

Guttman Scalability Confirms the Effectiveness of Reparative Therapy

Scott L. Hershberger, Ph.D.

Department of Psychology, California State University, Long Beach, 1250 Bellflower Blvd., Long Beach, California 90840; e-mail: scotth@csulb.edu.

Spitzer presents compelling evidence that a homosexual orientation can be changed to a heterosexual orientation by reparative therapy. The best of evidence is found in changes following reparative therapy in (a) homosexual sex, (b) homosexual self-identification, and (c) homosexual attractions and fantasies. In this commentary, I will focus on the more dramatic results for men, although the results for women also support the effectiveness of reparative therapy.

Most therapists would agree that it is easiest to lower or eradicate participation in homosexual sex, somewhat harder to change self-identification from homosexual to heterosexual, and hardest of all to lower or eradicate homosexual attractions and fantasies. In Spitzer's study, changes in sexual behavior, self-identification, and attractions and fantasies toward a predominately heterosexual orientation confirm this expected order. After reparative therapy, homosexual sex was lowered for 98% of those who had previously engaged in homosexual sex, heterosexual identity was affirmed by 78% more individuals, and homosexual attractions or fantasies were experienced by 47% fewer individuals. Spitzer also categorized individuals as having changed or not changed by dichotomizing each of the three measures "at a point that the author regarded as indicating more than a slight level of homosexuality" (p. 408).

We can quantify the close match between the expected pattern of change with the observed pattern of change by thinking of sex, self-identification, and attraction and fantasies as three items whose order conforms to that of a Guttman scale. For the items to form a Guttman scale, everyone who has significantly fewer homosexual attractions and fantasies should also be more likely to self-identify as a heterosexual, and all those who now self-identify as a heterosexual should be more likely to have reduced the number of their homosexual sex experiences. The coefficient of reproducibility (*CR*) can serve as a measure of goodness of fit between the observed and predicted change patterns. The *CR* is defined as:

$$CR = 1 - \frac{\sum e}{Nk},$$

where *e* is the number of individuals whose item order does not conform to that of a Guttman scale, *N* is the sample size, and *k* is the number of items (Dunn-Rankin,

1983). The *CR* ranges from 0 to 1, with 1 denoting perfect conformity to a Guttman scale.

In order to compute the *CR*, data at the individual level are required. Spitzer clearly states in his article that his raw data are available to others, but the data were not available at the time this commentary was written. Thus, the following procedure was used to obtain an estimate of the *CR* for these data. Data were first created for 133 individuals by requiring that, for each item, the simulated data have same proportion of individuals who changed as in the observed data. For example, 47% of the individuals in the observed data reduced their homosexual fantasies; therefore, the simulated data for this item was also defined to have 47% of the sample change. Although the simulated data can reflect the correct proportions of change for the items, it is impossible to specify for any one person the specific pattern of change. To overcome this difficulty, 1,000 bootstrap samples of 133 individuals each were created to provide an estimate of the true *CR*. From the 1,000 samples, the *CR* ranged from 0.83 to 0.95, with a mean of 0.92 and a *SD* of 0.03.

Does a mean *CR* of 0.92 indicate a good fit? No value of *CR* has been defined that is universally accepted as a dividing line between acceptable and unacceptable goodness of fit, although Guttman (1947) originally defined a value of 0.85 as the dividing line. Other authors (e.g., Torgerson, 1958) suggested that a *CR* above 0.90 is a better standard. Therefore, the mean *CR* of 0.92 indicates that the pattern of change among the measures does fit a Guttman scale well. The goodness of fit of the Guttman scale to the three items is even more impressive when one considers that the mean *CR* of 0.92 is certainly an underestimate of its true value in the real data: the data set from which each of the bootstrap samples was drawn was created by randomly specifying whether an individual changed or not on *each item separately*. Therefore, each individual's entire item pattern was random.

The orderly, law-like pattern of changes in homosexual sexual behavior, homosexual self-identification, and homosexual attraction and fantasy observed in Spitzer's study is strong evidence that reparative theory can assist individuals in changing their homosexual orientation to a heterosexual orientation. Now it is up to those skeptical of reparative therapy to provide comparably strong evidence to support their position. In my opinion, they have yet to do so.

Methodological Limitations Do Not Justify the Claim That Same-Sex Attraction Changed Through "Reparative Therapy"

Craig A. Hill, Ph.D., and Jeannie D. DiClementi, Psy.D.
Department of Psychology, Indiana University-Purdue

University Fort Wayne, Fort Wayne, Indiana 46805;
e-mail: hillc@ipfw.edu.

The study by Spitzer suffers from substantial limitations that render his conclusions virtually meaningless. The main problems are methodological and relate to demand characteristics, sampling, lack of control, and validity of measurement.

The problem of demand that pervades the study is a fundamentally confounding factor. The sample most likely consisted of individuals who have experienced intense anxiety and guilt to the extent that this sets them apart from a majority of other lesbians and gay men who are dealing with identity issues. Consequently, the reports of change provided by these participants may originate from a combination of erotophobia (negative emotional reactions to sexual issues) and remorse over perceived violations of religious doctrine or culturally and family-based values, rather than representing a self-enhancing change in erotic and romantic nature.

Individuals who are unhappy with their attraction to the same sex and who have gone to great lengths to change will be motivated for their attitudes to become consistent with their public behavior, in line with cognitive dissonance theory. Reports of attitude change may not be enduring over the long haul when they are related to such compelling issues as attraction and sexual desire.

The desire to change one's lesbian or gay sexual orientation is typically based on deeply entrenched negative attitudes about one's same-sex feelings, frequently called *internalized homophobia*, "the most insidious of the minority stress processes, . . . leading to a devaluation of the self and resultant internal conflicts and poor self-regard" (Meyer & Dean, 1998, p. 161). Meyer and Dean note that "men in the early stages of coming out and men who have sex with men but have not accepted their homosexuality are likely to have higher levels of internalized homophobia than their counterparts" (p. 179). The important role of traditional mainstream religions in promoting the internalization of homophobia is demonstrated in Meyer and Dean's study of 912 gay men. The men who were religious, but who were not associated with gay churches or synagogues, experienced higher levels of internalized homophobia than men who were religious, but who were associated with gay religious organizations; in fact, these latter men were equivalent in homophobia to nonreligious men, suggesting a beneficial influence of gay-affirming beliefs.

The effect of such conflict and anguish very likely distorts assessments made by individuals who have gone to great lengths to seek help. In response to antigay attacks, individuals with high levels of internalized homo-

phobia likely experience their orientation as a source of pain, rather than as a source of pleasure, love, and intimacy. Antigay attacks are therefore often interpreted as justified punishment for being lesbian or gay and this may contribute to the process of wanting to change their sexual orientation (Garnets, Herek, & Levy, 1990).

Contrary to claims made by Spitzer, bias could account for a substantial portion of the changes reported by his respondents. Sexuality researchers have been concerned for some time about the potential for biases that diminish the accuracy of information, both in self-report methods (Meston, Heiman, Trapnell, & Paulhus, 1998) and in interviews (Catania, 1999). Such biases are likely due to nonconscious self-enhancing or social desirability processes (Brown & Sinclair, 1999). Such distortion never occurs in a form where all individuals skew their reports perfectly in line with the desirable standard, an argument Spitzer employed to dismiss the probability that bias affected the results of his study. Rather, bias is identified by differences in group averages. The variability within groups is never close to zero, with all members of a group falling extremely close to one end of the dimension. This is true even in validity research in which participants are requested to distort their responses by the researcher (Holden & Jackson, 1981). Therefore, substantial distortion could have occurred in self-reports of respondents in the Spitzer study which would not be evidenced by all respondents rating themselves at the "perfect" (i.e., heterosexual) end of the scales.

In contrast to the argument advanced by Spitzer, which was also intended to discount the possibility of bias, gender differences can be found in distortion, *especially* with respect to sexuality. One example of this is that men report greater numbers of sexual partners than do women of the same age, which cannot be entirely or even substantially attributed to larger numbers of men having sex with a small number of women (Brown & Sinclair, 1999). Furthermore, according to the availability heuristic (Kahneman & Tversky, 1973), the salience of an event biases judgment about the frequency of the event. For example, adolescent same-sex experimentation is quite common, but for the individual who is horrified at the thought that she or he may be gay, one or two such contacts could easily be perceived as excessive, and consequently the perception, and subsequent reports, of homosexual activity are gross overestimates. Such a possibility suggests that the participants in the Spitzer study may not actually be gay to begin with, or they are bisexual; in either case, their reports of change therefore would not really reflect a change of sexual orientation among lesbians and gay men. The most plausible explanation is that a nonconscious cognitive distortion affects judgments about

sexual experiences for both sexes (Brown & Sinclair, 1999; Meston et al., 1998).

Given the possibility of distortion and the demand for attitude-behavior consistency, issues related to sampling and control should be elevated to the highest level. Simply locating people who claim to have changed does not provide convincing data. It would be possible to locate people who claim and sincerely believe any number of phenomena that are not easily verifiable empirically and about which many professionals are skeptical. All participants were obtained from organizations or therapists who are extremely committed to the efficacy of reparative therapy and only people who contacted the researcher were included in the study. These people are by definition those who believe in the effectiveness of the reparative techniques and earnestly need the technique to be effective both from a spiritual and an emotional perspective. Moreover, basic to any behavior change research is the inclusion of persons who attempted behavior change and failed.

Effort to employ the highest degree of control is incumbent upon researchers of such sensitive topics. Assignment of participants to experimental conditions is not the only means of enhancing control. Because of the problems with a hypothesis-biased population, it is critical to establish the nature of the sample obtained in terms of theoretically relevant characteristics; it is important as well to compare these individuals to other samples of lesbians and gay men to determine exactly the ways in which they are similar or different. Such critical characteristics should include homophobia, erotophobia, sex guilt, emotional stability, psychological disorders, self-esteem, social functioning, and religious guilt, to name but a few.

A comparison could involve matching therapy participants with nontherapy participants to determine the extent to which nontherapy participants had attempted sexual orientation changes in the past. With only one group purporting to experience change and with no comparison group, little confidence can be placed in claims that change occurred specifically due to therapeutic intervention and not due to some other factor. Moreover, change based on retrospective reports related to therapeutic progress are highly suspect in terms of validity, again especially given the incredible demand for change inherent in the life situation of these respondents, and given the fact that the pretherapy period was on average 12 years prior to the data collection period.

In addition to emotionally-based cognitive distortions, the issue of bias in assessment of current heterosexual relationships must be considered. When an individual who desperately wants to be heterosexual is finally involved in a heterosexual relationship and is asked, "Are you emotionally satisfied with your relationship?," what

the researcher may actually be measuring is relief at achieving this greatly desired goal and not necessarily what most individuals mean by satisfaction with a relationship. The same can be said for the question about physical satisfaction. What is not investigated is the effort that goes into becoming physically aroused, because anecdotal reports suggest that "ex-lesbians" or "ex-gays" often must spend a great deal of effort achieving levels of arousal sufficient to engage in heterosexual sexual behavior.

The only conclusion that is indisputable in Spitzer's study is that he has identified a subset of lesbians and gay men (who in fact may actually be more appropriately considered bisexual) who claim to have changed their overt sexual behavior; the nature of the change, and the process through which it occurred, has not been convincingly established. Given the importance of this issue for individuals struggling with their sexual orientation, to claim otherwise is misleading and dangerous.

Initiating Treatment Evaluations

Donald F. Klein, M.D.

New York State Psychiatric Institute, 1051 Riverside Drive, Unit 22, New York, New York 10032-2695; e-mail: donaldk737@aol.com.

Spitzer presents face valid evidence that changes in homosexual behavior and feelings of desire and satisfaction can be achieved by some, to varying degrees, via "reparative therapy."

These reports, which Spitzer quite logically argues are convincing, are necessary preliminaries to an open trial, which enrolls a series of appropriate subjects, treats them all and records the results. This allows an estimate of the proportion of good outcomes, although it does not establish that the therapy caused the benefit since patients may improve in spite of their treatment. Volunteers who claim successful outcomes do not yield an estimate of the proportion benefited. Also, just how common such reported successes are remains obscure.

Spitzer calls for a consecutive series who perceive homosexuality as a problem of theirs they wish fixed, who have been evaluated before and after treatment. I agree this is the correct next step. Apparently, Nicolosi and Byrd claim to have such data. If this is more than therapist self-serving, they have an obligation to present their data or to stop making such important claims. Such strictures also apply to the range of official groups that assert the uselessness and damaging effects of "reparative therapy" for homosexuality. Where are their data? If it is nothing but anecdotes and presumptions, how can they claim they are being professionally responsible? Relevant observations are needed to raise hopes that a treatment is worth