

orientation” or the belief that there is direct pathway from a gene (or complex of genes or mixture of chemicals) and the like to desiring to have sex with someone who has similar genitals, the answer is nothing. But, if you believe, as I do, that the complex of sex practices is learned in a particular historical and cultural situation, then the idea of comparing sexual practices with other social practices is not very mysterious.

It is clear from the Spitzer study that some persons who have long (and perhaps exclusive) histories of sex with same gender partners move, often with religious and other “therapeutic” supports, to lives in which they have sex with persons of the other gender. From prior studies, we know that substantial numbers of persons who have had long (and perhaps exclusive) lives with partners of the other gender have voluntarily (sometimes describing the experience as finding their true selves) and sometimes with therapeutic support to lives in which they have sex exclusively or nearly exclusively with persons of the same gender. This movement back and forth in partner preferences is not unexpected given the general movement of humans from one to another deeply seated and “un-chosen” life way or practice.

Most of our early life is “un-chosen.” We learn whatever is expected of us, sometimes well, sometimes poorly, sometimes we learn what adults want of us, sometimes we learn the opposite. Religious parents raise atheists as well as believers, parents raise children who speak English, who then learn French and emigrate, working class parents raise children who go to Harvard and learn to exploit the working class, rich parents raise children who become communists. These are all minority outcomes, but they happen. If you believe, as I do, that language, religion, and gender learning are as deeply embedded in a person (as un-chosen, if you will) as the specific sexual preference of the gender of the sexual partner, then in adulthood when some ways of life seem uncomfortable and choices are apparent (sometimes hard choices), it is not surprising that changes take place.

It is the politics of the change that shape most of the debate. The actual process is quite understandable and expectable. Whether such change should happen, what direction it should take, whether it should be encouraged, and who should decide whether it is a good or bad thing for either individuals or communities is the political question.

Too Flawed: Don’t Publish

Lawrence Hartmann, M.D.

Department of Psychiatry, Harvard Medical School,
147 Brattle St., Cambridge, Massachusetts 02138; e-mail:
lhartmann@worldnet.att.net.

I think Spitzer’s paper is too flawed to publish, and is likely to do harm. Yet, here it is, being published.

The area of change in sexual orientation is an interesting one, worth study, and one where little is known, but where many think they know, and where many have passionate commitments to what they think they know. The area is embedded in so much context, history, bias, passion, self-delusion, and even lying that it deserves a higher level than Spitzer provides of careful and detailed openness, skepticism, and multiple levels of scrutiny.

There are too many problems, major and minor, in the article to discuss in reasonable detail in a brief commentary. I will touch on only a few below. Some of the problems are technical; others are of emphasis or skepticism, of definitions, of numbers, and of ethics. Some of these problems were pointed out to Spitzer by many very critical peers when he presented the paper at a May 2001 panel at the meeting of the American Psychiatric Association. He has not fixed them. The article remains a failure at establishing what it says it establishes: that some gay people can be changed to straight.

One might, from this article (from some of its interesting and energetic outreach attempts, data, careful-looking psychiatric research apparatus, and discussion of possible problems), and from my numerical calculations below, legitimately conclude that a *minute* number of gay people *may* be able to change their sexual orientation somewhat in some way, *perhaps related to* interventions designed to help them change. Modesty and accuracy demand “may,” not “can,” and “perhaps related to,” not “caused by.” And a *minute* number is a key element in this potential legitimate conclusion—and that is very different from the loose common word “some.” (That the vast majority of gay people probably cannot change orientation is essentially left out of this study).

Spitzer rightly recognizes that credibility of his subjects is a basic and significant potential problem, but I think his statements are rather naive and under-concerned about that. Also, the selection of subjects is worth far more scrutiny than it gets. A study designed around the question “Can I find any people who say ‘x’ about themselves?” especially in an area people feel strongly about and have fought about—such a study is particularly vulnerable to (and pretty much invites) pressure, distortion, bias, self-delusion, and lying. Spitzer does not protect adequately against this, as even he seems to know.

Selection/inclusion criteria and procedures are not made clear enough. There was no prospective study. No controls. No independent observations or measurements at all: physiological, psychological, social. Spitzer relies wholly on self-reporting and on one 45-min telephone

interview. That is understandably convenient and cheap, but allows rather easy evasion, distortion, and lies.

A differently cogent matter: Spitzer nearly never puts “reparative therapy” in proper distancing quotation marks. That is, he repeatedly takes a strong side on the central question he is ostensibly studying. (To many colleagues, “reparative therapy” is neither reparative nor therapy; it is, rather, destructive pseudotherapy; but relatively neutral words would have been possible, such as “intervention aimed at changing sexual orientation,” or, simply, “reparative therapy” in quotes). Spitzer also seems implicitly to accept, uncritically, the high percentages of cure or change estimated by the rather notorious Bergler, Socarides, and Nicolosi (which are widely considered unreliable, probably wishful at best, and much exaggerated).

Another matter, and one of several ethical issues: Spitzer stated that he found no evidence of harm of “reparative therapy” in these 200 people. That is misleading and breathtaking in its blinkered view and its omission of the relevant groups: (1) the many more (10 times as many? 100 times? Far more?) who have had some form of “reparative therapy” and who do not think it helped them. In addition to recent less-than-pleased-or-changed ex-patients of “reparative therapists,” excluded by Spitzer, how many gay or bisexual people were in analysis or dynamic psychotherapy heavy with “reparative” elements from, say, 1930 to 1980 or 1990? These should be looked at carefully for harm. I know many, and I have treated several. And (2) the wider population at large, straight and gay. Most mental health professionals I know consider that the semisanctioned existence of “reparative therapy” probably harms millions of nontreated gay people.

There is inevitably continual spillover from “reparative therapy’s” narrowly defined into (1) many psychotherapies (much psychoanalysis and dynamic psychotherapy of the twentieth century contained significant doses of such spillover and “therapeutic” first cousins) and (2) public policy, law, values, definitions of illness, etc. Even if “reparative therapy” helps a few people in some ways, as I think it may, it nearly certainly harms a far larger number of people, and that is a major ethical issue relevant to Spitzer’s study but apparently not seriously considered by him.

Then numbers: One large aspect of the study pretty well *proves* just about the *opposite* of what it says it shows. Spitzer scoured the United States for several years (“actively for 2 or 3 years,” he told me in 2001). He is an experienced social-psychiatric researcher attached to a major university, and he went energetically and repeatedly to all the antigay groups he could find: to the National Association for Research and Therapy of Homosexuality (NARTH), to all the religious change-gay-to-straight pro-

grams he could find, all the “reparative therapists” he could find, in all the United States, to find all the gay-to-straight ex-patients he could find. That yielded about 200 people he felt could be used for his study. 200. Let us leave aside for the moment that, in my fairly educated view, I suspect that of the 200 many were heavily biased, and were probably distorting and/or lying, since there was clearly pressure on many to do so, and nearly certainly many had some coaching as to how to do so. Spitzer’s idea that it was probably proof of reliability that many said they had only partly changed seems to me wishful and naive. But even if I am very dubious about many of them, 200 is the number Spitzer offers.

Numbers matter. Two hundred out of how many? Taking the United States population over the age of 18 at about 210 million, and taking the gay population (never precisely definable but still estimatable) as about 3–10%, that gives us a figure of about 6–21 million gay people in the United States. Spitzer, after 2 or 3 years of energetic research seeking, found 200 possible sex-orientation-changed people out of 6–21 million gay people. That means we are talking about possible changes in not 1%, not one tenth of a percent, not one hundredth of 1%, but about one one-thousandth of 1% (.000009–.00003) of the adult gay population. If any form of cancer had a cure rate of one in 100,000, that would not be called evidence that that cancer is curable; rather, to call it curable on that basis would be considered a cruel delusion and false promise.

As a further relevant issue of numbers, consider bisexuals. There are probably millions of Americans who are, by generally acceptable if imprecise criteria (e.g., postpubertal behavior, feelings, fantasies), bisexual. For reasons little understood, a great many of them vary in their sexual behavior styles and/or enthusiasms from one point in their lives to another. Certainly many thousands or hundreds of thousands are more heterosexual now than they were a few years ago, and certainly many thousands or hundreds of thousands are more homosexual than they were a few years ago. Nearly certainly, many thousands of such people are in some form of psychotherapy, and of those, some probably attribute various changes in themselves to the therapy. That little understood area alone would be expected to produce far larger numbers of what may look like changes in sexual orientation than Spitzer’s 200. And Spitzer does nearly nothing to acknowledge or help understand this area or to note that it vastly overshadows his 200.

The context of possible changes of sexual orientation is heavy with the history of demonizing, criminalizing, pathologizing, scapegoating, guilt-inducing, and otherwise socially, economically, physically, and emotionally harming—not hundreds of people but millions. Much

of that, even if more subtly than in some past times and places, still goes on now in much of the world, and in much of the United States. That matters, and Spitzer largely ignores it.

Spitzer's article implies, without solid scientific support, something that has great and perhaps all-but-irresistible appeal to the popular press, to many politicians, and to many members of the public: that therapy can change gayness to straightness. Spitzer alerted the popular press before presenting part of this paper at the meeting of the American Psychiatric Association in 2001 and the popular press did some harm then. It is very likely to do more harm now, with the study's publication.

Spitzer's article, for all its dignified-looking data, scientific journal format, and partial disclaimers, is in essence irresponsible and unscientific. It does not constitute scientific evidence that gayness can be changed.

Evaluating Interventions to Alter Sexual Orientation: Methodological and Ethical Considerations

Gregory M. Herek, Ph.D.

Department of Psychology, University of California, Davis, California 95616-8686; e-mail: gmherek@ucdavis.edu.

Consider this scenario:

A pharmaceutical company claims its new dietary supplement can change left-handed people to right-handers. Medical associations oppose the supplement on the grounds that it harms many people who use it. Noting that there is no reason for left-handed people to try to change, they urge their members not to recommend or administer the product to their patients. To test the drug company's claim, a researcher conducts brief telephone interviews with self-proclaimed "ex-lefties." He recruits respondents mainly through the drug company, which promotes his study to individuals who have given public testimonials about the product's effectiveness. They say they tried the supplement because they felt miserable as left-handers in a right-handed world. Most claim they now function as right-handers, although many report occasional thoughts about using their left hand and some occasionally lapse into left-handedness. The researcher's findings are based entirely on the one-time interviews in which he asked the ex-lefties to rate their handedness prior to taking the supplement (12 years earlier, on average) and during the previous year. Respondents' ratings of their past and current handedness are significantly different. The researcher concludes that the supplement does indeed change left-handers to right-handers in some cases. Meanwhile, other researchers and clinicians report anecdotally that the food supplement does not change most left-handers to right-

handers, but many who tried the supplement report serious negative side effects.

The main questions raised by this hypothetical story concern whether the researcher's data are valid, whether the product's harmful effects would justify its use even if it is sometimes effective, and why left-handers should be encouraged to change in the first place. Similar questions arise from Spitzer's study of self-reported change from homosexuality to heterosexuality following participation in an intervention. Because of space limitations, this comment discusses only four of the many criticisms that can be made of Spitzer's article.

Reliance on Self-Report

Spitzer's data are ultimately the testimonials of a highly select sample of activists from groups whose *raison d'être* is to promote efforts to change homosexuals into heterosexuals. It is difficult to imagine how his recruitment strategy would have yielded anything other than reports of substantial shifts to a heterosexual orientation. Despite his acknowledgment of its serious methodological inadequacies, Spitzer asks readers to take it on faith that his respondents were both willing and able to report accurately on their past and current thoughts, feelings, and behaviors.

This represents a curious abdication of the scientist's obligation to design a study in a way to avoid known sources of bias. Recognizing that even subtle and unintentional biases can affect the data, researchers routinely adopt elaborate safeguards to prevent their own expectations and those of their research subjects from affecting a study's outcomes. Spitzer's study lacked such safeguards, despite the obvious threats to validity inherent in his sampling procedures.

Even if Spitzer's respondents sincerely tried to give true accounts of their feelings and daily behaviors from (on average) 12 years prior to the interview, their reports cannot be assumed to be reliable. People often are inaccurate when recalling earlier mental states, especially when their emotions, goals, or beliefs have changed in the interim (Levine & Safer, 2002). Memories of past beliefs, attitudes, and behaviors are affected by many factors, including personal theories about one's own behavior change over time (e.g., Ross, 1989). For this reason, asking research participants to recall their preintervention thoughts and feelings is always problematic, even when they are unaware of the study's purpose and have no ideological stake in its outcome. Given the inherently biased nature of Spitzer's sample, his failure to make even minimal attempts to assess the data's reliability (e.g., by assessing internal consistency within interviews and through follow-up interviews) and validity (e.g., through