Archives of Sexual Behavior, Vol. 28, No. 4, 1999

Guest Essay

Unresolved Issues in Scientific Sexology

Nathaniel McConaghy, D.Sc., M.D.[1]

A number of unresolved issues in sexology research and practice are reviewed. Penile volume assessment of sexual arousal has consistently proved more sensitive than penile circumference assessment and requires much shorter exposure to the erotic stimuli eliciting the arousal, reducing the subjects' ability to modify their responses. Failure to acknowledge this has allowed acceptance of evidence based on penile circumference assessment that behavioral treatments such as directed masturbation can increase the ability of sex offenders to be heterosexually aroused and aversive therapy can reduce their deviant urges whereas penile volume assessment indicates these procedures are ineffective. A randomized controlled trial of relapse prevention versus no treatment for sex offenders found more treated than untreated subjects reoffended after a mean follow-up period of 4 years. Researchers and therapists accepted that a post hoc *statistical manipulation of the results provided evidence of a treatment effect. Subsequently it has been recommended that randomized controlled evaluations of treatments of sex offenders be abandoned. Meta-analysis of outcome studies has been used uncritically. The majority of men and women who report homosexual feelings and/or behavior report predominant heterosexual feelings and behavior and do not identify as homosexual. These consistent findings remain ignored. Studies of the etiology and development of homosexuality and heterosexuality treat them as distributed categorically rather than dimensionally and investigate only self-identified homosexuals and heterosexuals. With this methodology the predominantly heterosexual majority are excluded or misclassified. The belief that the European concept of the homosexual is a late 19th-century invention is based on an inadequate reading of literature. Limitations of the DSM classification of sexual and gender identity disorders are pointed out. The validity of self-report of sexual behavior has been questioned on the basis that men report a markedly higher average number of sexual partners than women. Possible sex differences in reporting the number of partners who are of the same sex, casual, or perpetrators or victims of sexual coercion and child abuse have not*

[1] School of Psychiatry, University of New South Wales, High Street, Kensington, Australia, 2033.

0004-0002/99/0800-0285$16.00/0

been taken into account. Failure of sexology to progress due to lack of resolution of conflicting issues may contribute to the low impact factor of its journals.

KEY WORDS: penile volume and circumference assessment; sex offender treatment; predominantly heterosexual homosexuals; DSM classification of sexual disorders; self-report of number of sexual partners.

INTRODUCTION

The offer of Richard Green, the Editor, to board members to write a "think piece" has given me the opportunity to express my concern that the development of the areas of sexology in which I have been involved since the 1960s has been characterized by failure to resolve basic disagreements. A possible factor in this failure is that academics who have supported a particular position might consider their reputations weakened if they acknowledge they were incorrect (McConaghy, 1993). Whatever the reason, the failure to attempt to resolve many issues on the basis of available empirical findings is commonly accompanied by an apparent ignorance that such basic disagreements exist, although they have important consequences not only in the paralysis of theory but in unnecessary or inappropriate assessment and treatment of people with sexual problems. A number of these issues are reviewed, with the aim of encouraging an awareness of the need for their resolution.

VALIDITY OF PHALLOMETRIC ASSESSMENT OF INDIVIDUAL SUBJECTS

Evidence of the validity of penile volume assessment of the sexual orientation of individuals was provided by Freund (1963). All of 65 heterosexual men and 48 of 58 homosexual men were classified correctly. The finding was replicated. All of 11 men who identified as heterosexual were correctly classified and 17 of 22 men seeking treatment for compulsive homosexual feelings were classified as predominantly homosexual; prior to assessment, 3 of the 5 classified as predominantly heterosexual had reported that they were (McConaghy, 1967). Assessments were based on subjects' penile volume responses to films of nude men and women shown for 10–13 sec. When penile circumference assessment of sexual arousal was introduced, it was assumed that it was equivalent to penile volume assessment, though its use required much longer presentation of stimuli. In an attempt to validate the procedure, the penile circumference responses to 2-min video clips of a nude young woman failed to discriminate, even as groups, six homosexual from six heterosexual men. To discriminate them as groups, not as individuals, required 2-min presentations of videos of male homosexual or lesbian sexual activity (Mavissakalian *et al*., 1975). This finding was also replicated. Penile circumference responses to heterosexual, homosexual, and lesbian slides or

audiotapes did not distinguish eight homosexual from eight heterosexual men as groups (Sakheim *et al.*, 1985). Use of more sexually stimulating films with lesbian and homosexual, but not heterosexual, erotic content correctly identified seven of the eight heterosexual and all eight homosexual men.

This evidence of the superior validity of penile volume assessment of sexual orientation using brief duration stimuli of films of nude men and women compared with penile circumference assessment using longer duration stimuli remains unacknowledged. The need with penile circumference assessment to use longer duration erotic stimuli would appear to contribute to the greater ease of subjects to modify consciously or unconsciously their penile circumference compared to their penile volume responses. Laws and Rubin (1969) investigated four men's ability to modify their penile circumference responses. All four were able to reduce erections to erotic films 10–12 min in duration by at least 50% and to produce 30–90% of maximum erection without being shown any erotic stimuli. Freund (1971, p. 225) commented that the study did not "obey the rule of showing each picture for no more than a few seconds, to provoke only the minimal penis volume changes of which the subjects are often not aware." Freund asked subjects to attempt to produce sexual arousal by using fantasies which would be erotic for them, when shown pictures of members of the nonpreferred sex, and to attempt to diminish arousal to pictures of members of the preferred sex by imagining something disagreeable. With penile volume assessment and 13-sec presentation of stimuli, only 2 of 22 heterosexual and 1 of 9 homosexual men tested for the first time, and 3 of 20 heterosexual and 5 of 15 homosexual men previously tested were able to produce penile volume responses which misclassified them.

The reason penile circumference assessment requires longer exposure to erotic stimuli is that in many men the increase in blood flow which accompanies the development of penile tumescence is insufficient to maintain penile circumference while maintaining the rapid increase in penile length. Hence in the initial stage of tumescence in these men, as penile volume responses increase, penile circumference responses decrease. The graphic evidence that in these men the two responses were negatively related as mirror images for some seconds following presentation of the erotic stimuli (McConaghy, 1974) was questioned on the basis that no statistical comparisons were reported and the number of subjects was small (Rosen and Keefe, 1978). The irony of this methodologically formalistic questioning of the visual evidence of the negative relationship went unnoticed. At the time the use of penile circumference assessment in single-case studies in which no statistical comparisons were made and the number of subjects was small was accepted without question. Earls and Marshall (1982) subsequently used measures of penile length and circumference to demonstrate the negative correlation. No attempt was made to replicate these findings, which have remained ignored. McAnulty and Adams (1992a) commented that as penile volume is a function of circumference and length, from a physiological viewpoint, penile volume and circumference responses are necessarily highly correlated. The related failure to

acknowledge men's greater ability to modify their penile responses when longer presentations of erotic stimuli are employed has important implications for the evaluation of the evidence that behavior therapies can reduce deviant preferences of rapists and pedophiles, discussed subsequently.

The continued acceptance of the validity of penile circumference responses in assessment of individual men's sexual orientation resulted in the extension of their use to the assessment of individual men's paraphilic sexual interest. Despite the evidence from numerous studies of the limited ability of this assessment to distinguish rapists and pedophiles from controls as individuals rather than as groups (McConaghy, 1993), the assessment remains advocated and widely used both in the assessment and evaluation of response to treatment of individual sex offenders. Using meta-analysis of studies investigating the penile circumference responses of rapists Lalumiere and Quinsey (1994) demonstrated that the assessment did discriminate rapists from nonrapists, as groups. That it was necessary to combine statistically the responses of a number of groups of rapists and of nonrapists to obtain convincing evidence these groups could be distinguished at a statistically significant level meant that the assessment had little likelihood of distinguishing them as individuals. Nevertheless, the authors considered that the result supported the use of the assessment to identify individual offenders' treatment needs and risk of recidivism.

Marshall (1996, p. 168) commented, "For a test to have merit, it must be shown that it is in a standardized form that is broadly acceptable, that it is reliable and valid, and that either it is resilient to faking or faking can be reliably discerned. Unfortunately, the available data on phallometric assessments do not meet any of these empirical and technical requirements . . . the wisest course of action may be to withdraw its clinical use until more adequate data are available." However, in the New Zealand Kia Marama program for treating incarcerated child molesters in which Marshall took part (Hudson *et al*., 1995), all clients undergo phallometric testing to reveal the presence or absence of deviant attraction to children. Howes (1998) recently commented that in spite of doubts about the ability of (penile circumference) plethysmography to discriminate sexual offenders from nonsexual offenders, there is likely to be little disagreement with the observation that it has assumed the leading, if not the definitive, role in present-day assessment of deviant sexual arousal.

LACK OF VALID EVIDENCE THAT PHALLOMETRICALLY ASSESSED SEXUAL PREFERENCES CAN BE ALTERED BY TREATMENT

As Barbaree *et al*. (1995) pointed out, attempts to change pedophilia and other paraphilias were based on earlier work attempting to change homosexuality. They considered Adams and Sturgis' (1977) critical review of these studies

and accepted their conclusion that modest changes were found in the direction of increasing heterosexual arousal and decreasing homosexual arousal. Adams and Sturgis treated outcome measures of reported changes in sexual feelings and behavior as equivalent to changes in physiological measures of arousal. They did not consider the possibility that following treatment men who wished to change their sexual orientation would be motivated to reduce their homosexual and increase their heterosexual behavior. If successful, they could consider that their sexual feelings had changed in the same direction, without there being any meaningful change in their physiological sexual arousal. Even without treatment many men with homosexual feelings, if so motivated, have the potential to change their choice of same- to opposite-sex partners and, if they do, can consider that their feelings have changed. As discussed subsequently, most men who report homosexual behavior cease it as they age. Men who identify as homosexual make up only a minority of those who report past homosexual behavior and half or more of men who identify as homosexual give a history of heterosexual intercourse and arousal (McConaghy, 1993). Adams and Sturgis (1977) pointed out that, of the men treated in the studies they reviewed, a minimum of 45% had some prior experience with heterosexual dating and 30% reported prior attempts at heterosexual coitus, emphasizing that these were minimum figures, as many studies did not report the data.

Adams and Sturgis referred to a study (McConaghy, 1975) using the valid penile volume assessment of men's ratio of heterosexual/homosexual arousability, which concluded that the change in treated men's penile volume responses in the heterosexual direction was a placebo response. They gave no further attention to this possibility. In the study, which aimed to reduce homosexual and increase heterosexual arousability in men seeking this change, the men were randomly allocated to aversive therapy or to a positive conditioning procedure. The latter procedure proved ineffective and hence acted as a placebo therapy. However, the same degree of change of penile volume responses in the heterosexual direction occurred in men treated with aversive therapy and those who received the placebo procedure. It was concluded that the change was a placebo response with both procedures, resulting from the treated subjects' motivation, conscious or unconscious, to show this change following treatment. McConaghy (1976) subsequently reviewed the changes in penile volume responses of 150 men who had received treatment with an aim similar to that of the 1975 study. Thirty-one were the subjects of that study and 119 had been treated with aversive procedures in three earlier studies. Prior to treatment, 33 of the 150 men showed penile volume responses indicative of predominant heterosexuality; following treatment, 53 did. Fewer men showed this change, 20 (13%) of 150 subjects, than did the homosexual men requested by Freund (1971) to modify their penile volume responses so as to appear heterosexual without treatment. He found that 5 (20%) of 15 who had shown responses indicating a homosexual orientation when previously tested were able to do so. McConaghy (1976) concluded that the reduction in homosexual feeling and

behavior reported by men following aversive procedures was not accompanied by meaningful change in a valid physiological measure of individual men's sexual orientation.

Adams and Sturgis (1977) may have ignored the possibility that changes in reported sexual feelings and behavior did not reflect changes in physiological arousal because they were impressed by the findings of the single-case studies they reviewed which reported changes in physiological arousal using penile circumference measures. In a review of these studies, McConaghy (1977) pointed out that they were seriously flawed. With the single-case design used, a treatment was believed to be demonstrated effective if the measures of change it produced were reversed when the treatment was withdrawn and reoccurred when the treatment was reintroduced. With this design, to control for chance relationships the duration of the periods of treatment, withdrawal, and reintroduction should be fixed prior to the study, and not *post hoc* during the study. There is no control for expectancy effects. In the single-case studies reviewed by Adams and Sturgis, the treatments were not introduced and withdrawn at fixed intervals but were continued until the sought outcome (evidence of reduced homosexual or increased heterosexual arousal) occurred. The treatments were then withdrawn until the sought for reversal of the outcome occurred. The treatments were then introduced again and continued until the sought for outcome reoccurred. As the outcome measures (self-reports and penile circumference responses to pictures of female nudes) fluctuated markedly when they were assessed repeatedly throughout the single-case studies, it was inevitable that on a number of occasions they would be in the sought-for range, maximizing the possibility of their being due to chance. Inconsistencies between the outcome measures were ignored. In addition to failing to criticize the methodologically inappropriate used of single-case design, Adams and Sturgis made no reference to the lack of validity of the penile circumference measure of the sexual arousal of individual homosexual men used in the single-case studies, though as pointed out earlier, it had been demonstrated to be incapable of discriminating groups of homosexual from heterosexual men.

In an analysis of the changes following aversive therapy in men with homosexual feelings or behavior, McConaghy (1976) pointed out that they reported less preoccupation with homosexual thoughts and a weakening of compulsions to become involved in homosexual activity. To account for the paradoxical finding that increased control of compulsive activity was not associated with a meaningful change in sexual orientation as measured by the valid penile volume assessment, the hypothesis was advanced that aversive therapies did not act by reducing physiological sexual arousal. Rather they reduced the compulsivity of sexual behaviors, which was maintained by an independent physiological behavior completion mechanism (McConaghy, 1980). This mechanism was activated when initiation of a behavior was stimulated by thoughts or environmental cues. If the subject attempted not to complete the behavior, the behavior completion mechanism activated the reticular arousal system and the resulting arousal was experienced

as sufficiently aversive as to drive the subject to complete the behavior against his or her will. On the basis of the hypothesis a treatment was developed aimed not at modifying sexual feelings, but at increasing the treated subject's control of such feelings previously experienced as compulsive. The subjects were trained to relax and then instructed to visualize repeatedly, in a relaxed state, being in situations in which they had previously carried out the activities they experienced as compulsive. They were then instructed to visualize not completing the activities but carrying out some alternative activity (such as leaving the situation) while remaining relaxed. The treatment was termed imaginal desensitization, a possibly inappropriate term, as it failed to distinguish clearly the treatment from systematic desensitization. It was subsequently suggested it be termed alternative behavior completion (McConaghy, 1993).

If the behavior completion hypothesis was valid, it was considered that imaginal desensitizion would be at least as effective as aversive procedures in giving subjects control of activities previously experienced as compulsive. To investigate this, the procedure of covert sensitization was first shown in a randomized controlled study to be comparable in effect to electric shock aversion in subjects with compulsive homosexuality (McConaghy *et al*., 1981). In a second randomized controlled study imaginal desensitization compared to covert sensitization in treatment of subjects with compulsive sexuality produced a significantly greater reduction in the strength of compulsive sexual urges at 1 year following treatment, a reduction paralleled by a greater reduction in inappropriate behaviors (McConaghy *et al*., 1985). In a third randomized controlled study of treatment of men with compulsive sexuality, imaginal desensitization was shown to be in the same range of effectiveness as low-dose medroxyprogesterone therapy, another therapy aimed at giving treated subjects control rather than changing their sexual preferences (McConaghy *et al*., 1988). The behavior completion model was considered to account for the compulsivity not only of sexual but also of nonsexual behaviors. This belief was supported by further randomized controlled studies in which covert sensitization and other aversive procedures were shown to be significantly inferior to imaginal desensitization in the treatment of compulsive gamblers, the inferiority persisting for an average of 6 years following treatment (McConaghy *et al*., 1983, 1991).

The studies discussed were considered to provide evidence that it was not possible to modify homosexual or deviant sexual preferences and that treatment should be aimed at increasing subjects' control over the compulsive aspects of their sexual behaviors. This conclusion has gone unrefuted, and no replications of the studies' findings have been attempted. Instead McConaghy's earliest study was cited to support the use of aversive therapy, which was described as the primary procedure used to eliminate deviant sexual arousal (McAnulty and Adams, 1992b). McAnulty and Adams reviewed the evidence supporting the use of aversive and other therapies to modify sexual preference without referring to the major problem with this evidence. This is the ability discussed above for men to modify their penile arousal assessments without any change having occurred in their

sexual orientation. This ability is greater with penile assessments using exposure to stimuli of 30 sec or more, the duration used in almost all studies reporting changes following treatments. This greater ability has also been shown with assessment of deviant sexual preferences. When their penile volume responses to pictures of nude children and adults shown for 13 sec were measured, only 1 of 19 heterosexual pedophiles and 0 of 9 homosexual pedophiles were able to produce responses indicating a preference for adult females or males (Freund, 1971). Using penile volume assessment with 54-sec presentations of stimuli, Wilson (1998) found that 17 of 20 university students were able to fake a preference for female children or at least raise their level of arousal to the equivalent of their responses to women.

Howes (1998) compared the penile circumference responses of incarcerated male offenders to 30-sec presentations of consenting and deviant sexual stimuli. Deviant arousal was found in 86 of the total of 40 rapists and 50 nonsexual offenders. When asked to inhibit the deviant arousal to repetition of the deviant stimuli, all but 1 of the nonsexual offenders and 13 of the 40 rapists were able to do so. Howes considered the difference in these ratios meaningful and did not relate it to the fact that the IQ of the rapists was significantly lower than that of the nonsexual offenders. Nor did he report whether the rapists were to be involved in treatment which incorporated the use of penile circumference assessment to evaluate their response. If they were, they could be motivated to save their attempts to modify their responses until after treatment. Such motivation would be understandable if they were aware that if they were unable to modify their penile circumference responses in an appropriate direction following covert sensitization or electric shock aversive therapy, they could be given olfactory aversion using noxious substances (McAnulty and Adams, 1992b) or satiation therapy. With the latter therapy they would be instructed to masturbate continuously for 1 hr, whether or not they ejaculated during that time, while verbalizing every variation they could imagine concerning their deviant fantasies (Marshall *et al*., 1983). A change in subjects' penile circumference responses to paraphilic stimuli was considered acceptable evidence of the treatment's efficacy (Laws and Marshall, 1991). In relation to the use of repeated penile circumference assessments to assess response to procedures aimed at changing sexual preference, the opportunity for such repetitions to allow subjects to improve their ability to modify their responses has been given no attention. The data provided by Freund (1971) concerning the ability of heterosexual and homosexual men to modify their responses, referred to earlier, showed evidence of improved ability with repetition of the assessment.

The most commonly employed behavioral technique considered to increase paraphiliacs' heterosexual arousability is orgasmic or masturbatory reconditioning or retraining. In its original form the subject was asked to masturbate and to report when orgasm was imminent, when he was shown the picture of an attractive, scantily dressed woman until he reported ejaculation. Ten years after its introduction, Conrad and Winzce (1976) pointed out that the evidence of its

efficacy had not gone beyond the case-study level. In their own evaluation they relied upon the invalid penile circumference responses to pictures of nudes to assess individual subjects' heterosexual arousal. Laws and Marshall (1991) distinguished between the common form, in which the subjects used deviant fantasies initially, and directed masturbation, in which they were instructed to use exclusively heterosexual fantasies from the commencement of masturbation. Laws and Marshall considered that there were inadequate data to support the efficacy of the thematic shift procedure, but there was some evidence that directed masturbation might be effective. This evidence was based largely on changes in the invalid assessment of sexual arousal of individual men by their penile circumference responses to pictures of nude women. Contrary evidence indicating that orgasmic reconditioning was likely to be ineffective has remained ignored (McConaghy, 1978). Using the valid penile volume assessment, men's balance of heterosexual to homosexual feeling was investigated prior to treatment for compulsive homosexual feelings. Married men who had repeatedly experienced orgasm in the presence of female cues, *viz.*, their wives, showed no evidence of increased heterosexual arousability compared to single men without this experience. This finding would seem to have also required discussion in relation to the belief concerning pedophiles that "each time the offender has sex with a child, he obviously pairs heightened sexual arousal with vivid, realistic visions of children and the proprioceptive stimuli produced by his own actions. These contacts provide powerful conditioning trials, and if repeated often enough, should entrench a growing attraction to sex with children even in the absence of masturbating to children" (Marshall and Eccles, 1993, p. 135).

In pointing out that it is unclear whether treatment-induced changes in penile assessments should be thought of as changes in sexual preference or changes in men's ability to control arousal, Lalumiere and Harris (1998) considered that both possibilities were welcome. They presumably believed that the ability of sex offenders to learn to control their penile responses to deviant stimuli in a test situation would reflect their ability and motivation to control their sexual arousal in real situations of temptation. However, learning an ability to deceive their therapists about changes in their sexual feelings could be a negative rather than a positive therapeutic change. It was suggested (McConaghy, 1997a) that the use of behavioral reconditioning of men with deviant physiological sexual arousal patterns may have been one factor contributing to the poorer response of the treated compared to the untreated sex offenders in the Californian Sex Offender Treatment and Evaluation Project, discussed subsequently.

The failure of changes in penile circumference assessment of deviant sexual arousal to correlate with the treated subjects' behavioral outcome (Marshall and Barbaree, 1988; Rice *et al.*, 1991) has not yet led to serious questioning of the value of the procedures aimed at producing such change. These procedures such as aversive therapies or prolonged or repeated masturbation on instruction would

seem less acceptable than procedures aimed at increasing subjects' control such as imaginal desensitization. The evidence reviewed which indicates that the latter procedure has a better behavioral outcome would seem to justify replication. The conclusion of Laws and Marshall (1991) that the combination of directed masturbation and satiation needs to be evaluated in a systematic study hopefully could stimulate a randomized controlled comparison of the combination with imaginal desensitization.

FAILURE TO CONFRONT THE EVIDENCE OF LACK OF EFFICACY OF RELAPSE PREVENTION IN SEX OFFENDER TREATMENT

To my knowledge, the Californian Sex Offender Treatment and Evaluation Project (SOTEP) is the only published randomized large-scale comparison of relapse prevention with no treatment. At a mean follow-up period of 4 years, more of the 172 sexual offenders randomly allocated to the treatment sexually reoffended than did 184 offenders randomly denied treatment (Marques, 1995). Day and Marques (1998) considered that as the project was designed as a longitudinal study, definitive conclusions about the research could not be drawn until the completion of follow-up in 2000. However, Barbaree (1997) has pointed out that treatment effect is likely to be greatest immediately after treatment, and to diminish thereafter.

In a *post hoc* analysis of the study findings, Marques (1995) found that randomization had resulted in a higher percentage of mentally disordered offenders and single offenders being allocated to the treatment than to the control group. Both these groups had higher rates of reoffense than the nonmentally disordered and married offenders. Marques considered it appropriate to control for these factors in a survival analysis and concluded that this provided evidence of a treatment effect. To my knowledge, this is the only study in the literature in which the group randomly allocated to treatment had a worse outcome than that allocated to no treatment, yet it was concluded that the study showed a treatment effect. As the conclusion was based on a statistical analysis of a relationship selected *post hoc* from the examination of an unstated number of relationships, the possibility of its being due to chance is high. It would seem that the most optimistic conclusion was that the study had not excluded the possibility that treatment could be effective for married and nonmentally disordered offenders, but as the possibility was based on a *post hoc* data analysis, a further study evaluating it was required before it could be accepted. The study did provide strong evidence that for the larger group of single offenders, the treatment was at best ineffective.

In meta-analyses of the outcomes of psychological therapies, placebo therapies have consistently proved more effective than no treatment (McConaghy, 1990). That relapse prevention in this study was less effective than no treatment

and hence did not have even a placebo effect suggests that, at least for single and mentally disordered offenders, it had a negative effect. Possible reasons for relapse prevention having a negative effect have been discussed (McConaghy, 1997a). It might have been expected that the finding of the SOTEP study would have resulted in concern that this treatment, the most popular model for structuring the treatment of sexual offenders (Hudson and Ward, 1996), should continue to be used routinely for single offenders. At least, recognition of a need for the immediate initiation of additional attempts to evaluate its efficacy appropriately could be expected. No such recognition has been evident. In contrast, the *post hoc* statistical correction carried out by Marques has been considered convincing and the conclusion that the study showed a modest treatment effect accepted (Barbaree, 1997; Hanson, 1997).

The American Psychological Association Task Force on Promotion and Dissemination of Psychological Procedures (1993) considered that in establishing the efficacy of these procedures, it was insufficient to contrast them with a waiting-list control group. The reason given was that relying on such evidence would leave psychologists at a serious disadvantage vis-à-vis psychiatrists, who can point to numerous double-blind placebo trials to support the validity of their interventions. Since 1993 there has been growing acceptance by psychiatrists that treatments be "evidence-based" on randomized controlled trials (Geddes and Harrison, 1997). In contrast, in relation to sex offender treatment there has been an increasing number of recommendations that the methodological rigor of its evaluation be weakened. Marshall and Pithers (1994) considered that other than in unusual circumstances, such as those of the Marques *et al*. study, they could not see how any ethically concerned researcher would suggest a random design treatment outcome study for sexual offenders. Designs employing untreated controls were described as strong (Hanson, 1996). Though pointing out that group designs not employing randomization do not allow outcomes to be validly attributed to the treatments evaluated, Miner (1997) concluded that such group designs and single-program follow-up studies will continue to be the designs of choice for offender treatment effectiveness research. Barbaree (1997) was concerned with the possibility of false negative or Type 2 errors in the evaluation of treatment outcome and considered that nonrecidivism studies of outcome have not received the attention they deserved. Failure of nonrecidivism or process studies to provide valid evidence of treatment efficacy was discussed elsewhere (McConaghy, 1999). The belief that when authorities agree, empirical evidence is unnecessary recently received a clear endorsement: "Although Dr. McConaghy is correct in noting that I provided no empirical evidence to support the aforementioned recommendation about not administering an antiandrogen in isolation, it happens to be the position adopted by the ATSA Executive Board of Directors on February 7, 1997, persuant to the recommendations of an ad hoc committee comprised of . . ." (Prentky, 1998). The statement was completed by a list of 13 names of workers active in the area.

Meta-analysis of sex offender treatment evaluation studies has been used with insufficient rigor. Searles (1985) pointed out that comparison of the effect sizes of outcome studies with different subject types is not meaningful. Effect size is calculated by dividing the difference in the mean outcome score of the treated and comparison group by the standard deviation. If subjects are markedly homogeneous, a small treatment effect will result in a large effect size; if they are markedly heterogeneous, a large treatment effect will result in a small effect size. The conclusion that meta-analysis of the responses to treatment of markedly different groups of sex offenders provided evidence of a treatment effect (Hall, 1995) has been accepted without criticism (Barbaree, 1997; Hanson, 1997; Miner, 1997). It found the highest treatment effect sizes in studies with the highest base rates of recidivism in the comparison groups, reflecting the heterogeniety of the groups treated in the different studies. The lack of a significant relationship between the percentage of treated subjects reoffending in each study and the effect size of that study further indicated that offender characteristics rather than the nature of the treatment were a major factor determining effect sizes. Inaccurate transcriptions and arbitrary treatment of findings of the studies analyzed (McConaghy, 1999) went unnoticed. It is clearly a demanding task for journal reviewers of meta-analyses to check that the outcome data of all studies analyzed were accurately and meaningfully transcribed. However, this would seem necessary for appropriate review. Pointing out the discrepancies between the findings of meta-analyses of studies with relatively small subject numbers and those of later large-scale controlled randomized studies, Borzak and Ridker (1995) concluded the findings of meta-analyses must be considered hypothesis-generating rather than hypothesis-testing. They considered differing results from megatrials and meta-analyses should be seen less as discrepancies and more as the outcome of an experiment whose results (produced by the large trial) differ from the hypothesis (produced by the meta-analysis).

THE NONGAY MAJORITY WITH HOMOSEXUAL FEELINGS

From an analysis of the Kinsey data, Van Wyk and Geist (1984) concluded that the popular conception of heterosexuality and homosexuality as polar extremes was accurate. McConaghy (1987) pointed out that their analysis supported the contrary conclusion, that heterosexuality/homosexuality was not a dichotomy but a continuum. Of the men and women who reported some homosexual behavior, almost twice as many reported predominant heterosexual behavior as did those who reported exclusively homosexual behavior, indicating a dimensional rather than a dichotomous distribution. McConaghy (1987) reported the anonymous responses to a questionnaires of medical students enrolled over 3 years. One item requested students to report the degree to which they were currently sexually attracted to members of the same versus those of the opposite sex on an 11-point scale from

0/100 to 100/0. Over the 3 years 75 to 85% of all enrolled students completed the questionnaire; 48 to 51% of female and 33 to 45% of male medical students reported awareness of some sexual attraction to members of the same sex. The majority of students reporting this awareness reported predominant sexual attraction to members of the opposite sex. Consistently over the 3 years, the ratio of homosexual to heterosexual attraction in the male students correlated significantly with a number of opposite sex-linked behaviors, supporting the validity of their reports. The findings were consistent with Kinsey's concept of a heterosexual–homosexual continuum. Subsequently male twins educationally more representative of the general population were requested to complete the same questionnaire anonymously (McConaghy *et al*., 1994). Of the 419 (69% of those requested) who did so, 12.2% reported some current awareness of homosexual attraction. As in the previous study the majority of those with this awareness (7.5%) reported predominant heterosexual attraction, compared to 4.6% who reported bisexual or predomination homosexual attraction. Significant correlations were again found between the men's reported ratio of homosexual to heterosexual attraction and opposite sex-linked behaviors. These correlations remained present when the men reporting bisexual or predominant homosexual attraction were excluded, supporting the validity of the reports of homosexual feelings in the predominantly heterosexual men.

Laumann *et al*. (1994) investigated by interview and anonymous questionnaire a representative sample of 3432 men and women aged 18 to 59 in the United States (78.5% of the eligible persons selected). In a number of analyses they combined their data with those from General Social Surveys from 1988 to 1993. Bisexual behavior decreased with age. In the periods since puberty, since age 18, in the previous 5 years, and in the previous year, it was reported by 5.8, 4.0, 2.1, and 0.7% of men and 3.3, 3.7, 1.4, and 0.3% of women, respectively. Exclusive homosexual behavior increased with age. It was reported by 0.6, 0.9, 2, and 2% of men and 0.2, 0.4, 0.8, and 1% of women, respectively, in the same four periods. Only with regard to sexual behavior in the last year, when the percentage of subjects reporting any homosexual behavior was reduced from 6.4 to 2.7% for men and 4.6 to 1.3% for women, did the percentage with same-sex partners exceed those with partners of both sexes. Of the 6.3% of men aware of current homosexual attraction, 2.6% reported that they were mostly and 0.6% that they were equally attracted to women, with 0.7% mostly and 2.4% exclusively attracted to men. Of the 4.4% of women aware of current homosexual attraction, 2.7% were mostly and 0.8% equally attracted to men, with 0.6 mostly and 0.3% exclusively attracted to women. That is, a slightly higher percentage of the men and a markedly higher percentage of the women were aware of predominantly heterosexual or bisexual than of predominantly or exclusively homosexual feelings.

The total percentage of men aware of homosexual feelings of 6.3% in the Laumann *et al*. study was markedly less than the 12.2% of male twins (McConaghy *et al*., 1994). This could reflect the more representative nature of the subjects in

the study by Laumann *et al*., but the fact that the data were not collected totally anonymously may have contributed. Comparable percentages to those reported by McConaghy *et al*. were found in representative samples of the populations of the United States, the United Kingdom, and France aged 16 to 50 investigated by a questionnaire considered to provide anonymity (Sell *et al*., 1995). Nevertheless, the consistent finding of McConaghy *et al*. (1987, 1994) that homosexual feelings were distributed dimensionally, with the majority of subjects with such feelings being aware of equal or predominantly heterosexual feelings, was supported by the Laumann *et al*. study. A similar finding was also reported in the study by Johnson *et al*. (1994), which investigated in face-to-face interviews the sexual behavior of 18,876 men and women, a 63.3% representative sample of the population of Great Britain aged 16–59. Of the 5.5% of men and 4.5% of women who reported awareness of homosexual attraction, 73% of the men and 84% of the women reported they were mostly aware of heterosexual attraction. Of the 5.2% of men and 2.6% of women who reported homosexual experience, 75% of the men and 85% of the women reported mostly heterosexual experience. A reduction in homosexual behavior with age was also found; having at least one homosexual partner with whom there was genital contact, ever, in the last 5 years, and in the last year, was reported, respectively, by 3.5, 1.4, and 1.1% of men and 1.7, 0.6, and 0.4% of women.

The popular conception of heterosexuality and homosexuality as polar extremes (Van Wyk and Geist, 1984) can be accounted for by the finding that only a minority of men and women involved in homosexual behavior or aware of homosexual feelings identify as homosexual or bisexual. In the study by Laumann *et al*. (1994), 150 (8.6%) of 1749 women and 143 (10.1%) of 1410 men reported having same-gender sex partners, and/or being aware of same-gender desire since the age of 18, and/or having same-gender (homosexual or bisexual) identity. Of the three aspects, 59% of the 150 women reported only same-gender desire, 13% only same-gender behavior, 13% same-gender desire and behavior, 15% same-gender desire, behavior, and identity, and 1% same-gender desire and identity. Of the 143 men, 44% reported only same-gender desire, 22% only same-gender behavior, 6% same-gender desire and behavior, 24% same-gender desire, behavior, and identity, 1% same-gender desire and identity, and 2% only same-gender identity. Hence of those subjects reporting at least one of the three aspects of homosexuality, only 16% of the women (1.4% of the total sample) and 27% of the men (2.7% of the total sample) identified as homosexual or bisexual. Of the 65 men and 89 women who reported homosexual desire without behavior, only 2 men and one woman identified as homosexual or bisexual.

The evidence of the dimensional nature of heterosexuality/homosexuality, with a predominance of men and women aware of homosexual feelings being aware of stronger heterosexual feelings, remains widely ignored. This is probably most important in studies investigating possible biological determinants of homosexuality. Such determinants would seem likely to act by producing homosexual

feelings. Whether the subject with such feelings acted upon them would be strongly influenced by such factors as their readiness to conform to social and religious views about the acceptability of homosexual behaviors. Blanchard and Bogaert (1996) investigated birth order as an etiological factor in male homosexuality. They treated homosexuality and bisexuality as categorically opposed to heterosexuality, allocating their subjects to one or the other group, on the basis of their self-identification. In view of the finding of Laumann *et al.* (1994) that only a quarter of men with homosexual feelings or behavior identify as homosexual, the majority, particularly those with homosexual feelings only, would either classify themselves as heterosexual or not volunteer for the study. If biological factors contribute to the etiology of homosexual feelings, this methodology would result in failure to investigate its role in these men who make up a significant percentage, indeed probably the majority, of men with homosexual feelings, as they would be inappropriately classified or excluded. Blanchard and Bogaert suggested that their finding that homosexual men have more older brothers could be explained by a maternal immune reaction provoked only by male fetuses and becoming stronger after each pregnancy with a male fetus. Before this possibility is considered it will be necessary to demonstrate that the birth order is influencing men's homosexual feelings, not their willingness to identify as homosexual. Sulloway (1996) reviewed evidence indicating that relative to their younger siblings, eldest children were more conforming, conventional, defensive, and less likely to take risks. He considered that laterborns were more likely to experiment sexually, and this could help to explain the higher rate of homosexuality reported among laterborn males. Though Sulloway incorrectly accepted that the relationship with birth order was with homosexual behavior rather than with identification as homosexual, his explanation could equally account for men with homosexual feelings who were laterborn being more likely than those firstborn to identify as homosexual.

In reporting their search for the gay gene, Hamer and Copeland (1994) commented of the four Kinsey scale measurements (self-identification, fantasy, attraction, and behavior) of the male volunteers they studied that the striking feature was that almost all the men were easily categorized as either gay or straight, with few, if any, in between. The group studied consisted of 114 self-acknowledged homosexual men and 142 relatives over age 18, 99 male and 43 female. Assessment of the subjects' degree of attraction to members of the same and opposite sex was based on a structured interview conducted in private. A graph of the attraction scores of the males (Hamer *et al.*, 1993) showed that most men who reported any attraction to men reported predominant or exclusive attraction to them. The authors showed no awareness that this finding conflicted with the finding in representative population samples that most men reporting attraction to men report equal or predominant attraction to women. In concluding it was appropriate to treat sexual orientation as a dimorphic rather than a continuous variable trait, the authors cited two studies

(Pillard and Weinrich, 1986; Bailey and Pillard, 1991) which also reported similar bimodal distributions of Kinsey scores in males. The subjects of these studies were also men who identified as homosexual or heterosexual. The method of selection of the subjects in all these studies would appear to have resulted in their investigating not homosexuality but identification as homosexual.

Evidence that the majority of men and women with homosexual feelings are aware of predominant heterosexual feelings and those with homosexual experience have predominant heterosexual experience is also ignored in theories of the development of sexual identity. Failure of men and women with homosexual feelings or behavior to identify as gay has been considered indicative of pathology. Minton and McDonald (1983/1984), in pointing out that many people engage in same-sex acts without necessarily identifying as homosexual, considered that rejection of a homosexual identity could be a result of homophobia and the discrepancy experienced as identity confusion. Troiden (1989) likewise pointed out that only a small portion of all people who have homosexual experiences ever adopt lesbian or gay identities and corresponding lifestyles. He also considered that they experienced identity confusion, which they reduced by defining themselves as ambisexual: "I guess I'm attracted to both women and men." He stated that this might or might not reflect their actual sexual interests and considered that adolescent gay males and lesbians who are gender typical, heterosexually active, and homosexually inexperienced were more confused regarding their sexual identities because their characteristics were at variance with prevailing stereotypes. He did not suggest that this situation might be best dealt with by attempting to change the stereotypes to acknowledge that the majority of men and women with homosexual feelings or who have carried out homosexual acts are predominantly heterosexual in feelings and will become exclusively heterosexual in behavior. Minton and McDonald (1983/1984) commented that individuals with homosexual preferences usually have only limited opportunities during adolescence to explore and act on their homosexual feelings. They failed to relate this to the evidence that more males, the subjects of most studies of homosexual identity formation, act on their homosexual feelings in adolescence than subsequently. Both these analyses of sexual identity appeared to assume that heterosexuality and homosexuality are categorically opposed, so that the majority of persons with homosexual feelings or behavior who are predominantly heterosexual could be regarded as homophobic or confused homosexuals.

Another aspect of the development of sexual orientation and identity which would seem to require investigation is the reduction of the percentage of men and women engaging in homosexual behavior with age. A significant percentage of the medical students and male twins investigated by McConaghy and colleagues (1987, 1994) reported that they were not currently aware of homosexual feelings they experienced in adolescence, indicating that homosexual feelings diminish or disappear with age in a proportion of the population.

Rejection of the dimensional nature of heterosexuality/homosexuality can take the form of a rejection of the existence of bisexuality. Altshuler (1984) stated that if the heterosexual–homosexual range were truly a continuum, bisexuals at the midpoint should show an equal frequency and pleasure with either sex and, therefore, an equal preference and relatively random choice in the sex of their partners. He considered the majority of his 13 subjects who identified as bisexual to be clearly predominantly homosexual, and that bisexuality as he defined it did not exist. McConaghy (1987) considered this concept of bisexuality naive, equivalent to arguing that for a person to like herrings and caviar equally, he or she must eat equal quantities of both. At least in men, heterosexual compared to homosexual relations are socially approved and enhancing of self-esteem. Hence men who are bisexual in the sense that they enjoy equally sexual relations with men or women and who are not self-punishing would select more opposite-sex partners. Men who choose equal number of partners of both sexes and so identify as bisexual are likely to be those with much stronger homosexual than heterosexual feelings.

The findings of the study by Laumann *et al*. (1994) indicate that the tendency of subjects with bisexual feelings to choose more opposite-sex partners is shown less by women, presumably reflecting a greater social tolerance of lesbian relations and/or a greater readiness of lesbian women to act independently of social attitudes. Current feelings were exclusively homosexual, predominantly homosexual, equally bisexual, predominantly heterosexual, and exclusively heterosexual in 2.4, 0.7, 0.6, 2.6, and 93.8% of men and 0.3, 0.6, 0.8, 2.7, and 95.6% of women, respectively. Any same-sex behavior in the previous year was reported by 2.7% of men and 1.3% of women. Hence the percentage of men reporting homosexual behavior was less than the percentage of men whose feelings were predominantly or exclusively homosexual (3.1%), whereas the percentage of women reporting homosexual behavior was greater than the percentage of women with predominant or exclusive feelings (0.9%). These figures suggest that few, if any, men but some women with equally bisexual or predominantly heterosexual feelings continue to have sex with same-sex partners as they age. Two and seven-tenths percent of men and 1.4% of women in the study by Laumann *et al*. identified as homosexual or bisexual, virtually the same percentage that reported any same-sex behavior in the previous year. It is likely that they were mainly the same subjects. As pointed out earlier, the sexual identity of the subjects in the study was related to their behavior rather than their feelings, with extremely few men or women aware of homosexual feelings identifying as homosexual or bisexual if they had not been involved in homosexual behavior. If few men with equally bisexual feelings were involved in homosexual behavior, few would identify as bisexual. Altshuler (1984) would appear to be correct in believing that the majority of men who identify as bisexual are predominantly homosexual in their feelings but incorrect in concluding that bisexuality of feelings does not exist.

The existence of bisexuality has also been questioned on the basis of penile circumference assessment. As pointed out earlier this assessment was unable to discriminate groups of homosexual from heterosexual men using their responses to slides of nude women or of heterosexual, lesbian, or male homosexual activity. It was necessary to use the more arousing stimuli of films of lesbian or male homosexual sexual activity to discriminate them as individuals. Penile volume responses can make this discrimination using films of nude men and women. Tollison *et al.* (1979) were unable to distinguish groups of men who reported bisexual behavior and feelings from those who reported exclusive homosexual behavior and feelings by their penile circumference responses to explicit homosexual and heterosexual films. Rather than consider this possible further evidence of the limited sensitivity of penile circumference assessment, they questioned the existence of bisexuality. They further claimed there was to that date no physiological evidence for bisexual arousal except where this was a by-product of sexual reorientation therapy. In fact two studies (Barr and McConaghy, 1971; McConaghy, 1978) had provided such evidence, demonstrating a relationship between penile volume responses indicative of bisexuality to films of nude men and women and degree of experience of heterosexual intercourse in men reporting homosexual feelings or behavior. Subsequently the degree of bisexuality assessed by the penile volume responses of individual sex offenders was shown to correlate with their reported ratio of heterosexual to homosexual feelings (McConaghy and Blaszczynski, 1991).

Hershberger (1997) commented that defining homosexuality is not difficult, that "homosexuality" involves the experience of feeling sexually attracted to people of the same sex, that a "homosexual" is a person who has these sexual feelings, and that, in his opinion, most people would readily recognize this definition. He was critical of the belief of "social constructionists" that you are only a homosexual if society has defined you to be one. However, it does seem that the definitions which he points out are the standard dictionary definitions and not appropriate for the men and women predominantly sexually attracted to people of the opposite sex who make up the majority of those attracted to people of the same sex. As they show significantly less opposite sex-linked interests, behaviors, and/or sexual identity compared to those with predominant homosexual feelings (McConaghy, 1987; McConaghy *et al.*, 1994), their failure to define themselves as homosexual does not seem unreasonable. To so define them preserves the stereotype, which is in need of change. At the same time, it would seem unacceptable from a scientific, if not a social, perspective that they be written out of the discourse on homosexuality as they are in the studies of men who self-identify as heterosexual compared to men who self-identify as homosexual or bisexual. These studies exclude or treat as heterosexual this majority of men who have homosexual feelings or have carried out homosexual behavior. Such studies are investigating not homosexuality but self-identification as homosexual.

IS THE HOMOSEXUAL AN ELABORATION OF 19TH CENTURY MEDICAL DISCOURSE?

An independent writing-out of the homosexual was initiated by Foucault (1978). As Murray (1995) pointed out, Foucault (1978, pp. 42–43), in his examples of the elaboration of European medical discourse about sexualities during the closing decades of the 19th century, claimed that

> as defined by the ancient civil or canonical codes, sodomy was a category of forbidden acts; their perpetrator was nothing more that the juridical subject of them. The nineteenth-century (northern European and American) homosexual became a personage, a past, a case history, and a childhood in addition to being a type of life, a life form, and a morphology. . . . The sodomite had been a temporary aberration; the homosexual was now a species.

Murray commented that Foucault's statement seems to have clouded the minds of a goodly number of social historical theorists and researchers, who have produced a voluminous discourse notable for refusing to look at evidence of lexicalized conceptions of homosexual "species" earlier and/or elsewhere. In addition to reviewing terms for various male homosexual roles from a range of Islamic societies across a millennium, Murray pointed out that the absence of terms does not prove the absence of a phenomenon and that cultures frequently contain covert categories.

As a covert category the homosexual would be more likely to be referred to in covert literature. In *Fanny Hill*, first published in England in 1749, and described as a secret classic, the heroine reports spying on two young men when

> the elder began to embrace, to press and kiss the younger . . . as made me conclude the other to be a girl in disguise: a mistake that nature kept me in countenance for, for she had certainly made one, when she gave him the male stamp. . . . When I came home again, and told Mrs. Cole this adventure, she very sensibly observ'd . . . that, as to the thing itself, the less said of it the better . . . that among numbers of that stamp whom she had known, or at least universally under the scandalous suspicion of it, she would not name an exception hardly of one of them, whose character was not, in all other respects, the most worthless and despicable that could be, stript of all the manly virtues of their own sex, and fill'd up with only the worst vices and follies of ours: that, *in fine*, they were scarce less execrable than ridiculous in their monstrous inconsistence, of loathing and condemning women, and all at the same time apeing their manners, airs, lips, skuttle, and, in general, all their little modes of affectation, which becomes them at least better, than they do these unsex'd, male-misses. (Cleland, 1970, pp. 189–190)

Representations which seem covertly to suggest the existence of the homosexual as a person occur at times in socially accepted literature. Miss Wade, in Dickens' *Little Dorrit*, is presented as having characteristics which are unwomanly and therefore associated with evil. The girl nicknamed Pet who is loved by the hero is "dimpled and spoilt with an air of timidity and dependence which is the best weakness in the world, giving her the only crowning charm a girl so pretty and pleasant could have been without." The heroine, Little Dorrit is labeled womanly because she is happy to sacrifice her life to her mean-spirited and self-indulgent father, described as well-looking though in an effeminate style. When she falls in

love with the hero her bosom would joyfully throw itself before him to receive a mortal wound directed at his breast, with the dying cry, "I love him." She believes that if you love anyone, you would no more be yourself, but you would quite lose and forget yourself in your devotion to him. Miss Wade is independent, strong, and forceful, and travels alone. She is consistently described as handsome, in contrast to the feminine women, who are pretty, but her beauty has a wasted look. She is proud, a negative quality in her, while it is positive in men. The nobility of the lower-class admirer of Little Dorrit is indicated by his protestation that "in my poor humble way, sir, I'm too proud and honourable to do it," which moved her father (and presumably, Dickens hoped, the reader) to tears. In addition to the presentation of Miss Wade's characteristics as unfeminine, terms implying unacceptable feelings or behaviors are regularly used in relation to her. Her early statement about the inevitability of events is said to imply that the events will be necessarily evil. When she encounters Tattycorum, a maid also described as handsome with lustrous dark hair and eyes, sullen and passionate, who complains of her treatment by her employers, she looked at the girl "as one afflicted with a diseased part might curiously watch the dissection and exposition of an analogous case." Later she writes to Tattycorum offering that if she felt herself hurt, she could go to Miss Wade and be considerately treated. When Tattycorum tells her mistress, Pet, of this and that she had met Miss Wade, the mistress asks her to take her hands away, adding that she scarcely liked to think of Miss Wade being so near her without her knowing it. Her father says of Tattycorum that she was a girl who might be lost and ruined if she wasn't among practical people (themselves). Subsequently the maid goes to stay with Miss Wade, and her employer and the hero visit to ask her to return. Miss Wade points out to the ex-maid that she can return to accept her employer's condescension, to be again the foil to his pretty daughter, and again have the droll name, Tattycorum, he has given her, setting her apart and reminding her of her lowly birth. The employer asks Tattycorum to consider what is in Miss Wade's heart, adding "What can you two be together? What can come of it?" He tells Miss Wade that she cannot hide what a dark spirit she has within her and that "if it should happen that you are a woman who, from whatever cause, has a perverted delight in making a sister-woman as wretched as she is (I am old enough to have heard of such), I warn her against you, and I warn you against yourself." Several references are made to the power of Miss Wade over Tattycorum, and the scene ends with Tattycorum refusing to return to her employer and Miss Wade putting her arm about her as if she took possession of her for evermore. Miss Wade later writes that they have been together ever since, sharing her small means.

The unexplained statement, "I am old enough to have heard of such (women)," has a possible parallel in Antonio's unexplained response to Bassanio in Shakespeare's "The Merchant of Venice": "I am a tainted wether of the flock, meetest for death." Antonio, a wealthy older man, is introduced in the company of young male friends, complaining of feeling sad, provoking one to suggest that he

is in love, treated as a joke. His unexplained response is made in the context of his having to repay a debt with a pound of his flesh, which will result in his death. He has borrowed the money for the younger Bassanio, to whom he is represented as devoted. When Bassanio requests the loan, he says that he already owes Antonio "the most, in money and in love." Antonio offers "my purse, my person, my extremest means" to help. Bassanio reveals his plan to pay his debts by marrying a wealthy and beautiful woman and will need money to maintain an appropriate appearance. Antonio repeats that his love is such that he will do whatever Bassanio wants, and as his fortunes are all in overseas ventures, he will need to borrow the money, which he does on terms requiring the above fatal payment if the debt is not repaid by a fixed date. When he cannot and he is to make the payment, he says, "Pray God, Bassanio come to see me pay his debt, and then I care not!" When Bassanio arrives, Antonio asks him to tell his wife of Antonio's end and "say how I lov'd you, speak me fair in death; and when the tale is told, bid her judge whether Bassanio had not once a love. Repent not you that you shall lose your friend, and he repents not that he pay your debt; for if the Jew do cut but deep enough, I'll pay it instantly with all my heart." Bassanio replies that he would sacrifice his life, his wife, and all the world to deliver Antonio.

Jane Austen, at the age of 16, in her *History of England*, written for the private amusement of her family, indicated an awareness of the homosexual as a person rather than as someone carrying out a temporary aberration. In the chapter on James the 1st she comments, "Sir Henry Percy tho' certainly the best bred Man of the party, had none of that general politeness which is so universally pleasing, as his Attentions were entirely confined to Lord Mounteagle." She says of James, "His Majesty was of that amiable disposition which inclines to Friendships, & in such points was possessed of a keener penetration in discovering Merit than many other people." Her biographer, David Nokes (1997), considered that in using the word penetration, she was bold enough to emphasize the point with another double-entendre. Certainly in *Mansfield Park*, when she wished to indicate the damaging effect of being reared by adults lacking appropriate ethical and social values, she allows Miss Crawford reared, by her immoral Admiral uncle, to make an inappropriate joke concerning naval life: "Of *Rears* and *Vices*, I saw enough. Now, do not be suspecting me of a pun, I entreat." In her history Austen continued with a charade on the word carpet, with the explanation that one of the principal favorites of his Majesty, Car, was his pet, adding that the other was George Villiers, afterward Duke of Buckingham. Green (1945, p. 456), in *A Short History of the English People*, written in 1874, after commenting that James was held, though unjustly, to be a drunkard, and was suspected of vices compared with which drunkenness was almost a virtue, adds that "all real control over affairs was . . . entrusted to worthless favourites whom the King chose to raise to honour. A Scottish page named Carr was created Earl of Rochester. . . ." In relation to Buckingham, he quotes a statesman of the time as finding unprecedented Buckingham's rise to

power and fortune solely on the basis of his beauty and gracefulness of person. Green adds that "the selfishness and recklessness of Buckingham were equal to his beauty; and the haughty young favourite on whose neck James loved to loll, and whose cheek he slobbered with kisses, was destined to drag down in his fatal career the throne of the Stuarts" (p. 457).

In view of the less specific nature of these references in the accepted literature, their relevance to contemporary attitudes has to be interpreted with caution. At least in Elizabethan times, a distinction appears to have existed between homosexual love and desire. In one of his sonnets addressed to a young man, "the master-mistress of my passion," Shakespeare emphasized this, at least for himself: "And for a woman wert thou first created; Till Nature, as she wrought thee, fell a-doting, And by addition me of thee defeated, By adding one thing to my purpose nothing. But since she prick'd thee out for women's pleasure, Mine by thy love, and thy love's use their treasure." A clear refutation of Foucault's belief may be found only in such writings as that of Cleland.

LIMITATIONS OF THE DSM CLASSIFICATION OF SEXUAL AND GENDER IDENTITY DISORDERS

The sexological scientific literature compared with that investigating other areas of human behavior rarely employs standardized interviews and diagnostic criteria, in particular, those of the DSM (McConaghy, 1998a). It was suggested that clinicians may have learned that, with unstandardized interviews, they can modify the course of the interview to obtain information that the patient would be reluctant to reveal in the more inflexible standardized interview. However, use of standardized diagnostic criteria has no equivalent disadvantage and has the advantage that they make clear the nature and severity of the symptoms of patients given a particular diagnosis in clinical and research communications. Neglect of the DSM diagnostic criteria of sexual disorders was attributed not to any aversion to the employment of standardized diagnostic criteria, but to the failure of those in the DSM to provide appropriate or adequate operational definitions.

Major issues in diagnosis of dysfunctions were left to the clinician's judgment. In DSM-IV diagnosis of female orgasmic disorder, the woman's orgasmic capacity is to be judged in relation to her age, sexual experience, and the adequacy of sexual stimulation she receives. Diagnosis of male orgasmic disorder requires taking into account the man's age, and whether the stimulation is adequate in focus, intensity, and duration. Diagnosis of premature ejaculation has to take into account factors that affect duration of the excitement phase, such as age, novelty of the sexual partner or situation, and recent frequency of sexual activity. As no operationally defined criteria are provided to establish how these judgments are to be made, the reliability of the diagnoses made by different clinicians is likely to be poor.

The term paraphilia was introduced in the DSM-III-R to replace the older term sexual deviation, the reason given being that it emphasized that the deviation (para) lay in that to which the subject was attracted (philia), namely, sexual objects or situations which were not part of normative arousal–activity patterns. Use of the term was retained in the DSM-IV, though this reason was no longer given, presumably because it was recognized that such sexual objects or situations are frequently part of normative arousal–activity patterns. They are reported as sexual fantasies by a significant number, possibly the majority of normal subjects (McConaghy, 1993). In addition, many adolescents not only experience such attractions but express them in behaviors (Person *et al*., 1989; Templeman and Stinnett, 1991). The older term, sexual deviation, had the advantage that it indicated no more than that the behaviors concerned deviated from those considered acceptable at the time. In the last few decades, masturbation and homosexuality have ceased to be regarded as deviant. The reason for the change would justify a similar change for other paraphilias, including sadomasochism and transvestism, discussed subsequently.

The DSM-IV diagnosis of paraphilia requires both that the sexually arousing fantasies, sexual urges, or behaviors have been present over a period of at least 6 months and that they cause clinically significant distress or impairment in social, occupational, or other important areas of functioning. However, as the DSM-IV points out, many individuals with these disorders assert that the behavior causes them no distress and that their only problem is social dysfunction as a result of the reaction of others to their behavior. Investigations of the prevalence of child–adult sexual activity indicate that the majority of perpetrators have not been detected (McConaghy, 1993) and so have not been exposed to the reaction of others. When some are detected some time after the offense, they appear to have previously shown no clinically significant distress or impairment in social, occupational, or other important areas of functioning, so excluding them from the DSM-IV diagnosis of pedophilia. This exclusion would also apply to those adult sex offenders against children who have carried out an isolated act without awareness of recurrent, intense sexually arousing fantasies or sexual urges concerning prepubescent children. Marshall and Eccles (1991) commented that many rapists, incest offenders, exhibitionists, and a substantial number of nonfamilial child molesters do not display or report deviant sexual preferences and yet they persistently engage in sexually offensive behaviors, so that most clinicians tend to ignore DSM diagnoses.

Sexual assault, the offense which, along with child molestation, most commonly results in incarceration, is not classified in the DSM-IV as a paraphilia. The DSM-III-R classified as sadists those rapists who were considered to inflict suffering on the victims far in excess of that necessary to gain compliance and in whom the visible pain of the victim was sexually arousing. This was considered to apply to less than 10% of rapists. The DSM-III-R further stated that some rapists were sexually aroused by coercing or forcing a nonconsenting person to engage

in intercourse and could maintain sexual arousal while observing the victim's suffering, but unlike persons with sexual sadism, they did not find the victim's suffering sexually arousing. Presumably rape was excluded from the paraphilias in the DSM to accommodate the feminist theory criticized by Palmer (1988) that rape is not sexually motivated but is an expression of patriarchal power. Apart from the cogent criticisms he made, the theory appears totally implausible in view of the evidence that a significant percentage of normal men report experiencing sexually arousing rape fantasies and, in laboratory studies, experience sexual arousal and demonstrate genital arousal to descriptions of forceful rape (McConaghy, 1993). From their findings, Heilbrun and Leif (1988) concluded that there was a sadistic component to normal male sexuality. The DSM-IV stated with regard to sadists that some act on their sexual sadistic sexual urges with nonconsenting victims, possibly allowing the retention of the concept of sadistic rapists advanced in the DSM-III-R. However, there appears to be no empirical evidence to support the DSM-III-R distinction between sadistic and nonsadistic rapists, and Knight and Prentky (1990) were unable to substantiate it in a prison population. Some rapists who did not inflict severe physical damage on victims nevertheless appeared to be motivated by sadistic or angry fantasies. Equally there is no evidence that self-identified sadomasochists compared with men representative of the normal population show a greater likelihood of raping nonconsenting subjects. It would seem appropriate that rape be considered an independent paraphilia, rather than a form of sadomasochism or a non-sexually motivated act.

In relation to sadism, the DSM-IV states that usually the severity of sadistic acts increases over time, and when severe and especially when associated with antisocial personality disorder, individuals with sexual sadism may seriously injure or kill their victims. As subjects with masochism and/or sadism rarely seek treatment or are criminally charged, information concerning the condition is obtained largely by investigating members of "S and M" clubs. This reveals its benign nature, with beating, bondage, and fetishistic practices being common, and more extreme and dangerous practices rare (McConaghy, 1997b). Members of these clubs were of above-average intelligence and social status and most wished to continue sadomasochistic activities. In view of the rarity with which sadomasochists seek medical treatment, few must suffer significant physical damage. The statement in the DSM-IV that usually the severity of the sadistic acts increases over time may apply to the condition of serial or sadistic murderers, the rarity of whom is masked by the worldwide media attention they receive. However, the statement is made in relation to sadism generally. There is no evidence that serial murder is on a continuum with the sadism of the much greater number of people who identify as enjoying "S and M," and the two conditions appear sufficiently independent to be defined separately. In view of its lack of relationship with psychiatric pathology, the reason homosexuality was no longer considered a disorder, it would seem reasonable that sadomasochism should also not be classified as a DSM disorder.

Possibly this was the intention when the requirement that paraphilias cause clinically significant distress or impairment in social, occupational, or other important areas of functioning was included in the DSM criteria. However, this is not indicated in the DSM-IV descriptions of sexual masochism and sexual sadism, which deal with the common form of these conditions.

The abandonment in the DSM-IV of the terms transvestism and transsexualism, widely accepted by both clinicians and subjects with these conditions, seems likely to confuse rather than clarify their diagnosis. Transsexualism was included with cross-gender identification in children as gender identity disorder and the syndrome of adult transvestism could be diagnosed only as gender identity disorder or transvestic fetishism. In relation to the decision as to which diagnosis to make, DSM-IV states that transvestic fetishism occurs in heterosexual (or bisexual) men for whom the cross-dressing behavior is for the purpose of sexual excitement. However, it also points out that sexual arousal to cross-dressing diminishes or disappears in some transvestic fetishists whom it contrasts with those in whom gender identity disorder may emerge. The term emphasizing the fetishistic aspect of the behavior is objectionable to most adult transvestites. If they still experience sexual excitement with cross-dressing, they dismiss it as insignificant or distracting and do not consider it, but rather their periodic enjoyment of the female role, the reason they cross-dress (McConaghy, 1993). As they therefore do not meet the diagnostic criterion of the DSM-IV concerning the purpose of their cross-dressing, are these men, the majority of adult transvestites, to be diagnosed as having gender identity disorder? Alternatively, in view of their lack of clinically significant distress, are they to receive no DSM-IV diagnosis? If they are not, they could be described as transvestites without the term implying that they have a disorder. This would seem the preferable solution, as transvestite sexual fantasies and behaviors are not uncommon, particularly in adolescence, and like sadomasochists, most adult transvestites do not seek treatment (McConaghy, 1993). Certainly to pathologize adult transvestism by labeling it a gender identity disorder would appear to be regressive.

Abandonment of the term transvestism in the DSM-IV, at least as a descriptive term for the behavior of most adult cross-dressers, who do not wish sex conversion, has already led to controversy. Zucker (1997) considered that Bullough and Bullough (1997) had a limited clinical understanding of the term transvestism as it is used in the DSM. He considered that in the DSM-IV transvestic fetishism is used to characterize men who at least at times are sexually aroused when they cross-dress but not forms of cross-dressing that are not accompanied by sexual arousal, as can be observed in homosexual men and women with gender identity disorder or homosexual "drag queens." He apparently did not see any difficulty with the DSM-IV statement that sexual arousal to cross-dressing diminishes or disappears in some transvestic fetishists who are not gender dysphoric or that they could be heterosexual or bisexual. As discussed earlier, most men who classify

themselves as bisexual would appear to have predominantly homosexual desire. Zucker considered that the empirical evidence adduced by Bullough and Bullough in no way demonstrated that men with a homosexual sexual orientation engaged in cross-dressing that is at least at times accompanied by sexual arousal.

Bullough and Bullough cited studies by Buhrich and colleagues, pointing out that these studies followed the DSM definition and excluded people who had no sexual arousal related to cross-dressing. Buhrich and Beaumont (1981) investigated 139 American and 97 Australian members of transvestite clubs. All were male. The 13 American and 11 Australians who reported they had not experienced fetishism to women's clothes were excluded. The remaining subjects were classified on the basis of their reported sexual preferences and behavior. When dressed as men 2% of the Americans and 1% of the Australians were bisexual, and 5% of the Australians were predominantly and 1% of the Americans were exclusively homosexual. When cross-dressed, 9% of the Americans and 7% of the Australians were bisexual, 2% of the Americans and 3% of the Australians were predominantly homosexual, and 2% of the Americans and 5% of the Australians were exclusively homosexual.

Consistent with the earlier discussions as to whether sexual orientation is dimensional or categorical and the nature of bisexuality, Sell (1997) pointed out the diagnosis of these conditions is unclear and confusing to researchers. In view of this lack of clarity, a variety of assessments was employed in an earlier study (Buhrich and McConaghy, 1977). Twelve men, all of whom first experienced ejaculation while cross-dressed, were investigated. Of the nine who dressed as men, their stated object of sexual interest when so dressed was predominantly women in six, men in one, and none in two. When dressed as women the 12 reported that this object was predominantly women in 5, men in 6, and none in 1. Kinsey scores of the 11 who reported any sexual interest were 6 (exclusively homosexual) for one, 5 for three, 3 for three, 1 for three, and 0 (exclusively heterosexual) for one. Phallometric assessment using the valid penile volume method demonstrated that two subjects showed markedly greater responses to moving pictures of nude men than women, and two the reverse, with the remainder showing relatively equivalent responses to pictures of men and women, indicative of bisexual interest.

The paradox that gender identity disorder of childhood remains classified, unlike one form of its adult expression as homosexuality, has been pointed out (McConaghy and Silove, 1991). The DSM-IV stated that gender identity disorder of childhood was not meant to describe a child's nonconformity to stereotypic sex role behavior, as, for example, in "tomboyishness" in girls or "sissyish" behavior in boys. Rather, it represents a profound disturbance of the normal sense of identity with regard to maleness and femaleness. It would that seem no scientific terms were considered necessary to replace the pejorative lay terms sissy and tomboy. These are not restricted to extreme opposite sex-linked behaviors. Sissy is applied to boys who avoid rough-and-tumble play and contact sport and show interest in housework or artistic activities, and tomboy to girls who show the opposite

behaviors. The degree to which these opposite sex-linked behaviors are shown in childhood correlate in adolescence and adulthood with the degree to which boys and men experience homosexual feelings and girls and women report masculine personality traits (McConaghy and Zamir, 1995).

Another area of dispute concerning paraphilias is their incidence in sex offenders. Abel *et al*. (1988) interviewed 561 men seeking evaluation and/or treatment for possible paraphilias who were given protection from mandatory reporting by a Certificate of Confidentiality. Less than 30% of subjects other than transsexuals had confined their deviant behavior to only one paraphilia. The authors pointed out that the findings were at variance with the traditional view of the paraphiliac, as fixated on one type of paraphilia to the exclusion of other kinds. They considered that the traditional view resulted from the use of inadequate interviewing and the lack of guarantee of confidentiality. However, if the traditional view is incorrect it is difficult to understand why offenders who are repeatedly charged are usually charged with the same offense (Day, 1994; McConaghy, 1998a). Also, evidence was reviewed (McConaghy, 1997b) indicating that as many as half the male population have carried out occasional paraphilic practices. Many of the paraphilias reported by the offenders in the study by Abel *et al*. may have been equivalent in that they were carried out infrequently and mainly in adolescence. Marshall (1996) also considered there was a need for nonoffender comparisons to evaluate the finding of Abel *et al*. In the population of sex offenders he studied few had more than one paraphilia, and only one had three, though he took careful steps to ensure confidentiality and to instill confidence in the subjects.

Mention of the DSM classification of paraphilias should not overlook its most charming feature—the unsourced statement that while the frotteur fondles his nonconsenting victim he usually fantasizes an exclusive, caring relationship with her.

THE VALIDITY OF SELF-REPORT OF SEXUAL BEHAVIORS

Lewontin's (1995) major criticism of the study by Laumann *et al*. (1994) was its reliance on self-report to provide data concerning sexual behaviors. He considered this reliance forced the interviewer to pretend that people usually know and tell the truth about important issues, a pretense which invalidated social science as a natural science. He focused on the finding that men reported 75% more sexual partners in the most recent 5 years than did women, whereas the average number of sex partners reported by men and women should, discounting homosexual partners, be equivalent. Laumann *et al*. (1994) attributed the discrepancy largely to men exaggerating or women understating the number of their partners, leading Lewontin to comment "in the single case where one can actually test the truth, the investigators themselves think it most likely that people are telling themselves and

others enormous lies" (p. 29). This skepticism of the validity of self-report was presumably a factor in his scorn of Laumann and co-workers' acceptance "without the academic equivalent of a snicker" that 45% of men aged 80–84 were still having sex with a partner. A similar attitude was shown by Sennett (1995, p. 43), who was interested and cheered by the finding, "Even if the aged have confused fantasy with fact." These statements indicated a confident belief that few, if any, older people continue sexual activity: a belief considered widespread by men and women aged 50 to over 70 (Brecher, 1984). Findings demonstrating that the survey evidence to the contrary has some validity were cited by McConaghy (1998b).

Wiederman (1997) pointed out the consistency with which discrepancies between men's and women's self-reported lifetime numbers of sex partners has been found in human sexuality research and that failure to address the issue has led to the questioning of the validity of all sex research based on self-report. In their discussion of possible reasons for the discrepancy, Laumann *et al.* (1994) included the possibility that more men may have sex with other men than women have sex with other women. The evidence of course is consistent that this is the case, and in addition, many of the men compared to women report large numbers of same sex partners (McConaghy, 1993). The fact that Laumann *et al.* raise this issue suggest that some studies may not have sufficiently investigated the sex of the intercourse partner. Wiederman does not discuss this possibility and, in his Study 1 of 151 men and 173 women, does not report if he did so. Participants were asked if they had ever experienced sexual intercourse (penis in vagina) and, if so, with how many different partners. The issue of how men and women who have sexual intercourse with same sex partners report the number and sex of such partners needs to be examined. In view of the rejection of the label of homosexual by the majority of such men and women discussed earlier, it is possible that some may report the number of their same sex partners as if they were of the opposite sex. Others may do so because they regard questionnaires investigating only opposite-sex relationships as homophobic. In Wiederman's Study 1, which apparently was of this nature, men were twice as likely as women to admit some degree of inaccuracy in their self-reports. The men who did so reported relatively high numbers of partners and accounted for the gender discrepancy in lifetime number of sex partners. Could the nature of the inaccuracy of some of these men be that they answered the question by including same sex partners? In Wiederman's Study 2, participants were asked separate questions about the number of female and male partners. In reporting the results, the number of sex partners was reported without indicating whether they were of the same sex, the opposite sex, or both. It was not possible to determine whether the possibility that men had more male partners than women had female partners contributed to the discrepancy in the number of lifetime partners reported by men compared to women.

The total number of people who report having had same-sex partners is not great—9% of men and 5% of women in a U.S. study (Laumann *et al.*, 1994) and

5.2% of men and 2.6% of women in a British study (Johnson *et al*., 1994). However, in the former study of those who since age 18 had same-sex partners compared to those who had not, the mean number of partners were 44.3 versus 15.7 for men and 19.7 versus 4.9 for women. In Johnson and co-workers' study of subjects with same-sex partners, 3.9% of men had a mean of more than 100 partners; no women had more than 20 partners. Of subjects with opposite-sex partners, 0.8% of both men and women had means of over 100 partners. Even if only a percentage of the men include the number of their male in that of their female partners, this could contribute markedly to the discrepancy in the number of partners reported by men and women.

Wiederman found that 6% of men and 7.5% of women in his Study 1 did not include casual partners in their lifetime number of sex partners, but removing these individuals did not appear to diminish the discrepancy. If some men and women do not include casual partners as sex partners, it is likely that a number of other types of sex partner would not be included. Whether sexually coersed or coersive men and women include the perpetrator or victim as sexual partners requires investigation. If men and women differ in reporting these partners, this could significantly contribute to the discrepancy in the numbers of their lifetime partners. In a study of medical students 26% of men and 31% of women reported that a member of the opposite sex and 4% of both men and women reported that a member of the same sex made constant physical attempts to have sexual activity with them, 17% of men and 9% of women experienced a person of the opposite sex being so sexually aroused that they felt it useless to stop them when they themselves did not want sexual intercourse, and 4% of men reported that they were so sexually aroused they could not stop themselves when the partner of the opposite did not want intercourse (McConaghy and Zamir, 1995). In relation to unwanted sexual experiences of university students, Muehlenhard (1989) found that 88% of women but only 27% of men felt bad or very bad about the experiences, 27% of the men feeling good or very good about them. As men are more likely than women to regard the achievement of sexual intercourse as "scoring," it is possible that more men than women would include the partners in unwanted sexual activities in the total number of their sexual partners. Muehlenhard (1988) also pointed out that men were more likely than women to overestimate their dates' interest in sex and to justify their own sexually coercive behaviors. This could result in some men regarding partners as consenting who considered that they were coerced. In such relationships the men could be more likely than the women to include the partners in the total number of their partners.

In relation to the coercion of children, it is possible that some perpetrators could number them as partners, though it would seem unlikely that victims particularly female victims would number perpetrators in this way. It has been calculated that about 5% of men and 0.5% of women molest girls (McConaghy, 1993). In the study by Laumann *et al*. (1994) 6% of men reported being touched by an older

woman when they were prepubertal; many did not appear to regard the activity as sexual abuse. Finkelhor (1985) suggested the finding that a high percentage of men who experienced sexual activity with an adult in their childhood reported that their immediate response was positive could be due to the need to maintain the male ethics of self-reliance and the portrayal of youthful sexuality in adventuresome terms. This could lead to more men than women reporting partners in their prepubertal sexual activity as lifetime sexual partners.

As Weiderman pointed out, the discrepancy in self-reported number of sex partners of men and women is one of the most troublesome examples in relation to the potential for bias and unreliability in self-reported sexual experience and warrants additional research. Such research should include the investigation of how men and women report sexual partners additional to those who conform to a conventional stereotype.

SIGNIFICANCE OF UNRESOLVED ISSUES IN SEXOLOGY

It may be that the impression is incorrect that sexology research is unlike research in other areas of science in its failure to attempt to resolve major theoretical and practical issues. However, there would seem to be a problem with sexology research indicated by the low impact factor of the major journals of sexology compared to other behavioral journals. As the impact factor of a journal measures the likelihood of the articles it contains being cited by other researches, the low factor of sexology journals reflects the low value placed on their contents by researchers, including sexology researchers. Lack of engagement in resolving major issues in sexology means that published articles repeat the same opposing opinions, which could significantly contribute to their low evaluation.

REFERENCES

Abel, G. G., Becker, J. V., Cunningham-Rathner, J., Mittelman, M., and Rouleau, J.-L. (1988). Multiple paraphilic diagnoses among sex offenders. *Bull. Am. Acad. Psychiatry Law* 16: 153–168.

Adams, H. E., and Sturgis, E. T. (1977). Status of behavioral reorientation techniques in the modification of homosexuality: A review. *Psychol. Bull.* 84: 1171–1188.

Altshuler, K. Z. (1984). On the question of bisexuality. *Am. J. Psychother.* 38: 484–493.

American Psychological Association Task Force on Promotion and Dissemination of Psychological Procedures (1993). *A report adopted by the Division 12 Board*.

Bailey, J. M., and Pillard, R. C. (1991). A genetic study of male sexual orientation. *Arch. Gen. Psychiatry* 48: 1089–1096.

Barbaree, H. (1997). Evaluating treatment efficacy with sexual offenders: The insensitivity of recidivism studies to treatment effects. *Sex. Abuse J. Res. Treat.* 9: 111–128.

Barbaree, H. E., Bogaert, A. F., and Seto, M. C. (1995). Sexual reorientation therapy for pedophiles: Practices and controversies. In Diamant, L., and McAnulty, R. (eds.), *The Psychology of Sexual Orientation, Behavior, and Identity: A Handbook*, Greenwood, Westport, CT, pp. 357–383.

Barr, R. F., and McConaghy, N. (1971). Penile volume responses to appetititive and aversive stimuli in relation to sexual orientation and conditioning performance. *Br. J. Psychiatry* 119: 377–383.
Blanchard, R., and Bogaert, A. F. (1996). Homosexuality in men and number of older brothers. *Am. J. Psychiatry* 153: 27–31.
Borzak, S., and Ridker, P. M. (1995). Discordance between meta-analyses and large-scale randomized controlled trials. Examples from the management of acute myocardial infarction. *Ann. Intern. Med*. 123: 873–877.
Brecher, E. M. (1984). *Love, Sex and Aging*, Brown, Boston.
Buhrich, N., and Beaumont, T. (1981). Comparison of transvestism in Australia and America. *Arch. Sex. Behav*. 10: 269–279.
Buhrich, N., and McConaghy, N. (1977). Can fetishism occur in transsexuals? *Arch. Sex. Behav*. 6: 223–235.
Bullough, B., and Bullough, V. (1997). Are transvestites necessarily heterosexual? *Arch. Sex. Behav*. 26: 1–12.
Cleland, J. (1970). *Fanny Hill, Memoirs of a Woman of Pleasure*, Granada, St. Albans.
Conrad, S. R., and Winzce, J. P. (1976). Orgasmic reconditioning: A controlled study of its effects upon the sexual arousal and behavior of adult male homosexuals. *Behav. Ther*. 7: 155–166.
Day, K. (1994). Male mentally handicapped sex offenders. *Br. J. Psychiatry* 165: 630–639.
Day, D. M., and Marques, J. K. (1998). A clarification of SOTEP's method and preliminary findings: A reply to Nathaniel McConaghy. *Sex. Abuse J. Res. Treat.* 10: 162–166.
Earls, C. M., and Marshall, W. L. (1982). The simultaneous and independent measurement of penile circumference and length. *Behav. Res. Meth. Instr.* 14: 447–450.
Finkelhor, D. (1985). Sexual abuse of boys. In Burgess, A. W. (ed.), *Rape and Sexual Assault*, Garland, New York, pp. 97–103.
Foucault, M. (1978). *The History of Sexuality*, Pantheon, New York.
Freund, K. (1963). A laboratory method of diagnosing predominance of homo- or hetero-erotic interest in the male. *Behav. Res. Ther*. 12: 355–359.
Freund, K. (1971). A note on the use of the phallometric method of measuring mild sexual arousal in the male. *Behav. Ther.* 2: 223–228.
Geddes, J. R., and Harrison, P. J. (1997). Closing the gap between research and practice. *Br. J. Psychiatry* 171: 220–225.
Green, J. R. (1945). *A Short History of the English People*, J. M. Dent, London.
Hall, G. C. N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *J. Consult. Clin. Psychol*. 63: 802–809.
Hamer, D., and Copeland, P. (1994). *The Science of Desire. The Search for the Gay Gene and the Biology of Behavior*, Simon & Schuster, New York.
Hamer, D. H., Hu, S., Magnuson, V. L., Hu, N., and Pattatucci, A. M. L. (1993). A linkage between DNA markers on the X chromosome and male sexual orientation. *Science* 261: 321–327.
Hanson, R. K. (1996). Evaluating the contribution of relapse prevention theory to the treatment of sexual offenders. *Sex. Abuse J. Res. Treat.* 8: 201–208.
Hanson, R. K. (1997). How to know what works with sexual offenders. *Sex. Abuse J. Res. Treat.* 9: 129–145.
Heilbrun, A. B., and Leif, D. T. (1988). Erotic value of female distress in sexual explicit photographs. *J. Sex Res*. 24: 47–57.
Hershberger, S. L. (1997). Book review: Modern homosexualities: Fragments of lesbian and gay experience. *Arch. Sex. Behav*. 26: 559–561.
Howes, R. J. (1998). Plethysmographic assessment of incarcerated nonsexual offenders: A comparison with rapists. *Sex. Abuse J. Res. Treat.* 10: 183–194.
Hudson, S. M., and Ward, T. (1996). Introduction to the special issue on relapse prevention. *Sex. Abuse J. Res. Treat*. 8: 173–175.
Hudson, S. M., Marshall, W. L., Ward, T., Johnston, P. W., and Jones, R. L. (1995). Kia Marama: A cognitive-behavioural program for incarcerated child molesters. *Behav. Change* 12: 69–80.
Johnson, A. M., Wadsworth, J., Wellings, K., Field, J., and Bradshaw, S. (1994). *Sexual Attitudes and Lifestyles*, Blackwell Scientific, Oxford.
Knight, R. A., and Prentky, R. A. (1990). Classifying sexual offenders. In Marshall, W. L., Laws, D. R., and Barbaree, H. E. (eds.), *Handbook of Sexual Assault*, Plenum Press, New York, pp. 23–52.

Lalumiere , M. L., and Harris, G. T. (1998). Common questions regarding the use of phallometric testing with sexual offenders. *Sex. Abuse J. Res. Treat.* 10: 227–237.
Lalumiere, M. L., and Quinsey, V. L. (1994). The discriminability of rapists from non-sex offenders using phallometric measures: A meta-analysis. *Crim. Just. Behav.* 21: 150–175.
Laumann, E. O., Gagnon, J. H., Michael, R. T., and Michaels, S. (1994). *The Social Organization of Sexuality*, University of Chicago Press, Chicago.
Laws, D. R., and Marshall, W. L. (1991). Masturbatory reconditioning with sexual deviates: An evaluative review. *Adv. Behav. Res. Ther.* 13: 13–25.
Laws, D. R., and Rubin, H. H. (1969). Instructional control of an autonomic sexual response. *J. Behav. Anal*. 2: 93–99.
Lewontin, R. C. (1995). Sex, lies, and social science. *N.Y. Rev.* April 20: 24–29.
Marques, J. K. (1995, September). *How to answer the question, "Does sex offender treatment work?"* Paper presented at the International Expert Conference on Sex Offenders, Utrecht, The Netherlands.
Marshall, W. L. (1996). Assessment, treatment, and theorizing about sex offenders. *Crim. Just. Behav.* 23: 162–199.
Marshall, W. L., and Barbaree, H. E. (1988). The long-term evaluation of a behavioral treatment program for child molesters. *Behav. Res. Ther*. 26: 499–511.
Marshall, W. L., and Eccles, A. (1991). Issues in clinical practice with sex offenders. *J. Interpers. Violence* 6: 68–93.
Marshall, W. L., and Eccles, A. (1993). Pavlovian conditioning processes in adolescent sex offenders. In Barbaree, H. E., Marshall, W. L., and Hudson, S. M. (eds.), *The Juvenile Sex Offender*, Guilford, New York, pp. 118–142.
Marshall, W. L., and Pithers, W. D. (1994). A reconsideration of treatment outcome with sex offenders. *Crim. J. Behav.* 21: 6–9.
Marshall, W. L., Earls, C. M., Segal, Z., and Darke, J. (1983). A behavioral program for the assessment and treatment of sexual aggressors. In Craig, K. D., and McMahon, R. J. (eds.), *Advances in Clinical Behavior Therapy*, Brunner/Mazel, New York, pp. 148–174.
Mavissakalian, M., Blanchard, E. B., Abel, G. G., and Barlow, D. H. (1975). Responses to complex erotic stimuli in homosexual and heterosexual males. *Br. J. Psychiatry* 126: 252–257.
McAnulty, R. D., and Adams, H. E. (1992a). Validity and ethics of penile circumference measures of sexual arousal: A reply to McConaghy. *Arch. Sex. Behav*. 21: 177–186.
McAnulty, R. D.., and Adams, H. E. (1992b). Behavior therapy with paraphilic disorders. In Turner, S. M., Calhoun, K. S., and Adams, H. E. (eds.), *Handbook of Clinical Behavior Therapy*, John Wiley and Sons, New York, pp. 175–201.
McConaghy, N. (1967). Penile volume change to moving pictures of male and female nudes in heterosexual and homosexual males. *Behav. Res. Ther*. 5: 43–48.
McConaghy, N. (1974). Measurements of change in penile dimensions. *Arch. Sex. Behav*. 3: 381–388.
McConaghy, N. (1975). Aversive and positive conditioning treatments of homosexuality. *Behav. Res. Ther*. 13: 309–319.
McConaghy, N. (1976). Is a homosexual orientation irreversible? *Br. J. Psychiatry* 129: 556–563.
McConaghy, N. (1977). Behavioral treatment in homosexuality. In Hersen, M., Eisler, R. M., and Miller, P. M. (eds.), *Progress in Behavior Modification, Vol. 5*, Academic Press, New York, pp. 309–380.
McConaghy, N. (1978). Heterosexual experience, marital status and orientation of homosexual males. *Arch. Sex. Behav*. 7: 575–581.
McConaghy, N. (1980). Behavior completion mechanisms rather than primary drives maintain behavioral patterns. *Activitas Nerv. Super.* (Praha) 22: 138–151.
McConaghy, N. (1987). Heterosexuality/homosexuality: Dichotomy or continuum. *Arch. Sex. Behav*. 16: 411–424.
McConaghy, N. (1990). Can reliance be placed on a single meta-analysis? *Aust. N.Z. J. Psychiatry* 24: 405–415.
McConaghy, N. (1993). *Sexual Behavior: Problems and Management*, Plenum, New York.
McConaghy, N. (1997a). Science and the mismanagement of rapists and paedophiles. *Psychiatry Psychol. Law* 4: 109–123.
McConaghy, N. (1997b). Sexual and gender identity disorders. In Turner, S. M., and Hersen, M. (eds.), *Adult Psychopathology and Diagnosis*, 3rd ed., Wiley and Sons, New York, pp. 409–464.

McConaghy, N. (1998a). Assessment of sexual dysfunction and deviation. In Bellack, A. S., and Hersen, M. (eds.), *Behavioral Assessment: A Practical Handbook*, 4th ed., Allyn and Bacon, Boston, pp. 315–341.

McConaghy, N. (1998b). Sexual dysfunction. In Hersen, M., and Van Hasselt, V. B. (eds.), *Handbook of Clinical Geropsychology*, Plenum Press, New York, pp. 239–271.

McConaghy, N. (1999). Methodological issues concerning evaluation of treatment for sexual offenders: Randomization, treatment dropouts, untreated controls, and within-treatment studies. *Sex. Abuse J. Res. Treat.* 11: 183–193.

McConaghy, N., and Blaszczynski, A. (1991). Initial stages of validation by penile volume assessment that sexual orientation is distributed dimensionally. *Compr. Psychiatry* 32: 52–58.

McConaghy, N., and Silove, D. (1991). Opposite sex behaviours correlate with degree of homosexual feelings in the predominantly heterosexual. *Aust. N.Z. J. Psychiatry* 25: 77–83.

McConaghy, N., and Zamir, R. (1995). Heterosexual and homosexual coercion, sexual orientation and sexual roles in medical students. *Arch. Sex. Behav*. 24: 489–502.

McConaghy, N., Armstrong, M. S., Blaszczynski, A., and Allcock, C. (1981). Controlled comparison of aversive therapy and covert sensitization in compulsive homosexuality. *Behav. Res. Ther*. 19: 425–434.

McConaghy, N., Armstrong, M. S., Blaszczynski, A., and Allcock, C. (1983). Controlled comparison of aversive therapy and imaginal desensitization in compulsive gambling. *Br. J. Psychiatry* 142: 366–372.

McConaghy, N., Armstrong, M. S., and Blaszczynski, A. (1985). Expectancy, covert sensitization and imaginal desensitization in compulsive sexuality. *Acta Psychiatr. Scand*. 72: 1176–187.

McConaghy, N., Blaszczynski A., and Kidson, W. (1988). Treatment of sex offenders with imaginal desensitization and/or medroxyprogesterone. *Acta Psychiatr. Scand*. 77: 199–206.

McConaghy, N., Blaszczynski A., and Frankova, A. (1991). Comparison of imaginal desensitization and other behavioural therapies for pathological gambling. *Br. J. Psychiatry* 159: 390–393.

McConaghy, N., Buhrich, N., and Silove, D. (1994). Opposite sex-linked behaviours and homosexual feelings in the predominantly heterosexual male majority. *Arch. Sex. Behav*. 23: 565–577.

Miner, M. H. (1997). How can we conduct treatment outcome research? *Sex. Abuse J. Res. Treat.* 9: 95–110.

Minton, H. L., and McDonald, G. J. (1983/1984). Homosexual identity formation as a developmental process. *J. Homosex*. 9: 91–104.

Muehlenhard, C. L. (1988). Misinterpreted dating behaviors and the risk of date rape. *J. Soc. Clin. Psychol*. 6: 20–37.

Muehlenhard, C. L. (1989). Young men pressured into having sex with women. *Med. Aspects Hum. Sex*. 23: 50–62.

Murray, S. O. (1995). Southwest Asian and North African terms for homosexual roles. *Arch. Sex. Behav*. 24: 631–654.

Nokes, D. (1997). *Jane Austen, a Life*, Fourth Estate, London.

Palmer, C. T. (1988). Twelve reasons why rape is not sexually motivated: A skeptical examination. *J. Sex Res*. 25: 788–790.

Person, E. S., Terestman, N., Myers, W. A., Goldberg, E. L., and Salvadori, C. (1989). Gender differences in sexual behaviors and fantasies in a college population. *J. Sex Marital Ther*. 15: 187–198.

Pillard, R. C., and Weinrich, J. D. (1986). Evidence of familial nature of male homosexuality. *Arch. Gen. Psychiatry* 43: 808–812.

Prentky, R. A. (1998). In reply (letter to the Editor). *Sex. Abuse J. Res. Treat*. 10: 339–341.

Rice, M. E., Quinsey, V. L., and Harris, G. T. (1991). Sexual recidivism among child molesters released from a maximum security psychiatric institution. *J. Consult. Clin. Psychol*. 59: 381–386.

Rosen, R. C., and Keefe, R. J. (1978). The measurement of human penile tumescence. *Psychophysiology* 15: 366–376.

Sakheim, D. K., Barlow, D. H., Beck, J. G., and Abrahamson, D. J. (1985). A comparison of male heterosexual and male homosexual patterns of sexual arousal. *J. Sex Res*. 21: 183–198.

Searles, J. S. (1985). A methodological and empirical critique of psychotherapy outcome meta-analysis. *Behav. Res. Ther*. 23: 453–463.

Sell, R. L. (1997). Defining and measuring sexual orientation: A review. *Arch. Sex. Behav*. 26: 643–658.

Sell, R. L., Wells, J. A., and Wypij, D. (1995). The prevalence of homosexual behavior and attraction in the United States, the United Kingdom and France: results of national population-based samples. *Arch. Sex. Behav.* 24: 235–248.
Sennett, R. (1995). "Sex, lies and social science": An exchange. *N.Y. Rev. Books*, May 25: 43.
Sulloway, F. J. (1996). *Born to Rebel*, Pantheon Books, New York.
Templeman, T. L., and Stinnett, R. D. (1991). Patterns of sexual arousal and history in a "normal" sample of young men. *Arch. Sex. Behav*. 20: 137–150.
Tollison, C. D., Adams, H. E., and Tollison, J. W. (1979). Cognitive and physiological indices of sexual arousal in homosexual, bisexual, and heterosexual males. *J. Behav. Assess.* 1: 305–314.
Troiden, R. R. (1989). The formation of homosexual identities. *J. Homosex*. 17: 43–73.
Van Wyk, P. H., and Geist, C. S. (1984). Psychosocial development of heterosexual, bisexual, and homosexual behavior. *Arch. Sex. Behav*. 13: 505–544.
Wheeler, D., and Rubin, H. B. (1987). A comparison of volumetric and circumferential measures of penile erection. *Arch. Sex. Behav*. 16: 289–299.
Wiederman, M. W. (1997). The truth must be here somewhere: Examining the gender discrepancy in self-reported lifetime number of sex partners. *J. Sex Res*. 34: 375–386.
Wilson, R. J. (1998). Psychophysiological signs of faking in the phallometric test. *Sex. Abuse J. Res. Treat.* 10: 113–126.
Zucker, K. J. (1997). Letter to the Editor. *Arch. Sex. Behav*. 26: 671–672.