

the research but also claims that it provides support for the possibility of reparative therapy to change sexual orientation from homosexual to heterosexual. Spitzer cited some nonsexual benefits from this therapy. He noted that participants reported that after therapy they had a greater sense of masculinity in males and femininity in females. This needs more clarification in light of the many studies of the complexities of gender roles (Crawford & Unger, 2000; Kilmartin, 2000).

In conclusion, even the limited hypothesis that some individuals whose orientation is predominantly homosexual can become predominantly heterosexual following reparative therapy is not supported by this study. It may be possible that some of the research participants might have a more fluid sexual orientation, such as bisexuality (Bohan, 1996).

Are Converts to Be Believed? Assessing Sexual Orientation "Conversions"

Kenneth M. Cohen, Ph.D., and Ritch C. Savin-Williams, Ph.D.

Counseling & Psychological Services, Gannett Health Center, Cornell University, Ithaca, New York 14853; e-mail: kmc17@cornell.edu; Department of Human Development, Cornell University, Ithaca, New York; e-mail: rcs15@cornell.edu.

The realization that subject selection criteria significantly influence results when investigating socially stigmatized populations is one of the earliest and most frequently documented lessons learned by sexologists. Examples of the biasing influence of subject selection are legendary, including the early discovery that recruitment venue could determine outcomes, such as age of coming out, number of gay friends, and self-esteem level (Harry, 1986) and, more recently, that selecting gay youth from support groups (Savin-Williams & Ream, in press) or with sex-atypical behavior (McDaniel, Purcell, & D'Augelli, 2001) escalates levels of reported suicidality. Reviewing problems inherent in sampling "homosexuality," Sandfort (1997) argued that "findings in a specific study depend heavily on the definition and operationalization of homosexuality adopted, and on the way the sample has been put together" (p. 261). Yet, sample selection remains one of the "unresolved issues in scientific sexology" reviewed by McConaghy (1999) several years ago in this *Journal*.

No where is this issue more pertinent than in assessing the highly volatile issue of sexual orientation change following reparative therapy. As Spitzer noted in his literature review, the question of the effectiveness and consequences of "reorientation" therapy has been a source of vitriolic controversy. Because of this current debate, we

believe that sexologists must be extra vigilant to conduct the most methodologically sophisticated, conservative investigations possible, which implies being particularly attentive to subject selection. We do not believe that it is justifiable for contemporary research scholars to presume their conclusions by selection criteria: Including subjects most likely to agree with the author's hypothesis while excluding those most likely to give disconfirming results. To permit conclusions about causality, which Spitzer denies making, but nevertheless embraces ("change in sexual orientation following some kind of therapy does occur in some gay men and lesbians"), the scientific method must be fully embraced, for only then will it be possible to determine whether events (sexual orientation change) are the consequence of chance, untested variables, or the study's independent variable (reparative therapy). Scientific research also depends on the willingness of the researcher to question the truthfulness of subject claims. Given the failure of previous scientific attempts to demonstrate change in sexual orientation (versus sexual identity or behavior), the findings by Spitzer deserve close methodological scrutiny.

In this commentary, we ignore Spitzer's apparent conceptual misunderstanding about the purpose of his study (a title and literature review that promises a study about changes in sexual orientation from homosexual to heterosexual, but a proposed hypothesis that merely suggests shifts in sexual attractions, an event not uncommon among many sexual-minority individuals—see Diamond, 2003a). We also ignore several methodological procedures that create doubt about whether Spitzer used proper scientific methods to ensure the validity of his data, including his definition of sexual orientation, the lack of reliability and validity data on instruments, and the failure to behaviorally anchor response items. Instead, our focus is on subject selection biases that raise serious questions about the veracity of subject claims of reorientation. It is our contention that Spitzer selected a unique group of subjects who were decidedly invested in demonstrating the possibility and benefits of reparative therapy. This one fatal flaw seriously diminishes the internal and external validity of his study and necessarily precludes the very conclusions Spitzer offered: "The mental health professions should stop moving in the direction of banning sexual reorientation therapy" and that "many patients . . . can make a rational choice to work toward developing their heterosexual potential and minimizing their unwanted homosexual attractions."

Who were Spitzer's subjects and how could this collective affect the internal and external validity of his findings? In terms of venue, volunteers were recruited through "repeated" notices to ex-gay religious ministries,

therapies, and political organizations that promote biasing conditions. Subjects were clearly not blind to the study's hypothesis or purpose and most, if not all, had compelling motivations to provide data that would prove the hypothesis correct. Indeed, subjects *could not* participate in the study unless their perceived experience supported the study's hypothesis. Thus, subjects had a strong desire to change (including 19% who were directors of ex-gay ministries or mental health professionals), strong desire to witness to others (e.g., 78% publicly spoke in favor of efforts to change homosexual orientation, often at church functions), strong desire to affirm their religiosity (93% reported that religion was "very" or "extremely" important to them), and strong desire to believe that their own conversion was successful. These biasing conditions are not conducive or even normative to scientific investigations. The intent to eliminate or at least reduce social desirability as a potentially damaging influence on the veracity of results is standard fare for scientific research and yet it appears that Spitzer did everything within his power to *promote* if not *ensure* his intended responses. Thus, it is exceedingly difficult to take at face value the independence of the study's data.

Spitzer himself recognized this point by acknowledging that "subjects' and their spouses' high motivation to provide data supporting the value of efforts to change sexual orientation" was present, but he concluded that it was unlikely that their reports were "biased due to self-deception, exaggeration, or even lying." If subjects were biased, according to Spitzer, they would have reported a "rapid onset of change," a "complete or near complete change in all sexual orientation measures," marital adjustment scores "higher than that of the normative reference group," and no gender differences. Although the study's findings were sufficient "at least to the author" to rule out systematic bias, Spitzer provided no empirical or theoretical support for the reasoning underlying these assumptions. Furthermore, his reasoning is not self-evident. For example, when lie/faking scales are integrated into questionnaires (e.g., the MMPI-2), degree of deception falls along a continuum and only rarely are subjects distributed at either pole. Spitzer assumed that bias must be conscious, intentional, and deceitful, but we know of no evidence supporting this perspective which, naively, discredits the psychological sophistication of his subjects and ignores the possibility that unconscious defense mechanisms (e.g., repression, suppression, denial) are operative. Indeed, given the nature of the subject pool, it would be surprising if subjects did *not* engage in data manipulation.

This subject selection bias might also have contributed to a memory bias that could have further generated inflated reports of change. To the extent that PRE

therapy (average was 12 years before the interview) homoerotic feelings were emotionally disturbing, on recollection years later they would likely be remembered as greater or more omnipresent than they were at the time. When combined with a desire to minimize current homoerotic feelings, inflated PRE-POST differences might well be expected that would erroneously suggest greater reorientation change than that actually achieved. Also contributing to this expected inflation was Spitzer's reliance not on physiological measures of sexual orientation, which he recognized are rare in social science research, but on reports of behaviors and cognitions that are mostly under conscious control. Yet, comparing sexual behavior, attractions, masturbatory fantasies, and masturbatory orgasm fantasies, Cohen (1999) found that as same-sex attracted males progressed toward orgasm, and thus were less able to control fantasy content, greater quantity of homoerotic thoughts surfaced. This distinction becomes imperative when assessing subjects who find their homoeroticism exceedingly unacceptable and voluntarily attempt conversion by means such as intentional suppression. Thus, it is not surprising that Spitzer found the greatest change in those aspects of same-sex sexuality most under conscious control (behavior and identity), some change in sexuality under some control (attractions), and the least change in aspects least under conscious control (homoerotic content during masturbatory fantasies). Yet, even masturbation fantasy content must be carefully assessed. Spitzer accepted at face value claims of reduced homoeroticism and increased heteroeroticism during fantasies. Without probing, however, it is nearly impossible to verify the veracity of these claims because homoerotic content is easily obscured by heteroerotic images.

Finally, on another methodological note, it is curious that despite the reorientation power Spitzer ascribed to reparative therapy, he provided no credible evidence that therapy was actually the mechanism of reported changes. Although we contend that the study's design invalidates the data collected, if one were to believe the subjects' reports of changes in their sexuality, it would be just as valid (and perhaps more parsimonious, given prior research) to assert that numerous other uncontrolled, extraneous factors (e.g., time, history, maturation) significantly contributed to the "successful reorientation" changes during the *many* years subjects were in treatment. Given the myriad confounding variables that may have contributed to apparent therapy effects, including degree of homoeroticism, marital status, previous dissatisfaction with same-sex attractions, psychological vulnerability of subjects (depressed, suicidal, unhappily married, distressed by their sexuality, living a religious lie), church/therapy/ support group involvement, length of therapy, need to

proselytize, it is regrettable that Spitzer did not consider these possible contributors to reorientation. They might have explained more variance than the therapy itself.

In conclusion, a most basic and frequently documented lesson learned by sexologists investigating socially stigmatized populations cannot be overstated: Subject selection impacts research findings. As scientists, we must disbelieve Spitzer's data because they are so compromised by subject selection bias as to raise serious objections to any claims Spitzer might make about their meaning and generalizability. Research that cannot be applied to nonstudy participants is of limited utility in the social and behavioral sciences.

Reconsidering "Sexual Desire" in the Context of Reparative Therapy

Lisa M. Diamond, Ph.D.

Department of Psychology, University of Utah, 380 South 1530 East, Room 502, Salt Lake City, Utah 84112-0251; e-mail: diamond@psych.utah.edu.

Clarifying the Question

First things first: Is Spitzer's study really "about" changing sexual orientation? In order to answer this question, we need to agree on a definition of sexual orientation and its defining criteria. But, of course, these issues have long been topics of heated debate (e.g., Bailey, 1995; Bem, 1996; Ellis, 1996; Golden, 1987; Rust, 1992; Veniegas & Conley, 2000). Is sexual orientation an innate sexual predisposition or a learned behavioral pattern? Does it primarily influence sexual desire or does it also shape affiliative preferences, affectional feelings, and gender-typed behavior? Such debates might seem shopworn at this point, but they are neither resolved nor irrelevant. To the contrary, the more we learn about the diversity of same-sex sexuality across different populations and contexts, the more we must regularly reevaluate our implicit and explicit models of this phenomenon and the hypotheses they prompt us to test.

Spitzer's central question—whether homosexuals can change into heterosexuals—presumes a fairly reductionistic sexual taxonomy that has garnered increasing scientific skepticism over the years. Kinsey et al. (1948) were perhaps the first and most famous to caution that "The world is not to be divided into sheep and goats" (p. 639) and empirical data increasingly buttress this perspective (Blackwood, 2000; Murray, 2000). For example, representative studies of American adolescents (French, Story, Remafedi, Resnick, & Blum, 1996; Garofalo, Wolf, Wissow, Woods, & Goodman, 1999) and adults (Laumann, Gagnon, Michael, & Michaels, 1994) have

found that most individuals with same-sex attractions *also* report experiencing other-sex attractions, and both changes in, and disjunctures among, sexual behaviors, attractions, and identity are widespread (Baumeister, 2000; Diamond, 2000, 2003b; Golden, 1987; Pattatucci & Hamer, 1995; Rust, 1992; Stokes, Damon, & McKirnan, 1997; Weinberg, Williams, & Pryor, 1994).

Spitzer's research question and methodology do not acknowledge these complexities. Rather, he uncritically treats sexual attractions, fantasies, and emotional longings as coordinated indices of one's underlying "sheep" or "goat" status, despite the fact that (1) it is increasingly unclear whether these discrete types even *exist* as natural categories and (2) it is similarly unclear whether "sheepness" or "goatness" could ever be reliably diagnosed by coordinated and stable patterns of fantasy, desire, and affection. Given these problems, some researchers have argued that "it makes more sense to ask about specific aspects of same-gender behavior, practice, and feelings during specific periods of an individual's life rather than a single yes-or-no question about whether a person is homosexual" (Laumann et al., 1994, pp. 285–286). What, then, might we learn from Spitzer's study if we jettison extrapolations to "sexual orientation" and focus instead on domain-specific changes?

Interpreting the Findings: The Meaning and Experience of Desire

Unfortunately, a number of factors hamper interpretation of Spitzer's data, such as the significant and obvious problems of self-selection and self-report biases. Yet, I will leave aside these concerns, trusting that other commentators will address them in depth. Granting for the sake of argument that some of Spitzer's participants did, in fact, experience declines in their self-reported same-sex desires, how should we interpret such changes?

First of all, as noted above, the phenomenon of plasticity in sexual desire over time has already been documented in several prospective studies, and is not newsworthy in and of itself (Diamond, 2000, 2003b; Pattatucci & Hamer, 1995; Stokes et al., 1997; Weinberg et al., 1994). Spitzer, however, is more concerned with *effortful* changes effected through cognitive-behavioral strategies, such as "thought stopping," avoidance of situations that trigger same-sex attractions, and social support mobilization. Can these techniques actually alter one's subjective desires? Of course they can—just as attending Weight Watchers meetings and keeping "forbidden" foods out of the house can attenuate a dieter's natural, evolved cravings for salty, fatty, calorie-dense foods. Furthermore, any reader of