orientation dimensions sexual fantasies, partner preference and self-identification led to a substantial decrease of participants who were classified as bisexual (Tables 7 and 8). Some of the sexual orientation differences then became non-significant, especially for the sexual fantasies dimension.

Women who were unsure about their self-identification reported significantly higher rates of suicide attempts compared to heterosexuals. There was only one man who was unsure about his sexual identification and he stated to having attempted suicide.

Notably, in nearly all analyses where there was no statistical significance, odds ratios clearly exceeded 1.00. This hints at an increased suicide attempt rate among individuals with homosexual or bisexual fantasies, partner preference, behavior, or self-identification. Among men, non-significant sexual orientation differences occurred most frequently for sexual behavior since the age of 18 as dimension of sexual orientation. Among women, no significant differences occurred between homosexuals and heterosexuals for all dimensions of sexual orientation.

Because of some significant sexual orientation differences for age and education we inspected if the results for suicide attempts were influenced by these variables. Therefore, age and education were controlled for in logistic regression analysis with suicide attempts as dependent variables. Since the results remained comparable, they are not reported here. Thus, it can be assumed that sexual orientation differences in suicide attempts rates are not caused by sexual orientation differences with respect to age or education.

Table 4 Suicide attempts as a function of sexual orientation (women, $N = 726$)

Sexual orientation	N[a]	%	Suicide attempts											
			Gate question				With some intent to die				With strong intent to die			
			n	%	OR (95% CI)	*p*	*n*	%	OR (95% CI)	*p*	*n*	%	OR (95% CI)	*p*
Fantasies														
Heterosexual	471	(65)	14	(3)	Ref. cat.		9	(2)	Ref. cat.		7	(1)	Ref. cat.	
Bisexual	243	(33)	18	(7)	2.6 (1.3–5.4)	**	16	(7)	3.6 (1.6–8.7)	**	8	(3)	2.2 (0.8–6.6)	
Homosexual	9	(1)	1	(11)	4.5 (0.2–28.0)		1	(11)	7.0 (0.3–46.4)		1	(11)	9.0 (0.3–62.7)	
Bisexual/homosexual combined	252	(35)	19	(8)	2.7 (1.3–5.5)	**	17	(7)	3.7 (1.6–8.8)	**	9	(4)	2.4 (0.9–7.0)	
Preferred partner														
Heterosexual	579	(80)	19	(3)	Ref. cat.		14	(2)	Ref. cat.		10	(2)	Ref. cat.	
Bisexual	139	(19)	13	(9)	3.0 (1.4–6.3)	**	11	(8)	3.5 (1.5–7.9)	**	5	(4)	2.2 (0.6–6.3)	
Homosexual	7	(1)	1	(14)	5.4 (0.2–35.3)		1	(14)	7.4 (0.3–49.4)		1	(14)	10.3 (0.4–72.0)	
Bisexual/homosexual combined	146	(20)	14	(10)	3.1 (1.5–6.4)	**	12	(8)	3.6 (1.6–8.1)	**	6	(4)	2.5 (0.8–6.8)	
Self-identification														
Heterosexual	622	(86)	22	(4)	Ref. cat.		17	(3)	Ref. cat.		12	(2)	Ref. cat.	
Bisexual	69	(10)	7	(10)	3.1 (1.2–7.3)	*	7	(10)	4.1 (1.5–9.9)	**	2	(3)	1.6 (0.2–6.2)	
Homosexual	6	(1)	1	(17)	6.0 (0.2–41.2)		1	(17)	7.8 (0.3–54.6)		1	(17)	11.1 (0.4–79.8)	
Bisexual/homosexual combined	75	(10)	8	(11)	3.3 (1.3–7.4)	*	8	(11)	4.3 (1.7–10.1)	**	3	(4)	2.2 (0.5–7.2)	
Not sure	8	(1)	3	(38)	16.4 (3.0–74.3)	**	1	(13)	5.6 (0.2–35.3)		1	(13)	8.0 (0.3–52.0)	
Behavior (12 mos)														
Heterosexual	626	(86)	24	(4)	Ref. cat.		18	(3)	Ref. cat.		10	(2)	Ref. cat.	
Bisexual	12	(2)	3	(25)	8.5 (1.7–31.5)	*	4	(33)	16.9 (4.0–60.5)	**	2	(17)	12.7 (1.6–58.8)	*
Homosexual	8	(1)	1	(13)	4.0 (0.2–24.4)		1	(13)	5.3 (0.2–33.4)		1	(13)	9.6 (0.4–64.4)	
Bisexual/homosexual combined	20	(3)	4	(20)	6.4 (1.7–19.3)	**	5	(25)	11.3 (3.3–33.4)	**	3	(15)	11.1 (2.2–41.1)	**
No sex	67	(9)	4	(6)	1.6 (0.5–4.4)		2	(3)	1.1 (0.2–4.0)		2	(3)	2.0 (0.3–8.0)	
Behavior (prev.)														
Heterosexual	661	(91)	24	(4)	Ref. cat.		17	(3)	Ref. cat.		11	(2)	Ref. cat.	
Bisexual	37	(5)	6	(16)	5.2 (1.8–13.1)	**	8	(22)	10.4 (3.9–25.8)	**	4	(11)	7.3 (1.9–23.0)	**
Homosexual	5	(1)	1	(20)	7.2 (0.3–54.8)		1	(20)	10.3 (0.4–79.3)		1	(20)	15.9 (0.5–127.2)	
Bisexual/homosexual combined	42	(6)	7	(17)	5.4 (2.0–12.8)	**	9	(21)	10.3 (4.1–24.7)	**	5	(12)	8.1 (2.4–23.8)	**
No sex	11	(2)	1	(9)	3.0 (0.1–16.8)		0	(0)	–		0	(0)	–	

* $p < .05$; ** $p < .01$

[a] Numbers may not add up to the total sample size because of missing data on some sexual orientation dimensions

Discussion

In this study of Austrian adults, suicide attempts were more frequently reported by participants with homosexual or bisexual fantasies, partner preference, sexual behavior, and self-identification, compared to their heterosexually classified participants. Contrary to our study, previous Austrian and German studies recruited their participants from the LGB community with possible biases. However, high rates of suicide attempts occurred with both sampling approaches.

In our study, increased rates of suicide attempts among sexual minority individuals occurred for all dimensions of sexual orientation: sexual behavior, sexual fantasies, partner preference, and self-identification. Most of these differences were also statistically significant.

Among women, the risk differences were strongest for the behavioral indicators of a non-heterosexual orientation. For men, largest odds were found for homosexual or bisexual behavior in the past 12 months and for self-identification as homosexual or bisexual. Homosexual behavior may be a "social fact" that is especially associated with minority stress, as discussed by Wichstrom and Hegna (2003). In accordance with that hypothesis, non-heterosexual fantasies or attraction were the dimensions with the smallest odds for suicide attempts, perhaps because these dimensions remain hidden and are less associated with minority stress. For men, the odds for having attempted suicide were also high for those who identified as homosexual or bisexual. Perhaps these men more often came out and were more frequently exposed to minority stressors. They also may be more gender nonconforming with related

Table 5 Suicide attempts as a function of sexual orientation (men, $N = 656$)

Sexual orientation	N^a	%	Suicide attempts											
			Gate question				With some intent to die				With strong intent to die			
			n	%	OR (95% CI)	*p*	*n*	%	OR (95% CI)	*p*	*n*	%	OR (95% CI)	*p*
Fantasies														
Heterosexual	580	(88)	16	(3)	Ref. cat.		12	(2)	Ref. cat.		10	(2)	Ref. cat.	
Bisexual	65	(10)	7	(11)	4.3 (1.6–10.6)	**	6	(9)	4.9 (1.6–13.1)	**	5	(8)	4.8 (1.4–14.2)	*
Homosexual	11	(2)	3	(27)	13.4 (2.6–52.8)	**	3	(27)	17.9 (3.4–73.0)	**	2	(18)	13.1 (1.7–61.5)	*
Bisexual/homosexual combined	76	(12)	10	(13)	5.3 (2.2–12.2)	**	9	(12)	6.4 (2.5–15.7)	**	7	(9)	5.8 (2.0–15.8)	**
Preferred partner														
Heterosexual	605	(92)	18	(3)	Ref. cat.		13	(2)	Ref. cat.		11	(2)	Ref. cat.	
Bisexual	38	(6)	5	(13)	5.0 (1.5–13.6)	**	5	(13)	7.0 (2.1–20.0)	**	4	(11)	6.5 (1.7–20.4)	**
Homosexual	12	(2)	3	(25)	11.1 (2.2–41.8)	**	3	(25)	15.4 (3.0–60.3)	**	2	(17)	11.2 (1.4–50.8)	*
Bisexual/homosexual combined	50	(8)	8	(16)	6.2 (2.4–14.9)	**	8	(16)	8.7 (3.2–22.0)	**	6	(12)	7.4 (2.4–20.7)	**
Self-identification														
Heterosexual	597	(91)	16	(3)	Ref. cat.		11	(2)	Ref. cat.		9	(2)	Ref. cat.	
Bisexual	38	(6)	6	(16)	6.9 (2.3–18.1)	**	6	(16)	10.0 (3.2–28.5)	**	5	(13)	9.9 (2.8–31.0)	**
Homosexual	10	(2)	2	(20)	9.4 (1.2–42.8)	*	2	(20)	13.7 (1.7–65.2)	*	1	(10)	8.0 (0.3–51.4)	
Bisexual/homosexual combined	48	(7)	8	(17)	7.3 (2.8–17.8)	**	8	(17)	10.6 (3.9–28.1)	**	6	(13)	9.3 (2.9–27.6)	**
Not sure	1	(0)	1	(100)	–		1	(100)	–		0	(0)	–	
Behavior (12 mos)														
Heterosexual	589	(90)	17	(3)	Ref. cat.		12	(2)	Ref. cat.		10	(2)	Ref. cat.	
Bisexual	3	(0)	1	(33)	17.5 (0.5–227.1)		1	(33)	24.8 (0.8–327.6)		0	(0)	–	
Homosexual	16	(2)	3	(19)	8.0 (1.6–28.1)	*	3	(19)	11.3 (2.2–41.9)	**	2	(13)	8.6 (1.1–37.7)	*
Bisexual/homosexual combined	19	(3)	4	(21)	9.1 (2.3–28.8)	**	4	(21)	12.9 (3.2–43.0)	**	2	(11)	7.1 (0.9–30.5)	
No sex	42	(6)	4	(10)	3.6 (1.0–10.5)		4	(10)	5.2 (1.3–15.9)	*	4	(10)	6.2 (1.6–19.8)	*
Behavior (prev.)														
Heterosexual	613	(93)	20	(3)	Ref. cat.		15	(2)	Ref. cat.		13	(2)	Ref. cat.	
Bisexual	20	(3)	2	(10)	3.5 (0.5–13.4)		2	(10)	4.7 (0.6–18.6)		1	(5)	2.7 (0.1–15.0)	
Homosexual	7	(1)	1	(14)	5.5 (0.2–35.4)		1	(14)	7.3 (0.3–48.5)		0	(0)	–	
Bisexual/homosexual combined	27	(4)	3	(11)	3.8 (0.8–12.4)		3	(11)	5.1 (1.1–17.2)	*	1	(4)	2.0 (0.1–10.7)	
No sex	12	(2)	2	(17)	6.2 (0.8–26.2)		2	(17)	8.3 (1.1–36.1)	*	2	(17)	9.6 (1.3–42.5)	*

* $p < .05$; ** $p < .01$

[a] Numbers may not add up to the total sample size because of missing data on some sexual orientation dimensions

negative experiences. However, a homosexual or bisexual identity was not the strongest predictor for suicide attempts among women. Furthermore, among men (but not among women), non-significant differences most likely occurred if sexual orientation was based on sexual behavior since the age of 18. Maybe this is not a very valid dimension of sexual orientation. Similarly, for men but not for women, studies that used sexual behavior over a longer period tended to result in weaker sexual orientation differences for depression and suicidality than studies with other dimensions or a mixture of dimensions (Plöderl et al., 2006). Moreover, most studies report a stronger sexual orientation difference for depression and suicidality among men than among women, but this effect was reversed in two studies that used the behavioral dimension over an extended time (Frisell, Lichtenstein, Rahman, & Långström, 2010; Gilman et al., 2001). Thus, using lifetime homosexual behavior may obscure sexual orientation differences of suicidality among men. However, this may not apply to other mental health problems and, as mentioned in the introduction, current research gives no clear picture which dimension of sexual orientation is associated with the strongest risks for various mental health problems. Therefore, more research about the covariation of risk and protective factors with different dimensions of sexual orientation is needed. However, we again want to stress that an increased risk for having attempted suicide occurred for all dimensions of sexual orientation in our study. Similarly, there were no significant heterogeneities among the studies that entered in a meta-analysis and that used different dimensions of sexual orientation (Plöderl et al., 2006). Thus, it is likely that future studies will also find significant sexual orientation differences independent of the

Table 6 Suicide attempts with clear intent to die and with medical treatment as a function of sexual orientation (total sample, $N = 1{,}382$)

Sexual orientation	N[a]	%	Suicide attempts											
			Gate question				Clear expectation of death				With medical treatment			
			n	%	OR (95% CI)	*p*	*n*	%	OR (95% CI)	*p*	*n*	%	OR (95% CI)	*p*
Fantasies														
Heterosexual	1051	(76)	30	(3)	Ref. cat.		12	(1)	Ref. cat.		11	(1)	Ref. cat.	
Bisexual	308	(22)	25	(8)	3.0 (1.7–5.2)	**	3	(1)	0.9 (0.2–2.8)		8	(3)	2.5 (1.0–6.4)	
Homosexual	20	(1)	4	(20)	8.7 (2.3–25.7)	**	2	(10)	10.1 (1.4–41.6)	*	3	(15)	17.0 (3.4–61.9)	**
Bisexual/homosexual combined	328	(24)	29	(9)	3.3 (1.9–5.6)	**	5	(2)	1.4 (0.4–3.8)		11	(3)	3.3 (1.4–7.8)	**
Preferred partner														
Heterosexual	1184	(86)	37	(3)	Ref. cat.		13	(1)	Ref. cat.		12	(1)	Ref. cat.	
Bisexual	177	(13)	18	(10)	3.5 (1.9–6.3)	**	2	(1)	1.1 (0.2–4.1)		7	(4)	4.1 (1.5–10.3)	**
Homosexual	19	(1)	4	(21)	8.4 (2.3–24.9)	**	2	(11)	11.1 (1.5–45.4)	*	3	(16)	18.7 (3.8–67.5)	**
Bisexual/homosexual combined	196	(14)	22	(11)	3.9 (2.2–6.8)	**	4	(2)	1.9 (0.5–5.6)		10	(5)	5.3 (2.2–12.5)	**
Self-identification														
Heterosexual	1219	(88)	38	(3)	Ref. cat.		13	(1)	Ref. cat.		14	(1)	Ref. cat.	
Bisexual	107	(8)	13	(12)	4.3 (2.1–8.2)	**	0	(0)	–		3	(3)	2.6 (0.6–8.2)	
Homosexual	16	(1)	3	(19)	7.4 (1.6–24.5)	*	2	(13)	13.9 (1.9–58.0)	*	3	(19)	20.3 (4.1–73.4)	**
Bisexual/homosexual combined	123	(9)	16	(13)	4.7 (2.4–8.5)	**	2	(2)	1.6 (0.2–6.1)		6	(5)	4.5 (1.5–11.5)	**
Not sure	9	(1)	4	(44)	24.8 (5.7–101.4)	**	1	(11)	12.8 (0.5–80.0)		1	(11)	11.9 (0.5–73.7)	
Behavior (12 mos)														
Heterosexual	1215	(88)	41	(3)	Ref. cat.		10	(1)	Ref. cat.		14	(1)	Ref. cat.	
Bisexual	15	(1)	4	(27)	10.6 (2.7–33.0)	**	0	(0)	–		1	(7)	6.8 (0.3–38.5)	
Homosexual	24	(2)	4	(17)	5.9 (1.6–16.6)	*	2	(8)	11.5 (1.5–48.2)	*	3	(13)	12.6 (2.6–43.0)	**
Bisexual/homosexual combined	39	(3)	8	(21)	7.5 (3.0–16.6)	**	2	(5)	6.9 (0.9–27.8)		4	(10)	10.0 (2.6–30.0)	**
No sex	109	(8)	8	(7)	2.3 (1.0–4.8)		4	(4)	4.7 (1.2–14.5)	*	2	(2)	1.7 (0.2–6.3)	
Behavior (prev.)														
Heterosexual	1274	(92)	44	(3)	Ref. cat.		12	(1)	Ref. cat.		16	(1)	Ref. cat.	
Bisexual	57	(4)	8	(14)	4.6 (1.9–9.9)	**	2	(4)	4.1 (0.6–15.6)		3	(5)	4.5 (1.0–14.3)	
Homosexual	12	(1)	2	(17)	5.9 (0.8–23.7)		1	(8)	10.6 (0.4–63.1)		2	(17)	16.4 (2.2–70.6)	*
Bisexual/homosexual combined	69	(5)	10	(14)	4.8 (2.2–9.7)	**	3	(4)	4.9 (1.1–16.3)	*	5	(7)	6.2 (2.0–16.7)	**
No sex	23	(2)	3	(13)	4.4 (1.0–13.5)		1	(4)	5.4 (0.2–29.5)		0	(0)	–	

* $p < .05$; ** $p < .01$

[a] Numbers may not add up to the total sample size because of missing data on some sexual orientation dimensions

dimension used and how participants were classified as homosexual or bisexual.

A critical issue is where to set the cut-off between heterosexuality and bisexuality on the sexual orientation dimension sexual fantasies, partner preference, and self-identification. However, it made no substantial difference if predominantly heterosexuals were categorized as bisexual or heterosexual. Of course, it would be best to not collapse categories at all, but this would require much larger samples.

Notably, participants who were not sure of their self-identification seemed to have an increased risk for having attempted suicide. Similar findings were also found in other studies for depression, suicidality, and for some forms of substance abuse/dependency (Espelage et al., 2008; McCabe et al., 2009; Olshen, McVeigh, Wunsch-Hitzig, & Rickert, 2007). Espelage et al. discuss that questioning youth, who had the most problematic indices of mental health in their study, may have no support from other gay and lesbian peers and may use dysfunctional strategies such as alcohol abuse instead. Those who chose the "unsure" category may be still in the process of forming a homosexual identity, which was described by Cass (1979) as "identity confusion" phase. This assumption is undermined by the relatively lower mean age of the questioning participants. Furthermore, the unsure category may have been chosen by those who do not accept their homosexual feelings and this may possibly place them at risk for suicide because of the self-hate involved (Cass, 1979). Rejection of one's homosexual identity may result from a stronger exposure to homophobia, as found in the study of Espelage et al. (2008). Finally, confusion about or a change of sexual orientation is common

Table 7 Suicide attempts as a function of sexual orientation with a stricter cut-off for bisexuality (women, $N = 726$)

Sexual orientation	N^a	%	Suicide attempts											
			Gate question				With some intent to die				With strong intent to die			
			n	%	OR (95% CI)	*p*	*n*	(%)	OR (95% CI)	*p*	*n*	(%)	OR (95% CI)	*p*
Fantasies														
Heterosexual	675	(93)	29	(4)	Ref. cat.		22	(3)	Ref. cat.		14	(2)	Ref. cat.	
Bisexual	39	(5)	3	(8)	1.9 (0.4–5.8)		3	(8)	2.6 (0.6–8.0)		1	(3)	1.4 (0.1–7.3)	
Homosexual	9	(1)	1	(11)	3.1 (0.1–18.2)		1	(11)	4.1 (0.2–24.6)		1	(11)	6.5 (0.2–40.4)	
Bisexual/homosexual combined	48	(7)	4	(8)	2.1 (0.6–5.6)		4	(8)	2.8 (0.8–7.7)		2	(4)	2.2 (0.3–8.2)	
Preferred partner														
Heterosexual	697	(96)	28	(4)	Ref. cat.		22	(3)	Ref. cat.		14	(20)	Ref. cat.	
Bisexual	21	(3)	4	(19)	5.7 (1.5–17.0)	*	3	(14)	5.3 (1.1–17.3)	*	1	(5)	2.7 (0.1–14.9)	
Homosexual	7	(1)	1	(14)	4.4 (0.2–28.1)		1	(14)	5.6 (0.2–36.5)		1	(14)	8.9 (0.3–59.6)	
Bisexual/homosexual combined	28	(4)	5	(18)	5.3 (1.6–14.1)	**	4	(14)	5.2 (1.4–15.2)	*	2	(7)	4.0 (0.6–15.4)	
Self-identification														
Heterosexual	677	(93)	26	(4)	Ref. cat.		20	(3)	Ref. cat.		13	(2)	Ref. cat.	
Bisexual	14	(2)	3	(21)	7.0 (1.4–24.6)	*	4	(29)	13.2 (3.3–44.4)	**	1	(7)	4.4 (0.2–25.2)	
Homosexual	6	(1)	1	(17)	5.5 (0.2–37.6)		1	(17)	7.2 (0.3–49.9)		1	(17)	11.1 (0.4–79.6)	
Bisexual/homosexual combined	20	(3)	4	(20)	6.4 (1.7–19.1)	**	5	(25)	11.0 (3.2–32.2)	**	2	(10)	6.0 (0.8–24.3)	

* $p < .05$; ** $p < .01$

[a] Numbers may not add up to the total sample size because of missing data on some sexual orientation dimensions

Table 8 Suicide attempts as a function of sexual orientation with a stricter cut-off for bisexuality (men, $N = 656$)

Sexual orientation	N^a	%	Suicide attempts											
			Gate question				With some intent to die				With strong intent to die			
			n	(%)	OR (95% CI)	*p*	*n*	(%)	OR (95% CI)	*p*	*n*	(%)	OR (95% CI)	*p*
Fantasies														
Heterosexual	634	(97)	21	(3)	Ref. cat.		16	(3)	Ref. cat.		14	(2)	Ref. cat.	
Bisexual	11	(2)	2	(18)	6.8 (0.9–29.2)		2	(18)	8.9 (1.2–39.4)	*	1	(9)	4.9 (0.2–29.2)	
Homosexual	11	(2)	3	(27)	11.1 (2.2–42.8)	**	3	(27)	14.7 (2.8–57.8)	**	2	(18)	10.2 (1.3–45.8)	*
Bisexual/homosexual combined	22	(3)	5	(23)	8.7 (2.6–24.7)	**	5	(23)	11.4 (3.3–33.7)	**	3	(14)	7.2 (1.5–24.8)	*
Preferred partner														
Heterosexual	633	(96)	22	(3)	Ref. cat.		17	(3)	Ref. cat.		14	(2)	Ref. cat.	
Bisexual	10	(2)	1	(10)	3.4 (0.1–20.0)		1	(10)	4.5 (0.1–26.5)		1	(10)	5.5 (0.2–32.9)	
Homosexual	12	(2)	3	(25)	9.4 (1.9–35.1)	**	3	(25)	12.3 (2.4–46.6)	**	2	(17)	9.2 (1.2–40.3)	*
Bisexual/homosexual combined	22	(3)	4	(18)	6.3 (1.7–18.9)	**	4	(18)	8.2 (2.1–25.3)	**	3	(14)	7.2 (1.5–24.7)	*
Self-identification														
Heterosexual	624	(95)	20	(3)	Ref. cat.		15	(2)	Ref. cat.		12	(2)	Ref. cat.	
Bisexual	11	(2)	2	(18)	7.0 (0.9–30.3)		2	(18)	9.4 (1.2–41.7)	*	2	(18)	11.7 (1.5–53.7)	*
Homosexual	10	(2)	2	(20)	7.9 (1.0–35.0)	*	2	(20)	10.5 (1.4–48.1)	*	1	(10)	6.3 (0.2–38.6)	
Bisexual/homosexual combined	21	(3)	4	(19)	7.2 (1.9–22.0)	**	4	(19)	9.7 (2.5–30.6)	**	3	(14)	8.7 (1.8–31.0)	*

OR odds ratios, *CI* confidence interval

* $p < .05$; ** $p < .01$

[a] Numbers may not add up to the total sample size because of missing data on some sexual orientation dimensions

among individuals with a borderline personality disorder (Reich & Zaranini, 2008; Wilkinson-Ryan & Westen, 2000), and the association of borderline disorder with suicidality is well known (Paris, 2004). Clearly, more research about those who are not sure or question their sexual identity is needed and using this category would be of great importance in future studies.

A main outcome of this study was that the sexual orientation differences remained or even increased for more rigorous classifications of suicide attempts, i.e., including intent to die or injuries. This replicates previous studies that also used more rigorous definitions of suicide attempts (Bagley & Tremblay, 1997), including several large scaled Youth Risk Behavior Surveys (Bagley & Tremblay, 2000, Table 1; Tremblay & Ramsay, 2004, Table 4). It remains a debate if there is also an increased rate of completed suicides among sexual minorities, mainly because psychological autopsy studies did not find increased rates of homosexual or bisexual individuals among those who died by suicide. However, there are important caveats in these studies, for example the assessment of sexual orientation (McDaniel et al., 2001). An elevated risk for completed suicides was reported by a Danish study that used national registers (and thus overcame self-report problems): individuals in registered same-sex partnerships had about three-times greater risk for having committed suicide than those in opposite-sex marriages (Qin, Agerbo, & Mortensen, 2003). Moreover, a homosexual or bisexual orientation was a strong predictor for future suicides among psychiatric patients (Martin, Cloninger, Guze, & Clayton, 1985; Motto, Heilbron, & Juster, 1985) and a homosexual or bisexual self-identification was clearly overrepresented (25%) among failed suicides (Weyrauch, Roy-Byrne, Katon, & Wilson, 2001). Given these findings, it may be plausible to assume that the increased suicide risk of sexual minorities is not a result of weaknesses of research methods.

Another possible caveat in our study was that the sample differed with respect to age and education from the Austrian population, and there were significant differences for age and education on some dimensions of sexual orientation. However, controlling for age and education did not alter the associations between sexual orientation and suicide attempt rates. Reviewers of this article also questioned the representativeness of the sample because of the high rate of participants, especially women, with homosexual or bisexual fantasies or partner preference. Besides representativeness problems, this may have also resulted from a too liberal cut-off between heterosexual and bisexual participants. Indeed, using a stricter cut-off dropped the proportion of homosexual or bisexual participants to rates that are comparable to those of representative studies (e.g., Mosher, Chandra, & Jones, 2005). Furthermore, high rates of individuals with some degree of homosexual or bisexual orientation were also found in representative samples. For example, 16% of participants reported moderate or strong same sex attraction in a German study (Steffens & Wagner, 2004), and some potential of non-heterosexuality was found in 33% of women and 65% of men in a Finnish study (Santtila et al., 2007). Therefore, the rate of non-heterosexually classified participants seems to depend strongly on the dimensions and criteria used to assess sexual orientation. However, given our sampling procedure, biases are likely to be present in our study. Thus it remains speculative if our sample overestimated or underestimated the population suicide attempt rates across different dimensions of sexual orientation. For example, in snowball samples, popular individuals are likely over-sampled, isolated individuals tend to be neglected (Heckathorn, 2002). Isolation is a risk factor for suicide, and this may have lead to an underestimation of the true problem. However, the suicide attempt rates in our study are comparable to those of a German study (Weissman et al., 1999). Thus, given that suicide attempt rates are comparable in Austria and Germany, the suicide attempt rate in our sample seems to be quite representative. Another caveat could have been caused by the homophily bias: the recruiter and the recruited participants tend to be similar (Heckathorn, 2002), thus the results could be biased towards the results that would be obtained from a student sample. Real random samples that investigate the link between sexual orientation and suicidality could overcome these biases, but are still lacking in German speaking regions. However, such studies still may not remove all sampling biases (Meyer, 2009). Another possible caveat is that more rigorous classifications of suicide attempters were again based on self-reports. This may involve similar biases (e.g., memory distortions) already apparent in the simple gate question on suicide attempts. There were also some missing data on the follow-up items. Additional information going beyond self-reports would be beneficial but are hard to gather in epidemiological surveys. A possible solution for enhancing the validity of self-reported suicidal behavior would be structured interviews (e.g., Linehan, Comtois, Brown, Heard, & Wagner, 2006); however, a face-to-face approach may lead to underreporting of sexual orientation related information. For some dimensions of sexual orientation, our study was also limited because of the relatively small number of participants classified as homosexual or bisexual, especially homosexual women. This resulted in wide confidence intervals and, in a few cases, no statistical test was possible because no suicide attempter was in the subgroup. Given the previous findings, it is likely that all differences will become statistically significant with increased statistical power, but only larger samples can clarify this issue. Participants who were classified as bisexual had a comparable or smaller risk difference than homosexual participants in most cases when compared with heterosexuals. However, given the relatively small numbers of homosexually or bisexually classified participants, such comparisons should be interpreted with caution.

We did not assess confounding variables that likely influence the suicide attempt risk, such as depression, substance abuse, internalized homophobia, or coming out. It is very likely that controlling for such risk factors may decrease the association of suicide attempts and sexual orientation (e.g., Russell & Joiner, 2001; Safren & Heimberg, 1999). This was already the goal in a previous Austrian study (Plöderl & Fartacek, 2005). The focus of our study was on the importance of

different dimensions of sexual orientation, a detailed assessment of suicide attempts, and sampling not explicitly from the gay or lesbian communities. Furthermore, we wanted to keep the questionnaire as short as possible to increase the response rate.

The strengths of this study included that we did not sample LGB participants from gay or lesbian communities. Thus, the study is a unique contribution to the question if suicide attempts are increased among sexual minority individuals in German-speaking Europe. Moreover, our study is one of the few that used different dimensions of sexual orientation simultaneously; additionally, the response rate was relatively high. Finally, only few previous studies assessed suicide attempts in accordance with current taxonomies in suicidology.

The present study found increased lifetime rates of suicide attempts among Austrian men and women who reported homosexual or bisexual attraction, behavior, or self-identification, compared to their heterosexual counterparts. This increased risk remained for even more rigorous definitions of suicide attempts. The results were in line with previous studies and thus indicate that suicide risk is not a methodological artefact but really increased among sexual minority individuals. Therefore, suicide preventive efforts should include sexual minorities as target groups.

References

Aragon, T. (2007). epitools: Epidemiology tools [Computer software manual]. Available at http://www.epitools.net (R package version 0.4-9).

Bagley, C., & Tremblay, P. (1997). Suicidal behaviors in homosexual and bisexual males. *Crisis, 18*, 24–34.

Bagley, C., & Tremblay, P. (2000). Elevated rates of suicidal behavior in gay, lesbian, and bisexual youth. *Crisis, 21*, 111–117.

Bagley, C., Wood, M., & Young, L. (1994). Victim to abuser: Mental health and behavioral sequels of child sexual abuse in a community survey of young adult males. *Child Abuse and Neglect, 18*, 683–697.

Berliner Senatsverwaltung für Bildung, Jugend und Sport. (1998). *Sie liebt sie. Er liebt ihn. Eine Studie zur psychosozialen Lage junger Lesben, Schwuler und Bisexueller in Berlin*.

Biechele, U. (2001). *Schwule Jugendliche: Lebenssituation und psychosozialer Hilfebedarf. In Beratung von Lesben und Schwulen. Dokumentation der VII Fachtagung des Verbandes lesbischer Psychologinnen und schwuler Psychologen in Deutschland*. Berlin: Deutsche AIDS-Hilfe.

Cass, V. C. (1979). Homosexual identity formation: A theoretical model. *Journal of Homosexuality, 4*, 219–235.

Dannecker, M., & Reiche, R. (1974). *Der gewöhnliche Homosexuelle. Eine soziologische Untersuchung über männliche Homosexuelle in der Bundesrepublik*. Frankfurt: Fischer.

Espelage, D., Aragon, S., Birkett, M., & Koenig, B. (2008). Homophobic teasing, psychological outcomes, and sexual orientation among high school students: What influence do parents and schools have. *School Psychology Review, 37*, 202–216.

Frisell, T., Lichtenstein, P., Rahman, Q., & Långström, N. (2010). Psychiatric morbidity associated with same-sex sexual behaviour: Influence of minority stress and familial factors. *Psychological Medicine, 40*, 315–324.

Gilman, S. E., Cochran, S. D., Mays, V. M., Hughes, M., Ostrow, D., & Kessler, R. C. (2001). Risk of psychiatric disorders among individuals reporting same-sex sexual partners in the National Comorbidity Survey. *American Journal of Public Health, 91*, 933–939.

Heckathorn, D. D. (2002). Respondent-driven sampling II: Deriving valid population estimates from chain-referral samples of hidden populations. *Social Problems, 49*, 11–34.

Herek, G. M., & Garnets, L. D. (2007). Sexual orientation and mental health. *Annual Review of Clinical Psychology, 3*, 353–375.

King, M., Semlyen, J., Tai, S. S., Killaspy, H., Osborn, D., Popelyuk, D., et al. (2008). A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. *BMC Psychiatry, 8*, 70. doi:10.1186/1471-244X-8-70.

Knox, S., Kippax, S., Crawford, J., Prestage, G., & Van De Ven, P. (1999). Non-prescription drug use by gay men in Sydney, Melbourne and Brisbane. *Drug and Alcohol Review, 18*, 425–433.

Lewis, R. J., Derlega, V. J., Berndt, A., Morris, L. M., & Rose, S. (2001). An empirical analysis of stressors for gay men and lesbians. *Journal of Homosexuality, 42*, 63–88.

Linehan, M. M., Comtois, K. A., Brown, M. Z., Heard, H. L., & Wagner, A. (2006). Suicide attempt self-injury interview SASII: Development, reliability, and validity of a scale to assess suicide attempts and intentional self-injury. *Psychological Assessment, 18*, 303–312.

Mao, L., Kidd, M. R., Rogers, G., Andrews, G., Newman, C. E., Booth, A., et al. (2009). Social factors associated with major depressive disorder in homosexually active, gay men attending general practices in urban Australia. *Australian and New Zealand Journal of Public Health, 33*, 83–86.

Martin, R. L., Cloninger, C. R., Guze, S. B., & Clayton, P. J. (1985). Mortality in a follow-up of 500 psychiatric outpatients. II. Cause-specific mortality. *Archives of General Psychiatry, 42*, 58–66.

McCabe, S. E., Hughes, T. L., Bostwick, W. B., West, B. T., & Boyd, C. J. (2009). Sexual orientation, substance use behaviors and substance dependence in the United States. *Addiction, 104*, 1333–1345.

McDaniel, J. S., Purcell, D., & D'Augelli, A. R. (2001). The relationship between sexual orientation and risk for suicide: Research findings and future directions for research and prevention. *Suicide and Life-Threatening Behavior, 31*, 84–105.

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin, 129*, 674–697.

Meyer, I. H. (2009). Sampling lesbian, gay, and bisexual populations. *Journal of Counseling Psychology, 56*, 23–31.

Mills, T. C., Paul, J., Stall, R., Pollack, L., Canchola, C., Chang, Y. J., et al. (2004). Distress and depression in men who have sex with men: The Urban Men's Health Study. *American Journal of Psychiatry, 161*, 278–285.

Mosher, W. D., Chandra, A., & Jones, J. (2005). Sexual behavior and selected health measures: Men and women 15–44 years of age, United States. *Advance Data, 362*, 1–55.

Motto, J. A., Heilbron, D. C., & Juster, R. P. (1985). Development of a clinical instrument to estimate suicide risk. *American Journal of Psychiatry, 142*, 680–686.

Muehrer, P. (1995). Suicide and sexual orientation: A critical summary of recent research and directions for future research. *Suicide and Life-Threatening Behavior, 25*(Suppl), 72–81.

Olshen, E., McVeigh, K. H., Wunsch-Hitzig, R. A., & Rickert, V. I. (2007). Dating violence, sexual assault, and suicide attempts among urban teenagers. *Archives of Pediatrics and Adolescent Medicine, 161*, 539–545.

Paris, J. (2004). Half in love with easeful death: The meaning of chronic suicidality in borderline personality disorder. *Harvard Review of Psychiatry, 12*, 42–48.

Pathela, P., Hajat, A., Schillinger, J., Blank, S., Sell, R., & Mostashari, F. (2006). Discordance between sexual behavior and self-reported sexual identity: A population-based survey of New York City men. *Annals of Internal Medicine, 145*, 416–425.

Plöderl, M. (2005). *Sexuelle Orientierung, psychische Gesundheit und Suizidalität*. Weinheim: Beltz-PVU.

Plöderl, M., & Fartacek, R. (2005). Suicidality and associated risk factors among lesbian, gay, and bisexual compared to heterosexual Austrian adults. *Suicide and Life-Threatening Behavior, 35*, 661–670.

Plöderl, M., Kralovec, K., Yazdi, K., & Fartacek, R. (in press). A closer look at self-reported suicide attempts: False positives and false negatives. *Suicide and Life-Threatening Behavior.*

Plöderl, M., Sauer, J., & Fartacek, R. (2006). Suizidalität und psychische Gesundheit von homo- und bisexuellen Männern und Frauen. Eine metaanalytische Zusammenfassung internationaler Zufallsstichproben. *Verhaltenstherapie und psychosoziale Praxis*, *117*, 4–10.

Qin, P., Agerbo, E., & Mortensen, P. B. (2003). Suicide risk in relation to socioeconomic, demographic, psychiatric, and familial factors: A national register-based study of all suicides in Denmark, 1981–1997. *American Journal of Psychiatry, 160*, 765–772.

R Development Core Team. (2008). R: A language and environment for statistical computing [Computer software manual]. Available at http://www.R-project.org.

Reich, D. B., & Zanarini, M. C. (2008). Sexual orientation and relationship choice in borderline personality disorder over ten years of prospective follow-up. *Journal of Personality Disorders, 22*, 564–572.

Russell, S. T. (2003). Sexual minority youth and suicide risk. *American Behavioral Scientist, 46*, 1241–1257.

Russell, S. T., & Joyner, K. J. (2001). Adolescent sexual orientation and suicide risk: Evidence from a national study. *American Journal of Public Health, 91*, 1276–1281.

Safren, S. A., & Heimberg, R. G. (1999). Depression, hopelessness, suicidality, and related factors in sexual minority and heterosexual adolescents. *Journal of Consulting and Clinical Psychology, 67*, 859–866.

Santtila, P., Sandnabba, N. K., Harlaar, N., Varjonen, M., Alanko, K., & von der Pahlen, B. (2007). Potential for homosexual response is prevalent and genetic. *Biological Psychology, 77*, 102–105.

Savin-Williams, R. C. (2001). Suicide attempts among sexual-minority youths: Population and measurement issues. *Journal of Consulting and Clinical Psychology, 69*, 983–991.

Savin-Williams, R. C. (2006). Who's gay? Does it matter? *Current Directions in Psychological Science, 15*, 40–44.

Savin-Williams, R. C. (2008). Then and now: Recruitment, definition, diversity, and positive attributes of same-sex populations. *Developmental Psychology, 44*, 135–138.

Savin-Williams, R. C., & Ream, G. L. (2003). Suicide attempts among sexual-minority male youth. *Journal of Clinical Child and Adolescent Psychology, 32*, 509–522.

Silverman, M. M., Berman, A. L., Sanddal, N. D., O'Carroll, P. W., & Joiner, T. E. (2007). Rebuilding the tower of Babel: A revised nomenclature for the study of suicide and suicidal behaviors part 2: Suicide-related ideations, communications, and behaviors. *Suicide and Life-Threatening Behavior, 37*, 264–277.

Statistik Austria. (2006). *Demografisches Jahrbuch 2005*. Available at http://www.statistik.at/web_de/Redirect/index.htm?dDocName=0170.

Statistik Austria. (2009). *Bildungsstand der Bevölkerung*. Available at http://www.statistik.at/web_de/statistiken/bildung_und_kultur/bildungsstand_der_bevoelkerung/020912.html.

Steffens, M. C., & Wagner, C. (2004). Attitudes towards lesbians, gay men, bisexual women, and bisexual men in Germany. *Journal of Sex Research, 41*, 137–149.

Tremblay, P., & Ramsay, R. (2004). *The changing social construction of western male homosexuality: Associations with worsening youth suicide problems*. Poster presented at the conference of the Canadian Association for Suicide Prevention, Alberta. Available at http://www.youth-suicide.com/#the-changing.

Weissman, M. M., Bland, R. C., Canino, G. J., Greenwald, S., Hwu, H. G., Joyce, P. R., et al. (1999). Prevalence of suicide ideation and suicide attempts in nine countries. *Psychological Medicine, 29*, 9–17.

Weyrauch, K. F., Roy-Byrne, P., Katon, W., & Wilson, L. (2001). Stressful life events and impulsiveness in failed suicide. *Suicide and Life-Threatening Behavior, 31*, 311–319.

Wichstrøm, L., & Hegna, K. (2003). Sexual orientation and suicide attempt: A longitudinal study of the general Norwegian population. *Journal of Abnormal Psychology, 112*, 144–151.

Wilkinson-Ryan, T., & Westen, D. (2000). Identity disturbance in borderline personality disorder: An empirical investigation. *American Journal of Psychiatry, 157*, 528–541.

Arch Sex Behav (2010) 39:1415–1428
DOI 10.1007/s10508-009-9570-y

ORIGINAL PAPER

Correlates of Gender Dysphoria in Taiwanese University Students

Meng-Chuan Lai · Yen-Nan Chiu · Kenneth D. Gadow · Susan Shur-Fen Gau · Hai-Gwo Hwu

Received: 16 October 2008 / Revised: 24 April 2009 / Accepted: 19 September 2009 / Published online: 25 November 2009

Abstract There have been no published reports regarding the epidemiological and psychiatric features of gender dysphoria in non-clinical young adults. The current study aimed to investigate the demographics, co-occurring psychiatric symptoms, and perceived parenting style and family support in Taiwanese young adults with gender dysphoria. The sample consisted of 5010 university freshmen (male, 51.6%) with a mean age of 19.6 years ($SD = 2.7$) from a national university in Taiwan. The questionnaires used for this university-based survey included the Adult Self Report Inventory-4 for psychopathology (including gender dysphoria), the Parental Bonding Instrument for parenting style, and the Family APGAR for perceived family support. Results showed that gender dysphoria was more prevalent in females (7.3%) than males (1.9%). Young adults with gender dysphoria were more likely to meet a wide but specific range of co-occurring psychiatric symptoms. The most significantly associated symptoms for males were agoraphobia, hypochondriasis, manic episode, and pathological gambling, and for females dissociative disorder, hypochondriasis, and body dysmorphic disorder. Both males and females with gender dysphoria perceived significantly less support from their families and less affection/care from both parents. Findings suggest that gender dysphoria, associated with a specific range of psychopathology and family/parenting dissatisfaction (with both similar and dissimilar patterns between sexes), is not uncommon in Taiwanese university students, particularly in females. This implies the importance of attention and specific measures to offset psychiatric conditions and to promote mental well-being of this population.

Keywords Gender dysphoria · Gender identity · University students · Psychopathology · Parenting style

Introduction

Gender dysphoria (GD) describes a group of individuals with distress resulting from the sense of incongruity between one's gender identity and birth sex (Brown, 1990; Cohen-Kettenis & Gooren, 1999; Zucker, 2008). Different intensities of GD range from the "unease about one's own sex" (Zucker, 2007) to the clinical diagnosis of gender identity disorder (GID) (American Psychiatric Association, 2000) or transsexualism (World Health Organization, 1992). GD also describes a "perceived suffering and conflict around [the] sense of gender and self" (Lothstein, 2006), and can be conceptualized as the constitution of cognitive (e.g., wish to be of the opposite sex), affective (e.g., distress toward physical sexual characteristics), and behavioral (e.g., cross-dressing) dimensions, which are not invariably presented in a fixed pattern for every individual with GD. Clinically, GD is the core of the diagnostic construct of GID and transsexualism. One's dissatisfaction toward the assigned

M.-C. Lai · Y.-N. Chiu · S. S.-F. Gau (✉) · H.-G. Hwu
Department of Psychiatry, National Taiwan University Hospital, No. 7, Chung-Shan South Road, 10002 Taipei, Taiwan
e-mail: gaushufe@ntu.edu.tw

M.-C. Lai
Department of Psychiatry, University of Cambridge, Cambridge, UK

K. D. Gadow
Department of Psychiatry and Behavioral Science, State University of New York, Stony Brook, NY, USA

S. S.-F. Gau · H.-G. Hwu
Department of Psychiatry, College of Medicine, National Taiwan University, Taipei, Taiwan

gender role rather than the inborn biological sex could also mimic the presentation of GD; hence, this warrants careful differentiation.

To our knowledge, there have been no published large-scale epidemiological studies directly examining the prevalence of GID in children and adolescents. However, some large-scale community surveys examined the epidemiological features of various dimensions of the broader construct of GD (Bailey, Dunne, & Martin, 2000; Coolidge, Thede, & Young, 2002; Iervolino, Hines, Golombok, Rust, & Plomin, 2005; Knafo, Iervolino, & Plomin, 2005; van Beijsterveldt, Hudziak, & Boomsma, 2006; Zucker & Bradley, 1995; Zucker, Bradley, & Sanikhani, 1997). Based on two items on the Child Behavior Checklist (CBCL), 1.0% of boys and 2.3% of girls were reported by their mothers that the child "*behaves like opposite sex*" and 0% of boys and 1.0% of girls have "*wishes to be of opposite sex*" among children aged 4–11 (Zucker et al., 1997). A recent survey on a large Dutch twin sample (van Beijsterveldt et al., 2006) showed the prevalence of maternal reports of "somewhat true" or "very true" on "*behaves like opposite sex*" as 3.4 and 2.4% in boys, and 5.2 and 3.4% in girls, ages 7 and 10 years old, respectively; and "*wishes to be of opposite sex*" as 1.0% of boys, ages 7 and 10 years old, and 1.7 and 0.9% of girls, ages 7 and 10 years old, respectively. Assessed by a six-item DSM-IV-based GID scale, the prevalence of clinically significant GID symptoms was 2.3% (0.7% in males and 3.7% in females) among an American community twin sample aged 4–17 (Coolidge et al., 2002). These numbers suggest that gender nonconformity and GD are not rare in child and adolescent populations and are more prevalent in females, contrary to the male-predominance of adult transsexuals in clinical populations (Bakker, van Kesteren, Gooren, & Bezemer, 1993; De Cuypere et al., 2007).

No direct epidemiological study has reported the rate for GD in adults; however, the estimates can be inferred from rates of adult transsexuals receiving medical services. Earlier studies in the Europe, Australia, and the U.S. reported a prevalence of 1 in 100,000 to 24,000 for male-to-female and 1 in 400,000 to 100,000 for female-to-male transsexuals (Hoenig & Kenna, 1974; Pauly, 1968; Ross, Walinder, Lundstrom, & Thuwe, 1981; Walinder, 1971). Later European studies revealed higher rates: 1 in 11,900 for male-to-female and 1 in 30,400 for female-to-male in the Netherlands (Bakker et al., 1993), 1 in 130,000 for both sexes in a Swedish region (Bodlund, Kullgren, Sundbom, & Hojerback, 1993), 1 in 36,000 for male-to-female and 1 in 94,000 for female-to-male in the former Federal Republic of Germany (Weitze & Osburg, 1996), and 1 in 12,900 for male-to-female and 1 in 33,800 for female-to-male in Belgium (De Cuypere et al., 2007). An indirect survey through Scottish general practitioners reported that around 1 in 7,440 male and 1 in 31,150 female patients over 15-years-old had GD (Wilson, Sharp, & Carr, 1999). Nearly all these Western studies demonstrated a male-predominance, yet a Polish study reported a striking female predominance (M:F = 1:5.5) (Godlewski, 1988). For Asian countries, a Singapore study reported 1 in 2,900 for male-to-female and 1 in 8,300 for female-to-male transsexuals (Tsoi, 1988). One population-based epidemiological study in Taiwan (1981–1986), using the Chinese Modified Diagnostic Interview Schedule (Hwu, Yeh, & Chang, 1989), reported that 1 in 770–2,500 indicated "*wishes to be recognized as the opposite sex*" (cognitive dimension), 1 in 590–5,000 "*intended to cross-dress or receive sex-reassignment treatment*" (behavioral dimension), and the lifetime prevalence rates for DSM-III adult transsexualism (American Psychiatric Association, 1980) were 1 in 5,680 for males and 1 in 590 for females (Chang, Hwu, & Yen, 2006; Hwu et al., 1989). The female-predominance and higher prevalence of adult transsexualism and GD in Taiwan compared to those reported in the European countries and Singapore are of great interest though no additional research on this topic has been conducted.

Although some studies reported various concurrent psychopathologies with GID, in which condition GD acted as a primary (i.e., other concurrent psychiatric symptoms were secondary to the suffering resulting from GD) or secondary presentation (i.e., GD was an epiphenomenon of other mental illnesses) (a Campo, Nijman, Merckelbach, & Evers, 2003), other studies on adult transsexuals showed low levels (normal range) of concurrent psychopathology before or after medical treatments (Cole, O'Boyle, Emory, & Meyer, 1997; Haraldsen & Dahl, 2000), indicating transsexualism to be an isolated diagnosis. Moreover, for those who suffered from concurrent mood and anxiety disorders, these comorbid conditions usually disappeared after the transsexualism was treated (Cole et al., 1997; Meyer, 2004). The role of concurrent psychopathology to GD thus warrants further investigation. Children with GID are reported to be more likely to exhibit emotional/behavioral problems, particularly internalizing symptoms (Cohen-Kettenis, Owen, Kaijser, Bradley, & Zucker, 2003; Wallien, Swaab, & Cohen-Kettenis, 2007; Zucker, 2007; Zucker & Bradley, 1995), in which separation anxiety (Sreenivasan, 1985; Zucker, Bradley, & Lowry Sullivan, 1996) figures into the ontogenesis of GID according to some investigators (Coates, 1990; Coates & Person, 1985; Zucker & Bradley, 1995). Others, however, also report externalizing behavior problems (Coolidge et al., 2002; Wallien et al., 2007). On the other hand, adults with GID may exhibit high rates of DSM-IV co-morbidity, including mood disorders, substance-related disorders, anxiety disorders, and somatoform disorders (Hepp, Kraemer, Schnyder, Miller, & Delsignore, 2005), and 70% of Taiwanese adults with transsexualism reported at least one other DSM-III psychiatric diagnosis, particularly anxiety disorders (Chang et al., 2006). It is possible that childhood and adulthood GD share both similar (e.g., internalizing symptoms) and dissimilar co-occurring psychopathologies, but this needs further formal comparisons.

Parental influences and family dynamics have been proposed to explain the ontogenesis of GD (Levine & Lothstein, 1981; Meyer, 1982), yet empirical studies do not support these assumptions (Green, 1987; Zucker & Bradley, 1995). Based on limited data, adult male-to-female transsexuals perceived their fathers rather than mothers as less emotionally warm, more rejecting, and more over-controlling (Cohen-Kettenis & Arrindell, 1990; Parker & Barr, 1982); adult female-to-male transsexuals reported both parents as more rejecting and less emotionally warm, but only their mothers as more overprotective (Cohen-Kettenis & Arrindell, 1990). Nevertheless, the causal relationship of such associations with GD is not yet known.

Although previous Western studies have reported a relatively high rate of GD and its equivalents in child and adolescent populations, we know little about its epidemiological and psychiatric features in young adulthood, a critical period for the development of various dimensions of the self, including the gender aspects. Owing to the lack of epidemiological knowledge about GD in this age group, we conducted a school-based survey on Taiwanese first-year university students to obtain the prevalence and demographics of GD in young adulthood, and to examine its psychiatric co-occurring symptoms and associations with perceived parenting style and family support. This is one of the first studies in the world reporting the epidemiological and psychiatric features of GD in a non-clinical, large-scale sample of young adult university students.

Method

Participants

The participants consisted of 5015 first-year college students (male: 51.6%, M age, 19.6 years; $SD = 2.7$) recruited from the National Taiwan University (NTU). The Research Ethics Committee of the National Taiwan University Hospital approved this study prior to implementation. In July 2003 and July 2004, a research invitation letter was mailed to each of the newly accepted students by the NTU. They were informed that participation in the survey was completely voluntary, and the issue of confidentiality was assured. There was no information about the proportion of eligible students who received the letter. Out of a total of 7531 (3756 in 2003 and 3775 in 2004) new entrants in the two consecutive years, 5015 (participation rate, 66.6%) consented to the study and completed the questionnaires in the first weeks of the fall semesters. The school counselors provided clear instructions on self-administration before the participants started to complete the questionnaires. Trained research assistants then checked the questionnaires to minimize the missing data. There were 5010 students with complete information for the data analysis.

Measures

Adult Self-Report Inventory-4 (ASRI-4)

The ASRI-4 is a 136-item self-reported DSM-IV-referenced rating scale (Gadow, Sprafkin, & Weiss, 2004) examining a broad range of current or lifetime symptoms of DSM-IV diagnoses (see Appendix). Two scoring procedures can be applied: the Symptom Count score (number of DSM-IV-specified symptoms) and the Symptom Severity score (a dimensional scoring). Using a cutoff score for each category of symptoms (for a DSM-referenced diagnosis), the former provides an indication for "clinically significant symptoms" for a specified disorder, which we described in the present study as " symptoms of (name) disorder"; this indicates the presence of significant symptoms of a certain disorder, but does not signify a clinical diagnosis for the lack of information about additional diagnostic criteria (e.g., age of onset, duration, impairment). The latter provides quantitative descriptions of the severity of certain dimensions of specific syndromes. Researches on both clinical and community samples demonstrated that ASRI-4 scores have satisfactory reliability, along with convergent and discriminant validity with corresponding scales of other commonly-used instruments measuring psychopathologies. The ASRI-4 scores also differentiate clinic versus non-clinic subjects, and the severity scores are only minimally correlated with the subjects' education or occupation levels, socioeconomic status, and IQ (Gadow et al., 2004; Sprafkin, Gadow, Weiss, Schneider, & Nolan, 2007).

The Chinese translation of the ASRI-4 was prepared with culturally-relevant colloquial expressions by Gau and colleagues with the performance of a two-way translation, ensuring satisfactory linguistic and content validity. The detailed psychometric properties of the Chinese ASRI-4 were described elsewhere (Chien, Gau, & Gadow, 2009). On rating, each item was assigned to one of four responses: 0 = never, 1 = sometimes, 2 = often, 3 = very often. Most ratings assess *current* (recent 6 months) conditions, except those for impulse-control disorders, depression, and manic episodes, which are rated on a *lifetime* basis. Items related to suicidal behaviors were extracted to generate inquiries about past and current (past 6 months) suicidal ideations, plans, and attempts.

Definition of Gender Dysphoria (GD)

Based on item 29, "*I wish I was the opposite sex,*" of the ASRI-4, *GD* was defined if the response was "often" or "very often" and *definite GD* if "very often." Subjects reporting symptoms of schizophrenia/psychosis and item 29 concurrently were not defined as GD to minimize the false-positive rate.

Parental Bonding Instrument (PBI)

The PBI is a 25-item instrument with a 4-point Likert scale measuring parenting styles during the child's first 16 years with three principle dimensions–affection/care (12 items; e.g., "Could make me feel better when I was upset"), overprotection (7 items; e.g., "Did not want me to grow up"), and authoritarianism (6 items; e.g., "Invaded my privacy") (Cox, Enns, & Clara, 2000). A high score on the affection/care subscale reflects affection and warmth; overprotection reflects overprotective parenting and denial of the child's psychological autonomy; authoritarianism reflects the degree of parental authoritarian control over the child's behaviors. This three-factor structure was well fitted in both Western (Heider et al., 2005; Lizardi & Klein, 2002) and Asian populations (Sato et al., 1999). The reliability and validity of the Chinese version PBI were described elsewhere (Gau et al., 2006).

Family Adaptation, Partnership, Growth, Affection, and Resolve (Family APGAR)

The Family APGAR is a 5-item instrument assessing family function, rated on a 3-point Likert scale responding to five descriptions about the participant's perceived support from, ways of interaction with, and satisfaction toward one's family (Smilkstein, 1978). Higher scores indicate less perceived family support. The Chinese version has been widely used in clinical and community settings for decades in Taiwan (Liao et al., 2002).

Statistical Analysis

We conducted the statistical analyses using SAS 9.1.3 (SAS Institute Inc., Cary, NC). The comparisons of demographic characteristics between the participants with and without GD were presented as frequencies and percentages for the categorical variables with chi-square test for significant test, and mean and *SD* for continuous variables with analysis of variance (ANOVA) for significant test. To control for effects of the lack of independence within the same department derived from a university-based sample, linear and nonlinear multilevel models were used for the analysis of continuous (Symptom Severity scores, parenting styles, and perceived family support) and binary outcomes (Symptom Count scores, i.e., the presence of a particular psychiatric symptom cluster), respectively. We first conducted a nonlinear multi-level logistic regression using the PROC NLMIXED to test whether there was a significant random intercept effect, and did not find a significant correlation in the same department and school; therefore, we used logistic regression models, ignoring the random intercept effect, to investigate the association between GD and psychiatric co-occurring syndromes, and to compute the odds ratios (OR) and 95% confidence intervals (CI) for the OR.

In addition to a linear multi-level model, we used Cohen's *d* to compute the effect sizes for the group comparisons with small, medium, and large effect sizes as Cohen's *d* 0.25–0.5, 0.5–0.8, and $\geq$0.8, respectively (Cohen, 1988). A multiple logistic regression analysis was also performed to identify the most associated psychopathologies for GD using stepwise model selection. Under the *a priori* assumption that males and females may have different patterns of GD-associated characteristics, we conducted the analyses for the two birth sexes separately. All the analyses controlled for participant's age, parents' ages, job status, education, and marital status, and number of offspring in the family.

Results

Prevalence and Demographics

There were 225 (4.5%) participants with *GD*, 176 (7.3%) females and 49 (1.9%) males, with a female versus male risk ratio of 4.1 (CI = 2.9–5.6). There were 71 (1.4%) participants with *definite GD*, with higher rates in female (53, 2.2%) than male participants (18, 0.7%). There were no significant group differences in terms of participant's age, family size, and the parents' age, educational level, occupational status, and marital status (Table 1).

Co-Occurring Symptoms

In univariate exploratory analyses, if adjusted for the multiple tests by setting the pre-selected alpha level as $p < .005$, then for male students GD was significantly associated with symptoms of agoraphobia, panic disorder, obsession, hypochondriasis, body dysmorphic disorder, dysthymia, manic episode, hypersomnia, pathological gambling, dissociation, and schizoid personality ($p < .005$, Table 2, left columns); for female students, GD was significantly associated with symptoms of generalized anxiety disorder, compulsion, hypochondriasis, body dysmorphic disorder, major depression, dysthymia, insomnia, oppositional defiant disorder, intermittent explosive disorder, dissociation, borderline, and schizoid personality ($p < .005$, Table 2, right columns). Regarding symptoms of spectrum disorders, GD was significantly associated with anxiety disorders, obsession–compulsion, and sleep disorders for both sexes, with impulse control disorders for males, and with depressive disorders for females (Table 2). GD was also significantly associated with lifetime suicidal ideation (overall prevalence rate, 46.7%; OR = 2.2; 95% CI = 1.7–2.9) but not suicidal attempt. Participants with GD had more severe symptoms of a wide range of psychopathology with small to medium effect sizes

Table 1 Demographic characteristics by participants with and without symptoms of gender dysphoria

Sociodemographic variables	Gender dysphoria				Statistics
	Yes ($N=225$)		No ($N=4785$)		
	N or M	(%) or (SD)	N or M	(%) or (SD)	
Biological sex					
Male	49	(21.8)	2536	(53.0)	$\chi^2=83.87$, $df=1$,
Female	176	(78.2)	2249	(47.0)	$p<.0001$
Age (years)	19.5	(1.9)	19.8	(2.8)	$F=1.48$
Paternal age at birth of participant (years)	30.2	(4.2)	30.6	(4.4)	$F=2.13$
Maternal age at birth of participant (years)	27.3	(3.5)	27.7	(3.7)	$F=3.24$
Paternal education level					
Senior high or lower	71	(32.3)	1379	(29.8)	$\chi^2=0.59$, $df=1$
College or higher	149	(67.7)	3242	(70.2)	
Maternal education level					
Senior high or lower	104	(47.3)	2062	(44.1)	$\chi^2=0.84$, $df=1$
College or higher	116	(52.7)	2611	(55.9)	
Paternal job					
Professional	56	(25.3)	1224	(26.7)	$\chi^2=0.20$, $df=2$
Technical	123	(55.7)	2504	(54.6)	
Others	42	(19.0)	857	(18.7)	
Maternal job					
Professional	18	(8.2)	408	(9.0)	$\chi^2=0.25$, $df=2$
Technical	95	(43.6)	2023	(44.4)	
Others	105	(48.2)	2126	(46.6)	
Parental marital status					
Married and cohabited	209	(93.3)	4326	(90.9)	$\chi^2=1.55$, $df=1$
Others	15	(6.7)	435	(9.1)	
Number of total offspring	2.4	(0.8)	2.3	(0.8)	$F<1$
Birth order					
Single child	17	(7.6)	416	(8.7)	$\chi^2=1.81$, $df=3$
First child	106	(47.1)	2173	(45.6)	
Middle child	33	(14.7)	590	(12.4)	
Youngest child	69	(30.7)	1589	(33.3)	

(Cohen's d ranging from 0.41 to 0.77) in males and small effect sizes in females (Cohen's d ranging from 0.23 to 0.48, Table 3).

Modification Effects from Biological Sex

Further analysis testing the interactions between biological sex and GD demonstrated that the magnitude of association between GD and symptoms of agoraphobia, obsession, manic episode, trichotillomania, impulse control disorders, anxiety disorders, and obsessive–compulsive disorder were significantly greater in males than in females (Table 2).

Final Fitted Model for Most Associated Psychopathology for GD

Multivariate logistic regression showed that, for the whole sample, the most associated variables for GD were female sex, and symptoms of agoraphobia, hypochondriasis, pathological gambling, and dissociative disorder. For male students, the symptoms of agoraphobia, hypochondriasis, manic episode, and pathological gambling were most significantly associated with GD; for female students, those were symptoms of dissociative disorder, hypochondriasis, and body dysmorphic disorder (Table 4).

Perceived Parenting Style and Family Support

Both sexes of participants with GD perceived less family support, as indicated by higher scores on the Family APGAR than their counterparts with a small effect size; they also perceived significantly less affection and care from both parents with small effect sizes (absolute Cohen's d ranging from 0.36 to 0.49). For females with GD, they additionally reported more

Table 2 Numbers, rates, and odds ratios for significant co-occurring symptoms by participants with and without symptoms of gender dysphoria

Symptoms of	Male						Female					
	Gender dysphoria				Statistics		Gender dysphoria				Statistics	
	Yes ($N=49$)		No ($N=2536$)		OR	CI	Yes ($N=176$)		No ($N=2249$)		OR	CI
	N	%	*N*	%			*N*	%	*N*	%		
Generalized anxiety disorder	6	12.2	168	6.6	2.6	1.1–6.5[j]	24	13.6	138	6.1	2.7	1.7–4.4[m]
Specific phobia	11	22.5	271	10.7	2.5	1.2–5.5[j]	31	17.6	283	12.6	1.5	1.0–2.3
Agoraphobia[a]	10	20.4	135	5.3	5.7	2.6–12.4[m]	21	11.9	172	7.7	1.6	0.9–2.6
Social phobia	13	26.5	336	13.3	2.4	1.2–4.8[j]	25	14.2	221	9.8	1.6	1.0–2.6[j]
Panic disorder	6	12.2	121	4.8	4.0	1.6–9.9[k]	18	10.2	120	5.3	1.7	0.9–3.0
Obsession[b]	21	42.9	408	16.1	4.0	2.1–7.7[m]	33	18.8	246	10.9	1.6	1.0–2.5[j]
Compulsion	9	18.4	265	10.5	1.7	0.7–4.0	20	11.4	120	5.3	2.2	1.3–3.8[k]
Hypochondriasis	9	18.4	212	8.4	3.2	1.5–6.9[k]	22	12.5	120	5.3	2.5	1.5–4.2[k]
Body dysmorphic disorder	20	40.8	385	15.2	4.0	2.1–7.8[m]	62	35.2	454	20.2	2.3	1.6–3.2[m]
Major depressive episode	7	14.3	166	6.6	3.0	1.3–7.0[j]	25	14.2	164	7.3	2.2	1.4–3.5[k]
Dysthymic disorder	10	20.4	206	8.1	3.2	1.5–7.0[k]	28	15.9	203	9.0	2.0	1.3–3.2[k]
Manic episode[c]	7	14.3	103	4.1	4.6	1.8–11.3[k]	6	3.4	65	2.9	1.4	0.6–3.4
Insomnia	13	26.5	363	14.3	2.6	1.3–5.3[j]	38	21.6	334	14.9	1.8	1.2–2.6[k]
Hypersomnia	20	40.8	543	21.4	2.9	1.5–5.4[k]	56	31.8	476	21.2	1.6	1.2–2.3[j]
Narcolepsy	5	10.2	121	4.8	2.1	0.7–6.2	13	7.4	89	4.0	1.9	1.0–3.6[j]
Oppositional defiant disorder	9	18.4	222	8.8	2.8	1.3–6.2[j]	32	18.2	208	9.3	2.3	1.5–3.5[k]
Conduct disorder	11	22.5	241	9.5	2.5	1.1–5.6[j]	10	5.7	66	2.9	1.9	0.9–4.0
Antisocial personality	5	10.2	80	3.2	3.7	1.3–10.7[j]	7	4.0	35	1.6	2.8	1.2–6.5[j]
Intermittent explosive disorder	9	18.4	210	8.3	2.9	1.3–6.5[j]	27	15.3	207	9.2	2.0	1.3–3.2[k]
Kleptomania	9	18.4	159	6.3	3.0	1.2–7.5[j]	10	5.7	75	3.3	1.6	0.7–3.4
Pathological gambling	3	6.1	25	1.0	13.3	3.5–50.2[k]	1	0.6	2	0.1	10.3	0.7–158.4
Trichotillomania[d]	6	12.2	48	1.9	4.6	1.3–15.8[j]	11	6.3	94	4.2	1.6	0.8–3.1
Dissociative disorder	8	16.3	151	6.0	3.8	1.7–8.6[k]	25	14.2	97	4.3	4.1	2.5–6.8[m]
Borderline personality	4	8.2	84	3.3	3.4	1.2–10.1[j]	11	6.3	59	2.6	2.8	1.4–5.6[k]
Schizoid personality	11	22.5	188	7.4	4.7	2.2–9.9[m]	19	10.8	83	3.7	3.6	2.1–6.3[m]
Lifetime suicidal ideation	23	46.9	584	23.0	2.5	1.3–4.8[j]	82	46.6	670	29.8	1.9	1.4–2.7[m]
Lifetime suicidal attempt	5	10.2	52	2.1	2.5	0.6–10.8	9	5.1	106	4.7	1.1	0.5–2.4
Symptoms of "spectrum disorders"												
Impulse control disorders[e]	15	30.6	226	8.9	3.4	1.6–7.4[k]	22	12.5	165	7.3	1.8	1.1–2.9[j]
Anxiety disorders (including OC symptoms)[f]	39	79.6	927	36.6	6.1	2.8–12.9[m]	88	50.0	735	32.7	1.9	1.4–2.7[m]
Anxiety disorders (excluding OC symptoms)[g]	29	59.2	676	26.7	4.5	2.3–8.7[m]	73	41.5	608	27.0	1.9	1.4–2.7[m]

Table 2 continued

Symptoms of	Male						Female					
	Gender dysphoria				Statistics		Gender dysphoria				Statistics	
	Yes ($N = 49$)		No ($N = 2536$)		OR	CI	Yes ($N = 176$)		No ($N = 2249$)		OR	CI
	N	%	*N*	%			*N*	%	*N*	%		
OCD[h]	27	55.1	557	22.0	4.3	2.2–8.1[m]	45	25.6	319	14.2	1.8	1.2–2.7[k]
Depressive disorders	10	20.4	235	9.3	2.9	1.3–6.2[j]	34	19.3	230	10.2	2.2	1.4–3.3[k]
Sleep disorders	27	55.1	852	33.6	2.8	1.5–5.4[k]	88	50.0	785	34.9	1.9	1.4–2.7[m]

OR odds ratio; *CI* confidence interval; *OC* obsessive–compulsive; *OCD* obsessive–compulsive disorder

Note 1: Interaction between biological sex and gender dysphoria for

[a] Symptoms of agoraphobia, $\beta = 0.51$, $\chi^2 = 5.42$, $p = .020$

[b] Symptoms of obsession, $\beta = 0.37$, $\chi^2 = 4.20$, $p = .040$

[c] Symptoms of manic episode, $\beta = 0.60$, $\chi^2 = 3.94$, $p = .047$

[d] Symptoms of trichotillomania, $\beta = 0.78$, $\chi^2 = 7.58$, $p = .006$

[e] Symptoms of impulse control disorders, $\beta = 0.46$, $\chi^2 = 5.26$, $p = .022$

[f] Symptoms of anxiety disorders (including OC symptoms), $\beta = 0.59$, $\chi^2 = 9.30$, $p = .002$

[g] Symptoms of anxiety disorders (excluding OC symptoms), $\beta = 0.37$, $\chi^2 = 4.82$, $p = .028$

[h] Symptoms of obsessive–compulsive disorder, $\beta = 0.37$, $\chi^2 = 4.64$, $p = .031$

Note 2: Symptoms listed in this table were those significantly associated with gender dysphoria (except for "lifetime suicidal attempt") for either or both biological sexes. Significance levels were given by superscripts: j ($p < .05$), k ($p < .005$), and m ($p < .0001$). All logistic regression analyses were controlled for participant age and available socioeconomic variables (parents' ages, jobs, education levels, marital status, and the number of siblings of the participant)

Table 3 Severity Sum-Score for dimensional symptoms from ASRI-4 by participants with and without symptoms of gender dysphoria

Severity Sum-Score of dimensional symptoms	Male							Female						
	Gender dysphoria				Statistics			Gender dysphoria				Statistics		
	Yes ($N=49$)		No ($N=2536$)		F	p	ES[a]	Yes ($N=176$)		No ($N=2249$)		F	p	ES[a]
	M	SD	M	SD				M	SD	M	SD			
Generalized anxiety disorder	10.39	3.84	7.83	3.66	18.72	<.001	0.68	9.47	3.70	8.10	3.56	21.58	<.001	0.38
Social phobia	3.51	1.90	2.32	1.79	16.12	<.001	0.64	2.62	1.91	2.21	1.66	8.24	.004	0.23
Somatization disorder	2.56	2.00	1.71	1.25	17.80	<.001	0.51	2.21	1.56	1.80	1.12	18.63	<.001	0.30
Post-traumatic stress disorder	0.82	0.38	0.56	0.33	24.62	<.001	0.73	0.72	0.36	0.56	0.31	37.48	<.001	0.48
Major depressive episode	9.60	4.58	6.71	3.89	20.59	<.001	0.68	8.66	3.91	6.92	3.70	32.70	<.001	0.46
Dysthymic disorder	9.56	3.99	6.83	3.66	20.81	<.001	0.71	8.42	3.42	6.84	3.39	32.07	<.001	0.46
Manic episode	7.59	5.46	5.06	4.00	15.00	<.001	0.53	5.37	3.78	4.24	3.63	14.34	<.001	0.30
Oppositional defiant disorder	8.87	4.37	6.20	3.52	22.09	<.001	0.67	7.36	3.64	6.14	3.42	18.03	<.001	0.35
Conduct disorder	3.36	3.17	2.10	2.06	14.53	<.001	0.47	1.99	1.53	1.30	1.39	34.89	<.001	0.47
Antisocial personality	4.28	2.87	3.11	2.09	11.56	<.001	0.47	3.21	1.86	2.41	1.72	31.37	<.001	0.45
Dissociative disorder	1.23	1.44	0.71	1.09	8.40	.004	0.41	0.97	1.14	0.58	0.96	22.49	<.001	0.37
Borderline personality	8.56	4.23	5.51	3.67	25.78	<.001	0.77	6.76	3.82	5.14	3.47	31.86	<.001	0.44
Schizoid personality	3.23	2.26	2.35	1.57	11.90	<.001	0.45	2.47	1.65	2.03	1.47	13.16	<.001	0.28

Note: Only symptom Severity Sum-Scores significantly associated with gender dysphoria were listed. In all general linear models, participant age, affiliated school, and available socioeconomic variables (parents' ages, jobs, education levels, marital status, and the number of siblings of the participant) were modeled as covariates

ES effect size (Cohen's *d*)

[a] Effect size: positive values indicate higher scores in the group with gender dysphoria, and vice versa

Table 4 Strongest co-occurring symptoms for gender dysphoria

	OR	95% CI	*p*
Combined sample			
Agoraphobia	2.4	1.4–4.3	.003
Hypochondriasis	2.8	1.6–4.9	<.001
Pathological gambling	14.5	3.9–53.7	<.001
Dissociative disorder	3.8	2.1–7.0	<.001
Female sex	5.2	3.1–8.9	<.001
For males only			
Agoraphobia	6.7	2.2–20.6	<.001
Hypochondriasis	4.9	1.8–13.0	.002
Manic episode	3.8	1.1–12.9	.031
Pathological gambling	13.4	2.7–67.2	.002
For females only			
Hypochondriasis	2.0	1.0–4.1	.046
Body dysmorphic disorder	1.7	1.0–2.8	.044
Dissociative disorder	3.7	1.8–7.5	<.001

OR odds ratio; *CI* confidence interval

overprotection from their mothers, and more authoritarian control from both parents with very small effect sizes (absolute Cohen's *d* ranging from 0.18 to 0.26, Table 5).

Discussion

As one of the first studies reporting the epidemiological and psychiatric features of GD in young adults from a large non-clinical sample, this study found that GD was not uncommon and co-occurred with certain psychopathologies among Taiwanese young adult university students.

Definition of Gender Dysphoria

Similar to studies using the CBCL item "wishes to be of opposite sex" to estimate the prevalence of GD in child and adolescent populations (van Beijsterveldt et al., 2006; Zucker et al., 1997), our study also adopted a single ASRI-4 item to assess the cognitive aspect of GD. The major limitation of defining cognitive GD by this single item was that information about other dimensions (behavioral and affective) of GD, sexual orientation and identity, and gender role characteristics were absent; thus, it was difficult to make a definite differentiation of our "cognitive GD" from other possibly related or overlapping conditions (e.g., gender-role dissatisfaction, non-heterosexuality). We recognized that young adults in this study were possibly exposed to social stereotypes of the dichotomized gender roles in traditional cultures. In order to reduce false-positives, we tried to exclude those with mainly gender-role dissatisfaction by defining GD as rated "often" or "very often" for the index item (excluding "sometimes," which was defined in the original ASRI-4 scoring as a positive response), expecting to identify those primarily with a persistent and high level of cognitive cross-gender identification.

Prevalence and Demographics

We found a high prevalence (males, 1.9%; females, 7.3%) of GD in Taiwanese university students. The present rates were higher than those of Western studies in children (van Beijsterveldt et al., 2006; Zucker et al., 1997) and, not surprisingly, remarkably higher than the prevalence rates of adult transsexuals in Western (Bakker et al., 1993; Bodlund et al., 1993; De Cuypere et al., 2007; Weitze & Osburg, 1996; Wilson et al., 1999) and Asian (Chang et al., 2006; Hwu et al., 1989; Tsoi, 1988) studies. Yet, it is closer to that of a study in American twin children and adolescents (Coolidge et al., 2002). The different prevalence rates across studies may be explained by the various instruments and defining/diagnostic criteria applied, the different ethnic/cultural groups investigated, and the problems of comparability of childhood to adulthood GD/GID.

Regarding measurements, although the instrument assessing GD was contextually parallel to that of the CBCL (van Beijsterveldt et al., 2006; Zucker et al., 1997), the meaning of "wishes to be of opposite sex" might be different for parental reports of the children and the adult self-reports. Moreover, our assessment was less comprehensive than the use of a six-item GID scale (Coolidge et al., 2002), and somewhat different from the two cognitive GD items of the DIS-CM (Chang et al., 2006; Hwu et al., 1989). Regarding sample group differences, the present high rates echoed the gross comparison that the prevalence for adult transsexuals in Asian countries (Chang et al., 2006; Hwu et al., 1989; Tsoi, 1988) was higher than that in Western countries (Bakker et al., 1993; Bodlund et al., 1993; De Cuypere et al., 2007; Weitze & Osburg, 1996; Wilson et al., 1999). The sampling bias also limits the generalizability of our results. Although recruiting from a national university that enrolls freshmen from all over Taiwan, our subjects had relatively higher-than-average academic achievements. Other participant characteristics, such as teenage developmental experiences, accessibility to information on the internet, personality characteristics, and coping styles, may also influence prevalence rates. Further investigation on a randomly selected, nationally representative sample is needed to generate more accurate estimates for the general population. Regarding comparability of childhood to adulthood GD/GID, it was recognized that childhood GID/GD does not always persist into adulthood (Drummond, Bradley, Peterson-Badali, & Zucker, 2008; Wallien & Cohen-Kettenis, 2008), and adult GID/GD might not necessarily be preceded by childhood GID/GD. This makes cross-age comparison difficult with the available data.

Table 5 Perceived family support and parenting style by participants with and without symptoms of gender dysphoria

	Male							Female						
	Gender dysphoria				Statistics			Gender dysphoria				Statistics		
	Yes ($N=49$)		No ($N=2536$)		F	p	ES[a]	Yes ($N=176$)		No ($N=2249$)		F	p	ES[a]
	M	*SD*	*M*	*SD*				*M*	*SD*	*M*	*SD*			
Family support: APGAR score[b]	1.86	0.62	1.63	0.56	7.04	.008	0.39	1.79	0.61	1.53	0.55	30.99	<.001	0.45
Parenting style: PBI scores														
Affection and care from														
Mother	23.74	7.63	26.48	5.88	9.19	.003	−0.40	25.02	6.49	27.64	6.27	25.57	<.001	−0.41
Father	19.59	8.03	23.23	6.90	12.60	<.001	−0.49	22.25	7.28	24.79	6.97	19.99	<.001	−0.36
Overprotection from														
Mother	7.00	3.72	6.91	4.09	0.02	ns		7.39	4.34	6.60	4.04	6.11	.014	0.19
Father	4.21	3.43	4.68	3.93	0.42	ns		5.42	3.97	4.99	3.79	1.91	ns	
Authoritarian control from														
Mother	5.41	4.25	5.02	3.71	0.30	ns		5.49	4.08	4.80	3.78	4.72	.030	0.18
Father	5.72	4.75	4.64	3.87	3.05	ns		5.49	4.09	4.46	3.78	10.74	.001	0.26

Note: In all general linear models, participant age and available socioeconomic variables (parents' ages, jobs, education levels, marital status, and the number of siblings of the participant) were modeled as covariates

ES effect size (Cohen's *d*); *APGAR* adaptation, partnership, growth, affection, and resolve; *PBI* Parental Bonding Instrument

[a] Effect size: positive value indicates higher scores in the group with gender dysphoria, and vice versa

[b] Higher score indicates lower perceived family support

Several past studies (Chang et al., 2006; Coolidge et al., 2002; Godlewski, 1988; Hwu et al., 1989; van Beijsterveldt et al., 2006; Zucker et al., 1997) and ours all demonstrated a trend of female-predominance for GD in both childhood and adulthood. These consistent findings are contrary to the male-predominance of adult transsexuals receiving medical services (Bakker et al., 1993; Bodlund et al., 1993; De Cuypere et al., 2007; Weitze & Osburg, 1996; Wilson et al., 1999), which may partly be explained by the clinical referral bias for males (Zucker et al., 1997). The female-predominance in GD may actually echo the fluidity and diversity of female sexuality depicted by past research (Diamond, 2008; Savin-Williams & Diamond, 2000; Sophie, 1985), but it also possibly reflects more or less the gender-role dissatisfaction caused by the long-standing oppression from a patriarchic structure both in the Western and the Taiwanese/Chinese societies (Greer, 1971; Huang & You, 2007). From our own data, we were not able to make clear differentiation between strong gender-role dissatisfaction and a milder degree of GD. Nevertheless, it is reasonable to assume that the female-predominance of GD, compared to the male-predominance of clinical GID, reflects not only the natural aspect of female sexuality but also the nurtural part, including their appraisals towards their gender identity and gender role characteristics. The role of nature and nurture on the interface of gender role and gender identity is thus an interesting area for further investigation. Our finding of a markedly stronger female-predominance than those from Western studies (Coolidge et al., 2002; van Beijsterveldt et al., 2006; Zucker et al., 1997) implies the need for further biological and sociocultural explorations. The impact of Taiwanese/Chinese traditional cultural gender role socialization on the participants' reports of GD should also be considered (e.g., males being less willing to report gender-nonconformity and females being more likely to be dissatisfied with their gender roles) (Huang & You, 2007).

Co-Occurring Psychopathology

Corresponding to previous reports regarding the diverse nature of associated psychopathology both in children (Zucker & Bradley, 1995) and adults (Hepp et al., 2005) with GID, we also found increased rates of a wide range of concurrent psychopathologies in individuals with cognitive GD than those without. Similar to previous reports on children with GID (Cohen-Kettenis et al., 2003; Wallien et al., 2007; Zucker & Bradley, 1995), university students with GD were more likely than those without to have various anxious and depressive symptoms. Furthermore, our findings lend evidence to support that young adults with GD are more likely to have symptoms of somatoform disorders (Hepp et al., 2005); specifically, symptoms of hypochondriasis were most significantly associated in both sexes and body dysmorphic disorder in females, suggesting the possibility that the presence of high bodily concerns or pathological bodily worries may be related to the dissatisfaction toward their physical/sexual appearances in young adulthood. Similar to the Dutch twin study (van Beijsterveldt et al., 2006), our results also showed evidence of externalizing symptoms in addition to internalizing ones, especially those related to impulse-dyscontrol, in university young adults (particularly males) with GD. These findings are novel and have clinical implications and thus warrant further investigation.

Besides the aforementioned co-occurring psychopathology documented in prior research, our study also indicated certain novel ones, including symptoms of obsession–compulsion, agoraphobia, dissociation, dyssomnia, and borderline and schizoid personalities. The association of obsession–compulsion and agoraphobia to GD is unique, which indicates the presence of excessive ruminations and a distinct expression of anxiety related to GD; the association was stronger in males, suggesting a sex-specific concurrent anxiety pattern in males with GD. The association between GD and dissociative symptoms is significant and particularly strong in females, reminding us to be mindful of the possible psychoplastic evolvements related to experiences of trauma or maltreatment, as evidenced by past research showing higher dissociative symptoms associated with emotional maltreatment in transsexuals (Kersting et al., 2003). The increased co-occurrence with dyssomnia may be explained by concurrent anxiety, depression, and other psychiatric symptoms. Finally, despite that confusion with gender identity as one contributing symptom for borderline personality disorder (American Psychiatric Association, 2000), our results do not support borderline personality as a prerequisite for GD (for the low prevalence of borderline personality in subjects with GD). On the other hand, the co-occurrence of GD with symptoms of schizoid personality might be partly explained by their tendency of social isolation. Furthermore, the significant risk for suicidal ideation points to the seriousness of the condition and the need for careful clinical monitoring to prevent the potentially lethal consequences of suicidal behaviors.

The co-occurring psychopathologies were more apparent in males than in females for symptoms of agoraphobia, obsession–compulsion, manic episode, trichotillomania, impulse-control disorders, and anxiety disorders (Table 2). Moreover, the effect sizes of co-occurring symptom severities in participants with compared to without GD were generally higher in males than in females (Table 3). These findings could be relevant to the eventual understanding of both the biologic substrates and the sociocultural contributors to GD for different sexes in Taiwanese society.

It should be clear that although our initial univariate analyses identified a wide range of increased odds (even under a stringent critical value adjusted for multiple comparisons) for other kinds of psychiatric symptoms in subjects with GD, these should be considered exploratory. In addition, the findings

could also reflect that certain psychiatric symptoms (e.g., anxiety symptoms with other psychopathology) were highly comorbid with each other. It is thus plausible to assume that the multivariate analyses (Table 4) depict a more specific pattern of co-occurring psychopathology for GD. Furthermore, due to the limitation that GD defined in our study may more or less overlap with strong gender-role dissatisfaction, we should be mindful of the possibility that the associated psychopathologies partly reflect those related to gender-role dissatisfaction, as exemplified by past literatures showing, for instance, the association of undifferentiated sex-role orientation to depressive symptoms in working women (Napholz, 1994) and men with infertile female partners (Cook, 1993). Finally, since "concurrent psychopathology" revealed by questionnaire survey should be treated conservatively before further validation from direct interview, the readers should bear in mind that some of our findings (e.g., somatic concerns and anxiety in publics) may partially reflect normal responses for subjects with GD experiencing the pressure and disapproval from their surrounding environments.

Perceived Parenting Style and Family Support

Although there have been no studies examining the perceived parenting style and family support among individuals with GD, similar to the findings from adult transsexuals in Western society (Cohen-Kettenis & Arrindell, 1990; Parker & Barr, 1982), our results indicate that university young adult students of both sexes with GD perceived less support from their families and obtained less affection and care from their parents; specific for females with GD, they also perceived more overprotection from mothers and authoritarian control from both parents (Table 5). Our findings of higher variation in reporting parental affection and care, particularly in male students, may reflect the diverse responses of parental affection and care perceived by the male students.

It is interesting that since the "proximity to mother" is highly predictive of the presence of effeminacy in clinically-referred boys (Sreenivasan, 1985), and the trait of separation anxiety being considered important for childhood GID (Coates, 1990; Coates & Person, 1985; Sreenivasan, 1985; Zucker & Bradley, 1995; Zucker et al., 1996), the perception of parent–child relationship and family support of university students with GD in their young adulthood seems not merely an extension of the childhood ones (e.g., higher clinging to mother, marked separation anxiety). From clinical observations, it is not uncommon that the parental perceptions of cross-gender behaviors switch from "cute" at preschool age to averse later, thus explaining partially the non-continuity of parenting styles. It is of high interest to explore the temporal changes of the quality of and the associated factors influencing the family relationship along the developmental course of individuals with GD.

Methodological Considerations

The present study was limited by the cross-sectional design, the questionable generalizability of results from a university sample to other populations, and the sole employment of self-administered questionnaires instead of direct interviews. Moreover, though defining GD in terms of a single rating scale item is a commonly adopted research strategy, particularly for studies applying secondary data analyses like ours, the single-item approach constrains the understanding of the various aspects of sexuality and assesses merely the cognitive, instead of affective or behavioral, dimension of GD; information needed for clear differentiation of cognitive GD to non-heterosexuality or strong gender-role dissatisfaction is also lacking. Nevertheless, a recent study of the Gender Identity/Gender Dysphoria Questionnaire for Adolescents and Adults (GIDYQ-AA) showed that an item (No. 16, "wish or desire to be opposite sex") with similar wording as the ASRI-4 GD item demonstrated a very high factor loading (0.93) and corrected item–total correlation (0.91) with the GIDYQ-AA as a whole (Deogracias et al., 2007), implying that this single item is core to and corresponds well to the whole construct of GD.

The limitations were offset by several strengths. First, this was the first large-scale non-clinical study exploring aspect of GD in young adults. Secondly, the comprehensive evaluation of psychopathology and parent–child interaction provides a valuable opportunity illustrating the full psychiatric picture of GD. Finally, this is the only study investigating GD in Taiwanese and Asian population in the past two decades and may be regarded as the basis for cross-cultural comparisons.

Clinical Implications

The relatively high prevalence of cognitive GD revealed in the present study warrants clinicians to direct more attention to the presence and the commonality of the "variant gender identity" (Lothstein, 2006), and to be mindful of the potential sufferings and specific needs of university students with GD. The dysphoria in young adult university students with variant gender identity is possibly related to a wide range but specific pattern of concurrent psychiatric conditions, for instance symptoms of agoraphobia and hypochondriasis for males and dissociative symptoms for females. Perceived family supports and relationships are often less satisfying. We also need to be aware of the potential biological gender difference in the presentations of and associated psychiatric and family conditions with GD. School counselors and health care professionals should provide necessary supports and services to address these issues according to individual needs.

Acknowledgements This study was supported by grants from National Taiwan University Hospital (NTUH-92-S07) and from the Office of Student Affairs, National Taiwan University. The preparation

of this manuscript was supported by National Health Research Institute (NHRI-EX96-9407PC), Taiwan. The authors would like to thank the reviewers and the Editor for their insightful suggestions on the revision of the manuscript. Dr. Gadow is a shareholder in Checkmate Plus, publisher of the ASRI-4.

Appendix: Symptom Clusters Examined by the ASRI-4

Anxiety disorders: generalized anxiety disorder, specific phobia, social phobia, agoraphobia, and panic disorder
Obsessive–compulsive disorder
Tic disorders: motor and vocal tics
Somatoform disorders: somatization disorder, hypochondriasis, and body dysmorphic disorder
Stress-response disorders: post-traumatic stress disorder and adjustment disorder
Mood disorders: major depressive episode, dysthymic disorder, and manic episode
Sleep disorders: insomnia, hypersomnia, narcolepsy, and nightmare disorder
Eating disorders: anorexia nervosa and bulimia nervosa
Disruptive behavior disorders: attention deficit/hyperactivity disorder, oppositional defiant disorder, and conduct disorder
Impulse-control disorders: intermittent explosive disorder, kleptomania, pathological gambling, pyromania, and trichotillomania
Dissociative disorder
Gender identity disorder
Substance use
Schizophrenia
Personality disorders: schizoid personality disorder, antisocial personality disorder, and borderline personality disorder.

References

a Campo, J., Nijman, H., Merckelbach, H., & Evers, C. (2003). Psychiatric comorbidity of gender identity disorders: A survey among Dutch psychiatrists. *American Journal of Psychiatry, 160*, 1332–1336.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

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

Bakker, A., van Kesteren, P. J., Gooren, L. J., & Bezemer, P. D. (1993). The prevalence of transsexualism in The Netherlands. *Acta Psychiatrica Scandinavica, 87*, 237–238.

Bodlund, O., Kullgren, G., Sundbom, E., & Hojerback, T. (1993). Personality traits and disorders among transsexuals. *Acta Psychiatrica Scandinavica, 88*, 322–327.

Brown, G. R. (1990). A review of clinical approaches to gender dysphoria. *Journal of Clinical Psychiatry, 51*, 57–64.

Chang, S. S., Hwu, H. G., & Yen, Y. K. (2006). The community epidemiology of psychosexual dysfunctions and transsexualism. *Medicine Today Monthly, 33*, 993–1005.

Chien, Y. L., Gau, S. S., & Gadow, K. D. (2009). *Gender difference in psychiatric syndromes and psychiatric concurrent patterns in Taiwanese college students*. Manuscript submitted for publication.

Coates, S. (1990). Ontogenesis of boyhood gender identity disorder. *Journal of the American Academy of Psychoanalysis, 18*, 414–438.

Coates, S., & Person, E. S. (1985). Extreme boyhood femininity: Isolated behavior or pervasive disorder? *Journal of the American Academy of Child Psychiatry, 24*, 702–709.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen-Kettenis, P. T., & Arrindell, W. A. (1990). Perceived parental rearing style, parental divorce and transsexualism: A controlled study. *Psychological Medicine, 20*, 613–620.

Cohen-Kettenis, P. T., & Gooren, L. J. (1999). Transsexualism: A review of etiology, diagnosis and treatment. *Journal of Psychosomatic Research, 46*, 315–333.

Cohen-Kettenis, P. T., Owen, A., Kaijser, V. G., Bradley, S. J., & Zucker, K. J. (2003). Demographic characteristics, social competence, and behavior problems in children with gender identity disorder: A cross-national, cross-clinic comparative analysis. *Journal of Abnormal Child Psychology, 31*, 41–53.

Cole, C. M., O'Boyle, M., Emory, L. E., & Meyer, W. J. (1997). Comorbidity of gender dysphoria and other major psychiatric diagnoses. *Archives of Sexual Behavior, 26*, 13–26.

Cook, R. (1993). The relationship between sex role and emotional functioning in patients undergoing assisted conception. *Journal of Psychosomatic Obstetrics and Gynaecology, 14*, 31–40.

Coolidge, F. L., Thede, L. L., & Young, S. E. (2002). The heritability of gender identity disorder in a child and adolescent twin sample. *Behavior Genetics, 32*, 251–257.

Cox, B. J., Enns, M. W., & Clara, I. P. (2000). The Parental Bonding Instrument: Confirmatory evidence for a three-factor model in a psychiatric clinical sample and in the National Comorbidity Survey. *Social Psychiatry and Psychiatric Epidemiology, 35*, 353–357.

De Cuypere, G., Van Hemelrijck, M., Michel, A., Carael, B., Heylens, G., Rubens, R., et al. (2007). Prevalence and demography of transsexualism in Belgium. *European Psychiatry, 22*, 137–141.

Deogracias, J. J., Johnson, L. L., Meyer-Bahlburg, H. F., Kessler, S. J., Schober, J. M., & Zucker, K. J. (2007). The Gender Identity/Gender Dysphoria Questionnaire for Adolescents and Adults. *Journal of Sex Research, 44*, 370–379.

Diamond, L. M. (2008). Female bisexuality from adolescence to adulthood: Results from a 10-year longitudinal study. *Developmental Psychology, 44*, 5–14.

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

Gadow, K. D., Sprafkin, J., & Weiss, M. (2004). *Adult Self-Report Inventory-4 manual*. Stony Brook, NY: Checkmate Plus.

Gau, S. S., Shen, H. Y., Chou, M. C., Tang, C. S., Chiu, Y. N., & Gau, C. S. (2006). Determinants of adherence to methylphenidate and the impact of poor adherence on maternal and family measures. *Journal of Child and Adolescent Psychopharmacology, 16*, 286–297.

Godlewski, J. (1988). Transsexualism and anatomic sex ratio reversal in Poland. *Archives of Sexual Behavior, 17*, 547–548.

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

Greer, G. (1971). *The female eunuch*. New York: Bantam Books.

Haraldsen, I. R., & Dahl, A. A. (2000). Symptom profiles of gender dysphoric patients of transsexual type compared to patients with personality disorders and healthy adults. *Acta Psychiatrica Scandinavica, 102*, 276–281.

Heider, D., Matschinger, H., Bernert, S., Vilagut, G., Martinez-Alonso, M., Dietrich, S., et al. (2005). Empirical evidence for an invariant three-factor structure of the Parental Bonding Instrument in six European countries. *Psychiatry Research, 135*, 237–247.

Hepp, U., Kraemer, B., Schnyder, U., Miller, N., & Delsignore, A. (2005). Psychiatric comorbidity in gender identity disorder. *Journal of Psychosomatic Research, 58*, 259–261.

Hoenig, J., & Kenna, J. C. (1974). The prevalence of transsexualism in England and Wales. *British Journal of Psychiatry, 124*, 181–190.

Huang, S. L., & You, M. H. (Eds.). (2007). *Gender dimensions in Taiwanese society*. Taipei: Chuliu Publisher.

Hwu, H. G., Yeh, E. K., & Chang, L. Y. (1989). Prevalence of psychiatric disorders in Taiwan defined by the Chinese Diagnostic Interview Schedule. *Acta Psychiatrica Scandinavica, 79*, 136–147.

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

Kersting, A., Reutemann, M., Gast, U., Ohrmann, P., Suslow, T., Michael, N., et al. (2003). Dissociative disorders and traumatic childhood experiences in transsexuals. *Journal of Nervous and Mental Disease, 191*, 182–189.

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

Levine, S. B., & Lothstein, L. (1981). Transsexualism or the gender dysphoria syndromes. *Journal of Sex and Marital Therapy, 7*, 85–113.

Liao, S. C., Lee, M. B., Lee, Y. J., Weng, T., Shih, F. Y., & Ma, M. H. (2002). Association of psychological distress with psychological factors in rescue workers within two months after a major earthquake. *Journal of the Formosan Medical Association, 101*, 169–176.

Lizardi, H., & Klein, D. N. (2002). Evidence of increased sensitivity using a three-factor version of the Parental Bonding Instrument. *Journal of Nervous and Mental Disease, 190*, 619–623.

Lothstein, L. M. (2006). The scientific foundations of gender identity disorder. In T. G. Plante (Ed.), *Mental disorders of the new millennium* (Vol. 2, pp. 227–250). Westport, CT: Praeger Publishers.

Meyer, J. K. (1982). The theory of gender identity disorders. *Journal of the American Psychoanalytic Association, 30*, 381–418.

Meyer, W. J. (2004). Comorbidity of gender identity disorders [Letter]. *American Journal of Psychiatry, 161*, 934–935.

Napholz, L. (1994). Indices of psychological well-being and sex role orientation among working women. *Health Care for Women International, 15*, 307–316.

Parker, G., & Barr, R. (1982). Parental representations of transsexuals. *Archives of Sexual Behavior, 11*, 221–230.

Pauly, I. B. (1968). The current status of the change of sex operation. *Journal of Nervous and Mental Disease, 147*, 460–471.

Ross, M. W., Walinder, J., Lundstrom, B., & Thuwe, I. (1981). Cross-cultural approaches to transsexualism: A comparison between Sweden and Australia. *Acta Psychiatrica Scandinavica, 63*, 75–82.

Sato, T., Narita, T., Hirano, S., Kusunoki, K., Sakado, K., & Uehara, T. (1999). Confirmatory factor analysis of the Parental Bonding Instrument in a Japanese population. *Psychological Medicine, 29*, 127–133.

Savin-Williams, R. C., & Diamond, L. M. (2000). Sexual identity trajectories among sexual-minority youths: Gender comparisons. *Archives of Sexual Behavior, 29*, 607–627.

Smilkstein, G. (1978). The family APGAR: A proposal for a family function test and its use by physicians. *Journal of Family Practice, 6*, 1231–1239.

Sophie, J. (1985). A critical examination of stage theories of lesbian identity development. *Journal of Homosexuality, 12*, 39–51.

Sprafkin, J., Gadow, K. D., Weiss, M. D., Schneider, J., & Nolan, E. E. (2007). Psychiatric comorbidity in ADHD symptom subtypes in clinic and community adults. *Journal of Attention Disorders, 11*, 114–124.

Sreenivasan, U. (1985). Effeminate boys in a child psychiatric clinic: Prevalence and associated factors. *Journal of the American Academy of Child Psychiatry, 24*, 689–694.

Tsoi, W. F. (1988). The prevalence of transsexualism in Singapore. *Acta Psychiatrica Scandinavica, 78*, 501–504.

van Beijsterveldt, C. E., Hudziak, J. J., & Boomsma, D. I. (2006). Genetic and environmental influences on cross-gender behavior and relation to behavior problems: A study of Dutch twins at ages 7 and 10 years. *Archives of Sexual Behavior, 35*, 647–658.

Walinder, J. (1971). Incidence and sex ratio of transsexualism in Sweden. *British Journal of Psychiatry, 119*, 195–196.

Wallien, M. S., & Cohen-Kettenis, P. T. (2008). Psychosexual outcome of gender-dysphoric children. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 1413–1423.

Wallien, M. S., Swaab, H., & Cohen-Kettenis, P. T. (2007). Psychiatric comorbidity among children with gender identity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*, 1307–1314.

Weitze, C., & Osburg, S. (1996). Transsexualism in Germany: Empirical data on epidemiology and application of the German Transsexuals' Act during its first ten years. *Archives of Sexual Behavior, 25*, 409–425.

Wilson, P., Sharp, C., & Carr, S. (1999). The prevalence of gender dysphoria in Scotland: A primary care study. *British Journal of General Practice, 49*, 991–992.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: Author.

Zucker, K. J. (2007). Gender identity disorder. In A. Martin & F. R. Volkmar (Eds.), *Lewis' child and adolescent psychiatry: A comprehensive textbook* (4th ed., pp. 669–680). Philadelphia: Lippincott Williams & Wilkins.

Zucker, K. J. (2008). On the "natural history" of gender identity disorder in children [Editorial]. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 1361–1363.

Zucker, K. J., & Bradley, S. J. (1995). *Gender identity disorder and psychosexual problems in children and adolescents*. New York: Guilford Press.

Zucker, K. J., Bradley, S. J., & Lowry Sullivan, C. B. (1996). Traits of separation anxiety in boys with gender identity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 791–798.

Zucker, K. J., Bradley, S. J., & Sanikhani, M. (1997). Sex differences in referral rates of children with gender identity disorder: Some hypotheses. *Journal of Abnormal Child Psychology, 25*, 217–227.

Arch Sex Behav (2010) 39:1429–1437
DOI 10.1007/s10508-009-9551-1

ORIGINAL PAPER

A Five-Year Follow-Up Study of Swedish Adults with Gender Identity Disorder

Annika Johansson · Elisabet Sundbom · Torvald Höjerback · Owe Bodlund

Received: 30 April 2008 / Revised: 27 August 2009 / Accepted: 27 August 2009 / Published online: 9 October 2009

Abstract This follow-up study evaluated the outcome of sex reassignment as viewed by both clinicians and patients, with an additional focus on the outcome based on sex and subgroups. Of a total of 60 patients approved for sex reassignment, 42 (25 male-to-female [MF] and 17 female-to-male [FM]) transsexuals completed a follow-up assessment after 5 or more years in the process or 2 or more years after completed sex reassignment surgery. Twenty-six (62%) patients had an early onset and 16 (38%) patients had a late onset; 29 (69%) patients had a homosexual sexual orientation and 13 (31%) patients had a non-homosexual sexual orientation (relative to biological sex). At index and follow-up, a semi-structured interview was conducted. At follow-up, 32 patients had completed sex reassignment surgery, five were still in process, and five—following their own decision—had abstained from genital surgery. No one regretted their reassignment. The clinicians rated the global outcome as favorable in 62% of the cases, compared to 95% according to the patients themselves, with no differences between the subgroups. Based on the follow-up interview, more than 90% were stable or improved as regards work situation, partner relations, and sex life, but 5–15% were dissatisfied with the hormonal treatment, results of surgery, total sex reassignment procedure, or their present general health. Most outcome measures were rated positive and substantially equal for MF and FM. Late-onset transsexuals differed from those with early onset in some respects: these were mainly MF (88 vs. 42%), older when applying for sex reassignment (42 vs. 28 years), and non-homosexually oriented (56 vs. 15%). In conclusion, almost all patients were satisfied with the sex reassignment; 86% were assessed by clinicians at follow-up as stable or improved in global functioning.

Keywords Transsexualism · Gender identity disorder · Sex reassignment · Outcome

A. Johansson (✉) · T. Höjerback
Department of Psychiatry, Sexology Clinic,
St. Lars-omr, MC-huset, 221 85 Lund, Sweden
e-mail: annika.johansson@skane.se

A. Johansson · E. Sundbom · O. Bodlund
Department of Clinical Sciences, Division of Psychiatry,
University of Umeå, Umeå, Sweden

Introduction

Diagnosis

A transsexual person has a cross-gender identity in relation to his or her biological sex. Gender Identity Disorder (GID) is the diagnostic classification according to the most recent edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR) (American Psychiatric Association, 2000), describing persons with cross-gender identity, gender dysphoria, and concomitant persistent psychological distress or functional impairment. In Sweden, the ICD-10 (World Health Organization, 1992) diagnosis of "transsexualism" (TS) is used in clinical settings.

Early and late onset transsexualism are clinical classifications, constructed to divide the spectra of transsexualism into subgroups. Early-onset TS refers to people who early in life (childhood) have a strong wish to become the opposite sex, while late-onset TS refers to individuals whose cross-gender identification begins at puberty or later and who gradually develop a wish for sex reassignment (SR). The age of 12 has been suggested as the division point between early and late onset (Doorn, Poortinga, & Verschoor, 1994). Among male-to-female (MF) transsexuals with late onset, some have passed through a period of transvestic fetishism (Docter, 1988). From another theoretical framework, TSs can be categorized into subgroups on the basis of sexual orientation, i.e., homosexual

or non-homosexual in reference to biological sex (Blanchard, 1989).

Incidence

Since 1972, when Sweden passed legislation regulating surgical and legal sex reassignment, 10–15 people have completed the reassignment procedure annually, which means an incidence of 0.19 per 100,000 inhabitants over 15 years of age (Olsson & Möller, 2003). However, there has been a marked increase in the last 8 years. In 2006 and 2007, about 50 people applied and were accepted for SRS. This corresponds to a yearly incidence of approximately 1/100,000 in the age bracket of 18–65 (personal communication, Swedish National Board of Health and Welfare). This increase over the last few years concerns mainly younger MF.

Outcome of Sex Reassignment Surgery

In a review of studies from 1961 to 1991, covering 70 studies and 2,000 patients (Pfäfflin & Junge, 1998), the conclusion was that, despite differences in methodology and the choice of outcome variables, SRS basically resolved gender dysphoria problems. Early reports showed positive outcome for about 71% of the MF TSs and 89.5% for the female-to-male (FM) TSs and, in more recent studies, even more favorable outcome rates of 87 and 97%, respectively (Cohen-Kettenis & Pfäfflin, 2003). Disparities in study methodology lead to different conclusions concerning outcome and predictive factors, but most agree that female biological sex is a predictor for positive outcome. In several studies, early-onset TS, and consequently cross-dressing in childhood and younger age at SRS, are also considered predictors for positive outcome (Cohen-Kettenis & Pfäfflin, 2003; Lundström, Pauly, & Wålinder, 1984). Furthermore, homosexual orientation is a positive predictor for outcome (Blanchard, Steiner, Clemmensen, & Dickey, 1989).

Negative outcome includes the very few actual regrets (estimated at 1–2%) (Cohen-Kettenis & Pfäfflin, 2003; Kuiper & Cohen-Kettenis, 1998; Pfäfflin & Junge, 1998). In a previous Swedish study, 3.8% during the years 1972–1992 regretted their SRS. Factors associated with regrets showed that lack of support from the patient's family and belonging to the group of people with late-onset TS were predictive of regret of SRS (Landén, Wålinder, Hambert, & Lundström, 1998). Other risk factors for negative outcome often mentioned in studies are poor social support, severe psychopathology, unfavorable physical appearance, and poor surgical result (Cohen-Kettenis & Pfäfflin, 2003; Lawrence, 2003; Smith, van Goozen, Kuiper, & Cohen-Kettenis, 2005).

Lawrence (2003) concluded that results of surgery may be more important for global outcome than preoperative factors, such as transsexual typology or procedure compliance. Eldh, Berg, and Gustafsson (1997) also reported that the participants' dissatisfaction was associated with unsatisfactory physical and functional results after surgery.

These predictive factors for favorable vs. unfavorable outcome were pointed out by Swedish researchers in the 1970s and 1980s (Lundström et al., 1984; Wålinder, Lundström, & Thuwe, 1978; Wålinder & Thuwe, 1975), and later by Landén et al. (1998). In another Swedish study, focusing on personality factors, Bodlund and Kullgren (1996) described a group of 19 patients who had been approved for SRS. At follow-up after 5 years, 13 (70%) patients had improved in relation to social, psychological, and psychiatric aspects. One person regretted the SR and three (16%) had an unsatisfactory outcome in the sense that their psychosocial functioning had not improved or in some aspects had worsened. Prognostic factors associated with positive outcome were an absence of personality disorder diagnoses, having a positive self-image, and having a partner early in the process.

In a recent review, Gijs and Brewaeys (2007) report a favorable outcome in 96% of cases according to the patients themselves, and they argue that the patients' opinion regarding satisfaction may, in fact, be the decisive measure for outcome.

Aims of the Study

This study investigated outcome in terms of clinicians' and patients' evaluation of the process of sex reassignment. The study was prospective and longitudinal and focused on relief of gender dysphoria, satisfaction with the SR process, social functioning, work, relationships, and sexuality after a minimum of 5 years in the process and/or 2 years after SRS. The study also examined if outcome differed in regard to sex, age group, and diagnostic group (i.e., early vs. late-onset and homosexual vs. non-homosexual TSs).

The Swedish Procedure

According to Swedish national law, a person with transsexualism can apply for and get the necessary treatment for sex change within the public health care system. The procedure is often initiated by referral from a local psychiatric clinic to a specialized psychiatrist and gender team in order to obtain psychiatric assessment for a certificate of approval, following the patient's personal application for sex reassignment. There are six teams and two surgery clinics in Sweden for this purpose. The gender team provides a diagnostic evaluation, evaluates personal resources and the social situation, offers personal support, and authorizes the various examinations and treatments needed, such as physical exams, hair removal, speech training, hormone treatment, and surgery. After at least 1 year of so-called real-life experience, the patient's situation is re-evaluated and, if deemed suitable, hormone treatment is initiated, usually by the psychiatrist. At a minimum of 2 years, sex reassignment can be completed (Landén, Bodlund, Ekselius, Hambert, & Lundström,

2001). Most medical costs are subsidized by the Swedish government, so the total cost for the patient will not be more than about 300 USD per year.

The professional view in Sweden is that transsexualism is an authentic identity and that sex reassignment is the treatment of choice. Full treatment, in addition to hormones, involves sterilization and genital and plastic surgery. A complete sex change (according to the law) means that the patient has changed his or her first name, is on hormonal treatment, and has received a new national identity number indicating the new gender affiliation. The patient must also have been sterilized (or castrated), but not necessarily have had genital plastic surgery.

Method

Patients

The patients in this study were consecutively selected from two geographic regions of Sweden, one in the north and one in the south. At the time of inclusion, both centers had about five new referrals per year that were approved for SRS. About the same number of patients was excluded yearly from the SR process, due, for instance, to other diagnoses. The inclusion criteria for the follow-up study were fulfillment of the diagnostic criteria for transsexualism (or GID) and having been approved for sex reassignment since 5 years back or more, and/or completed sex reassignment since 2 years or more. Of the total population of 60 former patients (39 MF and 21 FM), 42 (70%) agreed to participate. Eighteen patients, 14 MF and 4 FM, were not included: one had died from complications of the SR surgery, eight were not reachable or did not respond to the request to participate, and nine said they were unwilling, mainly for integrity reasons. The dropouts were predominately MF (77.8%) and had a slightly lower score on the GAF-scale (Global Assessment of Functioning, Axis V in the DSM-IV) at index (63 vs. 71, $p = .001$), but did not differ from the patients as regards age, diagnostic subtype (early or late onset) or sexual orientation. Sixteen (88.9%) of the dropouts had completed SRS.

The follow-up group consisted of 25 MFs (59.5%) and 17 FMs (40.5%). Twenty-six (61.9%) patients were diagnosed by the clinician (based on the patients' own reports) as early onset and 16 (38.1%) as late onset.

As shown in Table 1, 32 (76.2%) patients had completed SRS, five were still in the process, and five (4 MF/1 FM) had discontinued further surgery. All were late-onset TSs and four of these last-mentioned five were still on hormones. Among the four MF who had discontinued surgery, two (see Table 2, No. 10 and 15) wanted only breast enlargement but not genital surgery, one (No. 12) was ambivalent, and the fourth (No. 14) had changed his mind about completing SRS since he had found a partner who accepted that he had double identities. The FM (No. 35) described herself as transgendered living with double identities and also very ambivalent about completing SRS. All of the discontinuers had stopped before castration and genital plastic surgery.

The FMs were significantly younger than the MFs at first assessment (index), follow-up, and at the time for SRS ($p < .05$). Among the MFs, late-onset TS dominated (56%) compared to 11.8% among FMs ($p = .003$). Half of the MFs (48%) were non-homosexually oriented compared to 5.9% of the FMs ($p = .003$). For the entire group, the time span between index and follow-up assessment was 4–16 years, with a mean of 9.0 years. Half of the group had completed SRS within 4 years before follow-up.

Procedure

At index, when the patients applied for treatment, all were clinically assessed by multiple methods: psychiatric, psychological, and somatic. They were interviewed, diagnosed according to ICD-10, and assessed regarding cognitive resources, personality traits, and disorders. Sociodemographic data were collected as well as data about partner experience, sexual orientation, and related issues. The patients also completed various self-report questionnaires.

At follow-up, after 5 years or more in the process or at least 2 years after SRS, the patients were contacted by the same psychiatrist (OB or TH) who had handled the index assessment and their application. The patients were interviewed and asked to complete the same questionnaires and tests as they did at index, which they complied with. This repeated procedure allowed us to compare status before and after SR.

Measures

The Interview

The semi-structured interview at follow-up focused on a variety of areas of functioning, with questions chosen by the researchers on the basis of their clinical experiences and from earlier Swedish studies (Bodlund & Kullgren, 1996; Eldh et al., 1997). The questions covered employment situation, financial situation, partner, family of origin and other significant relations, sexual functioning and orientation, and physical and mental health. There were also questions about how the subjects perceived their gender role and how they experienced the SR procedure as a whole and the outcome of hormonal treatment and surgery. There were a total of 55 pre-formulated questions about changes from index to follow-up. The patient's ratings were made on a 3- or 5-point ordinal scale, with 3 indicating "good" and 1 "poor" on the 3-point scale and 5 = "much better" and 1 = "much worse" on the 5-point scale. In addition, general health aspects and medication were rated by both the clinician and the patient. The five-point categories were then

Table 1 Description of the follow-up group

Description	Male-to-female (MF) $n = 25$ (59.5%) *M* (Range)	Female-to-male (FM) $n = 17$ (40.5%) *M* (Range)	Total $n = 42$ *M* (Range)
Age at index (in years)	37.3 (21–60)[a]	27.8 (18–46)	33.4 (18–60)
Age at SRS (in years)	38.2 (22–57)[b]	31.4 (22–49)	35.2 (22–57)
Age at follow-up (in years)	46.0 (25–69)[b]	38.9 (28–53)	43.1 (25–69)
	n (%)	*n* (%)	*n* (%)
Early-onset transsexualism	11 (44)[a]	15 (88.2)	26 (61.9)
Late-onset transsexualism	14 (56)[a]	2 (11.8)	16 (38.1)
Homosexual orientation	13 (52)[a]	16 (94.1)	29 (69.1)
Completed SRS	18 (72)[ns]	14 (82.4)	32 (76.2)

[a] Student's *t*-test $p < .01$

[b] Student's *t*-test $p < .05$

transformed into three categories: better/improved-unchanged-worse/impaired.

Standardized Rating Format

The format for the clinicians' evaluation of the outcome from index to follow-up was inspired by Hunt and Hampson's (1980) Standardized Rating Format, as well as the work of Wålinder and Thuwe (1975) and the modification of the format previously described by Bodlund and Kullgren (1996). The format was supplemented by a GAF estimation (Axis V in the DSM-IV) made by the clinician at both index and follow-up. An increase in GAF points of five or more was considered to be an improvement and a decrease of ≥five points was considered as worse. Each clinician collected all available information from medical records, personal knowledge of the patient over the years, and data from index and follow-up interviews, and used this information to make their evaluation of the outcome as regarded socioeconomic status, work/study status (including sick leave, disability pension, etc.), partner/family/friend relationships (i.e., changes in the quality of significant relations), actual use of psychiatric care, and global functioning according to the GAF scale and whether it was lower/worsened, stable/unchanged, or improved/higher. In addition, the patients' own opinion on the outcome was included in the format, which means that the patient's and the clinician's evaluations were not fully independent of each other. Global outcome was defined as positive/favorable if at least two of these six areas were improved and not worse in any functional area. Conversely, global outcome was defined as negative if at least two or more areas were worse/impaired and not improved in any. In our earlier study from 1996, we showed good inter-rater reliability (with a Cohen's kappa of about 0.80, unpublished data) between these two clinicians regarding their judgment of global outcome according to this rating format. Gathering all available data over several years and from different areas of functioning, a summary of the outcome of SR for each individual is shown in Table 2.

Results

At follow-up, 16 (38.1%) patients, 9 of the MFs (36%) and 7 of the FMs (41%), had a partner. Twenty-six (62%) patients were employed or involved in studies, compared to 21 at index. Among the other 16 patients, two were unemployed, two were retired, and 12 lived on disability pensions (compared to nine on disability pensions at index).

Clinicians' Evaluation of Global Outcome

As shown in Table 2, the clinicians rated 26 patients (62%) as globally improved, 10 (24%) as unchanged, and six (14%) as worse. Three out of five of those still in the process and all of those who had discontinued from further surgery were rated as improved.

According to the clinicians evaluation (Table 3), more MFs were rated as improved compared to FMs ($p = .04$), but there was no significant difference in outcome between the subgroups (early- vs. late-onset TSs and homosexual vs. non-homosexual sexual orientation).

Patients' Evaluation of Global Outcome

In Table 2, it is also evident that the patients' own assessments of global outcome were more positive than the clinicians'. Forty (95%) patients rated themselves as improved and only two viewed the outcome as negative. None of the patients who were still in the process rated the outcome as negative thus far. Also, those five who had interrupted the SR process seemed to be content with their decision to do so.

Table 2 Clinicians' and patients' evaluation of *global outcome of SR* according to the standard rating format

Patient no	Biol. sex	Years from index to follow-up	Diagnostic TS-typology	Sexual orientation	Soc-econ status	Work/ studies	Significant relations	Psychiatric care	GAF-difference ≥5 points	Global outcome-Patients' evaluation	Global outcome-Clinicians' evaluation
1	M	8	Late onset	Homo	+	+	0	0	−	+	0
2	M	12	Early onset	Non-HS	0	0	0	0	0	+	0
3	M	8	Late onset	Homo	+	+	+	+	+	+	+
4[a]	M	8	Late onset	Non-HS	0	+	0	+	0	+	+
5	M	16	Late onset	Non-HS	−	−	0	0	+	+	0
6[a]	M	7	Late onset	Homo	+	+	0	0	0	+	+
7	M	4	Late onset	Non-HS	−	0	0	−	−	+	−
8	M	5	Late onset	Homo	0	0	+	0	0	+	+
9[a]	M	14	Late onset	Non-HS	−	−	0	0	0	+	−
10[b]	M	5	Late onset	Non-HS	+	+	0	0	0	+	+
11	M	15	Early onset	Homo	+	+	0	0	+	+	+
12[b]	M	5	Late onset	Homo	0	0	+	0	0	+	+
13	M	15	Early onset	Homo	+	0	+	+	+	+	+
14[b]	M	5	Late onset	Non-HS	+	0	+	+	+	+	+
15[b]	M	8	Late onset	Non-HS	+	+	+	0	+	+	+
16	M	5	Early onset	Homo	0	0	0	0	−	+	0
17	M	15	Late onset	Homo	0	0	0	0	+	+	+
18	M	6	Early onset	Homo	+	+	0	0	0	+	+
19	M	8	Early onset	Homo	+	+	+	0	0	+	+
20	M	7	Late onset	Non-HS	0	0	+	0	0	+	+
21	M	13	Early onset	Non-HS	0	0	+	+	+	−	+
22	M	10	Early onset	Homo	0	0	−	0	+	+	0
23	M	4	Early onset	Homo	+	+	+	+	+	+	+
24	M	7	Early onset	Non-HS	+	0	+	+	+	+	+
25	M	5	Early onset	Non-HS	+	+	0	+	+	+	+
26[a]	F	10	Early onset	Homo	−	−	−	0	0	+	−
27	F	5	Early onset	Homo	0	0	+	0	0	+	+
28[a]	F	6	Early onset	Homo	0	0	0	0	+	+	+
29	F	9	Early onset	Homo	0	0	−	0	0	+	0
30	F	14	Early onset	Homo	0	0	0	0	+	+	+
31	F	13	Early onset	Homo	+	+	+	+	+	+	+
32	F	15	Early onset	Homo	0	0	−	0	−	+	−
33	F	16	Early onset	Homo	+	+	0	0	+	+	+
34	F	15	Early onset	Homo	+	+	0	0	0	+	+
35[b]	F	5	Late onset	Non-HS	+	+	+	0	0	+	+

Table 2 continued

Patient no	Biol. sex	Years from index to follow-up	Diagnostic TS-typology	Sexual orientation	Soc-econ status	Work/ studies	Significant relations	Psychiatric care	GAF-difference ≥5 points	Global outcome-Patients' evaluation	Global outcome-Clinicians' evaluation
36	F	11	Early onset	Homo	0	0	0	0	0	+	0
37	F	10	Early onset	Homo	0	0	0	0	0	+	0
38	F	7	Late onset	Homo	–	–	0	0	–	–	–
39	F	8	Early onset	Homo	0	0	–	+	Missing	+	0
40	F	10	Early onset	Homo	0	–	+	0	+	+	0
41	F	15	Early onset	Homo	0	–	–	–	0	+	–
42	F	5	Early onset	Homo	0	0	+	+	+	+	+
Sum	M 25	Mean 9.0	Early 26	Homo 29							
	F 17	Range 4–16	Late 16	Non-HS 13							
Improved	(+)				17 (40%)	15 (36%)	16 (38%)	11 (26%)	19 (46%)	40 (95%)	26 (62%)
Unchanged	(0)				20 (48%)	21 (50%)	20 (48%)	29 (69%)	17 (42%)	0 (0%)	10 (24%)
Worsened	(–)				5 (12%)	6 (14%)	6 (14%)	2 (5%)	5 (12%)	2 (5%)	6 (14%)

[a] Not completed SRS, still in the process [b] Discontinued from genital surgery (SRS). Non-HS = heterosexual orientation relative biological sex

Patients' Evaluation of Specific Aspects of Outcome

On the question of their gender identity at follow-up, there were no patients who identified themselves with their biological sex. Thirty-three (79%) patients identified themselves according to their preferred gender identity and felt that others also perceived them in the same way and not according to their biological sex. Nine (21%) identified themselves as just transsexuals, of which eight were MFs. These patients had not yet completed the SR procedure or had interrupted the process, and stated that they felt ambivalent due to a perceived lack of acceptance from other people or dissatisfaction with their physical appearance, which still indicated their biological sex.

Table 4 shows that the patients were extremely satisfied with the SR process as a whole, with no differences between the sexes. As many as 95.2% of the patients (40 out of 42) were satisfied. Two (one MF and one FM) were dissatisfied because of shortcomings in the genital surgery.

Few patients rated their work situation, partner relation or sex life as impaired at follow-up, with no significant differences between the sexes. Three experienced some impairment in their work situation. One, a MF, had a typically masculine job and experienced difficulties coping at work as a woman. Another MF had previously held a job but was unemployed at the time of the follow-up study. The third case, also a MF, was an unemployed former student.

Regarding impaired partner relations, two (MF) gave sexual reasons. One claimed that the hormonal treatment led to a negative effect on her sex drive and the other reported that her vagina was not functional for intercourse. The third patient (FM) expressed social problems in trying to live in a masculine role and was also slightly depressed.

Current sex life, compared to before SR and hormone treatment, was rated as improved or unchanged by 95% of the patients. Two MFs expressed "impaired sex life" because of lower sex drive and lack of partner, respectively.

There was, however, a significant difference between the sexes regarding their sexual orientation. Twelve of the MFs (48%) preferred a female partner compared to only one among the FMs (6%), who had a preference for a male partner ($p = .004$), i.e., non-homosexuality in relation to biological sex. There was also a significant correlation (Pearson $r = .45$, $p = .01$) between male sex, late onset, and non-homosexual orientation.

The initiation of hormonal treatment is a crucial part of the SR process and often functions as a confirmation of the prospective of the procedure. Thirty-five (89.7%) patients were, in general, satisfied with the hormonal treatment, three stated neither/nor, and one was dissatisfied because she was forced to cease treatment due to an allergic reaction.

The general evaluation of the surgery treatment showed that of 33 patients (32 with genital surgery and one with only mastectomy), 22 (66.7%) were satisfied, seven (21.2%) were

Table 3 Clinicians' evaluation of global outcome, in relation to sex and type of TS

	MF $n = 25$	FM $n = 17$	Early-onset TS $n = 26$	Late-onset TS $n = 16$	Homosex orientation $n = 29$	Non-homosex $n = 13$
Improved	18 (72%)[a]	8 (47%)	15 (57.6%)	11 (68.8%)	17 (58.6%)	9 (69.2%)
Unchanged	5 (20%)	5 (29.4%)	8 (30.7%)	2 (12.5%)	8 (27.6%)	2 (15.4%)
Worsened	2 (8%)	4 (23.5%)	3 (11.5%)	3 (18.7%)	4 (13.8%)	2 (15.4%)

[a] Student's *t*-test $p = .04$

Table 4 The patients' statements according to the follow-up interview concerning satisfaction with the SR process and outcome in regard to work, partner relationships, and sex life

	MF	FM	All
SR process as a whole	$n = 25$	$n = 17$	$n = 42$
Satisfied	24 (96%)	16 (94.1%)	40 (95.2%)
Neither/nor	0	0	0
Dissatisfied	1 (4%)	1 (5.9%)	2 (4.8%)
Work situation	$n = 23$	$n = 15$	$n = 38$
Better	9 (39.1%)	8 (53.3%)	17 (44.7%)
Unchanged	11 (47.8%)	7 (46.7%)	18 (47.4%)
Worsened	3 (13.1%)	0	3 (7.9%)
Partner relations	$n = 23$	$n = 14$	$n = 37$
Better	16 (69.6%)	7 (50%)	23 (62.2%)
Unchanged	5 (21.7%)	6 (42.9%)	11 (29.7%)
Worsened	2 (8.7%)	1 (7.1%)	3 (8.1%)
Sex life	$n = 24$	$n = 16$	$n = 40$
Better	16 (66.7%)	12 (75%)	28 (70%)
Unchanged	6 (25%)	4 (25%)	10 (25%)
Worsened	2 (8.3%)	0	2 (5%)

neither/nor, and four (12.1%) were dissatisfied. Three of these latter four were FMs.

The patients were also interviewed about changes from index to follow-up in their general health, psychiatric distress, and use of psychotropic medication. The majority (55.4%) rated their general health as improved, 28.5% stated "no change," and 16.1% felt that their overall health was impaired. There were no sex differences, but the impaired group was slightly older. A common response associated with improvement was better psychological well-being. Half of those who rated their health as "unchanged" meant they were still in good health. As for psychological problems, 30–50% stated they had suffered from insomnia, depression or anxiety in the last year, but only five out of 42 (12%) were on antidepressants or tranquillizers. Only three patients (7.1%) were receiving ongoing psychiatric treatment (for other reasons than the SR procedure).

When comparing the outcome (according to statements in the interview) for early onset vs. late-onset TSs, some significant differences were found, as shown in Table 5.

Most late-onset TSs were MF, much older when applying for SRS, and less satisfied with their sex life. Only 15.4% of TSs with early onset had a non-homosexual orientation, compared to more than half of those with late onset. The same pattern emerged when comparing homosexual to non-homosexual TSs: 92% of the latter were MF, older at index (37 vs. 32, ns), and significantly less satisfied with their sex life ($p = .003$). In addition, the majority (62%) were late-onset TS ($p = .005$). However, there were no differences within these two subgroups regarding satisfaction with current work situation, financial status, general health, or hormonal or surgical treatment. Over 90% in both groups were satisfied with the SR process as a whole.

Table 5 Significant differences between TSs with early versus late onset

	Early-onset TS $n = 26$ (61.9%)	Late-onset TS $n = 16$ (38.1%)	*t-test*
Male sex (MF)	11 (42.3%)	14 (87.5%)	$p < .01$
Age at index (years)	28	42	$p < .01$
Satisfied with sex life at follow-up	20 (76.9%)	8 (50%)	$p < .05$
Non-homosexual orientation (relative to biological sex)	4 (15.4%)	9 (56.2%)	$p < .01$

Discussion

The study aimed at describing the outcome of SR from both the patient's and the professional's perspective. According to the clinician, more than 60% of the patients were judged as globally improved, compared to 95% according to the patients' own judgment. In a previous study by Bodlund and Kullgren (1996), the corresponding findings, respectively, were 68 and 63%. In other studies, using a variety of outcome measures, the average positive outcome was between 71 and 97% (Cohen-Kettenis & Pfäfflin, 2003). The discrepancy between the clinicians' and the patients' evaluation can be explained by different standards when evaluating changes over several years and by the fact that the clinician took into account several more objective and, as we see it, relevant outcome factors, such as work situation, social relationships, financial situation, partnership, GAF value, etc. Furthermore, it is not easy for the patient to openly regret such an

irreversible decision as change sex. On the other hand, it can be argued that the only legitimate basis for judging outcome is the patient's satisfaction (Gijs & Brewaeys, 2007). However, the outcome was very encouraging from both perspectives, with almost 90% enjoying a stable or improved life situation at follow-up and only six out of 42 (according to the clinician) with a less favorable outcome.

No significant difference in outcome between early and late onset or between homosexual and non-homosexual oriented TSs was found, but there was a more positive outcome for MFs compared to FMs. These findings differ from most other studies and are difficult to explain. More FMs were impaired as concerned socioeconomic status, work, and partner situation, but they rated themselves as equally satisfied with the outcome as the MFs did. One explanation could be that FMs were already well functioning at index in many areas, which is in accordance with our clinical impression, and that MFs had more to gain in terms of social and psychological functioning during the SR process.

The clinician also rated the outcome as less favorable than the patients did themselves in the interview, especially concerning "partner relationship." This may be because the clinician's evaluation had a broader meaning, covering relationships with friends and family as well as partner relationships. A similar discrepancy was seen regarding the work situation, which we interpret as differing valuations of, e.g., a disability pension.

The vast majority of patients identified themselves with their preferred sex and the few who called themselves transsexuals were those who had interrupted or were still in the process. This finding supports the idea that the treatment goal is a complete SRS.

When interviewed about how content they were with the SR process as a whole, almost all the patients (95%) rated themselves as satisfied and no one regretted the SR. However, four patients were discontent with the surgery, three of which were FMs. For FMs, genital surgery is more complicated and the result is more cosmetic than functional. On the other hand, there were only two patients (5%) who said that their sexuality was impaired after SRS, and they were MFs. So the surgical result is not always decisive for sexual functioning or satisfaction.

The outcome data displayed few sex differences. FMs were younger at application, mainly reflecting the fact that 15 out of a total 17 were early-onset TS, which implies an earlier onset of cross-gender orientation. Only one of the FMs reported a non-homosexual orientation compared to half of the group of MFs. As in previous studies, we found an expected connection between male biological sex, late-onset TS, and non-homosexual orientation.

As stated above, we could not verify that late-onset TS would have a less favorable outcome despite the fact that the late-onset TSs were older, mostly MF, and more often single, non-homosexual oriented, and less satisfied with their sex life. This does not seem to interfere with their assessment of being as satisfied and well-functioning as early-onset TSs. However, these results are supported by Lawrence (2003) who showed that diagnostic typology is not entirely decisive for outcome. There is nonetheless a problem in comparing different studies according to the different views of how to categorize the spectra of transsexualism into subgroups. When dichotomizing into early/late onset, MF/FM, and homo/non-homosexual, there is always a risk of losing statistical power when the groups become small. As shown in Tables 2 and 3, there was no robust connection in this study between sexual orientation, age of onset, and outcome.

Of the total population of 60 individuals who had changed sex in these two regions of Sweden, 42 completed the follow-up procedure. This response rate of 70% allows for generalizations. As for the eight dropouts who declined to participate, we can only speculate on the following reasons, perhaps among others: (a) a long time had passed since SRS and they no longer had contact with the TS team, (b) they were dissatisfied with the procedure and/or result and/or clinician, or (c) they would like to preserve their anonymity and not risk being revealed to close relations or friends who may not be aware of their earlier life situation.

The strengths in this study were the prospective longitudinal design, which allowed the clinicians to develop extensive knowledge of their patients over several years of contact (M = 9 years), and the fact that both the clinician's and the patient's perspective were considered in the evaluation of outcome. On the other hand, the results may have been biased by the fact that we were investigating our own patients, which may have challenged our objectivity. Another possibility is that the patients may tend to want to please their clinician because of their dependency, for instance, on future support and prescriptions (of hormones). Another limitation was the small number of patients, which is a constant problem in this specific research area. However, it is noteworthy that we managed to follow-up and evaluate as many as 70% of all treated transsexuals in these two geographic areas.

In this study we did not aim at investigating the patients who were not approved for SR due to other diagnoses, instability in personality or situation etc. But clinically we are guided by a previous Swedish study by Lundström (1981) who made a follow-up of a group of patients not accepted for SR. His conclusion was that there was new strong evidence for the complexity of the male group regarding diagnoses, and that many patients in the rejected group had persistent cross-sex feelings and were more dissatisfied with life, yet some of them thought the decision was right. The study also pointed out which patients could benefit from SR and which could adjust without. The situation today in Sweden is that a much smaller proportion is excluded from SRS, mainly because a more rigorous evaluation is made before being referred to the gender team, and possibly also because an ongoing shift in diagnostic thinking. However, there is a need for renewed and recurring follow-up of these rejected patients.

In the past few years, the number of applicants for SR in Sweden has increased dramatically. Contributing factors to this increase are probably a higher tolerance in our society, extensive and more easily spread information about treatment options and legal aspects via the Internet and other mass media, and possibly also interacting cultural trends. A likely consequence of this increase will be a larger proportion of atypical cases applying for SR, which in turn will challenge our current clinical procedure for inclusion and exclusion. However, the ambition will still be to maintain the overall positive outcome of SRS and hopefully with no regrets.

Acknowledgments This study was supported by grants from the Sjöbring Foundation and the Lindhaga Foundation, the University Hospital of Lund, and the Department of Psychiatry at the University Hospital of Umeå, Sweden.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (Fourth ed., text rev.). Washington, DC: Author.

Blanchard, R. (1989). The classification and labeling of nonhomosexual gender dysphorias. *Archives of Sexual Behavior, 18*, 315–334.

Blanchard, R., Steiner, B. W., Clemmensen, L. H., & Dickey, R. (1989). Prediction of regrets in postoperative transsexuals. *Canadian Journal of Psychiatry, 34*, 43–45.

Bodlund, O., & Kullgren, G. (1996). Transsexualism: General outcome and prognostic factors. *Archives of Sexual Behavior, 25*, 303–316.

Cohen-Kettenis, P. T., & Pfäfflin, F. (2003). *Transgenderism and intersexuality in childhood and adolescence: Making choices*. Thousands Oaks, CA: Sage.

Docter, R. F. (1988). *Transvestites and transsexuals: Toward a theory of cross-gender behavior*. New York: Plenum Press.

Doorn, C. D., Poortinga, J., & Verschoor, A. M. (1994). Cross-gender identity in transvestites and male transsexuals. *Archives of Sexual Behavior, 23*, 185–201.

Eldh, J., Berg, A., & Gustafsson, M. (1997). Long term follow up after sex reassignment surgery. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery, 31*, 39–45.

Gijs, L., & Brewaeys, A. (2007). Surgical treatment of gender dysphoria in adults and adolescents: Recent developments, effectiveness, and challenges. *Annual Review of Sex Research, 18*, 178–224.

Hunt, D., & Hampson, J. L. (1980). Transsexualism: A standardized psychosocial rating format for the evaluation of results of sex reassignment surgery. *Archives of Sexual Behavior, 9*, 255–263.

Kuiper, A. J., & Cohen-Kettenis, P. T. (1998). Gender role reversal among postoperative transsexuals. *International Journal of Transgenderism, 2*. Available at http://www.symposion.com/itj/itjc0502.htm.

Landén, M., Bodlund, O., Ekselius, L., Hambert, G., & Lundström, B. (2001). Bytt är bytt, kommer aldrig igen. Könsbyte för närvarande bästa metoden för transsexuella [Done is done and gone is gone; sex change presently the best remedy for transsexuals]. *Läkartidningen, 30–31*(98), 3322–3326.

Landén, M., Wålinder, J., Hambert, G., & Lundström, B. (1998). Factors predictive of regret in sex reassignment. *Acta Psychiatrica Scandinavia, 97*, 284–289.

Lawrence, A. A. (2003). Factors associated with satisfaction or regret following male-to-female sex reassignment surgery. *Archives of Sexual Behavior, 32*, 299–315.

Lundström, B. (1981). *Gender dysphoria: A social-psychiatric follow-up study of 31 cases not accepted for sex reassignment*. Doctoral dissertation, Department of Psychiatry, University of Gothenburg, Gothenburg, Sweden.

Lundström, B., Pauly, I. B., & Wålinder, J. (1984). Outcome of sex reassignment surgery. *Acta Psychiatrica Scandinavica, 70*, 289–294.

Olsson, S., & Möller, R. (2003). On the incidence of transsexualism in Sweden, 1972–2002. *Archives of Sexual Behavior, 32*, 381–386.

Pfäfflin, F., & Junge, A. (1998). Sex reassignment: Thirty years of international follow-up studies. A comprehensive review, 1961–1991. *International Journal of Transgenderism.* http://www.symposion.com/ijt/books/index.htm.

Smith, Y. L. S., van Goozen, S. H. M., Kuiper, A. J., & Cohen-Kettenis, P. T. (2005). Sex reassignment: Outcomes and predictors for adolescents and adult transsexuals. *Psychological Medicine, 35*, 89–99.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: Author.

Wålinder, J., Lundström, B., & Thuwe, I. (1978). Prognostic factors in the assessment of male transsexuals for sex reassignment. *British Journal of Psychiatry, 132*, 16–20.

Wålinder, J., & Thuwe, I. (1975). *A social follow-up study of 24 sex-reassigned transsexuals*. Gothenburg, Sweden: Scandinavian University Books.

Arch Sex Behav (2010) 39:1439–1441
DOI 10.1007/s10508-009-9518-2

COMMENT

The Biogeography and Evolution of Female Homosexual Behavior in Japanese Macaques

Paul L. Vasey · Hester Jiskoot

Received: 29 December 2008 / Revised: 16 June 2009 / Accepted: 16 June 2009 / Published online: 18 August 2009

Abstract In certain Japanese macaque (*Macaca fuscata*) populations, females routinely engage in same-sex courtship, mounting, and consortship activity. Drawing on behavioral, biogeographic, and genetic research, we suggest that female homosexual behavior may be associated with genetically distinct free-ranging populations of Japanese macaques. In addition, we briefly discuss the implications of this research for the evolution of female homosexual behavior in this species.

Keywords Female homosexual behavior · Japanese macaques · Biogeography · Haplotypes

Female Japanese macaque (*Macaca fuscata*), in certain populations, are unusual in that, in addition to engaging in heterosexual behavior, they routinely engage in same-sex courtship and series mounting (with pelvic thrusting) during temporary, but exclusive relationships termed "consortships." Females compete with males for same-sex sexual partners and will often choose same-sex sexual partners even when given the simultaneous choice of a sexually motivated male alternative (Vasey, 1998). These behaviors do not appear to serve any sociosexual function including: attracting male mates, alliance formation, dominance demonstration, alloparental care acquisition, reconciliation, practice for heterosexual copulation, and inter-individual tension reduction (reviewed in Vasey, 2006).

Several lines of evidence indicate that female-female mounting, courtship, and consortship activity in Japanese macaques is sexually motivated. First, these interactions mirror male-female sexual behavior in many aspects of their expression. For example, the courtship behavior that females exhibit during homosexual consortships is virtually indistinguishable from that which occurs during heterosexual consortships (Vasey, Rains, VanderLaan, Duckworth, & Kovacovsky, 2008). Similarly, female-female mounting across the ovarian cycle parallels the pattern found for male-female mounting (O'Neill, Fedigan, & Ziegler, 2004). In addition, females exercise incest avoidance with close female kin (Chapais, Gauthier, Prud'homme, & Vasey, 1997) despite the fact that these same kin dyads engage in various forms of social affiliation together, such as grooming (Baxter & Fedigan, 1979). Finally, during most same-sex mounts (78.3%), female mounters engage in vulvar, perineal, and anal stimulation (Vasey & Duckworth, 2006).

This distinctive pattern of female-female consortship activity has been reported in three free-ranging Japanese macaque populations, which are located in the central region of the island of Honshu (Jigokudani: Enomoto, 1974; Mino-o: Perloe, 1989; Arashiyama: Wolfe, 1984). Japanese macaques range across the entire island of Honshu. As such, evidence from free-ranging populations indicates that female homosexual consortship activity is confined to a restricted geographic region within Japanese macaque habitat. This raises the possibility that the unique pattern of female-female consortship activity described above may be associated with genetically distinct free-ranging Japanese macaque populations.

Recent mitochondrial DNA (mtDNA) research on Japanese macaques indicates genetic differentiation into five

P. L. Vasey (✉)
Department of Psychology, University of Lethbridge, 4401 University Drive, Lethbridge, AB T1K 3M4, Canada
e-mail: paul.vasey@uleth.ca

H. Jiskoot
Department of Geography, University of Lethbridge, Lethbridge, AB, Canada

haplogroups all of which occur on the island of Honshu (Kawamoto et al., 2007). The three free-ranging populations in which female homosexual behavior has been observed fall within a single haplogroup (A1), which is exclusive to Honshu. Each of these free-ranging populations is characterized by separate haplotypes (e.g., locations: 36 Yamanouchi, haplotype JN15; 87 Arashiyama, haplotype JN21; 93 Mino-o, haplotype JN35; Kawamoto et al., 2007), of which there are 27 within the A1 haplogroup. Although limited, information on captive Japanese macaque colonies derived from additional areas of Honshu, and in which female homosexual behavior has been observed, corroborates the suggestion that this behavior is exclusively associated with the A1 haplogroup on the island Honshu (e.g., Wakasa-B colony [Primate Research Institute, Inuyama, Japan], location 96 Wakasa, A1 haplotype JN36, Kawamoto et al., 2007; Vasey & Reinhart, 2009). This information strengthens the inference that the distinctive pattern of female homosexual behavior observed in free-ranging Japanese macaque populations may indeed be associated with genetically distinct Japanese macaque populations on the island of Honshu.

Marked geographical variations and concentrations of particular haplogroups are thought to reflect the effects of genetic drift, particularly population bottlenecks and founder events. Japanese macaques are genetically closest to the eastern rhesus macaque (*M. mulatta*: Hoelzer & Melnick, 1996). Research suggests that in the Middle Pleistocene a small subgroup of eastern rhesus macaques migrated eastward from Korea to Japan (Marmi, Bertranpetit, Terradas, Takenaka, & Domingo-Roura, 2004) along a landbridge that resulted from the glacial sea levels being 130–150 m lower than at present (Millien-Parra & Jaeger, 1999). On the basis of palaeogeographic evidence and the propinquity of Japanese macaque genotypes, Marmi et al. (2004) suggested that the small founder population of eastern rhesus macaques dispersed quickly over the Japanese archipelago resulting in an initially homogeneous set of mtDNA genotypes.

These homogeneous genotypes then evolved independently during the late Pleistocene/Holocene in reduced and discrete local populations (Marmi et al., 2004). In light of their habitat requirements (Azuma, 1985), low-lying coastal glacial refugia with temperate vegetation likely facilitated the establishment of genetically distinct Japanese macaque populations. In support of this suggestion, we note that some present-day haplogroups straddle the waterways between islands (Kawamoto et al., 2007), lending weight to the conclusion that genetic divergence began in glacial refugia when sea levels were low, rather than from the geographic isolation of islands. Furthermore, we note that the core area of the A1 haplogroup is located in the central western region of Honshu, a location that is associated with glacial refugia (Tsukada, 1982; Kawamoto et al., 2007). Notably, this region coincides with specific vegetation (e.g. Japanese beech [*Fagus crenata*]; Tomaru et al., 1997) and mammalian fauna (e.g. Japanese sika deer [*Cervus nippon*]; Nagata et al., 1999), which Kawamoto et al. (2007) attribute to local selection pressures within a coastal glacial refugium on the western side of Honshu.

The distinctive pattern of female homosexual behavior observed in some Japanese macaque groups (i.e., sexually, not sociosexually, motivated same-sex series mounting with thrusting, courtship and consortships, along with inter-sexual mate competition and facultative same-sex sexual partner preference) may have evolved in populations living in the core A1 haplogroup area of central western Honshu (e.g., locations: 87 Arashiyama and 93 Mino-o; Kawamoto et al., 2007) once they became geographically isolated in glacial refugia. Females that were behaviorally capable of this distinctive type of homosexual behavior may have then spread from the core A1 haplogroup area following deglaciation and subsequently established populations in higher elevation, inland regions of Honshu (>800 m above sea level (asl), and in some places up to 3000 m asl; e.g., location: 36 Yamanouchi). Terrain and postglacial climate zonation during the late Pleistocene/Holocene may have subsequently facilitated further inter-group differences in the expression of female homosexual behavior through geographic isolation.

Acknowledgements We thank Mike Huffman, Jean-Baptiste Leca, Doug VanderLaan, Lesley Terry, and one anonymous reviewer. P.L.V. was funded by the University of Lethbridge and by a Natural Sciences and Engineering Research Council (NSERC) of Canada Discovery Grant. H.J. was funded by a NSERC Discovery Grant and by a NSERC University Faculty Award.

References

Azuma, S. (1985). Ecological biogeography of Japanese monkeys (*Macaca fuscata* Blyth) in the warm- and cold-temperate forest. In T. Kawamichi (Ed.), *Contemporary mammalogy in China and Japan* (pp. 1–5). Osaka: Mammalogical Society of Japan.

Baxter, M. J., & Fedigan, L. M. (1979). Grooming and consort partner selection in a troop of Japanese monkeys (*Macaca fuscata*). *Archives of Sexual Behavior, 8*, 445–458.

Chapais, B., Gauthier, C., Prud'homme, J., & Vasey, P. L. (1997). Relatedness threshold for nepotism in Japanese macaques. *Animal Behavior, 53*, 533–548.

Enomoto, T. (1974). The sexual behavior of Japanese monkeys. *Journal of Human Evolution, 3*, 351–372.

Hoelzer, G. A., & Melnick, D. J. (1996). Evolutionary relationships of the macaque. In J. E. Fa & D. G. Lindburg (Eds.), *Evolution and ecology of the macaque societies* (pp. 3–17). Cambridge: Cambridge University Press.

Kawamoto, Y., Shotake, T., Nozawa, K., Kawamoto, S., Tomari, K., Kawai, S., et al. (2007). Postglacial population expansion of Japanese macaque (*Macaca fuscata*) inferred from mitochondrial DNA phylogeography. *Primates, 48*, 27–40.

Marmi, J., Bertranpetit, J., Terradas, J., Takenaka, O., & Domingo-Roura, X. (2004). Radiation and phylogeography in the Japanese macaque, *Macaca fuscata*. *Molecular Phylogenetics and Evolution, 30*, 676–685.

Millien-Parra, V., & Jaeger, J.-J. (1999). Island biogeography of the Japanese terrestrial mammal assemblages: An example of a relict fauna. *Journal of Biogeography, 26*, 959–972.

Nagata, J., Masuda, R., Tamate, H. B., Hamasaki, S., Ochiai, K., Asada, M., et al. (1999). Two genetically distinct lineages of the sika deer, *Cervus nippon,* in the Japanese islands: Comparison of mitochondrial D-loop region sequences. *Molecular Phylogenetics and Evolution, 13*, 511–519.

O'Neill, A. C., Fedigan, L. M., & Ziegler, T. E. (2004). Ovarian cycle phase and same-sex mating behavior in Japanese macaque females. *American Journal of Primatology, 63*, 25–31.

Perloe, S. I. (1989). Monkeys of Minoo. Produced by S. I. Perloe, Department of Psychology, Haverford College [VHS; 45 min].

Tomaru, N., Mitutsuji, T., Takahashi, M., Tsumura, Y., Uchida, K., & Ohba, K. (1997). Genetic diversity in *Fagus crenata* (Japanese beech): Influence of the distributional shift during the late-Quaternary. *Heredity, 78*, 241–251.

Tsukada, M. (1982). *Crypotmeria japonica*: Glacial refugia and late-glacial and postglacial migration. *Ecology, 63*, 1091–1105.

Vasey, P. L. (1998). Female choice and inter-sexual competition for female sexual partners in Japanese macaques. *Behaviour, 135*, 579–597.

Vasey, P. L. (2006). The pursuit of pleasure: Homosexual behavior, sexual reward and evolutionary history in Japanese macaques. In V. Sommer & P. L. Vasey (Eds.), *Homosexual behavior in animals: An evolutionary perspective* (pp. 349–364). Cambridge: Cambridge University Press.

Vasey, P. L., & Duckworth, N. (2006). Sexual reward via vulvar, perineal and anal stimulation: A proximate mechanism for female homosexual mounting in Japanese macaques. *Archives of Sexual Behavior, 35*, 523–532.

Vasey, P. L., Rains, D., VanderLaan, D. P., Duckworth, N., & Kovacovsky, S. D. (2008). Courtship behavior during heterosexual and homosexual consortships in Japanese macaques. *Behavioral Processes, 78*, 401–407.

Vasey, P. L., & Reinhart, C. (2009). Female homosexual behavior in a new group of Japanese macaques: Evolutionary implications. *Laboratory Primate Newsletter, 48*(3), 8–10.

Wolfe, L. D. (1984). Japanese macaque female sexual behavior: A comparison of Arashiyama East and West. In M. F. Small (Ed.), *Female primates: Studies by women primatologists* (pp. 141–157). New York: Alan R. Liss.

Arch Sex Behav (2010) 39:1443–1447
DOI 10.1007/s10508-010-9645-9

ORIGINAL PAPER

Paraphilic Coercive Disorder in the DSM: The Right Diagnosis for the Right Reasons

Paul Stern

Published online: 22 June 2010

Abstract The recommendation to include a Paraphilic Coercive Disorder (PCD) diagnosis in the DSM-5 represents an improvement over current options and would lead to the shrinking of the pool of individuals considered for detention as Sexually Violent Predators. A precise description of the diagnostic criteria for PCD would permit psychologists and psychiatrists to use more specific and narrow criteria for those who seek sexual gratification by coercing others to engage in unwanted sexual behavior. This might permit mental health professionals to abandon the Paraphilia NOS designation in favor of the more defined PCD in appropriate cases. Various critics have attacked the proposal on what appears to be misplaced ideological grounds. Not only should ideological concerns not play a part in a scientific debate, but the critics' predictions of how the PCD diagnosis would play out in the legal arena are likely wrong. Paraphilic Coercive Disorder would give the judicial system the best opportunity to most accurately identify the small group of men who have previously committed, and are likely in the future to commit, this type of predatory sexual violence.

Keywords DSM-5 · Paraphilic Coercive Disorder · Paraphilia NOS (rape) · Paraphilia NOS (nonconsent) · Civil commitment · Sexually violent predator laws · Sex offenders

A primary goal of the criminal justice system is to achieve the most accurate determination of truth based upon the most reliable information available and without consideration of extraneous or irrelevant factors. It is vital that judges and juries are able to rely upon the most accurate and scientific information available. To do that, expert witnesses who testify in court must themselves be able to rely on the most accurate scientific information available.

In cases involving a mental health diagnosis of a party, the expert generally will rely on what is included in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR; American Psychiatric Association, 2000). If something is excluded from the DSM because of nonscientific concerns, then the expert witness is denied access to the most reliable information. In turn, the courts are denied access to the most scientifically reliable information, and faulty, even dangerous, decisions may be made.

Ideological and political views may properly be used to influence public policy. However, they should not be used to influence science. I write because I am concerned that ideology and politics—and not objective science—might be used to evaluate the recommendation of the DSM-5 Work Group on Sexual and Gender Identity Disorders to include Paraphilic Coercive Disorder (PCD) in the DSM-5.

One type of litigation in which the DSM plays a significant role is the civil commitment of sexually violent predators (SVP). Currently, 20 American states have laws that permit the detention of a select group of individuals if they have a specific history of sexual offending, plus a mental abnormality or personality disorder that causes the person serious difficulty in controlling his sexually violent behavior, and the mental abnormality or personality disorder makes the person likely to engage in predatory acts of sexual violence if not confined to a secure facility. These strict requirements mean that a tiny[1] percentage

P. Stern (✉)
Snohomish County Prosecutors Office, 3000 Rockefeller Avenue, M/S 504, Everett, WA 98201, USA
e-mail: pstern@co.snohomish.wa.us

[1] In the 6 years after the effective date of the Washington State SVP statute, only 5% of the sex offenders released from custody who met the statutory criteria for SVP commitment were even referred to prosecutors for consideration of filing a SVP petition. Prosecutors declined to file in two-thirds of those cases. Thus, the number of SVP petitions filed in Washington State was approximately 1.5% of the number of cases reviewed (Milloy, 2003).

of sex offenders released from prison every year are even considered for SVP assessment.

I am a lawyer. More specifically, I am a prosecutor who has spent more than 25 years working with issues related to sexual offending and other forms of interpersonal violence and, most specifically, with the interplay between mental health issues and the court system. I have been involved with issues regarding sexual offenders and the judicial system throughout my career. I served on the Board of Directors of the Association for the Treatment of Sexual Abusers, the Advisory Committee to the Washington State Sex Offender Treatment Provider Board, and on the Washington State Twin Rivers Sex Offender Treatment Program Advisory Committee. Those experiences, perhaps, are why the American Psychiatric Association (APA) was kind enough to invite me to serve as an Advisor to the Paraphilias subworkgroup of the DSM-5 Work Group on Sexual and Gender Identity Disorders.

I have handled more than a dozen SVP cases. I have worked on SVP cases and the issues relating to them since the first law was enacted in the state of Washington in 1991. I do not want to detain as SVPs those who should not be held nor do I want dangerous men released to the street because of irresponsible statements made by mental health professionals. I want accuracy.

That is why I support the subworkgroup's recommendation that PCD be added to the DSM-5 (see Table 1). I endorse it even though it will mean that some dangerous men currently considered under the SVP laws will not be able to be detained. I endorse it even though it will make my job harder. I endorse it because it will further my goal of allowing the court system to achieve the most accurate determination of truth.

Paraphilic Coercive Disorder gives the judicial system the best opportunity to most accurately identify the small group of men who have previously committed, and are likely in the future to commit, this type of predatory sexual violence.

This small group of men has long been recognized. A "deviant rape pattern" in which the male prefers raping over intimate sexual interaction has been documented in the sex offender literature for nearly 40 years (Freund, Seeley, Marshall, & Glinfort, 1972). This "deviant rape pattern" was more fully explained in 1986, nearly a half decade before the SVP laws were drafted. Freund, Scher, Racansky, Campbell, and Heasman (1986) found that a coercive preference is a type of courtship disorder, perhaps less severe than true sadism, but a distinct group. They noted that a preferential rape pattern is "connected with either an abnormally high tolerance of, or a preference for, an abnormally strong agonistic component in sexual interaction" (p. 34). Freund et al. concluded: "This outcome suggests that a substantial proportion of rape-prone males have a special erotic affinity for, or are erotically less inhibited than normals by, situations which are characteristic of the courtship disorders" (p. 30).

Table 1 Diagnostic criteria for Paraphilic Coercive Disorder proposed by the Paraphilias subworkgroup of the Sexual and Gender Identity Disorders Work Group for DSM-5

A. Over a period of at least 6 months, recurrent, intense sexually arousing fantasies or sexual urges focused on sexual coercion
B. The person is distressed or impaired by these attractions, or has sought sexual stimulation from forcing sex on three or more nonconsenting persons on separate occasions
C. The diagnosis of Paraphilic Coercive Disorder is not made if the patient meets criteria for a diagnosis of Sexual Sadism Disorder

In their study on multiple paraphilic diagnoses among sex offenders, Abel et al. (1988) identified 21 paraphilias, including one they labeled "rape."

Kafka (1991) described a specific rapist who met the criteria for a PCD when he wrote about successfully treating him pharmacologically. He noted that this individual could belong to a larger group of such men who might meet the current proposed DSM-5 criteria for PCD (M. Kafka, personal communication, March 16, 2010).

Given these findings, and his own analysis, David Thornton, Ph.D., another Advisor to the Paraphilias subworkgroup, concluded: "There is significant empirical support for the existence of a distinctive coercive paraphilia among men convicted of rape. This paraphilia involves preferential sexual arousal to forcing sex upon a woman in a way that she obviously experiences as coercive….Although conceptually related to sexual sadism, it represents a distinct paraphilia" (Thornton, 2010, p. 416).

Two other advisors have, for differing reasons, recommended against a PCD diagnosis. Quinsey's (2010) critique is that since coercive sex can have evolutionary (Darwinian fitness) benefits, it should not be called a pathology. Knight (2010) is concerned that PCD lacks a defining taxometric. However, academic conjecture about the reproductive fitness of psychological traits in the theorized evolutionary environment is not an actual or practical basis for excluding a condition from the DSM, and being unable to define a condition with perfection should not prevent the scientific field from acknowledging that the condition exists.

The treatment of sex offenders presupposes the value of mental health intervention for men with a clinical diagnosis of a mental illness. Presumably, those who provide mental health treatment to individuals convicted of forced sexual assaults are providing an appropriate diagnosis when they submit written reports or bill insurance companies.

In the absence of a PCD diagnosis, this "distinct paraphilia" is now being documented as Paraphilia Not Otherwise Specified (rape) or Paraphilia Not Otherwise Specified (nonconsent) or some similar terminology. This status quo can be improved.

The authors of the DSM-IV-TR noted that their highest priority was to make the DSM "practical and useful for

clinicians by striving for brevity of criteria sets, clarity of language, and explicit statements of the constructs embodied in the diagnostic criteria" (American Psychiatric Association, 2000, p. xxiii). The role of the DSM in forensic settings was explicitly understood: "…when the presence of a mental disorder is the predicate for a subsequent legal determination (e.g., involuntary civil commitment), the use of an established system of diagnosis enhances the value and reliability of the determination" (American Psychiatric Association, 2000, p. xxxiii).

The addition of PCD enhances this quest for both clarity and reliability. A precise description of the diagnostic criteria for PCD would permit psychologists and psychiatrists to abandon the Paraphilia NOS designation in favor of the more defined PCD in appropriate cases. Describing a disorder with more clarity is a specific and laudable goal of the DSM. The adoption of PCD accomplishes that goal. To a lawyer interested in getting the right answer, the application of a more specific PCD model would be an advance and one which would protect potential SVPs from a more general diagnosis.

Since the DSM-5 Work Group on Sexual and Gender Identity Disorders announced its proposals, there has been some productive feedback and discussion about their proposals. Many comments have spawned intelligent, reasoned debate. Others have made broad and unhelpful attacks on the character of the committee. Those gratuitous comments are ignored here. I write out of concern that some of the critical attacks reveal political and/or ideological concerns not about the proposal per se, but what the authors perceive as the legal ramifications of those proposals. Those concerns are wholly misplaced. Not only should ideological concerns not play a part in a scientific debate, but the critics' predictions of how the PCD diagnosis would play out in a legal arena are wrong.

One critic recently claimed: "The [SVP] statutes are a well meaning effort to reduce the threat to public safety posed by those recidivist sexual offenders who have received prison sentences that are judged to be too short" (Frances, 2010a). This statement is as inaccurate as it is cavalier. SVP statutes were drafted in an effort to respond to very rare and difficult situations: What should society's response be when the best available information indicates that a previously convicted particular individual is very likely to violently sexually assault a stranger if given that opportunity? This involves the balancing of the individual's rights with the equal obligation of authorities to protect innocent members of the public from that danger.

SVP laws are not efforts to impose additional punishment on an individual because a past sentence was "too short." They are meant to protect innocent, vulnerable members of the community. That erroneous comment, however, perhaps best illustrates the ideological view of the commentator.

This same critic (Allen Frances, M.D.) disparaged the recommendation to include PCD in another article. Most troubling is he *began* that argument with this: "Paraphilic Coercive Disorder would expand the pool of sex offenders who are eligible for indefinite civil commitment…" (Frances, 2010b). Again, an ideological view on how a diagnosis may be used in the legal arena is a misguided critique of whether it is scientifically appropriate to include it in the DSM-5.

In the first-mentioned article, Frances showcased the heart of his argument: "Although the SVP statutes have twice passed Supreme Court tests, they rest on questionable constitutional grounds and may sometimes result in a misuse of psychiatry" (Frances, 2010a).

In law, the standard of whether a statute is on sufficient constitutional grounds is what the United States Supreme Court says about it. SVP laws have three times[2] been upheld as constitutional by the United States Supreme Court. Yet, the critic insists in his belief that these laws are wrong or of "questionable constitutional grounds." He maintains that view as part of his argument attacking the validity of the PCD diagnosis. This merely demonstrates that the concerns raised about the inclusion of the PCD diagnosis are ideological, or perhaps legal, but not scientifically based. The phenomenon of psychiatrists and psychologists engaging in legal analysis (and letting that influence their professional roles) is dangerous, in part because they are apt to do exactly what Frances did: Get the law wrong. The truth is that Dr. Frances is mistaken in his concern that PCD might "expand the pool of sex offenders who are eligible for indefinite civil commitment." PCD will not expand the pool; it will shrink it.

Currently, the men who would fall within the PCD diagnosis are not being ignored by the SVP laws or by people like me. They are still being caught in that "pool." They are falling into the Paraphilia Not Otherwise Specified diagnosis with a (nonconsent) or (rape) descriptor.

Ironically, others have argued that the Paraphilia Not Otherwise Specified diagnosis is unfair to use in SVP cases because it is too vague or unreliable. These arguments were rejected by Packard and Levenson (2006), who demonstrated that psychologists' ability to reliably diagnose paraphilias, including Paraphilia Not Otherwise Specified, is no different than the ability of mental health professionals to diagnose psychiatric conditions in general.

Nevertheless, making the Paraphilia Not Otherwise Specified diagnosis more specific would respond to those criticisms. A PCD diagnosis would theoretically replace the Paraphilia Not Otherwise Specified diagnosis in many circumstances, and because it would be more specific, it would be apt to capture fewer men. Thus, the PCD diagnosis would *shrink* the pool of potential SVP candidates.

[2] A federal civil commitment statute was upheld by the United States Supreme Court on May 17, 2010. United States v. Comstock, 560 U.S. ___, (2010), 2010 WL 1946729 (US) (case 08-1224).

Which is precisely what I as a career prosecutor want. The civil commitment of sexually violent predators should involve the identification of only those who properly qualify under the law. The narrower the definition of those eligible for civil commitment, the better. PCD is a more precise—i.e., narrower—definition under which to consider those men who derive sexual pleasure by the use of force or coercion on their subjects. It is a more precise definition than Paraphilia Not Otherwise Specified. More precise means it is more narrow (shrinking the pool). It also means potentially making a diagnosis more accurate.[3] The result would be to increase the ability to properly identify those who should be considered as sexually violent predators.

My goal as a lawyer is to get the right answer. It doesn't take a taxonomic analysis to know that there are men in the world who seek sexual gratification by coercing others to engage in unwanted sexual behavior. It doesn't require years of study to document that there are men whose urges, fantasies, and behaviors satisfy the PCD descriptors. It is abundantly evident those people exist. A short period of time practicing in this field gives a disinterested person that knowledge.

The APA faces a moment of intellectual challenge. For more than a decade, the opinion of the APA about SVP laws has been unambiguous and unequivocal. A Task Force Report of the APA entitled *Dangerous Sex Offenders* (American Psychiatric Association, 1999) ended with these declarations:

"In the opinion of the Task Force, sexual predator commitment laws represent a serious assault on the integrity of psychiatry….In the opinion of the Task Force, psychiatry must vigorously oppose these statutes in order to preserve the moral authority of the profession and to ensure continuing societal confidence in the medical model of civil commitment" (American Psychiatric Association, 1999, p. 173).

To emphasize its position and its commitment to "vigorously oppose these statutes" the APA filed amicus briefs in two separate United States Supreme Court cases, urging the SVP statutes be stricken as unconstitutional. In the first case to reach the U.S. Supreme Court (*Kansas v. Hendricks*), the APA filed a friend-of-the-court brief arguing the SVP laws would be a form of "double jeopardy," and expressing concern that "only the prosecutor, representing the State, can bring the action, which then has all the outward appearance of a criminal trial" (American Psychiatric Association, 1996, pp. 7, 13).

The APA argued again in *Kansas v. Crane*: "What the State should not be permitted to do under the Constitution is to evade the criminal justice system, when dissatisfied with decisions it has made, by setting up an alternative regime broadly authorizing the indefinite locking up of individuals based on risk of future offenses—outside the *parens-patriae* based tradition of medically justified civil commitment for those suffering severe mental illness and in disregard for the tight limits on permitted preventive detention" (American Psychiatric Association, 2001).

Each time the APA position was rejected and the statutes at hand were upheld.

Now the APA faces the challenge of setting aside its ideological views and adopting the scientific proposals of its committee. One commentator has been unable to separate ideology from science. I write in the hope and anticipation that the APA will rise to the challenge.

Ironically, by accepting the DSM-5 committee's recommendation, they can adopt the best science and simultaneously limit the reach of the SVP laws.

Acknowledgments The author is an Advisor to the DSM-V Paraphilias subworkgroup of the Sexual and Gender Identity Disorders Workgroup (Chair, Kenneth J. Zucker, Ph.D.). The opinions expressed in this article are the thoughts and opinions of the author only, and do not necessarily represent the opinions or reflect any policies of the Snohomish County Prosecutor's Office. Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders V Workgroup Reports* (Copyright 2010), American Psychiatric Association.

References

Abel, G. G., Becker, J. V., Cunningham-Rathner, J., Mittelman, M., & Rouleau, J. (1988). Multiple paraphilic diagnosis among sex offenders. *Bulletin of the American Academy of Psychiatry and the Law, 16*, 153–168.

American Psychiatric Association. (1996). *Amicus Curiae Brief of the APA in support of Leroy Hendricks*, Kansas v. Hendricks, 95-1649, 95-9075.

American Psychiatric Association. (1999). *Dangerous sex offenders. A Task Force Report of the American Psychiatric Association*. Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

American Psychiatric Association. (2001). *Amicus Curiae Brief of the APA in support of Michael Crane,* Kansas v. Crane, 00-957.

Frances, A. (2010a, March 14). DSM5 sexual disorders make no sense. *Psychology Today*. http://www.psychologytoday.com/node/39514. Retrieved April 28, 2010.

Frances, A. (2010b, February 11). Opening Pandora's box: The 19 worst suggestions for DSM5. *Psychiatric Times*. http://www.psychiatrictimes.com/display/article/10168/1522341. Retrieved April 29, 2010.

Freund, K., Scher, H., Racansky, I. G., Campbell, K., & Heasman, G. (1986). Males disposed to commit rape. *Archives of Sexual Behavior, 15*, 23–35.

Freund, K., Seeley, H. R., Marshall, W. E., & Glinfort, E. K. (1972). Sexual offenders needing special assessment and/or therapy. *Canadian Journal of Criminology and Corrections, 14*, 3–23.

Kafka, M. (1991). Successful treatment of Paraphilic Coercive Disorder (a rapist) with fluoxetine hydrochloride. *British Journal of Psychiatry, 158*, 844–847.

Knight, R. A. (2010). Is a diagnostic category for Paraphilic Coercive Disorder defensible? *Archives of Sexual Behavior, 39*, 419–426.

[3] A second criticism raised of the PCD diagnosis is that it may be based exclusively on "behaviors." That critique is odd. A pattern of behavior is used as the basis for a variety of diagnoses in those who have committed criminal behavior, such as Antisocial Personality Disorder (301.7) and Conduct Disorder (312.8x) (American Psychiatric Association, 2000).

Milloy, C. (2003). *Six-year follow-up of released sex offenders recommended for commitment under Washington's Sexually Violent Predator Law, where no petition was filed*. Washington State Institute for Public Policy.

Packard, R. L., & Levenson, J. S. (2006). Revisiting the reliability of diagnostic decisions in sex offender civil commitment. *Sexual Offender Treatment, 1*, 1–15.

Quinsey, V. L. (2010). Coercive paraphilic disorder. *Archives of Sexual Behavior, 39*, 405–410.

Thornton, D. (2010). Evidence regarding the need for a diagnostic category for a coercive paraphilia. *Archives of Sexual Behavior, 39*, 411–418.

Arch Sex Behav (2010) 39:1449–1452
DOI 10.1007/s10508-009-9506-6

CLINICAL CASE REPORT SERIES

Effect of Sildenafil on Penile Plethysmography Responding: A Pilot Investigation

Nathan J. Kolla · Ray Blanchard · Philip E. Klassen · Michael E. Kuban · Thomas Blak

Received: 13 January 2009 / Revised: 29 April 2009 / Accepted: 1 May 2009 / Published online: 27 May 2009

Abstract Despite the utility of penile plethysmography in helping diagnose sexual offenders, some men respond minimally to the test stimuli. This pilot investigation examined the effect of sildenafil on phallometric responding in a non-forensic sample ($n = 12$) of middle-aged males. Participants underwent phallometric testing on two separate occasions and were administered 50 mg of sildenafil prior to one of these sessions. Study data indicated that pre-treatment with sildenafil produced a significant mean percentage increase (28%) in peak response compared with the untreated condition. The implication of this result is discussed in view of the selected sample.

Keywords Penile plethysmography · Sildenafil · Sexual offenders

Introduction

The phallometric test (penile plethysmography) is a psychophysiological tool used to assess the erotic age and gender preferences of adult males. In volumetric phallometry, the penile blood volumes of patients are measured in a laboratory setting while they are exposed to a standardized set of sexually themed stimuli depicting adults and children. Increases in penile blood volume correspond to the patients' relative attraction to the different categories of persons. Several reviews addressing the use of phallometric testing in clinical diagnosis exist (Harris & Rice, 1996; Lalumière & Harris, 1998; Launay, 1999). About 10% of males undergoing phallometric testing show no clinically meaningful response to any of the gender or age categories (Blanchard, Klassen, Dickey, Kuban, & Blak, 2001). Lack of response in these individuals may be due to a constellation of factors, including neurological, vascular, or hormonal abnormalities. Sildenafil (Viagra) has been shown to increase the potential for erection in sexually stimulated males through indirect relaxation of penile erectile tissue (Seftel, 2005). We therefore hypothesized that administration of sildenafil to participants prior to phallometric testing would produce quantitative increases in penile blood flow measured during phallometry compared with testing without sildenafil.

N. J. Kolla (✉) · R. Blanchard · P. E. Klassen · M. E. Kuban · T. Blak
Law and Mental Health Program, Centre for Addiction and Mental Health, 250 College St., Toronto, ON M5T 1R8, Canada
e-mail: nathan.kolla@utoronto.ca

Method

Participants

This study was approved by our institution's research ethics board. Men between the ages of 40 and 65 years who responded to recruitment advertisements posted in the hospital were eligible to participate. The recruitment poster indicated that the study investigators were investigating male patterns of sexual arousal in response to a single dose of sildenafil. Potential male participants between 40 and 65 were requested to contact the lead investigator (NK) by telephone or email for further information. When potential participants contacted the lead investigator, they were given more detailed information on the study protocol, including a description of the phallometric method. The pre-screening procedure was limited to asking potential participants whether they currently lived with children, which was an exclusion criterion of the study. One individual was excluded from participation on this basis.

Candidates were excluded from participation on the following bases: (1) non-fluent in English; (2) presence of psychotic disorder; (3) legal blindness or deafness; (4) IQ < 70; (5) serious cardiac, hepatic, ophthalmic, or renal illness; (6) history of priapism; (7) current use of sildenafil or similar medications; (8) allergy to sildenafil citrate or similar medications; (9) current use of nitrates or other medications known to interfere with the metabolism of sildenafil citrate; and (10) current residence in a household with children. We imposed the last criterion to preclude any obligation on the part of the study investigators to contact authorities should participants' self-reports and phallometric data suggest they present a danger to children. The presence of psychotic disorders was ruled out clinically during the history. Diagnoses of mental retardation (e.g., IQ < 70) were similarly excluded on clinical grounds. Individuals with intellectual disabilities manifest difficulties in communication, display impaired social skills, and present with compromised daily living skills. As the overwhelming majority of developmentally delayed adults either live with or are supervised by family members (Seltzer, Begun, Seltzer, & Krauss, 1991), it is unlikely that these individuals would present for assessment alone. All participants received a complete history and physical by the lead investigator in addition to routine bloodwork, urinalysis, and electrocardiogram. Candidates with potentially serious medical disease revealed during the history were excluded from participation if they had grossly abnormal physical examination findings or laboratory results.

Twenty-one participants began the study. Four individuals were disqualified based on study exclusion criteria: two had poorly controlled hepatitis C infection, one was recovering from a major cardiovascular surgery, and one participant was already taking a pro-erectile medication related to sildenafil. Five participants were lost to follow-up leaving a final sample size of 12. Among those who completed the study, four participants had received previous inpatient and outpatient treatment for alcohol abuse or dependence; two individuals admitted to continued heavy use of the substance. Three participants suffered from mood disorders with two taking antidepressant medication. One individual had well-controlled hepatitis C with near normal liver enzymes, while another had a previous gastrointestinal cancer.

Measures

Each participant served as his own control for the study. Participants were administered 50 mg of sildenafil 1 h prior to either the first or second phallometric testing sessions; the order of drug and non-drug conditions was balanced among participants. Testing sessions were separated by at least 48 h to ensure that drug metabolites were sufficiently cleared from the participants' systems.

The test stimuli were audiotaped narratives presented through headphones and accompanied by slides of nude models. There were seven categories of narratives, which described sexual interactions with prepubescent girls, pubescent girls, adult women, prepubescent boys, pubescent boys, and adult men as well as solitary, nonsexual activities or "neutral" stimuli. All narratives were written in the second person and present tense and were approximately 100 words long. Readers are referred to Blanchard et al. (2001) for specific details on procedures used in this study for the quantification of phallometric response.

Results

Twelve males participated in this open label, pilot investigation. The mean age of the sample was 50.5 $\pm$ 5.5 years (range 42–58). There were 11 self-identified heterosexuals and one homosexual. All participants denied a history of sexual offending and claimed a sexual preference for the mature adult physique. Participants' penile responses to each stimulus (e.g., picture of nude model) were quantified as the apogee of the curve of blood volume change in cubic centimetres occurring during each exposure. The mean of the participants' three highest cc scores for all test stimuli was calculated as the output index (Freund, 1967). The overall mean output index for all participants in the untreated condition was 5.8 $\pm$ 4.2 cc, while the overall mean output index for participants in the drug condition was 6.7 $\pm$ 4.7 cc. Individual results with and without sildenafil treatment are presented in Table 1. To control for between-subjects variability in output index, we expressed treatment effect as the percentage

Table 1 Mean output indices for participants in sildenafil and untreated conditions

Subject	OI sildenafil	OI untreated	Percent change[a]
1	6.31	5.86	7.68
2	6.26	6.75	−7.26
3	6.23	3.29	89.36
4	2.76	2.37	16.46
5	18.38	9.00	104.22
6	2.69	2.70	−0.37
7	8.34	10.37	−19.58
8	12.82	15.83	−19.01
9	5.07	3.14	61.46
10	6.41	5.76	11.28
11	3.73	2.81	32.74
12	1.86	1.20	55.00
M	6.74	5.76	27.67
SD	4.71	4.24	41.33

[a] (Sildenafil − untreated)/untreated × 100

increase in response in the sildenafil condition relative to the untreated condition. Comparing the mean percentage increase against an expected value of zero using a one-sample t test, we found a 28% increase in responding with sildenafil, $t(11) = 2.32, p = .041$, two-tailed. There was only a 4% increase in responding from the first to second testing sessions, which was not significant, $t(11) < 1$.

Discussion

The results of this preliminary investigation provide evidence that low dose sildenafil has the ability to increase penile blood flow in males during phallometric testing. As forensic patients referred for assessment before or after sentencing comprise the majority of individuals presenting for phallometric testing, poor responding becomes problematic, especially when such data are required for sentencing or treatment recommendations. Administering sildenafil to patients prior to testing could potentially ameliorate inadequate or non-response. Our sample contained older males with high levels of medical comorbidity and substance misuse, which is relevant because sildenafil appears to be less effective in treating older versus younger men with erectile dysfunction (ED) due to various physical etiologies (Monga, Bettencourt, & Barrett-Connor, 2002). Researchers have also observed that sildenafil can increase erectile activity in young, healthy, potent males who would be more closely matched to patients presenting for phallometric testing (Salonia et al., 2005). Our sample was, therefore, biased against detecting a positive effect of sildenafil.

Very few placebo-controlled studies of sildenafil have examined objective measures of the drug's effectiveness. For example, the majority of studies investigating the clinical efficacy of sildenafil in ED have used standardized, self-reported measures of erectile function as study endpoints (Carson, 2003). One randomized, double-blind, placebo-controlled crossover study of sildenafil in men with ED used circumferential penile plethysmography to test whether sildenafil could increase the mean duration of erections with $\geq$60% rigidity (Gingell, Sultana, Wulff, & Gepi-Attee, 2004). Penile erectile activity was measured at several time points after administration of sildenafil; participants selected from a variety of sexually explicit films to view while their erections were being monitored. At 1 h post-administration of sildenafil, the mean duration of erections in the drug condition was 26 min compared with 3 min in the placebo condition ($p \leq .0001$). Although these investigators tested a higher dose of sildenafil (100 mg), a dose-dependent relationship has been observed for sildenafil and penile rigidity in men with ED: another placebo-controlled cross-over study which employed circumferential phallometry found that the mean duration of rigidity >80% at the base and tip of the penis increased in incremental fashion to increasing doses of sildenafil (10, 25, and 50 mg). All results were significantly greater than those obtained under placebo (Boolell, Gepi-Attee, Gingell, & Allen, 1996). These findings suggest that placebo responding is minimal in studies of sildenafil where relatively objective techniques are used to measure penile erectile activity. However, as the samples in the foregoing studies were confined to men with ED, these results may not be entirely generalizable to our study which included only one participant with ED. To the best of our knowledge, there have been no phallometric investigations that have studied the effect of sildenafil in heterogenous, non-clinical samples.

There are two main scenarios where an intervention like sildenafil that could potentially convert phallometric non-responders to responders would be most beneficial. The first would be in situations where men who were low responders at baseline were falsely labeled as being pedophiles. Administering sildenafil to these individuals prior to phallometric testing could increase their responding to a level where their sexual preference for the mature physique became readily apparent. The second scenario would be insufficient phallometric responding in men who are true pedophiles. In this instance, sildenafil pre-treatment might improve the phallometric test's sensitivity for detecting pedophilia. While the theoretical possibility of sildenafil increasing the rate of false positives exists, it must be emphasized that phallometric testing is only one component in the global assessment of sexual offenders that includes a thorough clinical assessment and detailed sexual history. Without compelling evidence to make a clinical diagnosis of pedophilia, a phallometric diagnosis of pedophilia that was obtained with or without sildenafil would be easily discounted.

While our study suffered from a small sample size and an open design, the results are very encouraging given that they likely constitute an underestimation of the drug's potential effect. These initial findings suggest that administering sildenafil to patients who are unlikely to respond (e.g., older males) could be a useful intervention in phallometric testing. We anticipate conducting future randomized and placebo-controlled trials of sildenafil in phallometric testing with larger samples.

References

Blanchard, R., Klassen, P., Dickey, R., Kuban, M. E., & Blak, T. (2001). Sensitivity and specificity of the phallometric test for pedophilia in nonadmitting sex offenders. *Psychological Assessment, 13*, 118–126.

Boolell, M., Gepi-Attee, S., Gingell, J. C., & Allen, M. J. (1996). Sildenafil, a novel effective oral therapy for male erectile dysfunction. *British Journal of Urology, 78*, 257–261.

Carson, C. C. (2003). Sildenafil: A 4-year update in the treatment of 20 million erectile dysfunction patients. *Current Urology Reports, 4*, 488–496.

Freund, K. (1967). Diagnosing homo- or heterosexuality and erotic age-preference by means of a psychophysiological test. *Behaviour Research and Therapy, 5*, 209–228.

Gingell, C., Sultana, S. R., Wulff, M. B., & Gepi-Attee, S. (2004). Duration of action of sildenafil citrate in men with erectile dysfunction. *Journal of Sexual Medicine, 1*, 179–184.

Harris, G. T., & Rice, M. E. (1996). The science in phallometric measurement of male sexual interest. *Current Directions in Psychological Science, 5*, 156–160.

Lalumière, M. L., & Harris, G. T. (1998). Common questions regarding the use of phallometric testing with sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 28*, 227–237.

Launay, G. (1999). The phallometric assessment of sex offenders: An update. *Criminal Behavior and Mental Health, 9*, 254–274.

Monga, M., Bettencourt, R., & Barrett-Connor, E. (2002). Community-based study of erectile dysfunction and sildenafil use: The Rancho Bernardo study. *Urology, 59*, 753–757.

Salonia, A., Barbieri, L., Chen, J., Briganti, A., Dehò, F., Naspro, R., et al. (2005). Comparison of the effect of sildenafil and apomorphine SL on nocturnal erections in healthy volunteers: A placebo-controlled study. *European Urology, 47*, 524–529.

Seftel, A. D. (2005). Phosphodiesterase type 5 inhibitors: Molecular pharmacology and interactions with other phosphodiesterases. *Current Pharmaceutical Design, 11*, 4047–4058.

Seltzer, G. B., Begun, A., Seltzer, M. M., & Krauss, M. W. (1991). Adults with mental retardation and their aging mothers: Impacts of siblings. *Family Relations, 40*, 310–317.

Arch Sex Behav (2010) 39:1453–1455
DOI 10.1007/s10508-009-9578-3

COMMENT

Café Discussions on Oral Sex, Oral Cancer, and HPV Infection: Summative Report

Mario Augusto Brondani

Received: 10 February 2009 / Revised: 22 July 2009 / Accepted: 26 July 2009 / Published online: 7 January 2010

Abstract Recent emphasis has been placed on the potential links between oral sex, HPV infection, and oral cancer development. Such links were addressed by researchers, clinicians, and the community during two Café Scientifique discussions in October and November 2008, in Vancouver, Canada. The Cafes gathered panels of experts on oral pathology, dentistry, oncology, social work, and community-based research who interacted with an audience of policy makers, health care administrators, sociologists, sexologists, pharmacists, clinical and social researchers, social workers, technicians, and graduate, undergraduate, and high school students. This commentary summarizes the main points discussed during these two events to encourage a worldwide open dialogue about potential risks for oral cancer beyond tobacco smoking and excessive alcohol consumption as such malignancies have high mortality and morbidity, but are yet preventable diseases.

Keywords Oral sex · Oral cancer · Human papillomavirus

Introduction

Worldwide, more than half of the 650,000 patients diagnosed with head and neck cancer each year will die within 12 months (Ferlay, Bray, Pisani, & Parkin, 2004). Oral cancer significantly impacts quality of life, physically and psychosocially, and the genetic links between some types of oral-pharyngeal carcinomas and ano-genital malignancies have been established via the human papillomavirus (HPV), particularly the strains 16 and 18 as found by Herrero et al. (2003) and others (Anhang, Goodman, & Goldie, 2004; Cox, 2000; D'Souza et al., 2007; Ha & Califano, 2004). Since HPV 16 is the most predominant strain in ano-genital carcinomas, its transmission to the oropharyngeal region can occur through the mouth, especially during oral sex in sexually active individuals (Scully, 2005). Such assertive supposition, however, does not exist without controversy among both scientific and lay communities (Brondani, 2008, 2009).

During the Spring of 2008, Vancouver hosted two Café Scientifique discussions sponsored by the Canadian Institutes of Health Research. The Cafes gathered highly experienced presenters with backgrounds in oral pathology, dentistry, oncology, social work, and community-based research who shared their ideas with the public on October 31[1] and on November 20, 2008. The discussions critically highlighted four main areas related to the triad HPV-Oral Sex-Oral Cancer:

1. *Oral malignancies*: Incidence, morbidity, risk factors, individuals at higher risk, and screening programs were highlighted. Within this area, the following topics were discussed: transmission of HPV during oral sex and its potential association to oral cancer, and promoting dialogue between dental professionals and patients about oral sex as a risk factor for oral cancer were stressed.
2. *Heterosexism*: The potential reasons for health care providers not discussing sexuality/sexual practices and oral sex with their patients were discussed.
3. *Belief systems*: Current understanding of the triad HPV-oral sex-oral cancer from a community perspective was explored through the results from a provocative pilot

M. A. Brondani (✉)
Department of Community Dentistry and Prosthodontics and Dental Geriatrics, Faculty of Dentistry, University of British Columbia, 2199 Wesbrook Mall, Vancouver, BC V6T 1Z3, Canada
e-mail: brondani@interchange.ubc.ca

[1] This Café occurred as a parallel event within the 4th Gay Men's Health Summit in Vancouver, which is organized by the Community Based Research Centre (http://summit.cbrc.net/).

study on the values and beliefs men and women hold in regards to this triad.

4. *Community-based research*: The role of community based approaches in understanding sexual behavior. This area led to the discussion of oral sexual behavior in regards to HPV transmission and oral cancer development.

Café Scientifique

The two cafes began with a summary of some of the available scientific information on oral cancer, risk factors, and screening, on perceived risks for oral cancer development, as well as on the potential aspects hindering discussions about the triad HPV-oral sex-oral cancer during regular medical and dental appointments. After that, presenters promoted a discussion to engage the audience on the issues above. Each Café lasted for an average of 90 minutes and an assistant took field notes compiling the main statements and ideas debated and shared.

The goal of the Cafes was to bring together researchers, clinicians, and the public into a free conversation. In total, one oral pathologist, two general dentists, one oncologist, two social workers, and one community-based researcher participated as speakers to an audience of 67 participants in the two Cafes. Attendees included social workers, policy makers, health care administrators, sociologists, sexologists, pharmacists, technicians, clinical and social researchers, graduate students, undergraduate students, and high school students. This comment presents the main points addressed during the two Cafés Scientifique, from the perspectives of the presenters and from the perspectives of the audience. It also stressed the need in fostering dialogue about the potential role of oral HPV

Table 1 Amalgamating scientific and lay knowledge in regards to oral sex, HPV infection, and oral cancer development

HPV is the most frequent and common sexually transmitted infections worldwide	HPV is transmitted by skin and mucosal contact
There are more than 120 different strains of HPV (200 according to some studies)	One of the common clinical manifestations of HPV is a skin wart caused by HPV 1, 2, 3, and 4. Warts are skin growths, not cancers
It is estimated that about 75% of all sexually active men and women are infected with HPV. The majority will remain asymptomatic, but infectious	Most of HPV infections will resolve spontaneously, but it is unknown if the virus is eliminated completely or just suppressed
More than 40 types of HPV can be isolated in the mouth and genital tracts concomitantly (e.g., they are the same strain)	Genital warts are caused by HPV 6 and 11, different from those causing skin warts. Genital warts are not considered cancers
HPV 16, 18, 31, 35, and 45 usually do not cause visible warts in genital areas, but have been associated with pre-cancerous lesions (dysplasia) in the cervix and anus mucosa	95% of cervical cancers are associated with HPV 16 and 18. In British Columbia, 6 cases of anal cancer in men are diagnosed every year
Oral sex appears to be a potential venue for HPV transmission even though it is difficult to prove	Oral sex is defined as an intimate contact of teeth, gums, lips, and tongue with genital (vagina, scrotum, penis), groin, and anal area
Some oral cancers (oropharyngeal area) have been associated with HPV infection	HPV 16 and 18 are risk factors for the development of oral cancer, the same strain associated with ano-genital malignancies
Between 12,000 and 15,000 cases of oral cancer are diagnosed every year in the U.S. About 3,000 in Canada	Almost half of the individuals diagnosed with oral cancer will die in less than a year
There is no single cause for oral cancer. It is a multifactorial disease associated with risk factors, not caused by these factors	People infected with HPV are 30 times more likely to develop oral cancer than those who are not. Those who smoke tobacco are 3 times more likely and those who drink alcohol 2.5 times more likely to develop oral cancer than those who do not smoke or drink, respectively
High risk individuals include those with more than 5 different oral sexual partners during a life-time and earlier oral sexual experience	High frequency of oral sex, about 3–5 times a week in the past 30 days, increases 9 times the risk for developing oral cancer
Generalist health care providers focus mostly on tobacco and alcohol when talking about risk factors for oral cancer	In average, less than 5% of dentists ask about oral sexual practices during routine medical and dental history taken. The % among physicians is not much higher
Patients might not feel comfortable in discussing oral sex practices with a dentist	Patients might feel that oral sex practices are meant to be discussed in a dental office
Gardacil (MERK laboratories) and Cervarix (GSK laboratories) are used to PREVENT HPV 6, 11, 16, and 18 infection in girls, before the first sexual experience. Most of women at high risk for HPV infection do not develop cervical cancer	Gardacil and Cervarix have not been tested against oral HPV. Such vaccines have been advocated to be used in boys, before their first sexual experience

Note: The facts listed in this table might or might not reflect the current knowledge on the topic as it combines both expert opinion and lay understandings. It does not necessarily concur with the published literature

infection to the health of individuals of all sexual orientations. Table 1 outlines the main issues highlighted by the presenters and the attendees.

From Table 1, it remains clear the need for encouraging discussion between health professionals and the public about issues related to oral sex practices and HPV transmission. As the author emphasized, "the mouth and its related structures are subject to harm from smoking, high sugar intake, and lack of proper brushing and flossing, [but yet] we only focus our inquiries on tobacco, dietary history, and oral hygiene practices [and] avoid asking about oral sex practices and all of the potential pathogens associated with it" (Brondani, 2008).

In 2008, two websites of public access in Canada—the Canadian Cancer Society (2008) and the British Columbia Cancer Agency (2008)—finally acknowledged HPV infection as a potential risk factor for oral cancer development. Such acknowledgment followed the American Cancer Society (2008) statement about HPV being a potential risk factor for the development of almost 30% of oral and oropharyngeal cancers. In this regard, oral sex can be a venue for HPV transmission and should be discussed accordingly since fewer than one-third of the sexually active individuals infected by HPV actually heard about this virus (Anhang et al., 2004). As shared by presenters and attendees of the Cafes, and by worldwide researchers (Scully, 2005), it is time to open our eyes about other potential risks for oral cancer beyond tobacco smoking and excessive alcohol consumption. Oral cancer has high mortality and morbidity, but still is a preventable disease as emphasised by Syrjänen (2007).

From Here and Beyond: Next Steps

The two Cafes offered the opportunity for discussing scientific evidence in the realm of lay public understanding. For the Cafes to make a difference outside their realm, however, the ideas and points brought up for discussion need to be disseminated widely and support further discussions, research, and knowledge transfer worldwide. The author recently published an editorial (Brondani, 2008) from which the idea for these Cafes emerged. Although this above mentioned editorial has been used in studies and researches worldwide, much more needs to be accomplished. This comment emerges as another venue for dissemination, but it is up to the professionals and the public to determine and take the next steps to nurture open and friendly discussions of the potential implication of oral HPV infection and vaccine to sexually active individuals. Readers with any questions, comments or ideas are encouraged to contact the author.

Acknowledgments The author would like to acknowledge all the attendees who engaged into lively discussions during these two events. The author is also thankful to Drs. Allan Hovan, Catherine Poh, Brian O'Neill, and Francisco Ibanez Carrasco for their enthusiastic participation as presenters. Special thanks goes to Cheryle Colombe as an engaging volunteer. Support for the Cafes came from the Canadian Institutes of Health Research funding collaboration agreement # CA 2008-032.

References

American Cancer Society. (2008). What are the risk factors for oral cavity and oropharyngeal cancers? Retrieved January 29, 2008, from http://www.cancer.org/docroot/CRI/content/CRI_2_4_2X_What_are_the_risk_factors_for_oral_ cavity_and_oropharyngeal_cancer_60.asp?sitearea/CRI.

Anhang, R., Goodman, A., & Goldie, S. J. (2004). HPV communications: Review of existing research and recommendations for patient education. *Cancer Journal for Clinicians, 54*, 248–259.

British Columbia Cancer Agency. (2008). *Care & research*. Retrieved July 12, 2008, from http://www.bccancer.bc.ca/PPI/Screening/oral/questionsanswers.htm.

Brondani, M. A. (2008). HPV, oral sex, and the risk of oral cancer: Food for thought [Editorial]. *Special Care in Dentistry, 28*, 183–184.

Brondani, M. A. (2009). Oral sex and oral cancer: Does pleasure need caution? *Journal of the Canadian Dental Association, 75*, 249.

Canadian Cancer Society. (2008). Retrieved August 19, 2008, from http://www.cancer.ca/ccs/internet/standard/0,3182,3172_10175_272695_langId-en,00.html.

Cox, S. (2000). Oral cancer in Australia: Risk factors and disease distribution. *Annuals of the Royal Australian College of Dental Surgeons, 15*, 261–263.

D'Souza, G., Kreimer, A. R., Viscidi, R., Pawlita, M., Fakhry, C., Koch, W. M., et al. (2007). Case–control study of human papillomavirus and oropharyngeal cancer. *New England Journal of Medicine, 356*, 1944–1956.

Ferlay, J., Bray, F., Pisani, P., & Parkin, D. M. (2004). *Globocan 2002: Cancer incidence and mortality worldwide. IARC CancerBase 5, version 2.0*. Lyon, France: IARC.

Ha, P. K., & Califano, J. A. (2004). The role of human papillomavirus in oral carcinogenesis. *Critical Reviews in Oral Biology and Medicine, 15*, 188–196.

Herrero, R., Castellsagué, X., Pawlita, M., Lissowska, J., Kee, F., Balaram, P., et al. (2003). Human papillomavirus and oral cancer: The international agency for research on cancer multicenter study. *Journal of the National Cancer Institute, 95*, 1772–1783.

Scully, C. (2005). Oral cancer: The evidence for sexual transmission. *British Dental Journal, 199*, 203–207.

Syrjänen, S. (2007). Human papillomaviruses in head and neck carcinomas. *New England Journal of Medicine, 356*, 1993–1996.

Arch Sex Behav (2010) 39:1457–1465
DOI 10.1007/s10508-010-9665-5

INVITED ESSAY

Robert Stoller's *Sex and Gender*: 40 Years On

Richard Green

Published online: 12 August 2010

Abstract By its title, *Sex and Gender* announced a conceptual breakthrough in distinguishing basic elements of human experience. In Robert Stoller's first book, patients illustrating this divergence were lucidly presented. Transvestites and the newly publicized transsexuals were two examples. Clinical and dynamic distinctions between the two formed a basis for Stoller's criteria for patient selection for "sex change." They remain current. The complex identity of the intersexed was described with sensitivity and insight. It, too, remains timely. An innovative description of the genesis of boyhood transsexualism was presented in considerable detail. This finding is less commonly reported today but is also not looked for. Stoller was sympathetic to the request for sex change. He credited a biological contribution to the development of masculinity and femininity. Both stances were remarkable for a psychoanalyst. Robert Stoller introduced the term "gender identity." It is now our vocabulary when we articulate this bedrock of personhood.

Keywords Gender identity · Transsexualism · Transvestism · Intersexuality · Psychoanalysis

Introduction

The significance of the title of Robert Stoller's first book was announced in the Preface:

> …we have split off "gender" as a distinguishable part of "sexuality." The word sex in this work will refer to the male and female sex and the component biological parts that determine whether one is male or female. [There remain] tremendous areas of behavior, feelings, thoughts, and fantasies that are related to the sexes and yet do not have primarily biological connotations. It is for some of these psychological phenomena that the term gender will be used…Thus, while sex and gender seem to common sense to be practically synonymous, and in everyday life to be inextricably bound together, one purpose of this study will be to confirm the fact that…sex and gender are not inevitably bound…each may go in its quite independent way. (Stoller, 1968, pp. vi–vii)

And, introduction of a seminal term: "While the work of our research team has been associated with the term *gender identity*, we are not militantly fixed either on copyrighting the term or defending the concept as one of the splendours of the scientific world" (Stoller, 1968, p. vi).

Gender Identity

I do not know exactly when Stoller began using the term. I am confident that it was the name of the program when I arrived at UCLA to begin my psychiatry training in 1962. The term was formally introduced to the psychiatric/psychoanalytic worlds in 1964 when he published "A Contribution to the Study of Gender Identity" (Stoller, 1964). What is this thing called "gender identity"?

> [It] starts with the knowledge and awareness, whether conscious or unconscious, that one belongs to one sex and not the other, though as one develops, gender identity becomes much more complicated so that…one may sense himself as not only a male but a masculine man… who fantasizes being a woman. (Stoller, 1968, p. 10)

R. Green (✉)
Psychological Medicine, Imperial College,
St. Dunstan's Road, London W6 8RP, UK
e-mail: richard.green@ic.ac.uk

Stoller (1968) wrote of the

> …essentially unalterable core of gender identity (e.g., I am male) to be distinguished from the related but different belief, *I am manly* (or masculine)….The sense of core gender identity…is derived from three sources: the anatomy and physiology of the genitalia; the attitudes of parents, siblings and peers toward the child's gender role; and a biological force that may more or less modify the attitudinal (environmental) forces. (p. 40)

His rationale for clinical study: "It is not easy to study the relative importance of each of these factors in normals because one factor cannot be dissected from the others. However, certain rare patients provide such an opportunity…" (Stoller, 1968, p. 40).

Let the games begin.

Transvestism

Transvestism was more the perversion in 1968 than today. It was less the socially accepted behavior and in some jurisdictions was a crime. In consequence of the societal stigma, more transvestites were troubled then to bring their cross-dressing issues to a therapist. Today, in the UK, we have celebrity transvestites who are not stage performers but, for example, include an award winning sculptor and the recent male partner of a female TV celebrity and best-selling author. (The guy is also a cage fighter. This is a brutal sport fought by two men in close proximity. No sissy, this chap.)

In addition to more societal acceptance of cross-dressing, diminishing clinical contact resulted from the dismal rate of success of treatments. Aversion therapists had their day pairing electric shock or emesis with cross-dressing. That had only some short-term behavioral change. Aversion therapy has become a bete noir. Stoller commented on it in 1968: "…who defines what is anti-social and how much does this behavior endanger society or its individuals. How much pain should be inflicted on a patient to make him conform? For the homicidal, a great deal. How much for the transvestite?" (p. 243).

Psychoanalysis has not fared much better, although it is often difficult to fathom the results of therapy. For example, in 1997, the psychoanalytic panel on transvestism and transsexualism in Barcelona reported on three cases. All contained fascinating insights. But no data one way or the other on behavioral change (Chiland, 1988).

How does transvestism come about? Why do some males experience sexual arousal to dressing in womens' clothes? For Stoller, the essence was early trauma and its threat to masculinity:

> Transvestism is…a defensive structure raised to protect a threatened but desired sense of masculinity and maleness, and…to preserve a badly threatened potency. One should not be fooled by the apparent paradox that he does this via the detour of dressing like a woman. (Stoller, 1968, p. 180)

So, sexual arousal (erection), accompanying cross-dressing, bolsters damaged masculinity. For Stoller, the prototypic instance of the scarring humiliation early in life for the young boy is being dressed as a girl by a female.

How often is this the origin? For the researcher, there are obvious, perhaps insurmountable, obstacles in asking an adult to describe accurately the events of early childhood. How much of ancient history is re-inscribed by the pen of contemporary experience and attempts to come to grips with personal oddities? So, unless we are fortunate enough to welcome a parade of old mothers or adult sisters of the transvestite confessing to their earlier sins, we are often crippled in our race to the truth.

Stoller helped clarify some distinctions between transvestism and the recently emerged topic of interest—transsexualism. His insights on their relevance to patient selection for sex-change surgery remain timely. Study of the transvestite helped elucidate the distinction between sex and gender:

> A man with a sense of being feminine while cross-dressed is excitedly aware of being a male. Essential to his perversion are two aspects of gender identity, the latter one, *I am feminine*, and the earlier core identity, *I am (nonetheless) a male.* (Stoller, 1968, p. 40)

Stoller collaborated with Virginia Prince for three decades. Virginia was the most significant transvestite activist. She edited the magazine *Transvestia* for 100 editions (nearly as many as I edited *Archives of Sexual Behavior*), and, like me, founded an international organization (but, probably, with more cross-dressers). She died only recently, age 96.

Stoller recorded 2000 sessions of their dialogue. When he would be on holiday, the task of seeing Virginia would befall me. I would sometimes cringe. I would be subjected to her constant hectoring. But she had good points, as well. She delighted in informing us that there is no such thing as male or female clothing. A 3-piece suit does not have a penis.

The psychodynamics of the women in the lives of transvestites were also explored. As a clinician, his small sample may not have been representative. Two subgroups were described: the male-hater and the succoror. In writing of the male-hater, he was incubating his thoughts on the later exposition of hostility in erotic excitement (Stoller, 1979). The succorer (as a wife) encouraged the cross-dressing "with an air of innocent enthusiasm…" (Stoller, 1968, p. 207). "She does not know yet that as he becomes a more successful transvestite, her enthusiasm will wane…they will divorce" (Stoller, 1968, p. 212).

Other published reports are all over the place. One looked at 20 wives seen in a clinical setting and found all to be "moral masochists" (Wise, 1985). Relevant to Stoller's description of the male-hater was a subgroup termed "eye-for-an-eye." This wife would "extract punishment" by forcing the husband to do

housework while cross-dressed (symbiotic salvation?). Another interviewed wives obtained from the cross-dressing organizations of their husbands (Brown & Collier, 1989). Any former wives who might fit Stoller's subgroup would be long gone. Another characterized the primary reaction of wives as "stigma management" (Cairns, 1997). What will the neighbours say? What about the kids? Indeed, what about the children of transvestite fathers? Do they cross-dress more than most children? When they do, is it sexually arousing? Are the kids like the kids of homosexual parents, perhaps visiting, but not emigrating? Finally, a study with hundreds of subjects. Here, wives were "an unremarkable group…[who] unexpectedly found themselves in a marriage with a transvestite husband" (Docter, 1988).

I characterize the psychoanalytic, psychiatric, and psychologic descriptions of the wives of transvestites as an example of the blind men groping the elephant to provide its description. (Consider the description given by the one who grabbed the male's prehensile penis.)

Stoller saw society as the primary source of conflict for the transvestite (Stoller, 1968, p. 241). The International Classification of Diseases (ICD), the endless catalogue of illnesses published by the World Health Organization, includes transvestism. However, three countries in Europe (Denmark, Norway, and Sweden) have recently deleted transvestism from the version operant in their country. This because of the associated stigmatization (www.revisef65.org; www.pinknews.co.uk/). Both the ICD and the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) of the American Psychiatric Association (APA) are undergoing revision. Is transvestism to remain a perversion, a paraphilia, or some less pejorative euphemism to designate sexual mental disorder?

Transsexualism

In 1968, sex-change surgery in the U.S. was in its infancy. The Johns Hopkins Hospital announced its program, inspired by my other mentor, John Money, in 1966. The first multidisciplinary text on the subject was co-edited by John and me in 1969 (Green & Money, 1969).

Research conducted in the mid-1960s by myself with Stoller and MacAndrew revealed that most physicians, including psychiatrists, were opposed to the surgery, even after a patient had had years of psychotherapy and was probably suicidal if surgery were to be denied (Green, Stoller, & MacAndrew, 1966). So, for a psychoanalyst in 1968 to endorse sex-change, albeit begrudgingly, was brave. Stoller (1968) wrote: "…if one does not assist transsexual patients, they are deeply unhappy" (p. 248); "…most transsexuals are less depressed and anxious, more sociable and affectionate…after 'the change'" (p. 249).

He knew the objections:

> It has been suggested that "no psychotherapeutic procedure less than intensive, prolonged, classic psychoanalysis would have any effect. If properly done, it could probably reduce the patient's agitation and the level of unhappiness. It is not impossible that his major symptoms may decrease in frequency and urgency." (Ostow, personal communication, Stoller, 1968, p. 249)

Stoller (1968) responded: "This statement has the vigorous ring of sober caution; it also must have been written by someone who has never tried to get such a patient into psychoanalysis" (p. 249).

I interviewed a patient who did undergo psychoanalysis. One of the first transsexuals I saw was when I was seeing some of Harry Benjamin's patients in New York in the mid-1960s. Harry was the "father of transsexualism," published the first text in 1966, and had more patients than anyone (Benjamin, 1966). This patient had just completed ophthalmology residency training. For years, he had wanted to be a woman. His family, which contained psychoanalysts, dispatched him into analytic therapy for 8 years. There he was, in Benjamin's office, requesting sex change. (He, now she, has been the subject of a film starring Vanessa Redgrave, titled *Second Serve*. The patient was also a tennis star.)

Stoller argued in *Sex and Gender* that sex-change should be utilized as a research technique. Patient selection was crucial. It should be limited to those males who had been very feminine in childhood, had never lived acceptably in a masculine role, and who had not derived pleasure from their penis. He termed these "true transsexuals" (Stoller, 1968, p. 251).

I found such a patient at about the time *Sex and Gender* was published. Barbara had lived for 10 years as a woman. She was employed as a woman receptionist at a prominent restaurant. She was living with a man as husband and wife. The man reportedly considered her penis to be a congenital anomaly that would be corrected surgically. She reported that their sex life did not include her penis. She denied ever being sexually aroused by cross-dressing.

Will Goodwin was a urological surgeon at UCLA who had also emigrated from Johns Hopkins. He was willing to perform the surgery. I proposed the procedure to the psychiatry department chair, Henry Work. Stoller stood aside, providing me with enough rope.

A concern was raised about legal action, prosecution for mayhem. This common law crime derives from medieval England to punish anyone who would cut off the limb of one of the King's fighting men. For clarification regarding possible prosecution, I contacted the University of California legal counsel in Berkeley. This was a couple of years after the Free Speech Movement protest against the Vietnam War. Then, the university president was Clark Kerr and placards were all over Berkeley lettered FUCK. That stood for Freedom Under Clark Kerr.

I suspect that my query was a welcome change from that tumultuous period. I have never received such a swift reply from an attorney. He conceded that I would be vulnerable to prosecution for mayhem. The penalty if convicted was up to 10 years

in prison. Furthermore, since this proposed act involved two of us, a psychiatrist and a surgeon, it constituted conspiracy. The penalty was now up to 14 years. But, the university counsel reassured me, the University would pay our legal bill. Barbara had her surgery. We were not prosecuted.

The stringent criteria Stoller set for patient selection for sex-change was too influential. Subsequently, nearly every patient presented to psychiatrists with precisely the winning psychosexual history. When it dawned on our profession that perhaps we were being presented with a song and dance, the criteria for patient selection changed. It was no longer where you had been, but where you were going.

Money coined the term "Real Life Test." Patients needed to live full-time successfully in their aspired for gender role for at least a year to be eligible for surgery. Now patients acknowledged some masculinity in their life, fetishistic cross-dressing, pleasurable penile activity, and even heterosexuality. Stoller's diagnostic selection criteria have remained valid. Current research demonstrates that males from non-true transsexual backgrounds do not fare as well post sex-change (Smith, van Goozen, Kuiper, & Cohen-Kettenis, 2005).

Notwithstanding his "passive" endorsement of sex change, Stoller remained sceptical of the extent to which the procedures accomplished all the patients hoped for:

> Their anguish before the procedures is intense and genuine…nonetheless, they are left more or less dissatisfied, feeling that although the necessary procedures have feminized some of their appearance and functions, the results are far from complete. (Stoller, 1968, p. 247)

In the unique use of words that was a precious jewel in Stoller's writing, he later characterized the post-operative male as a "Near Miss" (Stoller, 1982).

Stoller's acknowledgement of the difficulty in engaging a pre-operative transsexual in psychoanalysis or even psychotherapy was echoed in 1997 by psychoanalyst Colette Chiland of Paris.

> We see a certain number of transsexuals before they undergo hormone treatment, but as soon as they start the process of so-called sex change, they are as good as lost to psychotherapy. It is very rare to get a transsexual in real analysis because they have very strong narcissistic defences and resistance to transference; when they begin to experience transference dependence, they run away. [However], we see more and more transsexuals coming for psychotherapy or even analysis after surgical sex reassignment. They do not regret the surgery, they feel they could not have done without it, but they still feel unhappy, too isolated, nether completely a man or a woman. (Chiland, 1988)

In Bob's words, a "near miss."

Chiland is indebted to Stoller: "All the work I have been doing with transsexuals for thirty years has its roots in the contributions of Stoller" (C. Chiland, personal e-mail communication, 29 October 2009).

But have all psychoanalysts given up the ghost of curing the preoperative transsexual? The Portman Clinic in London is one of the last psychoanalytic strongholds in the UK and carries a long history of treatment of sexual disorders. Here is their Letter to the Editor with 8 signatories published in the *Telegraph* newspaper in 2002:

> The recent judgment in the European Court of Human Rights in which a post-operative transsexual person was granted permission to marry in his [sic] adopted gender role, is a victory of fantasy over reality…Through years of psychoanalytic psychotherapy, some patients begin to understand the origins of their painful conflicted feelings and can find new ways of dealing with them, other than by trying to alter their bodies. The recent legal victory risks reinforcing a false belief that it is possible to actually change a person's gender [confusion here, perhaps, between sex and gender]. It might also strengthen the view that the only solution to psychic pain is a legal or surgical one. (Telegraph, 15 July 2002)

Those who forget the past…

Cross Sex and Gender Children

There is excited debate today on the significance of cross-gender identity and behavior in prepubertal boys and girls. Television shows feature families in which grade school age kids are attending school and living as a child of the other sex. There is vitriol from activists against treatment programs where the goal is enabling the child to live comfortably in his or her birth sex. There is demand that the diagnosis of Gender Identity Disorder be removed from the list of mental disorders. In 1968, Stoller wrote:

> The condition is pathological….If these boys are the adult transsexuals of future years, with their demands for sex-transformation procedures and the reportedly hopeless prognosis for psychiatric treatment, then the time to help them is in childhood, when their gender identity is still forming….The goal of treatment should be to make the child feel that he is a male and wants to be a masculine boy….The first step in treatment is to establish that one is in fact dealing with a childhood transsexual. Next one must start treatment immediately. If one waits until five or six or seven, the undoing is more difficult. (p. 252)

In the mid-1960s, I saw a very feminine 5-year-old boy and his mother was seen by Stoller. She entered analysis and was the source for a seminal theory of cross-gender development in a

male child, described below. I was seeing the boy twice weekly when I was a psychiatry resident, but, as described in a paper by Ralph Greenson, "When the resident was called into military service [sic] some six months later I decided to undertake the treatment" (Greenson, 1966). This resident took the boy to Greenson's home where they enthusiastically formed a treatment alliance.

A couple of years later, in *Sex and Gender*, Stoller (1968) wrote: "The first successfully treated case of a childhood transsexual is that of Greenson. A report written after the treatment was ended gives a vivid and warm account of this boy's rescue" (p. 254).

(I will digress to an autobiographical coincidence. The one other psychoanalytic paper on treating a "boy transvestite" published before Greenson's was two years earlier. It was by Melitta Sperling (Sperling, 1964). Melitta's daughter, Susan, and I were grade school classmates at PS 241 in Brooklyn. What was in the Brooklyn drinking water to inspire that psychoanalyst and this classmate of her young daughter to pursue an interest in cross-gender behavior in boys? When I published my first paper on the subject in 1960, did she know that I had been a friend of her daughter decades earlier?)

Four years after *Sex and Gender*, Larry Newman, Stoller, and I published "Treatment of Boyhood 'Transsexualism'" (Green, Newman, & Stoller, 1972). Transsexualism was in quotes as we could not be certain that these boys were pre-transsexual. We wrote: "We have previously described the behavior of very feminine young boys. Based on retrospective reports of childhood behavior given by adult males who want sex-change surgery, there is reason to believe these boys are pre-transsexual" (Green et al., p. 213). Why treat?

> As long-term follow-up data are not yet available, we cannot state with certainty that these boys, if untreated, would mature into adult transsexuals. However, because 1) there are no reports of adolescent or adult transsexuals who have successfully reoriented their profound cross-gender identity…(and) 2) because males with an extremely feminine identity in our culture undergo considerable social hardship, at all ages, we have attempted… intervention during boyhood. (Green et al., 1972, p. 213)

But, should we change society or change the child? "While privately, one might prefer to modify society's attitudes toward cross-gender behavior, in the consultation room with an unhappy youngster, one feels far more optimistic about modifying the behavior of that one child than the entire of society" (Green et al., p. 217).

Our caution about the adult outcome of such boys was wise. I had begun a long-term study of dozens of very feminine boys at about the time *Sex and Gender* was published. I had received a grant from the NIMH and Stoller was my supervisor. The study continued for about 15 years. Some boys entered therapy with myself or others; others were only periodically assessed.

In 1974, I considered the boys to be probably pre-transsexual. Their histories were juxtaposed with those of adult transsexuals in *Sexual Identity Conflict in Children and Adults* (Green, 1974). In 1987, it was clear that most were homosexual. Perhaps one was transsexual. The book title of the sequel was *The "Sissy Boy Syndrome" and the Development of Homosexuality* (Green, 1987).

In addition to the analytic treatment reports of Greenson and Sperling, Haber (1991) published a detailed report of five years' analysis of a boy from age 3. At last follow-up, when he was in his early 20 s, his mother reported that he was not dating and was interested in a career in acting. Regarding Haber's other similar patients, he has informed me, "I saw many boys and one girl in long-term intensive therapy and all seem(ed) to be gay" (C. Haber, personal e-mail communication, 26 October 2009).

How did these boys catch GID?

It was on a hot summer day in a public square in Madrid in 1966 at a psychiatric congress, with his wife Sybil standing by. Stoller proudly told me that he had unravelled the first psychodynamic explanation for the etiology of a psychiatric disorder. This was the ontogeny of extreme cross-gender behavior in a young boy. Deriving primarily from his analysis of the mother of the 5-year-old boy I had referred to Greenson, he had an extraordinary insight into the forces and influences yielding boyhood transsexualism:

> There had been excessive sharing of each other's anatomy by identification, made possible by continuous skin contact all through the day….he was never separated from the sight of her for more than a few minutes. Since he never slept more than an hour and a half at a time during the first year, her nights were spent in exhaustedly patting him, feeding him, or singing him back to sleep. (Stoller, 1968, p. 111)

"Another prime factor is a disturbance in (mother's) gender identity, her bisexuality…a heavy proportion of sensed and observable thoughts, feelings, and behavior reflecting both masculine and feminine identifications" (Stoller, 1968, p. 112).

The mother's parents had accepted her earlier cross-dressing. She was described as having experienced intense penis envy of her brothers and father. The boy was physically beautiful. He had been fascinated by women's clothes from 8 months and began cross-dressing as soon as he could walk. The behavior was admired by both parents.

> While she was able to permit him some chance of separation, she greatly slowed the process down, a fixation produced especially by the primordial joys of the continuous skin-to-skin contact…. (Stoller, 1968, p. 125)

> The boy was (the phallus) of her flesh…*He was his mother's feminized phallus*. (Stoller, 1968, p. 120)

Another case in which Stoller had been hoodwinked was consistent with these findings. This involved a patient who con-

vinced him, a sociologist, endocrinologists, and surgeons that, although a normal appearing male until puberty, the feminization that then occurred was the result of a biological variant. UCLA was not yet performing sex change surgery on transsexuals. But this patient was "intersex." So genital reconstruction was performed to conform to her very ample breast development resulting from her "intersex" condition. Five years later, she confessed that the whole story was a scam. She was transsexual. She had been swallowing her mother's estrogen tablets, secretly, from age 12.

Converting major embarrassment, and some anger, into victory, Stoller interviewed the patient's mother. The child had been beautiful. He was dressed in girls' clothes. Significantly, mother recalled that

> …from 2 to 8 years she and her son played a nighttime game. It was "mother hen and baby chick." Every night they both went to bed together at the same time, and the mother would curl herself up in such a way that she completely surrounded the little boy within the curve produced by her bent head, her torso, her arms, and her curled-up thighs and legs. (Stoller, 1968, p. 137)

As for the mother's gender identity, as a child, she thought of herself as a boy. "All of her interests were those of a boy….She dressed in boys' clothes" (Stoller, 1968, p. 138).

Here, Stoller obtained the retrospective history of an adult transsexual and mother that matched that obtained contemporaneously for a boy transsexual and his mother. He acknowledged that these "findings and speculations" did not completely exclude some biological force. "For instance, these mothers might not have wished to 'over-love' infants who struggled away from their excessive body contact" (Stoller, 1968, p. 139).

Chiland, for all her praise for Stoller, reports being unable to confirm some of his findings with children. "What he described was only a case among others, though he thought it was the typical constellation." She did, however, have a family that was just as he had written (C. Chiland, personal e-mail communication, 29 October 2009).

Psychoanalysts have not queued to propose theories of gender identity since 1968. An exception is Person and Ovesey (1983). They critiqued those of Freud, Horney, Jones, and Stoller and flogged their own. They faulted Stoller for his emphasis on a non-conflictual origin of male femininity and his analogy to the ethological concept of imprinting. Their proposal was a defence against separation anxiety. However, as their critique of Stoller was in response to a subsequent volume (Stoller, 1975), I will avoid this fray over how many angels dance on the head of the gender identity pin.

Non-analytic psychiatric critics snipe that Stoller's insights have not been confirmed. Yet, how many clinicians have replicated his methodology? And, for current researchers, this is not where the grant money is. Follow the money. The search for the "smoking gene."

Intersex

Treatment of the intersexed person has undergone radical change from 1968. How these persons are referred to has changed. How they should be treated in infancy and early childhood has changed. Whether their gender identity derives primarily from nature or nurture is in fresh dispute.

Stoller used the term intersex and hermaphrodite. At the time, patients were often referred to as pseudohermaphrodites if they, as was usually the case, did not have both ovarian and testicular tissue. Some thought that term made them seem like imposters.

Both the terms intersex and hermaphrodite were adopted by a patient activist group successful in bringing problems with their medical care to professional and public attention. This was the Intersex Society of North America. They were also known as "Hermaphrodites with Attitude." A major thrust of the early intersex activism was the genital conforming surgery carried out on infants and young children with ambiguous genitalia. Often these were females with congenital virilizing adrenal hyperplasia (CAH) where excessive prenatal androgen virilizes the female genitalia (New, 2001).

When these young children were correctly identified as 46,XX female, there was concern over the masculinized genitalia, notably the clitoris, which can appear as a small penis. To avoid social embarrassment with peers, and to avoid an androgynous sex signal to the child, the clitoris was cut down to size. This was decried by many now adult intersex activists as it often resulted in major loss of erotic sensation. The Intersex Society in 1994 stated that any genital surgery not required for strictly medical reasons should be postponed until the intersex person is old enough to provide informed consent (ISNA Recommendations for Treatment, www.isna.org).

To which sex should the infant be designated? In 1968, the prevailing theory was derived from the seminal work of John Money. Thirteen and 11 years before *Sex and Gender*, Money, working with the Hampsons, published three classic papers (Money, Hampson, & Hampson, 1955a, b, 1957). They showed that gender identity (not yet so named) followed the sex to which the infant was designated. This ascription would trump any biological variables, including gonads, genital appearance, internal reproductive structures, chromosomal configuration, etc. This dictum was engraved on tablets handed down from Mt. Hopkins.

Importantly, if it became apparent that a diagnostic error had been made in assigning sex to the infant, reassignment to the correct sex was feasible up until only about 18 months of age. Thereafter, identity as male or female was fixed.

Stoller did not challenge these tenets. But from his clinical experience, he enriched understanding, especially with his exposition on a possible hermaphroditic identity. Here, there was inconsistency, confusion, ambivalence over the "correct" sex designation for a child. Consequently, neither clear cut male nor female core gender identity was set in place. Years later, that person, well beyond the 18 month cut-off point could want

to, and be able to, change (Stoller, 1968, p. 36). "If from the start, the child's elders are not sure enough, because the child is anatomically ambiguous, then a defect in character structure can develop in which the patient knows only the gender identity of a hermaphrodite" (Stoller, 1968, p. 23).

Examples of late sex change in the intersex are more recently reported. Meyer-Bahlburg et al. (1996) described four patients who had CAH with gender change from female to male. The process was gradual and extended into adulthood. Factors appeared to be gender-atypical behavioral self-image, a gender atypical body image, and the development of erotic attraction to women.

I asked the lead author about Stoller's concept of a hermaphroditic identity in these patients. He replied "I did not discuss it with them, so I don't know. However, I later added a 'third gender' question to my follow-up protocol for XY (a different group of intersex patients) and the result was 15% (of 59 patients) agreed" (H. F. L. Meyer-Bahlburg, personal e-mail communication, 13 October 2009). Those persons with ambiguous genitalia reported more uncertainty about gender and more endorsed themselves with a third gender.

Stoller was sensitive to the issue of when, if at all, to do genital surgery. He reported a case where the penis was removed, leaving only a stump 3/4 inch at erection. This was in a chromosomal male, living as a woman, whose undescended testes had been previously removed. "At the time, she was relieved to have the operation; it shut others up about her abnormal genitalia….the patient now profoundly regrets the operation" (Stoller, 1968, p. 32).

He understood the conflicts of the intersexed and their suffering from secrecy and poor clinical management. He respected the work of Money but recognized the variance attributable to the parent–child experience of sex and gender that could yield a whole new ball game. There was a complexity of sex and gender requiring a different attention and a more nuanced understanding.

Biological Force

Proposing a biological force behind gender identity was heroic in 1968. As noted, between 1955 and 1957, the stone tablets handed down from Mt. Hopkins decreed the ultimate triumph of nurture over nature. As radical as it would be for any clinician to question that mandate, it was ever more so for a psychoanalyst. But, Stoller was confronted with an extraordinary case:

> This person had been born appearing as a normal female. During childhood, she was very much the tomboy. "In all games with other children the child seemed to take male roles.…She could scarcely be forced into female clothes. …Her companions were boys, with whom she played boys' games." (Stoller, 1968, p. 68)

At the onset of puberty, she developed a hoarse voice, presumably from laryngitis. The hoarseness continued. Her body began to virilize, not feminize. Endocrinological and surgical exploration revealed an intersex condition, namely 17 B-hydroxysteroid dehydrogenase-3 deficiency, and intra-abdominal testes (Liakopoulov, Keramydas, Dracopoulov, & Dacou-Voutetakis, 2009). The enzyme defect interferes with production of testosterone from androstenedione. Consequently, external anatomy does not differentiate prenatally as male. But, then, at puberty, androgen production and its effects can be sufficient to virilize.

> When the patient was informed of the discovery, the switch to living as a male was immediate. "She acted as though she were being told something of which she had always been dimly aware, of which she had no doubt. Her attitude was as if to say, 'Yes. Very good. Thank you. I am not surprised.'"! (Stoller, 1968, p. 70)

(As I write this, here in England we are in the midst of another 3 day postal strike. I see this intersex case as a further example of delayed male.)

Consider this note by Stoller, on this case, to stoke the debate on nature versus nurture in academic performance. This debate was fuelled in recent years by Larry Summers, then President of Harvard. He was forced to fall on his sword for suggesting perhaps the wrong explanation, an innate one, for why males excel in science and maths. (He has since been resuscitated as Director of the White House National Economic Council by the new President of his entire country). Stoller (1968) wrote: "After the sex reassignment, this patient came to be among the first in his class in mathematics; a subject in which he had done very poorly when he thought he was a girl" (p. 70).

Today, debate continues over whether a person with 44,XY karyotype and normal prenatal androgen, but born with defective genitalia, should be socialized as male or female. The "John/Joan" case challenged the intersex model of nurture trumping nature. Essentially, a monozygotic male twin lost his penis traumatically in infancy, was reassigned as a girl, raised as a girl, but in adolescence insisted on continuing life as a male. An explanation is normal prenatal androgen (Diamond & Sigmundson, 1997).

Cloacal exstrophy was not mentioned in *Sex and Gender*. Here, there is a horrific pelvic birth defect obscuring the genitalia. In consequence of the genital defect where phallus construction for an XY male child would be extremely difficult, many male infants were socialized as girls. However, many later requested to live as male. The explanation is that their prenatal androgen levels, presumed normal for a male, organized their brain as male (Reiner & Gearhart 2004). Stoller's biological force is alive and well.

Conclusion

Stoller's sensitivities and insights to the clinical issues in sex and gender cannot be fairly assessed by current practice or research. There is no comparable corpus of psychoanalytically derived data on the class of patients reported in *Sex and Gender*. There is no one clinician or researcher amassing and publishing such a corpus.

Sex research published in science journals typically reports extensive samples, assessed by questionnaire, and analyzed statistically. To this, Stoller complained, "Sex research is camouflaged by large samples, hidden in tabulated columns, and dissected beyond significance by statistical packages. Although statistical techniques may enable us to corroborate or deny a hypothesis, they do not produce one" (Stoller, 1973).

Even that strategy of impersonal personal research has latterly become even less personal. Sex research may be via the Internet. Large samples of almost every sexual oddity can be generated online. Human contact is limited to clicking "Send" or "Inbox."

As for the term "gender identity," Stoller demurred that it was not to be defended as one of the splendours of the scientific world. What has been its impact?

It has become the diagnostic term for extensive cross-sex identification and cross-gender behavior. For children, the term Gender Identity Disorder of Childhood entered the DSM, in 1980, promoted by an advisory committee on which I served. Then, in 1994, the diagnostic term "Transsexualism" was replaced by Gender Identity Disorder in the DSM, recommended by a subcommittee on which I served.

From 1994 to 2006, in London, I was research director and consultant psychiatrist at the world's largest treatment program for persons who want to "change sex." The surgeons perform three sex-change operations a week. The program is named the Gender Identity Clinic.

Google the term "gender identity" and you will retrieve 4 million responses. If you narrow the search terrain to Google Scholar, you will get 1.5 million responses. You need not scroll through all 1.5 million to find *Sex and Gender*. It is number 11.

Sex and Gender, with its wide clinical scope buttressed by extensive verbatim interviews, was a landmark contribution to the study and understanding of the bedrock of our personhood. Four decades have not dimmed its luminescence.

Acknowledgment This essay is modified from an invited address at the Robert Stoller Foundation, Los Angeles, CA, 20 November 2009.

References

Benjamin, H. (1966). *The transsexual phenomenon*. New York: Julian Press.

Brown, G. R., & Collier, L. (1989). Transvestites' women revisited: A nonpatient sample. *Archives of Sexual Behavior, 18*, 73–83.

Cairns, K. (1997). Counseling the partners of heterosexual male cross-dressers. *Canadian Journal of Human Sexuality, 6*, 297–306.

Chiland, C. (1988). Transvestism and transsexualism. *International Journal of Psycho-Analysis, 79*, 156–159.

Diamond, M., & Sigmundson, H. K. (1997). Sex reassignment at birth: Long-term review and clinical implications. *Archives of Pediatrics and Adolescent Medicine, 151*, 298–304.

Docter, R. (1988). *Transvestites and transsexuals: Toward a theory of cross-gender behavior*. New York: Plenum.

Green, R. (1974). *Sexual identity conflict in children and adults*. New York: Basic Books.

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

Green, R., & Money, J. (Eds.). (1969). *Transsexualism and sex reassignment*. Baltimore, MD: The Johns Hopkins Press.

Green, R., Newman, L. E., & Stoller, R. J. (1972). Treatment of boyhood 'transsexualism'. *Archives of General Psychiatry, 26*, 213–217.

Green, R., Stoller, R. J., & MacAndrew, C. (1966). Attitudes toward sex transformation procedures. *Archives of General Psychiatry, 15*, 178–182.

Greenson, R. R. (1966). A transvestite boy and a hypothesis. *International Journal of Psycho-Analysis, 47*, 396–403.

Haber, C. H. (1991). The psychoanalytic treatment of a preschool boy with a gender identity disorder. *Journal of the American Psychoanalytic Association, 39*, 107–129.

Intersex Society of North America. (1994). Recommendations for treatment. Available at www.isna.org.

Liakopoulov, M., Keramydas, D., Dracopoulov, M., & Dacou-Voutetakis, C. (2009). The dilemma of sex reassignment in an adolescent with 17B-HSD-3 deficiency raised as a female: Ten year follow-up [Letter to the Editor]. *Archives of Sexual Behavior, 38*, 615–618.

Meyer-Bahlburg, H. F. L., Gruen, R. S., New, M. I., Bell, J. J., Morishima, A., Shimshi, M., et al. (1996). Gender change from female to male in classical congenital adrenal hyperplasia. *Hormones and Behavior, 30*, 319–332.

Money, J., Hampson, J. G., & Hampson, J. L. (1955a). Hermaphroditism: Recommendations concerning assignment of sex, change of sex, and psychologic management. *Bulletin of the John Hopkins Hospital, 97*, 284–300.

Money, J., Hampson, J. G., & Hampson, J. L. (1955b). An examination of some basic sexual concepts: The evidence of human hermaphroditism. *Bulletin of the Johns Hopkins Hospital, 97*, 301–319.

Money, J., Hampson, J. G., & Hampson, J. L. (1957). Imprinting and the establishment of gender role. *Archives of Neurology and Psychiatry, 77*, 333–336.

New, M. I. (2001). Congenital adrenal hyperplasia. In A. Margioris & G. Chrousos (Eds.), *Adrenal disorders* (pp. 263–271). New York: Humana Press.

Person, E. S., & Ovesey, L. (1983). Psychoanalytic theories of gender identity. *Journal of the American Academy of Psychoanalysis, 11*, 203–226.

Reiner, W. G., & Gearhart, J. P. (2004). Discordant sexual identity in some genetic males with cloacal exstrophy assigned to female sex at birth. *New England Journal of Medicine, 350*, 333–341.

Smith, Y. L. S., van Goozen, S. H. M., Kuiper, A. J., & Cohen-Kettenis, P. T. (2005). Sex reassignment: Outcomes and predictors of treatment for adolescent and adult transsexuals. *Psychological Medicine, 35*, 89–99.

Sperling, M. (1964). The analysis of a boy with transvestite tendencies. *Psychoanalytic Study of the Child, 19*, 470–493.

Stoller, R. J. (1964). A contribution to the study of gender identity. *International Journal of Psycho-Analysis, 45*, 220–226.

Stoller, R. J. (1968). *Sex and gender: Vol. 1. The development of masculinity and femininity*. New York: Science House.

Stoller, R. J. (1973). *Splitting: A case of female masculinity*. New York: Quadrangle Books.

Stoller, R. J. (1975). *Sex and gender: Vol. 2. The transsexual experiment*. London: Hogarth Press.
Stoller, R. J. (1979). *Sexual excitement*. New York: Pantheon.
Stoller, R. J. (1982). Near miss: "Sex change" treatment and its evaluation. In M. R. Zales (Ed.), *Eating, sleeping, and sexuality* (pp. 258–283). New York: Bruner/Mazel.
Telegraph. (2002, 15 July). *The psychiatry of transsexuality*.
Wise, T. N. (1985). Coping with a transvestitic mate: Clinical implications. *Journal of Sex and Marital Therapy, 11*, 293–300.

Arch Sex Behav (2010) 39:1467–1468
DOI 10.1007/s10508-010-9686-0

BOOK REVIEW

Human Sexuality and Its Problems (Third Edition)

By John Bancroft. Churchhill Livingstone Elsevier, Edinburgh, 2009, 546 pp., $99.00

J. Dennis Fortenberry

Published online: 19 October 2010

Palimpsest refers to a sheet of parchment where the original ink is scraped away and replaced by new writing. Outlines of the original text are often still identifiable. In more expanded use, *palimpsest* refers to modified structures in which traces of the original can be identified.

The third edition of *Human Sexuality and Its Problems* is a literary palimpsest, appearing 20 years after the second edition and 26 years since the first. Traces of the earlier editions are easily detected, but the originals have been substantially replaced. Rightly so, because in a quarter-century, the realm of study of human sexuality has expanded, morphed, digitalized, molecularized, and globalized.

The great advantage of this volume's still-visible history, however, is the careful guidance given the reader by its author. Bancroft has been in the midst of much of this transformation of sexuality and sexuality research. As director of the Kinsey Institute for Research in Sex, Gender, and Reproduction from 1995 to 2005, he respectfully but firmly spoke against the demonization of Alfred Kinsey and reacquainted many with the importance of Kinsey's research. He played a key role in organization of the U.S. Surgeon General's *Call to Action for Sexual Responsibility* in 1999. Mostly ignored during the presidential administration of George Bush, the products of that *Call to Action* represented a remarkably civil and democratic dialog about the role of sex and sexuality in American public health. Bancroft alludes to the *Call to Action* at several points, most tellingly in the Introduction, where he also frames the new edition in terms of his personal growth during the years between first and third editions.

J. Dennis Fortenberry (✉)
Section of Adolescent Medicine, Indiana University School of Medicine, 410 W. 10th St., Room 1001, Indianapolis, IN 46202, USA
e-mail: jfortenb@iupui.edu

During Bancroft's tenure as director of the Kinsey Institute, he also had to deal with attempts within the American House of Representatives to withdraw funding for peer-reviewed sexuality research funded by the National Institutes of Health. Aspects of these and other details relevant to sexuality research are addressed throughout the book. Such details provide an invaluable breadth of perspective that should remind everyone of the importance—but ever-present fragility—of our work.

The book's layers take on additional richness through the occasional emergence of Bancroft's own voice. Bancroft uses that voice to step away from authorial objectivity to comment (sometimes sternly), to express doubt, or to congratulate the field for its progress. From time to time, he weighs in about ongoing debates. I found the presence of these interjections to be both helpful and refreshing. Like having the opportunity to ask questions of a mentor: "What do *you* think?" The questions and answers shape the relations of student and mentor. Or, in this case, they shape the dialog between author and reader. Texts should stimulate thought as well as teach, and a mental dialog with Bancroft filled many quiet mornings as I was reading, or as I moved into my own writing for the day.

Human Sexuality and Its Problems is organized in 16 chapters. Roughly half of the book (Chapters 1–10) present detailed summaries of a variety of topics. Chapters 11–16 generally address a variety of sexual problems and treatment/management approaches. Four areas of the book deserve special emphasis.

Chapter 2 addresses models of sexuality and the role of theory. A short discussion of epistemology is followed by critiques of two popular contemporary theories—Sexual Strategy Theory and Sexual Script Theory. Bancroft sharply dissects essential limitations of these theories, leaving the reader with negative impressions of both. Bancroft then presents the Dual Control Model, developed with colleagues at the Kinsey Institute. In essence, the Dual Control Model is a metaphor for the

summed excitatory and inhibitory central nervous system inputs that control sexual behavior. The Dual Control Model exemplifies Bancroft's commitment to the idea that sexuality is centered in human bodies and not in human social organization. This is not to say that social influences are unacknowledged. Bancroft is far too wide-ranging in his thought and experience to believe otherwise. Rather, neurophysiologic and anatomic issues are given primary attention with attempts to integrate other influences into one primary model.

A very long chapter, entitled "Sexual Arousal and Response—The Psychosomatic Circle," is in many ways the centerpiece of the book. Although present in the second edition, this chapter is substantially new and gives careful attention to the brain's function as a sexual organ. Subsequent sections address sexual arousal, genital responsiveness, and orgasm within the context of a host of psychological and hormonal influences. The chapter ends with a short but very useful exploration of gender differences in sexuality and sexual behavior. This chapter alone justifies purchase of the book.

A subsequent chapter is entitled "Heterosexuality." This chapter is rich in detail. However, I found the chapter to be profoundly disturbing. In addition to details about heterosexual behavior, the chapter also contains an essay on the history of sex surveys, and discussions of masturbation, sexual desire/fantasy, Internet sexuality, and erotica use. All of these topics would seem to apply with equal importance and usefulness to sexualities other than "hetero." In fact, this strikes me as an area where a different approach to a traditional heterosexuality/homosexuality dichotomy could have proven of great value. The use of such dichotomy—now well into the first decade of the twenty-first century—in a text on *human* sexuality seems sadly elegiac evidence that our socially constructed science continues to reify differences in sexuality that contribute to justifications for all manner of oppressions that plague our best intentions. Careful reading of the text shows that Bancroft understands this, but I believe this is an area where the book's historical structure harms its contemporary contribution.

I would also call attention to the section entitled "The Victim of Child Sexual Abuse" found in Chapter 16. Bancroft acknowledges that no other section of the book required so much updating of both the science and the attitude of the section. He attempts to give context and meaning to the vicious debates about the long-term effects of child sexual abuse that arose in the late 1990s. Fallout from this debate continues to have chilling effect on the kinds of questions asked about child sexual abuse, and the kinds of papers that might be accepted by most journals.

In an era of rapid access to almost any publication from any part of the world, accompanied by a nearly continuous electronic babble, this book stands out for its mastery of an immense topic, its contemplation of the difficulties of understanding, and its firm conviction that sexual understanding through sexual science is important. Having read it, I know more than I did, and questioned things I had assumed. It is a lovely resource. Most importantly, however, I think this book, like no other, represents a measured but essential tour of the science of sexuality and sexual behavior. I cannot imagine anyone committed to the field who would not find it provocative and continuously useful.

Arch Sex Behav (2010) 39:1469–1470
DOI 10.1007/s10508-010-9684-2

BOOK REVIEW

The Sex Lives of Saints: An Erotics of Ancient Hagiology

By Virginia Burrus. University of Pennsylvania Press, Philadelphia, 2004, 216 pp., $42.50

Peter W. Barlow

Published online: 20 October 2010

According to Burrus, the title of this volume is not intended to be oxymoronic. Together with an intriguing dusk jacket showing a sepia reproduction of Papety's "The Temptation of Saint Hilarion," the arresting title, and especially the sub-title, do indeed convey the book's message—that all may not be quite right with the conventional view of saintly life, and that the idea of there being a repressive, sepia-hued perception of sexuality during the early days of Christianity can be challenged by a new exploration of the hagiographical texts, leading, as the present book purports to show, to the uncovering of an exuberant eroticism.

The saints who are the subject of enquiry are those who comprise the "Desert Fathers"—devout Christians of the 4–5th centuries who fled the liberal life-styles of towns and cities to find spiritual retreat in the deserts of Palestine and North Africa. The most notable lives were recorded by St Jerome (d. 420) in his "Vita S. Pauli, prima eremitae" (374–379) (and herein is also included the life of Paul's saintly friend and desert companion, Antony), as well as the "Vita Malchi, monachi captivi" and "Vita S. Hilarionis," of 391–392, all of which were collected into a massive folio by Fr Heribert Rosweyde in the 17th century as "Vitae Patrum." Various English translations have followed (see http://www.vitae-patrum.org.uk/page2.html). The solitary lives of these men (and occasionally women, such as Pelagia and Mary of Egypt, both reformed harlots) were located at the very edge of existence—both geographically and biologically speaking. The sun scorched their bodies black, their hair was unkempt, they shunned contact with other eremites, they were a prey to madness and the visions that accompany extreme bodily privations: "Then febyll ther wyttis and fallen to fondness, Sum into despeyer and sum to madnes…. Who clymyt hye, hys fall grete ys" (see p. 181 in Waddell, 1998).

The desert was the perfect receptacle for the life which they wished to cast off; and here, too, arose the opportunity of re-finding and restoring the uncorrupted, essential human being within. Following such extended purgation, the desert fathers may well have concurred with William Blake who, in his "Proverbs of Hell," had the insight that "The road of excess leads to the palace of wisdom." In the desert was given the possibility of accomplishing the first step towards sainthood (it was not an easy step for, as St Jerome himself remarks in his epistle to the virgin Eustochium, "How often, when I was living in the desert…did I fancy myself among the pleasures of Rome"), and of setting desires above non-desires whence it might be possible, at a second step taken upon return to society, to equilibrate the non-desires with the desires; that is, to know and, at the same time, not to know those excesses which now have the appellation of the "seven deadly sins." Six of these (anger, covetousness, envy, gluttony, pride, sloth) relate to the immediate material needs of the body and the expressions of the emotional center. They are aspects of life as it is ordinarily led; they fall away in the contemplative silence of the desert.

But what of the seventh deadly sin, lust? This is an experience (it might simply be called "libido") arising in the sex center. Not only does it relate, as do the other "deadly sins" in their more usual everyday form ameliorated by the internal "buffers" of social propriety, to the well-being of the body and the psyche, but also (and more importantly) to a transcendent state of existence where sexual energy is revealed as the manifestation of a more fundamental generative energy (see also Bennett, 1975). This energy penetrates successive generations of living organisms and finds its most complex vehicle in man himself. Whereas the six bodily sins may be quelled by asceticism, the transcendent "sin" of lust may be a harder nut to crack (it would take an ascetic to speak here with authority!) since the

P. W. Barlow (✉)
School of Biological Sciences, University of Bristol,
Woodland Road, Bristol BS8 1UG, UK
e-mail: P.W.Barlow@bristol.ac.uk

sexual impulse is intrinsic to the fabric of each individual. To repress such a life-force—the desire to couple and reproduce being one of the very fundamental attributes of life forms with both "plus" and "minus" sexual characteristics—would be against Nature. Thus, the desert would enable ascesis to the point where the turning towards this "last temptation" engendered by an ineradicable life-force is balanced by a turning away from it: desire in equilibrium with non-desire contained within the perfection of sainthood. This indeed may have been the conclusion reached by St Jerome himself, though stated in a different, more oblique syntax, that "desires wither without the deferrals of sublimation; freedom loses sweetness when utterly released from constraint."

Burrus detects a sublime sensuality in St Jerome's writings on Malchus, Hilarion, and Paul. It is as though Jerome were so enchanted by the loving, crypto-erotic aspects of his heroes' lives that, for his own gratification, his style becomes not one of closure, but one of ambiguity. Jerome's writing, so Burrus believes, suppresses desire on the one hand while on the other hand intensifies it: magically, this brings about the dispersion of the very object of desire. Jerome thus displays great narrative skill in balancing the blissful agony of temptation with the acerbic fruits of resistance, evoking this dynamic in his narrative and finding a corresponding resonance in his readers.

In Chapter 3, "Hybrid Desire: Empire, Sadism and the Soldier Saint," crypto-eroticism is set aside and Burrus overtly displays her own ability to eroticize a text. Ostensibly, the chapter concerns the saintly adventures of Martin of Tours, as related by his friend, the ascetic Sulpicius Severus (d. ca. 425), in "De vita S. Martini" (ca. 404). Martin was a Pannonian (hailing from what is now Szombathely in Hungary) who, in his early years, served in the Roman army, but is later credited with the conversion of Gaul to Christianity. Sulpicius, born of noble parents in Aquitaine, becomes overwhelmed by, if not enamored of, the powerful personality of Martin, and through him becomes converted to the monastic life. Reading this chapter, the suspicion grows that Burrus has herself become eroticized by the tale that Sulpicius tells of Martin's deeds. She believes that the text of "De vita S. Martini" contains a "covert eroticism" and she sets out to extract, at every turn, alternative meanings to apparently simple statements in the furtherance of her argument. For example, while Sulpicius, speaking of Martin's relation with his servant, says, "he (Martin) often acted as though, while really master, he had been inferior," Burrus sees this as tantamount to a homosexual relationship—"Strikingly, he prefers to play bottom to his servant's top" (p. 96); and of his (Martin's) patience and humility, which "surpassed what seemed possible to human nature" (Sulpicius), this becomes "his capacity for humiliation is unnatural" (Burrus)—Martin's passiveness is held to be active eroto-masochism. Could there not be another reading both here and elsewhere in the descriptions (in Chapter 1) of the friendship between Antony and Paul, as related by St Jerome?—that of a higher love between man and fellow men which transcends the more "base" erotics, as usually understood. Even linking phrases of Burrus' text have become ambiguously charged: after Martin has given a series of commands to Sulpicius that he should submit to a new regime, "Sulpicius by now can do no more than *ejaculate*, 'What power and dignity were in Martin's words…'" (p. 102). This is surely an intriguing position for an author to position herself: readers become aware not only of a possible counter-erotics embedded within the original texts of the "Vitae Patrum," but also of the erotic predelictions of the author/commentator. By her own argument, Burrus places herself in the same situation as Jerome, of luxuriating in the tales that he/she has been telling. As Burrus says (p. 155) in a passage which follows an apposite quotation from Annie Sprinkle, that well known former porn star and present sex guru, "Reading the 'Lives of Harlots' with unrepentant pleasure is risky business…. I cannot deny the seductive allure of the sexiest of saints, the holy whore…" And, later, "Reading these 'harlots' without shame, even shamelessly *reading as a harlot* (emphasis in original)—once I got over the initial shock, I enjoyed the idea." Rarely do readers come so close to the author of a scholarly text, sharing not only the ideas and interpretations proposed, but also the emotional/erotic charge which the subject matter has generated in the author.

References

Bennett, J. G. (1975). *Sex*. Sherborne: Coombe Springs Press.
Waddell, H. (1998). *The desert fathers*. New York: Vintage Books.

Arch Sex Behav (2010) 39:1471–1472
DOI 10.1007/s10508-010-9685-1

BOOK REVIEW

Sexualities: Identities, Behaviors, and Society

Edited by Michael S. Kimmel and Rebecca F. Plante. Oxford University Press, New York, 2004, 487 pp., $59.95

Miriam Ehrenberg

Published online: 16 October 2010

This collection of articles is intended, according to the editors, to be an "ideal human sexualities text" for sociology and women's studies courses. As such, it needs to be evaluated in terms of its academic rigor and suitability for college students. Unfortunately, the book is somewhat remiss in its attention to important details and thus does not set a standard for students. Partly this seems to be a problem in copy editing. The table of contents, for instance, lists Part One and Part Four but omits Parts Two and Three, and the reader obtains a notion of the book's organizing themes only by turning the pages and finding Parts Two and Three and their titles.[1] A more serious problem is the lack of a bibliography. The publication sources of the articles included (47 in all, 45 of which were previously published and two that were written for the present collection) are referenced solely on the bottom of their first page and the index does not differentiate between names of the contributors and names of people cited in their articles. The extract from Freud, which appears in a chapter called "Classical Inquiries," is cited as copyrighted by the translator James Strachey in 1964 and l965. Nowhere is the original date of Freud's contribution noted, a glaring omission in a section on "classical" inquiries. The reference for the selection from Krafft-Ebing does not provide the translator, publisher or date. According to the introduction, some boundaries needed to be set in presenting evidence and therefore only data from recent American studies were considered. Yet, it appears that four of the studies included were conducted elsewhere, i.e., in Finland, Canada, and Australia, although the sample descriptions are not clear (the fault of the article authors, not the editors, but not corrected by the latter).

M. Ehrenberg (✉)
Department of Psychology, John Jay College of Criminal Justice, City University of New York, New York, NY 10019, USA
e-mail: mehrenberg@jjay.cuny.edu

The introduction provides a comprehensive orientation to the study of sexuality and clearly sets forth the editors' basic premises and arguments: first, that diversity of sexual expression and identities should be acknowledged and respected, given mutual consent between sexual partners; second, that the labeling of some behaviors as normal to the exclusion of others is a political decision; third, that more sexual information is better than less and leads to safer sex; and fourth (erroneously labeled "third"), that the great sexual divide is not between homosexuals and heterosexuals but between women and men. The four sections of the book include: Introduction to Sexualities, Sexual Identities, Sexual Behaviors, and Sexuality as Politics. The articles are grouped into 10 chapters, of approximately four articles each. The selections cover a wide range of topics, including new approaches to intersexed individuals, the emergence of the concept/term homosexuality, children's gender relations, adolescent masturbation, sexual fantasies, casual sex, same-sex issues and relationships, transgender issues, S & M practices, vibrator use, cyberspace sex, sex clubs and sex tourism, pornography, sexual violence and sexual health, among other topics. Despite this wide range of issues, the book contains some redundancies that could have been avoided by covering other important but ignored issues. For example, there are two chapters dealing with sexual issues among Asian-Americans. Why two devoted to this particular ethnic group instead of at least one dealing with issues in a different ethnic community? The Latina population, which numerically far exceeds the female Asian-American population in the U.S., often suffers from sexual exploitation in the form of machismo, and certainly deserves coverage. Sexual behaviors among African-American teenagers, who have the highest proportion of teen pregnancies among the female population, are worth

[1] Editor's note: On the Oxford University Press website advertising the book, the complete table of contents is shown.

exploring as is the connection between rap lyrics and the degradation of females.

The book contains two articles on intersexed individuals—an excellent article by Fausto-Sterling on the possibility of classifying the sexes into five, rather than two, divisions, and the other on "cosmetic" surgery to "repair" the genitals of such individuals. It would have been instructive to include an article on the epidemic of cosmetic surgery being fostered by the medical profession and plastic surgeons, not only for intersexed individuals, but for women at large, to enhance or reduce breast size, reshape lips, thighs, faces, and also reshape labia and clitorises to meet some media-generated ideal form. Morgan's (1991) essay on "Women and the Knife" would have been a welcome addition to this collection.

The section on transgendered individuals deals with the difficulties of coming out of the closet, but missing is any attempt to present the issues regarding such transformations in the first place. It would have been useful to explore the rise of the "transsexual empire" (Raymond, 1994) and the tacit acceptance that if one can medically aid in the transition process it is acceptable to do so. Such issues as why the transitions are overwhelmingly from M to F and what it means to live "successfully" as a member of the opposite sex are not explored—that is, is the ability to suppress hair growth, develop breasts, change one's hair style, wear dresses and so on all that is involved in being a woman? Does a successful transition to womanhood involve only outward appearance?

In its focus on women's issues, the book neglects important aspects of male sexuality and, more importantly, for a book designed for women's studies courses, does not treat the rampant sexualization of women's bodies and the continuing sexual double-standard. Other issues that warrant attention in a book on sexuality and women's studies are contraceptive choices and use, abortion, and the medicalization of women's reproductive processes (Kohler Reissman, 1983), among other topics.

Some of the organizational concepts underlying the various chapters are not always clear. For example, the first chapter on "classical inquiries" into sexuality includes Freud, Havelock Ellis, and Krafft-Ebing, but also, surprisingly, Kinsey et al. and Masters and Johnson. An excellent conspectus of the social origins of sexual development, an extract from Gagnon and Simon's book on *Sexual Conduct,* is also included in this chapter but it is a stretch to classify this as a classical inquiry. The chapter on "contemporary interrogations" contains articles that could easily have been subsumed in other chapters and are not necessarily more questioning or current than the remaining articles in the collection. The chapter on Sexual Behaviors includes articles on S & M, faking orgasms, vibrator use, and internet sex (as distinct from internet porn, which is in a different chapter), while casual sex on spring break is included in the chapter on Sexual Identities and Behavior. It would have been edifying in a chapter on sexual behaviors (and/or in a chapter on contemporary interrogations) to look at some of the recent inquiries into the pathologizing of sexual behaviors by the *DSM* that question which sexual behaviors can be considered "mental diseases and disorders" and which "normal." The chapter on "Same-Sex Sexualities" includes an article on "M/F Boxes" which more fittingly might have been included in the chapter on "Sexual Identities" while the article on "Dating and Romantic Relationships among Gay, Lesbian and Bisexual Youths," which appears in the chapter on "Being Sexual," might have been more logically included in the chapter on "Same-Sex Sexualities."

The chapter on "Pornograpy" contains two articles written expressly for the current collection. The article by Dines and Jensen presents a radical feminist critique of pornography that highlights its implicit violence against women. The authors argue against the civil libertarian position, noting that the only anti-pornography movement that gets serious attention is anti-feminist, grounded in conservative religious objections. The article by Abbott on performing sex looks at the pornography industry, covering production and performance as well as the impact on sex workers' personal sex lives. Abbott concludes that participants in the porn industry, despite beliefs to the contrary, are able to find pleasure in their work, including satisfaction from a sense of competency in their jobs.

While *Sexualities* is not the "ideal" sexualities text that the editors envisioned, it is still a thought-provoking collection of articles covering a broad array of topics that are tied together in a coherent framework. It would better serve a course on sexuality if more attention were given to issues of male sexuality and would benefit a course on women's issues if it addressed the sexual objectification of women that, despite the so-called sexual revolution, is becoming in many respects increasingly manifest.

References

Kohler Reissman, C. (1983). Women and medicalization: A new perspective. *Social Policy, 14*(Summer), 3–18.

Morgan, K. P. (1991). Women and the knife: Cosmetic surgery and the colonization of women's bodies. *Hypatia, 6*(3), 25–35.

Raymond, J. G. (1994). *The transsexual empire: The making of the she-male*. New York: Teachers College Press.

Arch Sex Behav (2010) 39:1473–1474
DOI 10.1007/s10508-010-9687-z

BOOK REVIEW

AIDS: Setting a Feminist Agenda

Edited by Lesley Doyal, Jennie Naido, and Tamsin Wilton. Taylor & Francis, Bristol, England, 1994, 208 pp., $75.00 (hardback), $27.00 (paperback)

Shelly Grabe

Published online: 19 October 2010

In this volume, the editors bring together writings aimed at a crucial link between feminism and HIV/AIDS work. The criticism offered of conventional HIV/AIDS work is that much of the attention to the AIDS epidemic has been rooted in the ideological agenda of Western scientific medicine which problematically reflects the hierarchical nature of a patriarchal medical profession that marginalizes the health needs of women. For this reason, the essays chosen for the edited volume discuss women's disempowerment from a gendered and sexualized lens that begins to explain why women have been marginalized in the discourse and work surrounding HIV/AIDS. While these goals are lofty and inarguably imperative to include in the discourse, the essays do not always live up to the aims set forth in the introduction. Furthermore, the reader should beware that the analysis and statistics are dated and situated in the context of early 1990s United Kingdom.

The book is divided into four sections: (I) HIV/AIDS and women, (II) The Disputed Body, (III) Masculinities/Femininities, and (IV) Live Issues for a Feminist Agenda. The essays in the first section attempt to address the ways in which women are united in their vulnerabilities to HIV and AIDS. For example, women experience greater biological vulnerability to sexually transmitted diseases that is reinforced by cultural, social, and economic factors. As such, the essays address how the hierarchical nature of patriarchy adversely affects women's ability to advocate for and protect themselves, which negatively impacts their susceptibility to living with HIV/AIDS. The authors also highlight that the social factors that lead to the spread of HIV, as well as good models of practice for working with individuals affected by HIV, are different for women—there is a sharp discussion of the documented male bias in access to quality care, the unique issues imposed by pregnancy and breastfeeding, and the ineptness of prevention programs riddled with gender bias (e.g., require women's trust in the faithfulness of male partners). Also in this section are simple facts about transmission aimed at myth-busting for the novice reader. Despite that, Section I is dated and repetitive, and many of the issues raised are still relevant today. In particular, problems related to the relative lack of control women have over their sexuality and reproductive care have not been resolved in the years since the publication of this book.[1]

In Section II (The Disputed Body), three essays address how normative objectification of the woman's body creates a context in which women's body health is at elevated risk due to the problem of estranging women from their bodies. In this section, the philosophical nature of women's bodies as a contested site of control and power, from which women's risk of HIV is elevated, is fleshed out. The idea of disembodied femininity—how femininity estranges women from their bodies by encouraging them at once to control and be preoccupied with the surface image of their bodies and how it appears to others *and* lack control and knowledge over how the same body—its desires, safety, and care are influenced by male control—is presented as a challenge to safe sex—having it and educating about it. These essays highlight the challenge in promoting safe sex by arguing that the sexual objectification of women's bodies in safe-sex educational material caters to the exact culture of objectification that originally put women's bodies at risk. If women continue to be led to believe that their body is an object for the desire of men, ensuring one's safety in

S. Grabe (✉)
Department of Psychology, University of California, Santa Cruz, 1156 High St., Santa Cruz, CA 95064, USA
e-mail: sgrabe@ucsc.edu

[1] Editor's note: When this book was sent out for review, the Editor mistakenly had read the year of publication as 2004, not 1994! Professor Grabe kindly reviewed the book nonetheless. Perhaps the reader in 2010 can reflect on whether or not the situation has changed 16 years on.

sexual encounters may take lower priority over male pleasure. Ultimately, these sections highlight the role of the body as power in the prevention and treatment of AIDS.

There were only two essays in Section III (Masculinities/Femininities). The first addressed the mental health needs of women with AIDS, making the argument that the social and medical construction of HIV as a health problem for men has direct implications for the mental health of women living with HIV. While this could be a compelling argument, I found the thesis only moderately supported and, furthermore, the topic has likely received considerable more attention in the past decade. The second essay addressed the power inherent in heterosexual masculinity through a discussion of young males' struggles to be "masculine" or, in other words, demonstrators of sexual prowess. While the empowerment of women surrounding issues of sexuality and femininity cannot be truly realized without addressing the challenges posed to both men and women as a result of masculinity, this essay did not seem to compliment the others in the volume very well. Indeed, the two essays in this section seemed to be placed together because they did not fit well any where else.

The final section (Live Issues for a Feminist Agenda) includes essays on specific groups of women: prostitutes, lesbians, women who have learning difficulties, and women who use drugs in the context of HIV/AIDS. Again, while a discussion of the risk and safety involved in these subgroups of women is important in the AIDS discourse, this material is dated and, in its narrow focus, does not lend to the major contribution of this book as a work that challenges how the hierarchical nature of patriarchy adversely affects women's ability to advocate for and protect themselves.

In sum, the citations and references to "present" research and prevalence rates would pose a dilemma for the reader interested in learning about the current state of feminist analysis in the context of the AIDS epidemic. As such, I would not recommend most essays in this book for a reader who is seeking a solid grounding in current issues. However, the theoretical analysis that situates women's bodies as "contested sites" is still relevant today and provides a timeless conceptual analysis of the body as a source of empowerment.

Arch Sex Behav (2010) 39:1475–1476
DOI 10.1007/s10508-010-9689-x

BOOK REVIEW

Before Homosexuality in the Arab-Islamic World, 1500–1800

By Khaled El-Rouayheb. University of Chicago Press, Chicago, 2005, 210 pp., $32.50

Steven Wozniak

Published online: 19 October 2010

This short book is an exploration of same-sex sexual and cultural practices in some Arabic-speaking parts of the Ottoman Empire between the 16th and 19th centuries. El-Rouayheb uses multiple sources, including poetry, biographical texts, travel journals of both European and Imperial authors, and religious and legal interpretations of the Koran and Sharia law. Much of the primary source material has been translated from manuscripts by the author himself. One consequence of the source material is that it tends to be written by and about an urban elite of educated and generally influential men, thus limiting the generalizability of the findings. From the outset of the book, El-Rouayheb contends that the modern concept of "homosexuality" did not exist during the period under review. Yet, he finds that much earlier scholarly work evaluating same-sex desire/behavior fails to recognize this simple fact. As a result, it leads to confusion and seeming paradoxes (e.g., the Koran appears to ban "homosexuality," yet the culture of the time appears to celebrate it quite openly, at least among a literate elite). By presenting a more nuanced view of the expressions of same-sex desire, El-Rouayheb aims to remedy this perplexity. In this modest goal, he is largely successful.

The book begins with a brief review of what the author calls the essentialist versus the constructionist views of homosexuality. The basic idea of the former is that the way in which sexuality is expressed and understood does not vary from culture to culture or from time period to time period. This is in contrast to the latter perspective, which finds that each culture and time period develops or constructs its own forms of sexuality. These may overlap with the forms of other cultures but, nevertheless, are determined uniquely by each one. Though stating "adjudication of the dispute between constructionists and essentialists should of course be based on a careful investigation of the historical evidence" (p. 6), it is clear that El-Rouayheb is a proponent of the constructionist viewpoint.

Much of the argument that unfolds throughout the book derives from English interpretations (and misinterpretations) of various Arabic words. The modern term *homosexuality* tends to be a somewhat blurry word that fails to capture many of the distinctions that significantly shaped the understanding of same-sex behavior in the pre-modern Middle East. One of the main ones relates to the role of each individual man in anal intercourse. Men who enjoyed being in the so-called passive role were commonly called *ma'bun*. They were thought to suffer from a medical illness (*ubnah*) for which physicians might recommend various treatments. This is in contrast to those men in the so-called active role, called *luti*, who are simply performing a specific behavior, *liwat* (p. 20). This latter group did not suffer from a medical illness and the sexual activity itself may or may not have been related to some pattern of behavior that a 21st century person might think of as homosexuality.

Further distinctions can be made based upon the reason that a man might choose to penetrate another sexually. Though the vast majority of examples provided have to do with passionate or romantic causes, sexual humiliation of another threatening male is suggested as an alternative. In the case described in the text, an outsider male is essentially gang-raped by a group of men after he threatened to sexually assault their wives. Instead of being punished or pathologized, these men were depicted as the saviors of their village.

The age of the two men involved in the sexual act seems to have been of great significance. For reasons that are not made completely clear, most of the literature discussed here involves adult men, generally thought of in the active role, with adolescent boys, usually thought of in the passive role. This may be due to the strict gender segregation practiced in these places

S. Wozniak (✉)
South of Market Mental Health Services, 760 Harrison Street,
San Francisco, CA 94107, USA
e-mail: stw_hs@yahoo.com

and the idea that an adolescent is "not completely a 'man' in the social and cultural sense" (pp. 25–26). Therefore, *luti* might more accurately be translated as "pederast" (p. 18). The majority of the book is an exploration of this type of relationship, always from the point of view of adult male admirers of adolescent boys as well as commentaries by these men's detractors. El-Rouayheb discusses love poetry extensively. It tends to idealize and romanticize these pederastic relationships. The poetry moves away from the language of *luti*, *liwat*, and *ma'bun* to metaphorically describe, for example, adolescents as "gazelles" and to offer heart-wrenching, passionate expositions of the development of "*idhar*, the first down of facial hair on the beloved's cheek" (pp. 62–63). A dreamy and, perhaps surprisingly, considering the content, charming image of this "love" emerges. With that, though, comes a certain tension. Presumably, these adolescent boys are not actually *ma'bun* themselves. So, ultimately, this passionate desire has no hope for fulfillment or perhaps can only be consummated by force or deception. Moreover, El-Rouayheb shows that a lot of debate existed at the time among the admirers of adolescents and their critics about the appropriateness and legality of this interest.

This leads to the final sections of the book, in which the various legal/religious traditions about *liwat* are analyzed. Though generally illegal with very strict penalties (often various types of execution), it seems that, in practice, these cases rarely went to trial. To go forward, four witnesses were required. Yet, there seems to have been a lot of social pressure not to witness for this kind of case. Lastly, leniency seems to have been practiced by the jurists. Religious arguments hinged on the question of whether or not beauty, in this case, the beauty of the adolescent male, had some relevance for the contemplation of God.

There is a short conclusion, which primarily summarizes the events and cultural changes occurring after the period covered by the book. El-Rouayheb argues that with the spread of Western ideas and values in the 19th and 20th centuries, the constructions of same-sex desire gradually changed to the point that they essentially mirrored the Western idea of homosexuality.

The book is well written, rich in detail, elegantly produced, but ultimately only descriptive in nature. There were several missed opportunities to introduce and explore ideas from feminist and queer theory perspectives. Though the author refers to Foucault, unlike his work on sexuality in the West, El-Rouayheb never quite places this cultural thread of same-sex desire/behavior into the larger context of pre-modern Ottoman social, political, and power structures. Moreover, the author's somewhat myopic view of male-male relationships appears to subvert his stated assumption that homosexuality did not exist for these men. The book offers no significant discussion of relationships with women or even with other adult men for that matter. A more far-reaching analysis of these men's private and public sexual lives seems to be required to develop a more complete picture. Perhaps El-Rouayheb has a sequel in mind? I look forward to it enthusiastically.