

# **CONDEMNATION OF A SCIENTIFIC ARTICLE: A CHRONOLOGY AND REFUTATION OF THE ATTACKS AND A DISCUSSION OF THREATS TO THE INTEGRITY OF SCIENCE**

**Bruce Rind**

*Department of Psychology, Temple University, Philadelphia, PA  
19122 (rind@vm.temple.edu)*

**Philip Tromovitch**

*Graduate School of Education, University of Pennsylvania*

**Robert Bauserman**

*Department of Health and Mental Hygiene, State of Maryland*

On July 12, 1999, our meta-analysis on child sexual abuse published in *Psychological Bulletin*, one of the American Psychological Association's (APA) premiere journals, was condemned by the U.S. Congress (*H. Con. Res. 107*). The condemnation followed months of attacks on the article, the APA, and us by various social conservatives and psychoanalytically-oriented clinicians. The American Association for the Advancement of Science (AAAS) was asked by the APA to independently review our article. After considering criticisms of it and the article itself, AAAS declined, but commented that it was the criticisms, not our methods or analyses, that troubled them because these criticisms misrepresented what we wrote. The current article chronicles this whole affair. First, we provide background, explaining why an article such as ours was needed. Then we accurately summarize the article, given that it has been so widely misrepresented. Next we present a chronology of the events leading up to and following the condemnation. We then present and refute all the major criticisms of the article, which have included both methodological and conceptual attacks. Next we discuss the threat to science that these events portend. We conclude by discussing the need to separate moral judgments from scientific research, the conflation of which formed the basis for the distortions and condemnation.

## Introduction

For over one hundred years the American Psychological Association (APA) has been publishing psychological theory and research in its various journals. July 12, 1999 may come to be seen as one of the most extraordinary dates in APA's publication history. On this day, the United States Congress, in an unprecedented move, condemned an article published by the APA in one of its most prestigious journals, *Psychological Bulletin*. The vote for House Concurrent Resolution 107 (*H. Con. Res. 107*) was 355-0, with 13 voting "present" (i.e., abstaining). On July 30, 1999, the Senate unanimously passed it as well. The article, entitled "A Meta-Analytic Examination of Assumed Properties of Child Sexual Abuse Using College Samples," was published exactly one year earlier and was authored by us (Rind et al., 1998). In short, the resolution proclaimed our study to be "severely flawed," condemning and denouncing all suggestions in the article indicating that "sexual relationships between adults and 'willing' children are less harmful than believed."

In response to pressure from Congress, the APA pledged to seek independent review of the article, another unprecedented event in the long history of APA publications. The APA asked the American Association for the Advancement of Science (AAAS) to conduct the review. AAAS is the largest science organization in America and is publisher of the prestigious journal *Science*. In an October 1999 letter, the AAAS Committee on Scientific Freedom and Responsibility informed the APA that after "considerable deliberations taking into account the views of...two consultants and extensive background materials on reactions to the published article," the Committee would not formally review the article (McCarty, 1999). They commented that they saw "no reason to second-guess the peer review process" used by *Psychological Bulletin* in its decision to publish the research. Importantly, the Committee went on to state that "[a]fter examining all the materials available to the Committee, we saw no clear evidence of improper application of methodology or other questionable practices on the part of the article's authors." Additionally, they commented that:

The Committee also wishes to express its grave concerns with the politicization of the debate over the article's methods and findings. In reviewing the set of background materials available to us, we found it deeply disconcerting that so many of the comments made by those in the political arena and in the media indicate a lack of understanding of the analysis presented by the authors or misrepresented the article's findings. All citizens, especially those in a position of public trust, have a responsibility to be accurate about the evidence that informs their public statements. We see little indication of that from the most vocal on this matter, behavior that the Committee finds very distressing.

Although the AAAS Committee ended its letter by noting that its decision not to review the article "should not be seen as either endorsement or criticism of it," the letter's other comments clearly point to the criticisms of the article, rather than the article itself, as the problem area. In an interview with the *Philadelphia Inquirer* (November 17, 1999, A20), the chair of the Committee, physicist Irving Lerch, commented that "[s]ome of the political statements were clearly self-serving. I think some politicians tried to inflame or cash in on public sentiment by purposely distorting what the authors said."

In the current article we will document the distortions and misrepresentations made by our most vocal critics and discuss the threat posed to sexual science by these attacks. First, we will provide background on why a study such as ours was needed. Second, because of the numerous distortions, we will present an accurate summary of it. Third, we will detail the chronology of events that led to congressional condemnation. Fourth, we will present and refute the major methodological and conceptual criticisms of our review. Fifth, we will discuss the problems created by the attacks and distortions, which include not only the threat to science, but to individuals as well. Finally, we will conclude by arguing for the need to separate moral judgments from scientific research as a prerequisite for achieving a valid understanding in science in general and sexual science in particular.

## Background

In the 1970s, the feminist movement raised consciousness first about rape and then incest, two problem areas they believed society had not taken seriously enough up to that time. Feminist ideas

#### 4 Sexuality & Culture

about the nature of rape served as the model for conceptualizing incest, which in turn came to serve as the model for understanding all forms of age-discrepant sexual relations involving minors (Finkelhor, 1984; Okami, 1990). Thus, the anti-incest campaign was a central factor in the evolution of thinking about child sexual abuse (CSA).

Anti-incest advocates helped to revive Freud's seduction theory, which had lain dormant for seven decades (*Frontline*, 1995a; Nathan & Snedeker, 1995; Pendergrast, 1996). Seduction theory held that childhood sexual seduction is the cause of all adult neuroses (Crews, 1998; Esman, 1994). Freud claimed that seduction is so disturbing that the psyche tends to repress memories of it. Though submerged, he maintained, these memories nevertheless express themselves in the form of symptoms, which can only be relieved by bringing the memories back to conscious awareness. Freud quickly abandoned this theory, however, claiming instead that wishes for seduction rather than actual seduction produce repressed memories. By the mid-1970s, feminist advocates challenged his decision as lacking in moral courage, vowing that victims would no longer be silenced (Crews, 1998; *Frontline*, 1995a). As a consequence of this advocacy, Freud's unitary concept of pathogenesis reemerged, and therapists increasingly came to view the evocation and abreaction of repressed memories as the *sine qua non* of therapeutic success (Esman, 1994).

But advocacy is not science. Freud's unitary concept, though politically expedient in the 1970s and beyond, was extreme at best and completely invalid at worst (Crews, 1998; *Frontline*, 1995a; Tavis, 1993). A major problem with this advocacy was the exaggerated claims about the effects of CSA that it was prone to fostering (Best, 1997). Such exaggeration has iatrogenic potential, as Brown and Finkelhor (1986) concisely cautioned:

...it is important that advocates not exaggerate or overstate the intensity or inevitability of these consequences. In addition to policymakers, victims and their families wait for research findings on the effects of sexual abuse and they may be further victimized by exaggerated claims about the effects of sexual abuse. It is not possible to maintain two sets of conclusions about the effects of abuse: a dire one for political purposes, and a hopeful one for family members. Thus the presence of both audiences requires that those who conduct and interpret research in this field maintain a posture of objectivity and balance (p. 178).

Other researchers similarly cautioned against hasty generalizations regarding CSA effects because secondary victimization stemming from overreaction is a real and serious threat (Baurmann, 1983; Seligman, 1994), and labeling someone as a victim and then treating him or her accordingly can be self-fulfilling (Constantine, 1983).

Many victimologists and other commentators on CSA effects did not heed these warnings, however. A sort of tunnel vision developed, in which researchers and others uncritically interpreted all data as evidence for CSA's pathogenicity. Numerous examples can be cited (see Okami, 1990, 1994; Tavis, 1993); here we provide three. In a widely cited study, Burgess et al. (1984) found that the majority of their child and adolescent subjects showed vague or no symptoms while their CSA was occurring, but they developed many symptoms after the intervention of the criminal justice and social service systems. Rather than parsimoniously attributing these symptoms to the intervention, they interpreted them as posttraumatic stress stemming from the CSA. Kelly (1993) reported that nearly half of his child subjects fell in the clinical range of traumatic stress after being sexually and ritually abused. However, his subjects were children in the McMartin preschool case, which has been thoroughly discredited. The stress responses were iatrogenic, stemming from therapeutic and criminal justice intervention, as well as parental anxiety and panic (Nathan, 1990; Nathan & Snedeker, 1995). In their often-cited *Psychological Bulletin* literature review, Kendall-Tackett et al. (1993) used these two studies among others to argue that posttraumatic stress disorder (PTSD) is a relatively frequent consequence of CSA. They also inflated the impression that CSA is pathogenic by including sexualized behavior as a symptom—it was the most frequently reported “symptom” in the studies they reviewed. This designation as a symptom, however, is a value judgment (cf. Constantine, 1983), whose scientific validity is not supported by cross-species and cross-cultural data. In their seminal review of these data, Ford and Beach (1951) observed that “[a]s long as the adult members of a society permit them to do so, immature males and females engage in practically every type of sexual behavior found in grown men and women” (p. 197) and concluded that “tendencies toward sexual behavior before maturity and even

before puberty are genetically determined in many primates, including humans” (p. 198).

The overstatement of child abuse researchers and advocates regarding CSA effects on psychological adjustment and memory led to profound iatrogenesis in the 1980s and 1990s (Nathan & Snedeker, 1995; Pendergrast, 1996). Between 1983 and 1993, staff members in dozens of day care centers around the country were accused by preschoolers of bizarre acts of CSA often accompanied by satanic ritual abuse (SRA). The appearance of a rash or ambiguous remark made by a single preschooler might ignite a full-scale police investigation of all the children at the center. Initially, the children would deny any abuse. But after repeated interrogations by therapists, the children would begin speaking of being molested in tunnels or magic rooms, sodomized with curling irons or swords, forced to consume feces or drink the blood of sacrificed babies, and made to watch animals being ritually tortured and killed. Children’s initial denials were attributed to PTSD, wherein it was believed that they failed to reveal the truth because of fear or amnesia. The complete absence of physical evidence to corroborate claims of brutal abuse was ignored. Dozens upon dozens of staff members were put on trial—sometimes the lengthiest and most expensive in their state’s history—and were frequently convicted and sentenced to lengthy prisons terms (*Frontline*, 1991, 1993, 1998; Nathan, 1990; Nathan & Snedeker, 1995).

Academic research has essentially debunked the validity of these cases. Researchers have empirically demonstrated that false memories can readily be implanted in preschoolers (e.g., Ceci & Bruck, 1993; Garven et al., 1998). While Ceci and his colleagues have focused on simple suggestion as a powerful implantation tool when suggestion is repeated over time, Garven et al. showed that the full range of techniques employed by the social workers in the McMartin case is substantially more powerful. Aside from suggestion, these techniques included social influence (e.g., misinforming children that others had already made allegations), positive and negative consequences (e.g., praise for “correct” answers, disappointment for “incorrect” ones), asked-and-answered questions (i.e., asking a child a question again that he or she has just unambiguously an-

swered, which conveys to the child that his or her first answer was wrong), and inviting speculation (i.e., asking the child to pretend or imagine what might have happened). Using preschoolers as subjects, Garven et al. showed that the full package of techniques more than tripled false memories (58% error rate) compared to suggestibility alone (17% error rate) during a single session. Investigative reports (e.g., Nathan, 1990; Nathan & Snedeker, 1995) and television documentaries (e.g., *Frontline*, 1991, 1993, 1998) have detailed the coercive techniques employed by therapists and prosecutors that elicited false allegations.

From the mid- to late 1980s and continuing well into the 1990s, thousands of mostly female adult patients began “recovering” memories of CSA while in therapy. Typically, the patient entered therapy complaining of unhappiness, but not knowing why. After being told that CSA might be the cause, encouraged to remember CSA experiences, and advised to read books such as Bass and Davis’ (1988) *The Courage to Heal*, vague memories of abuse might develop, which would later coalesce as vivid, detailed memories after many more therapy sessions. In thousands of such cases, patients proceeded to sue their assumed perpetrator (e.g., father), wreaking havoc in as many families (Jaroff, 1993; Ofshe & Watters, 1993; Prendergrast, 1996). *Frontline*’s (1995a) documentary “Divided Memories” provided numerous vivid examples. In one, a young woman not knowing how to relate to her mother went into therapy. After several months, with the encouragement of her therapist, she developed memories that eventually crystallized into recollections of satanic rituals involving her parents and grandparents. She was diagnosed as suffering from multiple personality disorder with twenty-seven alters. With her therapist’s support, she sued her parents for \$20 million. In a deposition, she claimed that her mother inserted a broomstick into her vagina, as well as spiders, wires, and vegetables, and that her father sexually assaulted her with tools from a hardware store. When reminded that she had perfect school attendance and pediatric reports, the daughter reacted, “Oh, that doesn’t matter.”

Although many clinicians continue to believe that recovered memories in adult patients are valid, academic researchers have

## 8 Sexuality & Culture

generally concluded that recovered memories are not supported by empirical data (e.g., Brandon et al., 1998; Crews, 1998; Kihlstrom, 1997; Loftus, 1993, 1997; Loftus & Polage, 1999; Ofshe & Watters, 1993; Pope & Hudson, 1995). In a recent comprehensive review of the literature, Brandon et al. (1998) found that opinions have often been expressed with great conviction but without evidence; they concluded that there is a high probability that memories “recovered” after long periods of amnesia are false. Their conclusions formed the basis of the official report issued by the Royal College of Psychiatrists in Great Britain in 1997. *Frontline*’s (1995a, 1995b) in-depth documentaries for PBS television have vividly portrayed the highly questionable nature of recovered memory therapy and the extensive iatrogenic damage it has engendered.

These dramatic events stemmed from social beliefs about the effects of CSA based on advocacy rather than empirical science. Some critics have referred to these events as the third wave of hysteria in America, the first two being the Salem witch trials and McCarthyism (e.g., Gardner, 1993). These occurrences highlight the importance of Brown and Finkelhor’s (1986) caution that exaggeration and overstatement have iatrogenic potential, and thus objectivity and balance are imperative in research on CSA. Our *Psychological Bulletin* article was in part a response to the problems just discussed. Our goal was to strive for greater objectivity and balance by strict adherence to sound methodological principles and analytical techniques, notably missing in many previous treatments of this issue.

### Summary of Our Two Meta-Analytic Literature Reviews

Since the early 1980s, several dozen literature reviews examining the relationship between CSA and psychological adjustment have been conducted. Most have been qualitative reviews, providing narrative summaries of results from studies based mainly on clinical and legal samples. Consistent with the pathogenic view of CSA, these reviews have generally concluded that CSA causes such diverse problems as depression, anxiety, low self-esteem,



sexual dysfunction, dissociation, suicidal ideation, and PTSD, and that these effects are pervasive in the general population of persons who have experienced CSA. However, these conclusions and the methodology used to achieve them are problematic. First, given that CSA has been consistently shown in the primary studies to be confounded with other problems such as physical abuse and emotional neglect, it cannot safely be assumed that CSA-symptom associations reflect the *effects* of CSA. Second, it cannot be safely assumed that findings from clinical and legal samples generalize. They may be a biased sample of CSA experiences in the general population, overestimating negative effects, or correlates. Third, narrative reviews are imprecise and subjective. It is not enough to know that CSA subjects are more poorly adjusted than controls; it is also important to know by how much. The narrative approach is also more vulnerable to confirmation bias, meaning that a reviewer with an initial belief may invalidly confirm the belief by emphasizing supportive findings while minimizing or ignoring unsupportive findings.

To address these shortcomings, we conducted two literature reviews examining the psychological correlates of CSA (Rind & Tromovitch, 1997; Rind et al., 1998). We focused on nonclinical/nonlegal samples for better generalizability. To obtain greater precision and objectivity we analyzed the data in a quantitative fashion by using meta-analysis, a technique that statistically summarizes results across studies. In our meta-analyses, each result represented the magnitude of association, as measured by Pearson's  $r$ , between CSA and adjustment, with small, medium, and large effect sizes corresponding to  $r$ s of .10, .30, and .50, respectively (Cohen, 1988). We analyzed results separately for males and females, and examined whether CSA-adjustment relations could be causally construed. Each literature review served as a test of four hypotheses: (a) CSA causes harm; (b) this harm is pervasive in the population of persons with a history of CSA; (c) this harm is likely to be intense; and (d) CSA is equivalent for boys and girls in terms of its widespread and intensely negative effects. These four hypotheses were derived directly from victimology and trauma theory, originating in turn from Freud's seduction theory.

*Meta-Analysis of National Probability Samples*

In our first meta-analysis (Rind & Tromovitch, 1997), we included only studies based on national probability samples. Because these samples were selected to be representative of national populations, their results are clearly more relevant to inferring effects in the general population than are clinical or legal samples, which represents a major strength of this review.

Prevalence rates for CSA for males and females were 11 percent and 19 percent, respectively. Although statistically significant, the association between CSA and adjustment was small: effect size  $r = .07$  for males and  $r = .10$  for females. Females reported more negative effects and reactions than males did. Although 68 percent of females reported some sort of negative effect at some point in their lives since the CSA occurred, only 42 percent of males did. This difference was statistically significant and of medium-small magnitude,  $r = .23$ . One study inquired about perceptions of permanent harm, finding it rare for both females (13%) and males (4%) (Baker & Duncan, 1985). Another study asked subjects about emotional reactions at the time of the CSA (López et al., 1995). Whereas only a minority of males reacted with negative emotions such as fear, disgust, and hostility at the time of the CSA, the majority of females did. Conversely, more males than females reacted with pleasure (27% vs. 10%). Across all emotions, there was a statistically significant sex difference of medium magnitude,  $r = .31$ . Finally, CSA was confounded with other environmental problems. One study found that sexually abused (SA) girls tended to have disruption in their family, school, and social environments *before* their CSA experience (Ageton, 1988).

Results contradicted or failed to support the four hypotheses. The confounding between CSA and third variables weakened assumptions about causality. The self-reported effects suggested that lasting harm is far from pervasive in the general population. The small effect sizes indicated that effects are not typically intense. Finally, consistent sex differences in self-reported effects and emotional reactions indicated that males and females, on average, experience CSA quite differently. The importance of these

findings is bolstered by use of national probability samples and precise statistical techniques. Weaknesses included the small number of studies used and scant data relevant to assessing causality.

### *Meta-Analysis of College Samples*

What about the review of college studies—the research that led to the congressional resolution? The small number of national probability samples suggested the need to extend this research by examining data from a larger collection of nonclinical/nonlegal samples where a greater amount of data relevant to assessing causality could be obtained. We chose to focus on studies based on college samples because: (a) these studies comprise the largest body of nonclinical/nonlegal research on CSA; (b) they are more representative of CSA in the general population than are clinical or legal samples because about 50 percent of adults in the U.S. have had some degree of college experience; (c) they provide the most extensive data on moderators of CSA-adjustment relations, which are relevant to assessing causality; and (d) they provide a rich amount of data on self-reports of effects and reactions, which are relevant to assessing pervasiveness of effects and sex differences. Another important advantage of these studies is the large amount of data available on males; most previous literature reviews have focused on females, including two meta-analytic reviews (Jumper, 1995; Neumann et al., 1996).

Potential shortcomings of using college samples are that students may be too young for symptoms to appear and students' CSA experiences might be less intense or severe than in the general population. The first possibility is doubtful; Neumann et al. (1996) showed that women younger and older than 30 were the same in terms of CSA-adjustment relations. As we demonstrated in our *Psychological Bulletin* article, the latter possibility is also unsupported. Prevalence rates were at least as high in the college population (14% for males and 27% for females) as in the general population, as indicated by results taken from national probability samples. Types of CSA (i.e., exhibitionism, fondling, oral sex, and

intercourse) in the college and national samples were similar, as were prevalence rates of incest and the frequency of CSA. Meta-analysis of the college data produced the same effect size estimates as the national probability data:  $r_s = .07$  and  $.10$  for males and females, respectively. Across eighteen different symptoms, effect sizes were generally homogeneous and were all small, ranging from  $r = .04$  for self-esteem to  $r = .13$  for anxiety. The magnitude of CSA-adjustment relations varied across studies as a function of the interaction of subjects' gender and reported willingness to participate, but not as a function of whether the CSA involved physical contact. Males who were involved in willing CSA were not different from controls in terms of adjustment. Willing and unwilling females and unwilling males, however, were somewhat more poorly adjusted than controls. Symptoms, reactions, and/or self-reported effects were worse when incest or the use of force was involved, but were unrelated to the occurrence of penetration or the duration and frequency of the CSA.

Positive, neutral, and negative reactions to CSA were, respectively, 37, 29, and 33 percent for males and 11, 18, and 72 percent for females. Negative reactions were significantly more frequent for females than males, with a medium effect size,  $r = .31$ . The vast majority of males and females reported no negative effects on their sex lives. Reports of lasting general effects were uncommon but were more frequent for females, and temporary negative effects were reported by a minority of males and a majority of females. Overall, reports of negative effects were significantly more frequent for females than for males, with a small-to-medium effect size,  $r = .22$ . Effect sizes for sex differences in reactions and self-reported effects were virtually identical to those in the national probability samples ( $r_s = .31$  and  $.23$ , respectively).

To investigate causality, we examined family environment (FE) as a third variable that might account for some or all of the association found between CSA and adjustment, which was  $r = .09$  across male and female subjects. We meta-analyzed FE-CSA relations and found that the two variables were reliably confounded with a small effect size,  $r = .13$ . Next, we meta-analyzed FE-adjustment relations, finding that they were statistically significantly related

with a medium effect size,  $r = .29$ . These latter two findings suggested that the CSA-adjustment relation would be even smaller if FE were statistically controlled for. Thus, we examined all studies that employed statistical control and found that 41 percent of CSA-adjustment relations were significant before statistical control but only 17 percent were significant afterwards—a 59 percent reduction of statistical significance when factoring out FE.

The small CSA-adjustment associations indicated that CSA does not produce lasting, intense effects on average in the college population. The self-reported effects and reaction data indicated that negative effects are far from pervasive and the two sexes respond quite differently to CSA. Analyses of confounding and statistical control indicated that causality is not well supported in the typical case. Nevertheless, CSA's potential for harm was indicated in a portion of the cases by self-reports of negative effects and the fact that some measures remained significant after statistical control. Results suggested that negative effects are increased by incest and use of force and, conversely, are lessened or eliminated by willingness of participation. Finally, the consistent similarities between results from the college and national samples suggest that these findings may be generalizable. Not only were prevalence rates, types of CSA, extent of incest, and frequency of episodes similar, but so were the magnitude of CSA-adjustment relations, self-reported effects and reactions, and sex differences in self-reported effects and reactions. These were the methods, findings, and conclusions that led to the later attacks.

### **Chronology of the Attacks**

Our *Psychological Bulletin* article appeared in print in July 1998. In December of that same year, the National Association for Research and Therapy of Homosexuality (NARTH), an organization with psychoanalytically-oriented clinicians at its core dedicated to the cure and prevention of homosexuality, offered the first critique, which appeared on its web page (<http://www.narth.com>). This critique, although scientifically superficial and morally based, was highly instrumental in initiating a chain of events that led to con-

gressional condemnation of our article. It formed the basis of numerous subsequent attacks by social conservatives, the first of which appeared in early March 1999 in *The Wanderer*, a conservative Catholic newspaper. This attack, drawn heavily from NARTH's critique, called our review a "pseudo-professional, pseudo-academic analysis." It claimed that "a team of academics from Temple University has endorsed the view that adult-child sexual relations are beneficial . . . and recommends overhauling and euphemizing the language of sexual abuse." It expressed regrets that homosexuality was depathologized and feared the same would now happen to pedophilia.

A listener to a Philadelphia talk station sent *The Wanderer* article to one of the station's hosts, who contacted one of us (Tromovitch), who agreed to appear on his radio show to discuss the study. After this appearance, the host launched an attack on Temple University, the paper, and us along the lines of the critique presented in *The Wanderer*. Another listener who heard this interview and the host's ensuing attacks wrote a letter to "Dr. Laura" Schlessinger, who runs a nationwide syndicated radio talk show, broadcast daily by about 485 radio stations in the U.S. and Canada, reaching about 20 million listeners. "Dr. Laura" is a social conservative who staunchly espouses family-values positions and has been dubbed "the poster girl of the Christian fundamentalists" by *Vanity Fair* magazine (*Der Spiegel*, August 2, 1999). The writer of the letter began by thanking "Dr. Laura" for her "constant fight against the assault on our families" and then characterized our study as a "new assault" (Schlessinger, 1999). On March 22, "Dr. Laura" began a series of attacks against our article and the APA that lasted for months. In developing her criticisms, she "had three renowned, licensed clinical psychologists and a scientist" review the article, who unanimously described it as "junk science" (Schlessinger, 1999). Three of these reviews, authored by clinical psychologists Samenow and NARTH members Nicolosi and van den Aardweg, appeared on her web page (<http://www.drlaura.com/monologue>). Finally, she made extensive use of criticisms provided to her by Paul J. Fink, former president of the American Psychiatric Association and cur-

rent president of a new group called the Leadership Council for Mental Health, Justice, and the Media.

The Leadership Council consists largely of persons whose psychological thinking is based on the psychoanalytic paradigm, who practice or advocate recovered memory therapy and the diagnosis and treatment of multiple personality disorder. They represent themselves as a nonprofit organization composed of “many of the nation’s most prominent mental health leaders” whose mission it is to “insure the public receives accurate information about mental health issues” (press release, May 24, 1999). Fink, writing on behalf of his organization, sent a letter to “Dr. Laura” asserting that we “loaded” our analysis with data involving primarily mild adult-child interactions with no physical contact, and that 60 percent of our data came from one single study done over forty years ago. Because these erroneous claims had a major influence on the fate of our paper, we will discuss them in detail later.

Shortly after “Dr. Laura” began attacking our study, the Family Research Council (FRC), a socially conservative lobbying group in Washington, D.C., joined in on the attack. Its stated purpose is to “reaffirm and promote nationally...the traditional family unit and the Judeo-Christian value system upon which it is built” (<http://www.frc.org>). To achieve this, it “keeps watch over political and cultural forces that threaten the traditional family, with a special focus on the homosexual agenda.” As such, the FRC has much common ground with NARTH and “Dr. Laura.” Attacks by the FRC, in turn, prompted the Alaska State Legislature to react, the first governmental body to do so. In mid-April 1999 a press release announced House Joint Resolution 36, which rejected the conclusions of our study, noting that “peer review has identified several questionable assumptions and methodologies in this research paper” (the only peer review performed on our article was that by *Psychological Bulletin*; presumably this “peer review” referred to “Dr. Laura’s” three clinical psychologists and a scientist). This resolution in turn served as the blueprint for numerous subsequent state resolutions as well as *H. Con. Res. 107*. It stated that the legislature “condemns and denounces all suggestions...[that] sexual relation-

ships between adults and willing children are less harmful than believed and might even be positive for 'willing' children," and concluded by encouraging "competent investigations to continue to research the effects of child sexual abuse using the best methodology so that the public and public policymakers may act upon accurate information."

On May 12, 1999, the FRC held a press conference in Washington, D.C. to demand that the APA repudiate our study. The FRC's press release for this event stated that our study was "based on the premise that a child can actually consent to sex with an adult." Participants included, among others, "Dr. Laura" via satellite, a representative from NARTH, and three conservative Republican congressmen (Reps. Delay-TX, Salmon-AZ, and Weldon-FL). Tom DeLay said the "lack of judgment shown by the American Psychological Association in publishing it absolutely confounds me" and "challenged" the APA to publicly admit it erred in publishing the article (*Washington Times*, May 13, 1999). According to the FRC's web page, Salmon called on the APA to denounce the study as "sick and twisted." Two days after the press conference, Raymond Fowler, CEO of the APA, appeared on MSNBC with Congressman Weldon. Weldon said that our study was "a very, very bad study . . . based on some very, very bad data" and that it should never have been published. Fowler replied that, "Well, with all due respect, it isn't a bad study. It's been peer-reviewed by the same principles as any kind of scientific publication. It's been examined by statistical experts. It's a good study." Weldon disagreed, saying that our study was based on "what they call meta-analysis, where they take a whole bunch of studies and put them together. But a whole bunch of studies that they put in this study were never peer-reviewed, and 60% of them were based on one study done over 40 years ago in 1946 [sic]." Weldon then read a quote from Fink, the source of the latter criticism.

Shortly after the MSNBC debate, Fowler contacted us, asking about the meaning of the 60 percent figure, adding that members of Congress were using it as "major data for discrediting" both the APA and us. We provided a complete refutation of Fink's comments (detailed below), but to no avail. The APA found itself in the



middle of a growing firestorm thanks to a coalition of social conservatives, two sorts of psychoanalysts—anti-homosexual and repressed memory advocates—and conservative Republican congressmen. As Fowler commented to us on June 8, he was “in hand to hand combat with congressmen, talk show hosts, the Christian Right, and the American Psychiatric Association.” This pressure, especially from Congress—with whom the APA must negotiate for political support and funding for both clinical treatment and behavioral research—proved to be too great. On June 9, Fowler wrote a letter to Congressman Delay that started by commending Delay for his strong stand against sexual abuse. Later he wrote that our article “included opinions of the authors that are inconsistent with APA’s stated and deeply held positions.” He then specified the “inconsistencies”:

It is the position of the Association that sexual activity between children and adults should never be considered or labeled as harmless or acceptable. Furthermore, it is the position of the Association that children cannot consent to sexual activities with adults.

Finally, Fowler offered a series of unprecedented concessions, among them that the APA would seek independent evaluation of the scientific quality of our article and that its journal editors would be asked to “fully consider the social policy implications of articles on controversial topics.” As a result, APA was congratulated rather than condemned in *H. Con. Res. 107* for “clarifying its opposition to any adult-child sexual relations...and for resolving to evaluate the scientific articles it publishes in light of their potential social, legal, and political implications.”

Critics of our review hailed Fowler’s concessions and Congress’ condemnation as a major victory. In an August 2, 1999 newsletter, the FRC listed its recent accomplishments, one being its role in Congress’ condemnation of our article. The Leadership Council also took credit for Congress’ condemnation, as well as the APA’s decision to have our paper reevaluated. In an unpublished letter written to the *Los Angeles Times*, Fink, Silberg, and Dallam (1999) wrote that

Our in depth analysis of the study was presented to the APA and members of the United States Congress. Soon after reading our analysis, APA took the unprecedented step of sending the study to an independent scientific organization for review. And Congress passed a bill condemning the study...

They later emphasized that “Congress passed the bill only *after* receiving our analysis. Thus they had firm data to support their position” (*italics in the original*).

But our critics’ celebration is not the end of the story. As discussed previously, AAAS rebuked their misunderstandings and distortions, finding no fault in our methods or analyses. AAAS’s opinion was far from isolated. Once Congress got involved, prompting news coverage in the mainstream media, numerous supportive commentaries began to emerge and continue as of this writing. These include, among others, articles in *The New Republic* (May 28, 1999, by J. Zenergle, feed service), *Der Spiegel* (August 2, 1999, by R. Paul, pp. 190-91), *National Journal* (August 7, 1999, by J. Rauch, pp. 2269-70), *New York Times Magazine* (November 7, 1999, by A. Sullivan, pp. 38-40), the *Skeptical Inquirer* (January/February 2000, by Berry & Berry, p. 20), *The Public Interest* (Winter 2000, by G. Zuriff, pp.29-39), and *Lingua Franca* (February, 2000, by S. Cole, pp. 12-14). Supportive newspaper columns or op-ed pieces have included, among many others, the *San Francisco Chronicle* (June 17, 1999, by J. Carroll, E12), *Los Angeles Times* (July 19, 1999, by C. Tavis), and *Boston Globe* (August 1, 1999, by S. Lamb, E1). The essence of these commentaries was that our review has been seriously distorted by social conservatives, exploited by political opportunists, was properly done, and has made important and valid distinctions. For example, in his popular syndicated column “Savage Love” (July 29, 1999), Dan Savage commented:

Why is this controversial? Speaking as a survivor of CSA at 14 with a 22-year-old woman; sex at 15 with a 30-year-old man—I can back the researchers up; I was not traumatized by these technically illegal sexual encounters; indeed, I initiated them and cherish their memory. It’s absurd to think that what I did at 15 would be considered “child sexual abuse,” or lumped together by lazy researchers with the incestuous rape of a 5-year-old girl.

### **Refutation of Criticisms Concerning Methodology and Data Analysis**

The AAAS opinion released in the fall of 1999 characterized the vocal critics of our article as lacking understanding of the analyses,

misrepresenting the findings, and/or failing in their responsibility to be accurate about the evidence that informed their public statements. In this section, we document the misunderstandings and misrepresentations concerning our methods and analyses. In the next section, we will document the misunderstandings and misrepresentations concerning two conceptual issues that led to especially harsh denunciations: our recommendations regarding terminology and our use of the construct of willingness or consent.

### *NARTH*

NARTH's critique, as mentioned earlier, was highly influential in igniting the attacks on our article. Its specific criticisms, however, did not deal directly with our methodology or data analysis—points with which any valid scientific rebuttal ought to be concerned. Instead, NARTH's critique focused on the fear that pedophilia would be normalized in the same way that they allege homosexuality was—by suggesting the use of value-neutral terms. NARTH dealt with our findings not by challenging them, but by presenting its own view as universal truth—i.e., CSA is invariably intensely harmful.

Citing mostly clinical research, NARTH claimed that the consequences of childhood seduction include guilt, shame, anxiety, lowered self-esteem, depression, vulnerability to drug and alcohol abuse, increased risk for suicide, and sexual problems (e.g., homosexuality). In making these claims, NARTH repeated the very same errors we described in the introductions in our two meta-analytic reviews. To reiterate, we argued for the need to conduct yet another literature review because previous ones predominately focused on clinical samples, reported extensive lists of symptoms without qualifications, assumed CSA was the cause of all problems despite confounding, and inappropriately generalized from clinical samples to all experiences of CSA. We discussed the need for more representative samples and caution in causal interpretation, and in our own reviews used national probability and college samples and paid careful attention to causal interpretation through statistical analysis and control. We demonstrated the usefulness of college samples by their high degree of correspondence with na-

tional samples in all relevant measures. NARTH completely ignored our logic, methods, and analyses. As if in “denial,” NARTH covered its eyes and just reiterated the prevailing dogma. Remarkably, although especially superficial, its critique had a major impact.

*“Dr. Laura” and the Criticism of Unpublished Studies*

One of “Dr. Laura’s” sources (Samenow) complained that our review was weak because 38 percent of the studies were never peer reviewed. “Dr. Laura” popularized this criticism, and Congressman Weldon used it to attack our study in his debate with Fowler on MSNBC. However, this criticism is methodologically and empirically without merit.

As a methodological issue, using unpublished studies in a meta-analysis is not only accepted practice, but is actually encouraged, provided the studies are well conducted (Rosenthal, 1994). Including unpublished studies is important because journals may have a bias toward publishing studies with statistically significant results, which in turn inflates estimates of the magnitude of relationships between variables (Rosenthal, 1984). To overcome this bias, referred to as the “file drawer problem” (i.e., studies with non-significant results are stashed away in file drawers rather than being published), meta-analysts who use only published studies and find an overall statistically significant relationship often compute the number of unpublished studies with null results that it would take to make the result nonsignificant. An alternative is simply to locate and use unpublished studies, as we did. Most (21 out of 23) of the unpublished studies we used were doctoral dissertations and two were master theses. As most academics know, doctoral dissertations are generally well supervised by a group of Ph.D.s through all stages of the study, and must be defended before a panel of Ph.D.s who may question the methods, analyses, and conclusions. Thus, dissertations appear to meet Rosenthal’s (1994) “well-conducted” criterion.

The argument just made, however, may not satisfy critics. But there is a more compelling argument—an empirical one. If the un-

published studies are anomalous compared to published studies—in this case, vastly underestimating CSA-symptom associations—then there should be a difference in CSA-symptom associations in the published versus unpublished studies. We conducted such an analysis and presented its results in our review (Rind et al., 1998, p. 34). We found that these associations did *not* statistically differ. The mean effect size was  $r = .11$  for 27 published samples and  $r = .08$  for 24 unpublished samples. From a practical point of view, these mean effect sizes are the same—both small and both hovering around  $r = .09$ , the mean effect size obtained in the published studies using national probability samples. Additionally, among the 54 effect sizes that we meta-analyzed, all but three were homogeneous—i.e., statistically consistent with the overall mean effect size for the college samples, which was also  $r = .09$ . The three outliers, two above the mean and one below, were all published. Thus, all unpublished studies were consistent with the overall trend, showing they were not anomalous and that they did not bias the results.

*The Leadership Council and its Criticism that We “Loaded our Analysis”*

As noted previously, the Leadership Council and its president, Paul J. Fink, were highly influential in characterizing our study as flawed and having it condemned. Congressman Weldon cited Fink’s comments as authoritative in showing that our study was based on very bad data, should never have been published by the APA, and should be repudiated by the APA. In an opinion piece, which we will discuss later in more detail, psychologist Romy Cawood defended our article as sound in all respects (*Greensboro News & Record*, July 4, 1999, H7). In response, one letter argued that Cawood left out some important facts, foremost of which was that “Dr. Paul J. Fink, past president of the American Psychiatric Association, has condemned the study” (*Dayton Daily News*, June 26, 1999, 13A). Another, written by Keven Bellows, the vice president and general manager of the “Dr. Laura” program, complained that Cawood’s ridiculing of “Dr. Laura” was unfounded because she, rather than relying on her own analysis to debunk our study,

consistently referenced “an analysis refuting the study by the Leadership Council....It was Fink’s scathing critique of the study that prompted Congress to pass a near-unanimous resolution condemning it” (*Dayton Daily News*, July 1, 1999, 11A). Because of the enormous influence wielded by Fink, his group, and their critique, we now devote considerable attention to their criticisms.

Right after he appeared on MSNBC, Fowler of the APA asked us for clarification on Fink’s assertions that, of the 59 studies we used, over 60 percent of the data came from one single study done 40 years ago, and that we “loaded” our analyses with data that were based on primarily mild interactions involving no physical contact. Fink was referring to the Landis (1956) study. The following is a summary of the points we made to demonstrate that these assertions are specious.

First, the Landis study was *not* used in our meta-analyses of CSA-symptom associations, which constituted the primary and most important analyses of our study. We used 54 samples comprised of 15,912 participants in these analyses, not one of whom came from Landis. Thus, we did not “load” our primary analyses with the Landis study.

Second, we did use the Landis data for our analyses of self-reported reactions to CSA. Our summary table (Rind et al., 1998, p. 36) shows that data were drawn from 9 male and 9 female samples. The Landis samples comprised more cases than all other samples (493 of 1,421 female cases; 183 of 606 male cases), but far less than 60 percent (in fact, Landis represented 35% of female and 30% of male cases). Among all samples, the Landis reaction data were the most negative. In computing mean reactions across samples, we employed *weighted* means, thereby giving the most weight to Landis’ negatively skewed data. This approach maximized reports of negative reactions relative to alternative methods. Using the weighted approach, mean positive and negative reactions for males were 37 and 33 percent, respectively. Using an *unweighted* approach would have yielded more positive and less negative reactions: 43 and 30 percent. Dropping the Landis data completely would have yielded even more positive and less negative reactions: 50 and 24 percent. The critics’ implication that we

used Landis' data to minimize reports of negative reactions is in fact the opposite of what happened; we handled them in such a way as to maximize negative reports. Notably, anyone could use the data from our table to determine the exact effect of Landis' study on the overall findings; the critics apparently failed to make these simple calculations.

Third, we analyzed self-reported effects from the 6 male and 5 female samples that had this information. Here, the Landis data made up 53 percent of the total *N* for males and 68 percent for females (combined was 63%). It appears that Fink and his group derived their criticism specifically from this analysis—but that was *not* implied in Fink's letter to "Dr. Laura" or subsequent public representations of it, in which all our analyses were implicated, which is clearly false. In our review, we first examined self-reported negative effects on subjects' current sex lives or attitudes. For males, negative effects ranged from 0.4 percent (Landis, 1956) to 16 percent (Condy et al., 1987). We computed their *unweighted* mean, obtaining 8.5 percent. Given that the Landis data were the least negative but by far most numerous in terms of cases, our approach *maximized* self-reported negative effects. Using a *weighted* approach would have yielded 4.4 percent. For females, the Landis data were similarly the least negative (2.2%) and most numerous. The unweighted mean that we reported was 13 percent, whereas the weighted mean would have been only 3.8 percent. Once again, our approach maximized rather than minimized negative reports. Additionally, we examined general reports of lasting negative effects using three male samples (0% for Landis; 0% and 27% for the other two) and three female samples (3% for Landis; 20% and 25% for the other two). In these cases, we did not give means, but instead concluded properly that lasting self-reported negative effects occurred for only a minority of the students—a conclusion that holds *independently* of the inclusion of the Landis study.

Fourth, the claim that we "loaded" our analysis "primarily with mild adult-child interactions involving no physical contact" is false. We included all the studies that were available at the time, 16 of which included cases exclusively involving physical contact. We examined whether CSA-symptom associations differed between

contact and non-contact CSA. They did not (see Rind et al., 1998, p. 33). The studies we included, in contradiction to Fink's claim, represented well the severity of CSA found in the general population, as we demonstrated in our review (Rind et al., 1998, pp. 29-31).

Fifth, we were criticized for using the Landis study in the first place. This criticism is selective, however, given that abuse researchers examining college samples have often cited this study for background (e.g., Finkelhor, 1979; Fromuth & Burkhart, 1989). Fink's implication that the Landis study is invalid because of high rates of non-contact CSA is itself misleading. Victimologists have expanded the definition of CSA to include non-contact experiences ranging up to age 18 in some studies. Even with this expanded definition, they still have characterized all CSA as pathogenic. We simply tested assumptions about CSA using the researchers' own definitions of CSA. Even if we assume that all self-reports of negative effects in the Landis study involved contact CSA, given that about 10 percent of male and 41 percent of female experiences were contact CSA, reports of lasting sexual harm would involve only 4 percent of male and 6 percent of female contact cases and reports of lasting general harm would involve 0 percent of male and 7 percent of female contact cases. There is no practical difference between these figures and the ones we used in our review; they are all in the single digits.

Fink's group has also asserted that Landis' data should be dismissed because, in the case of males, they consisted mainly of "mild" non-contact CSA—teenagers rejecting unwanted approaches. The implication is that if the teenage males had been involved in contact CSA, then self-reported effects would have been highly negative. Fromuth and Burkhart (1987) provided data directly relevant to this implication. They found that approximately 30 percent of their SA male students had experienced oral sex and another 25 percent intercourse—a fair degree of "severe" contact CSA. Instead of primarily negative self-reported effects, 39 percent reported a positive effect on their lives, 46 percent a neutral effect, and only 15 percent negative. Restricting the results to teenagers (aged 13 to 16 at the time of the CSA) resulted in an even



more positively skewed distribution: 60 percent positive, 37 percent neutral, and 3 percent negative. These results, missing from our original review where we relied on Fromuth and Burkhart's 1989 report, flatly contradict Fink's contention.

In short, the Landis data played no role at all in the key analyses of our study—the meta-analyses of CSA-symptom relations. For self-reported reactions and effects, we analyzed the Landis data in ways that maximized reporting of negative outcomes. Overall, the studies we used accurately represented levels of contact CSA in the general population, and restricting Landis' results to contact CSA left our basic conclusions untouched. The Leadership Council's claim that we "loaded" our analyses fails at every turn.

#### *Additional Criticisms by the Leadership Council*

In addition to the false claim that we loaded our analyses, which lay critics easily absorbed and widely cited, the Leadership Council made a number of more "technical" criticisms that have not been widely cited but nevertheless played a key role in our article's fate, as Fink et al. (1999) claimed in their unpublished letter to the *Los Angeles Times*, discussed previously. To reiterate, they claimed that their group's in-depth analysis, authored by Dallam et al. (1999), prompted the APA to seek an independent review of our article and Congress to condemn it. We recognize that this analysis was in draft form. But because it, and not some future version, had the alleged impact of prompting two historical events—i.e., re-review of an already rigorously peer-reviewed article in a premiere psychological journal and congressional condemnation—a careful scrutiny of this critique is in order for archival purposes. As we will show, this critique has a patina of scientific rigor to it, but it is in fact severely flawed. Because of its significance, we present and then address nearly all its points.

*Sample bias.* The authors claimed that our argument that college samples are representative of the general population because the prevalence of abuse was similar is unconvincing. They argued that similar prevalence does not necessarily mean similar experience

(e.g., severity) or outcomes and that those who made it to college may have had less severe CSA, been less distressed, better copers, and so on. As described in our *Psychological Bulletin* article and in the summary presented above, not only did we find similar prevalence rates between the college and national probability samples, but we also found similar experiences in terms of severity (i.e., types of CSA; degree of incest; frequency) and outcomes (e.g., identical mean effect sizes, similar reactions and self-reported effects). Given our documentation of these extensive similarities in the original article, the authors' speculation about lesser severity, distress, or better coping is not only empirically baseless but negligent as well.

*Inclusion of non-contact CSA.* The authors criticized our use of studies including non-contact experiences, claiming that this minimized effect sizes, and cited our use of the Landis (1956) and Risin and Koss (1987) studies as particularly problematic because of their large sample sizes. First, this criticism is better directed at victimologists who established the inclusion of non-contact experiences in definitions of CSA, leading the researchers of the studies we reviewed to use such overbroad definitions. We were testing victimological assumptions and therefore followed their definitions. Second, their citation of Landis (1956) and Risin and Koss (1987) is highly misleading, given that these two studies were *not* included in the meta-analyses of CSA-symptom relations. Third, despite this definitional weakness, we found that over all the studies, effect sizes did not differ between studies including only contact CSA and those with broader definitions. We note further that in Laumann, Gagnon, Michael, and Michaels' (1994) face-to-face interview study using a national probability sample, arguably the best and most important study ever conducted for understanding sexual behavior and its correlates in the general American population, CSA-adjustment effect sizes were only  $r = .07$  for males and  $r = .05$  for females. This study included only contact CSA occurring before puberty. Moreover, symptoms were uncorrelated with CSA severity (e.g., touching vs. oral sex or penetration). These results further support the conclusion that we were not reducing effect size estimates by including non-contact cases. These conceptual and empirical points show this criticism to be without merit.

*Poor measurement.* The authors claimed the studies we included were not uniform with respect to issues such as purpose, questions asked, and definition of CSA. In fact, the studies did have a common purpose: to examine outcomes of CSA. True, the studies varied widely in questions asked and definitions, but this problem applies to all CSA literature reviews. Nevertheless, the authors selectively applied this criticism to our review. More importantly, from an empirical view, had the nonuniformity been problematic, then heterogeneous effect sizes should have emerged. But across 94 percent of the samples, the effect sizes were homogeneous, which empirically invalidates this criticism.

The authors also claimed that identification of CSA based on retrospective self-report, as done in the college studies, is biased toward underreporting CSA because of factors such as shame, denial of victimization, or amnesia. Once again, this is a selective criticism that could be applied to all reviews. More importantly, the college studies were almost all anonymous and usually asked about “experiences” rather than “abuse” to minimize nonresponding, which weakens shame or denial as a biasing factor. Given that claims for amnesia for CSA are based more on conviction than scientific evidence, and that recovered memories after long periods of amnesia are most likely false (Brandon et al., 1998), amnesia cannot be accepted as a valid threat to self-reporting CSA.

The authors then claimed that our reporting different outcomes separately was a “divide and conquer” technique that attenuated effect sizes. Ignoring the obvious, that one should separately analyze different symptoms because CSA might have an impact in one area but not another, given the homogeneity of effect sizes across samples and the uniformity across symptoms, this criticism is without empirical support. It is also selective, in that analyzing and reporting outcomes separately is standard practice in CSA literature reviews. The authors claimed that only some CSA victims will have any given symptom (but all will have some symptom), minimizing assessed effects of CSA. Again, this represents selective attention to our review, when it applies to all others as well. This criticism also suffers from unfalsifiability (“data picking” and “multiple endpoints,” in which researchers seize upon any observed

symptom and attribute it to CSA), which can produce spurious confirmation of cause and effect (e.g., Gilovich, 1991). The authors also argued that the college studies used general measures of symptoms rather than CSA-specific ones, most notably PTSD. This is a tired criticism that has already been answered in other forums (e.g., Nash et al., 1993). Because CSA has come once again to be seen as the cause of all adult neuroses (Esman, 1994), it is valid to examine all types of CSA-symptom relations. Moreover, given that PTSD implies severe psychopathology, it is unreasonable to argue that measures such as anxiety and depression are not valid for examining CSA effects. Finally, "sexual problems" is often viewed as a CSA-specific outcome. Yet its effect size ( $r = .09$ ) was completely consistent with other CSA-symptom relations rather than being larger.

The authors went on to complain that some of our coding of data was misleading. As an example, for neutral reactions, they claimed that we coded Condy et al.'s (1987) "mixed" category and Long and Jackson's (1993) "low" and "ambivalent" responders as neutral reactions, yet nowhere did we reveal this to our readers. This claim is simply false; in Table 7 (Rind et al., 1998, p. 36), we used superscript *a* for our neutral percents for both of these studies, clarifying the coding procedure. They also complained that we coded Fromuth and Burkhart's (1989) "surprised" category as neutral. We had good reason to do so: these researchers based their measure of reactions on that of Finkelhor (1979), which in turn reflected that of Landis (1956), both of whom interpreted "surprised" as a neutral response. In our recent reading of Fromuth and Burkhart (1987)—their other publication on the same data—we found that neutral reactions were 30 percent, as opposed to the 28 percent we reported in our review. Using the coding of neutral data as an example, the authors suggested that all our conclusions regarding the reaction data were invalid. But even if we removed all three of the disputed studies, the overall percents from the remaining studies are virtually identical to what we reported in our review; none would differ by more than 2 percent. This did not stop the authors from claiming that all our conclusions regarding the proportion of negative reactions were invalid, charging that our statements based on

these data were “ill responsible [sic] and seem to be intent on misleading the reader.”

*Poor choice and interpretation of effect sizes.* The authors noted that Cohen (1988) described  $d = .8$  as a large effect size and then claimed that such a value, when converted to  $r$ , can fall anywhere between .00 (when CSA prevalence is 0%) and .39 (when CSA prevalence is 50%). They argued that given the low prevalence of CSA in the college population, the use of  $r$  rather than  $d$  reduced the appearance of the effect sizes—claiming that  $rs$  of .1 or .2 may still be large effect sizes, that is, if expressed in  $d$ . Their assertion, however, is purely speculative. Furthermore, this speculative criticism is negligent in that they completely ignored the fact that we directly addressed this issue in our review (Rind et al., 1998, p. 41) in a section entitled “statistical validity.” We noted that effect size attenuation is quite small for a 27–73 split (CSA prevalence for females). It is somewhat larger for a 14–86 split (CSA prevalence for males), but is still small in absolute magnitude for small effect sizes—i.e., an  $r = .07$  based on a 14–86 split increases at most by .03 (to  $r = .10$ ) in a 50–50 split. Furthermore, for the current article we examined their speculation empirically by converting all  $rs$  to  $ds$  using Rosenthal’s (1984) formula that takes into account population prevalence rates—we converted using sample prevalences from the individual samples. For the 14 male samples, with mean  $r = .07$ , we obtained mean  $d = .22$ ; for the 33 female samples, with mean  $r = .10$ , we obtained mean  $d = .25$ . Both of these mean  $ds$  are small, not large, according to Cohen’s (1988) designation in which  $d = .2$  is small and  $d = .8$  is large, which empirically flatly contradicts the authors’ speculation. As a side note, they further claimed, without providing evidence or a citation, that converting  $d$  to  $r$  does not mean that the converted  $r$  has the properties of the Pearson correlation coefficient. As we obtained all  $rs$  from  $ts$ ,  $Fs$ , or chi-squares, not from  $ds$ , this point is irrelevant.

*Deck stacked against CSA compared to FE.* The authors argued that the deck was stacked against CSA and in favor of family environment (FE) in accounting for symptoms variance, because CSA was measured dichotomously while FE was measured continuously. However, we fully addressed this issue as well in the section on statis-

tical validity. We reviewed several studies where CSA was constructed on a continuous scale and was then compared with FE in terms of accounting for adjustment. Based on this, we concluded that “[r]esults from these studies in which CSA was constructed to be continuous are consistent with results from studies in which CSA was treated dichotomously in terms of pointing to family environment, rather than CSA, as a significant contributor to current adjustment” (p. 41).

The authors also argued that the deck was stacked because measures of CSA were unreliable compared to measures of FE. Again, we directly addressed this issue in our review, providing evidence that allowed us to conclude that “[t]hese results point to acceptable reliabilities for measures of CSA, which are comparable to reliabilities for family environment measures” (p. 41). Once again, the authors ignored our discussion of this issue. The authors then argued that the statistical control we used was suspect because CSA and FE were confounded, citing Briere (1992) to argue that such statistical control is invalid under certain circumstances. In our treatment of this issue, we thoroughly addressed this point (pp. 43-44), showing why Briere’s criteria, although relevant in the clinical population, do not obtain in the college population.

*No way to infer causality or lack of it.* The authors then argued that there is no way to infer causality or the absence of it from retrospective observational data. All we should have been doing, they argued, was to examine the magnitude of relations. Once again we see selective criticism of our review when this criticism could, and *should*, be applied to all other reviews. Considering their later causal assertions about the “growing and well documented” literature on “the harmful and long-lasting *effects* of adult-child sexual contact” (italics added), we find this criticism disingenuous and indicative of extreme bias. In addition, the authors’ reasoning is erroneous. Correlation alone can never prove causality, but a demonstrable lack of correlation disproves causality.

*Double standard of interpretation.* The authors next argued that we claimed that young people (a) cannot be counted on to determine if their experience was harmful, but (b) can be relied upon to determine if it was not harmful. Supposedly, we supported the first claim by citing Nisbett and Wilson (1977), who concluded that

people are often unaware of the causes of their behavior when causal relations are ambiguous or complex. And we supposedly supported the second by suggesting that a willing encounter experienced positively be labeled adult-child sex, which they referred to as a “rather bizarre recommendation.” Once again, the authors misrepresented what we wrote. First, we cited Nisbett and Wilson (1977) to suggest that students may have been harmed even though they perceived no harm—the opposite of what the authors charged. Second, we explicitly stated in the Discussion that self-reports of lasting harm imply genuinely negative effects of CSA (p. 44). Below we address in detail our so-called “bizarre recommendation.” Given that we acknowledged that perceptions of harm imply negative effects, it is completely consistent to assume that perceptions of positive effects or reactions do not imply harm. It is the authors and victimologists in general who have the double standard of interpretation, because they accept the validity only of perceptions of harm.

*Misleading interpretation of misrepresented data.* The authors began by charging that we “loaded” our analysis with the Landis (1956) data, which they stated “represents a serious misreporting of data.” This characterization is itself a serious misrepresentation, as we have already shown (see above). They next once again attacked us for our “misuse” of data we coded as neutral in a few of the studies. Based on this, they claimed that all our conclusions about reactions were faulty, calling us “ill responsible [sic]” with seeming “intent on misleading the reader.” It is thus instructive to show precisely what happens when we remove the disputed studies and re-compute mean reactions. For males, positive, neutral, and negative reactions become 36, 31, and 34 percent (we reported them in our review as 37%, 29%, and 33%); for females, they become 12, 16, and 72 percent (we reported them as 11%, 18%, and 72%). These virtually identical results highlight Dallam et al.’s bias.

We are also excoriated for not including the types of disclaimers found, for example, in West and Woodhouse (1993). According to the authors, these researchers wrote in their conclusion that most of the incidents reported by male subjects were either trivial approaches that were soon rejected or minor indecencies that the boys saw as unimportant. The authors’ selective paraphrase is what is in fact

misleading. West and Woodhouse's (1993) most important conclusions coincide with ours. They stated that their "research and others like it have shown that boys' sexual encounters with older males and females are far from rare and for the most part *fairly innocuous*" (p. 126, italics added), while "the more serious kinds of adult-child sexual interaction are unusual" (p. 127). Furthermore, although many cases in West and Woodhouse's survey involved trivial approaches that were soon rejected, many other cases involved sexual contact perceived positively by the boys, which Dallam et al. seem incapable of acknowledging. West and Woodhouse (1993, pp. 32-88) reported lengthy interviews of 24 students, of whom 7 had oral sex or intercourse with adults. Of these, 6 experienced these encounters positively (cases 008, 050, 249, 261, 297, 337), while only one reacted negatively (case 239). Two of these students had heterosexual encounters, four had homosexual encounters, and one had both. Negative or neutral reactions occurred almost exclusively in the less "severe" cases involving approaches or fondling. Our own discussion of West and Woodhouse accords well with these findings.

*Flawed conclusions not well supported by the data.* The authors then claimed that research does not support our contention that boys are not typically harmed by CSA. To support this, they argued that boys who do not perceive harm are committing misattribution. They cited Myers (1989) to argue that sexually abused boys tend to deny or normalize their abuse. Their claim of misattribution is circular in that harm is assumed to occur, and lack of perceived harm is therefore inferred to be cognitive distortion, thereby "proving" the initial assumption. Myers' (1989) research was a clinical study involving unusual cases of unambiguous and severe abuse; their use of this citation shows an obstinacy in insisting that the typical case is validly characterized by the extreme. Our narrative literature review of nonclinical cases of boy-adult sex (Bauserman & Rind, 1997) coincided in its conclusions with those of West and Woodhouse (1993) just discussed: serious cases are rare and typical cases are fairly innocuous. The authors then used their misattribution claim and citation of Myers to argue that "we cannot assume that boys were not harmed based solely on their self-reports." This is a double standard, because the authors clearly ac-



cept reports of harm as evidence for harm. Ironically, the authors falsely accused us of having a double standard (see above), while exhibiting one themselves.

*Ignored majority of sample in arguing for lack of harmfulness.* The authors claimed that we minimized evidence for harm in our discussion, despite having found that all CSA-symptom associations but one were significant. What the authors ignored is that we demonstrated that all of these associations were small; FE was confounded with CSA; FE accounted for nearly ten times as much adjustment variance as CSA; and statistical control tended to reduce CSA-symptom associations to nonsignificance. They then stated that it was a “small minority—mainly older, *willing* adolescent males—who appeared to have suffered less harm” (italics added). First, in view of their very next criticism, where they dismissed the possibility of consent in age-discrepant sexual relations involving minors, it is remarkable that they use the term “willing” in the current criticism. Second, their phrase “suffered less harm” shows that their basic position is unscientific because it is unfalsifiable—harm is a premise that data cannot dispute.

### *Summary*

Methodological or statistical criticisms were absent in NARTH’s critique, which instead focused on reiterating clinically-based findings and dogma. “Dr. Laura’s” major criticism of the use of unpublished studies was flawed on both conceptual and empirical grounds. The Leadership Council’s claim that we “loaded” our analyses to minimize harm was blatantly wrong. Its in-depth critique was flawed by such problems as patently false assertions, speculations shown to be faulty by empirical examination, and bias in terms of selective attention and misrepresentation.

### **Refutation of Criticisms Concerning Conceptual Issues**

The major attacks on our review have not been methodological or statistical, however. They have concerned our suggestion that the term “child sexual abuse” is not always appropriate and our use

of the construct of willingness or “consent.” In this section, we present detailed discussions of these conceptual issues, showing that our treatment of them was scientifically sound.

### *Terminology*

NARTH was the first to attack the suggestion in our discussion that certain types of CSA should be relabeled by researchers with the value-neutral terms “adult-child sex” or “adult-adolescent sex” (see Rind et al., 1998, p. 46). NARTH misrepresented what we wrote, falsely claiming that we recommended that psychologists should stop using terms such as “sexual abuse” and should use the phrase “level of sexual intimacy” instead of “severity of abuse.” Regarding the latter point, what we actually wrote, in discussing the progression from exhibitionism to masturbation to intercourse, was that “many authors referred to this increasing level of sexual intimacy as ‘severity’” (Rind et al., 1998, p. 29). This distortion was repeated numerous times in opinion pieces around the country spreading a false impression of irresponsibility and lack of sensitivity. NARTH also attacked our view that science should separate itself from moral language, and complained that replacing the term “abuse” with neutral terms is “a repetition of the steps by which homosexuality was normalized.” Their logic resonated with many subsequent critics.

*The Wanderer* article asserted that we recommended “overhauling and euphemizing the language of sexual abuse.” Dallam et al., writing for the Leadership Council, charged that the “insistence on ‘value-neutral’ terminology normalizes child sexual abuse.” Steven Mirin, Medical Director of the American Psychiatric Association, wrote in a letter to the FRC that “academic hair-splitting over whether the act should be considered adult-child sex or child sexual abuse... is not in the public interest and obfuscates the moral issue involved.” And the APA agreed with them all, stating in its letter to Congressman DeLay that some of “the language in the article... is inflammatory.” In the end, the APA characterized this usage as our opinion, inconsistent with APA’s position, and then added that sexual activity between children and adults should never be “labeled as harmless.”

In fact, our suggestions for terminology resulted directly from the editorial and peer review process with APA's premiere journal. In our original two drafts, we did not suggest the use of value-neutral terms. This recommendation came only after the action editor requested an in-depth discussion of the "child sexual abuse" term. In response, we added comments in the Introduction (pp. 22-23), wrote a new section in the Discussion entitled "Child sexual abuse as a construct reconsidered" (pp. 45-46), and added recommendations in the "Summary and Conclusion" section (pp.46-47).

In accepting our article for publication, the action editor wrote that the "major and most difficult issue to address is the central one raised by Reviewer A concerning the conceptual and definitional issues . . . [that] might need to be reconsidered in light of your findings." After noting that 37 percent of males reacted positively to their CSA experience at the time, and 42 percent in retrospect, he wrote:

Although these experiences might meet legal and social definitions, the data suggest that the operationalizations employed might not sufficiently contextualize the events in such a way that adequately captures the essence of "abuse." Please note that I am not condoning behaviors that meet current definitions of CSA any more than I condone illicit substance use in minors. Indeed, both types of behaviors are legally and socially proscribed. Both, however, need to be contextualized in order to carefully assess their pathogenicity. . . . [P]erhaps we need to be more thoughtful about how we define CSA at a psychological level. That is, current definitions may not be sufficiently probing. I base this conclusion on the data regarding the extent that such experiences were positive, and the extent that such experiences correlate with outcome in men if the CSA was unwanted. . . . I'm not encouraging a conceptual definition that requires harm as an effect . . . but one which captures the essence of "abuse" . . . . With respect to the "big picture," I think you need to . . . spend more time in your discussion elaborating the conceptual and operational implications of your review. I believe that, in doing this, you can make the substantive contribution sufficient to warrant publication in *Psychological Bulletin*.

Reviewer A had noted that definitions for CSA have been too diffuse and inconsistent, resulting in "poor predictive utility." The idea is that differences in adjustment would likely be better accounted for if the term CSA were restricted to a subset of the very wide range of experiences currently labeled CSA, which would

advance understanding of CSA and prediction of its effects. In fact, another reviewer on a later draft reinforced the notion that the term “abuse” is problematic, recommending that we change “child sexual abuse” in our title to “child and adolescent sexual experiences”—a step we chose not to take.

Hence, our assignment was to reconceptualize the term “child sexual abuse.” We carefully outlined the problems caused in the past by the mixing of morality and science in other areas of sexuality, such as the seventeenth century transformation of masturbation from sin to sickness and medical representation of it as “self-abuse,” hindering scientific understanding of this behavior and creating iatrogenic victims in the process. We noted how several researchers in the college studies came to question their broad definition of the term “child sexual abuse” after gathering empirical data. We then discussed the problems the term CSA created for scientific validity, based on the results of our review. We noted that overinclusive definitions of CSA (combining experiences perceived as willing and positive with those perceived as coercive or negative) produced poor predictive validity—that is, CSA was not particularly predictive of negative outcomes, which it should have been had it been properly conceptualized. Because terms or constructs are scientifically useless if they lack validity, we argued “[t]o achieve better scientific validity, a more thoughtful approach is needed by researchers when labeling events that have heretofore been defined sociolegally as CSA” (p. 46)—note the use of the word “thoughtful,” which was taken directly from the action editor’s letter.

*Our recommendations.* Finally, we made our recommendations, which were directed to scientific use. We suggested “adult-child sex” or “adult-adolescent sex” whenever the minor was willing *and* had positive reactions, and “child sexual abuse” or “adolescent sexual abuse” whenever the minor felt that he or she did not freely participate *or* experienced negative reactions to it. We then commented that, “[b]y drawing these distinctions, researchers are likely to achieve a more scientifically valid understanding of the nature, causes, and consequences of the heterogeneous collection of behaviors heretofore labeled CSA” (pp. 46–47). From an empirical perspective, the data from our college samples as well as

many other non-clinical samples (cf. Bauserman & Rind, 1997) make it clear that these recommendations do improve predictive validity. Bogan's (1992) college study provides one example: all participants had experienced CSA, but those who did not label their experience as "abuse" were significantly better adjusted on all measures than those who did. From a methodological perspective, emotionally loaded terms are primes that can lead to invalid inferences (Rind & Bauserman, 1993). From a conceptual perspective, drawing distinctions between children and adolescents, as we did, is long overdue (see discussion below on consent for elaboration). All these points imply that our recommendations were completely defensible from a scientific point of view.

*Other sexologists concur that the term "abuse" is problematic.* The attacks on our recommendations concerning terminology may have created the false impression that our suggestions were radical and unprecedented in the scientific arena. The fact is that many other researchers have expressed concern with CSA terminology because of problems of scientific validity (e.g., Fishman, 1991; Fromuth & Burkhart, 1987; Green, 1992; Kilpatrick, 1987; Long & Jackson, 1993; Money & Weinrich, 1983; Nelson, 1989; Okami, 1990; Sandfort, 1992; West, 1998). For example, D. J. West (1998), a prominent criminologist and sexologist from Cambridge University (emeritus) and co-author of two of the college studies in our review, recently commented that professional use of terms such as abuse, perpetrator, victim, and survivor has incorrectly reinforced the idea that any kind of sexual incident with a child is likely to cause great and lasting harm. He noted further that this usage has "introduced a tone of moral revulsion alien to scientific inquiry" (p. 539). Richard Green (1992), psychiatrist, lawyer, and editor of the *Archives of Sexual Behavior* for the last quarter century, after qualitatively reviewing the CSA literature reached the same basic conclusions that we did. He succinctly described the terminology problem:

Ultimately, scientists, if no one else, must be objective in their approach to this emotional issue. Judgmental terminology regarding intergenerational sexuality is more dramatic than that in the earlier psychiatric literature on homosexuality. There, patients were labelled perverts and psychopaths.

Here, the experience is always *abuse*, the children are invariably *victims*, the adults are *perpetrators*, and those who later report childhood sexual experiences are, without apology to victims of the Nazi Holocaust, *survivors*. (p. 175)

### *Consent*

We were also repeatedly attacked for using the construct of consent. "Dr. Laura's" clinical consultants asked rhetorically "what child or adolescent can make an informed decision to consent to sex" or just asserted that "consent" is a misnomer when applied to minors. "Dr. Laura" asserted that minors are never willing in sexual contacts with adults. The FRC claimed in a press release for its May 1999 press conference that our study was "based on the premise that children can actually consent to sex with an adult." Its spokeswoman, Janet Parshall, added later that "children cannot consent to sex and any study that does not accept this premise should be dismissed." The Leadership Council's Dallam et al. wrote that our "study makes an artificial distinction between forced and consensual adult-child sex," adding that our "study suggests that children have the capacity to consent to sex with adults." Congress rejected the notion of consent by enclosing *willing* in quotation marks, and denounced the notion that willingness moderates outcomes. And Fowler of the APA, in his letter to DeLay, wrote that "it is the position of the Association that children cannot consent to sexual activity with adults."

*Simple vs. informed consent.* In *Webster's 3rd New International Dictionary*, the first definition of consent is: "compliance or approval especially of what is done or proposed by another." This definition can be termed "simple consent," of which children and adolescents are both capable. For example, ethical guidelines for research with children and adolescents typically require researchers to obtain the agreement or assent of the participant. The second definition is: "capable, deliberate, and voluntary agreement to or concurrence in some act or purpose implying physical and mental power and free action." This second definition is "informed consent," which the law takes into account and which is the typical ethical and social definition. Thus, the term "consent" clearly does not always or in-

evitably imply informed consent. All our references to “consent” or “willingness” centered on the first definition. From a scientific viewpoint, the issue is whether simple consent predicts reactions or outcomes successfully. If it does, then it is scientifically valid for use in research, irrespective of moral or ethical objections.

*Empirical validation of the consent construct.* Nearly all studies in our review distinguished between willing and forced acts. These studies generally defined CSA either as a child or adolescent’s sexual experience that was unwanted regardless of partner’s age, or as wanted or unwanted experiences with someone older—typically, at least 5 years older. We statistically compared these unwanted only and mixed consent groups to evaluate the predictive utility of this variable. In the case of males, clear predictive utility emerged: the association between CSA and adjustment problems was greater in the unwanted than mixed group, and only the unwanted group showed poorer adjustment than controls. These results demonstrated empirically the scientific validity of using the consent construct.

The procedure we used, however, was not completely satisfactory, because the second category (which included both wanted and unwanted experiences) overlapped with the first (which only included unwanted experiences). Unfortunately, researchers in the college studies rarely analyzed outcome data as a function of consent, although a few did. For example, Bergdahl (1983) used 5 categories from willing to sadistic to classify her female students’ CSA; she found a strong association between level of consent and self-reports of reactions and effects,  $r = .49$ . Bogan (1992) found in her SA female sample that lack of coercion was strongly associated with not labeling oneself as abused, which was in turn associated with better adjustment. These two college results, based on separating CSA by consent, provide additional empirical support for using the consent construct. Even more impressive evidence was provided in two studies based on non-clinical samples, in which CSA-adjustment associations were directly assessed as a function of consent.

In the first of these studies, Sandfort (1992) specifically examined CSA correlates as a function of consent. He used a Dutch

sample of 283 young male and female adults aged 18 to 23 consisting of both students and working people. Each subject was interviewed face-to-face and responded to standard measures of adjustment. CSA was restricted to sex involving physical contact occurring before 16 with someone over 16 and at least 5 years older. Most males viewed their CSA as consenting (71%), while only about a quarter of females did. Consenting subjects were as well adjusted as controls in both psychological and sexual areas. In the second, Coxell, King, Mezey, and Gordon (1999) examined a nonclinical sample of 2,474 men aged 18 to 94 in Great Britain recruited from general medical practices. Subjects were asked about unwanted and wanted CSA occurring before 16 with someone at least 5 years older; 5.3 percent had non-consenting CSA while 7.7 percent had consensual CSA. Subjects were asked whether they ever had a psychological problem of at least two weeks duration sometime in their life. Non-consenters had more problems than controls ( $r = .10$ ) or consenters ( $r = .15$ ). Controls and consenters did not differ ( $r = .02$ ). These results, especially the Coxell et al. study where both levels of consent were compared to controls, provide strong empirical support for the utility of the consent construct in accounting for adjustment variance. Noteworthy is that Coxell and his co-authors, who are “mainstream” anti-abuse researchers, repeatedly distinguished between consensual and non-consensual sex in their article, published in the prestigious *British Medical Journal*. Based on this, the FRC would have this study summarily dismissed as invalid (see Parshall’s quote given above).

*Children vs. adolescents.* It should also be made clear that when Congress, the Leadership Council, the FRC, or even the APA refer to “children” in the context of sexual relations with adults, they are not referring simply to biological children but instead to minors under the age of consent, which is generally from 16 to 18 in the U.S. Thus, they are talking not only about prepubescent children, but also adolescents. It is thus informative to review what the APA has had to say in the past about adolescents’ ability to provide informed consent in a different context. In an October 1989 amicus curiae brief to the U.S. Supreme Court, the APA argued, based on a review of the developmental literature, that pregnant girls do not



need parental consent to obtain abortions, because they are capable, in an informed consent sense, to decide for themselves. They wrote:

Psychological theory and research about cognitive, social and moral development strongly supports the conclusion that most adolescents are competent to make informed decisions about important life situations. . . . In fact, by middle adolescence (age 14–15) young people develop abilities similar to adults in reasoning about moral dilemmas, understanding social rules and laws, and reasoning about interpersonal relationships and interpersonal problems. . . . By middle adolescence most young people develop an adult-like identity and understanding of self. . . . Thus, by age 14 most adolescents have developed adult-like intellectual and social capacities including specific abilities outlined in the law as necessary for understanding treatment alternatives, considering risks and benefits, and giving legally competent consent. . . . [Additionally,] there are some 11-to-13-year-olds who possess adult-like capabilities in these areas.

In view of these conclusions, which are based on the developmental literature, it seems inconsistent for the APA to now reject even simple consent as a moderating variable in a rigorously peer-reviewed article, given that many of the CSA episodes analyzed involved adolescents.

Critics might argue that medical consent is different from sexual consent, such that adolescents below 16 or 18 cannot possibly give informed consent. But Western societies have no consensus in this regard, thus this argument reflects culture-bound American thinking. Graupner (1997; 1999; in press) documented age of consent laws throughout the current 57 European jurisdictions, finding that the median age of consent is 14, with a range from 12 to 17. Most jurisdictions (73%) have an age of consent somewhere between 12 and 15. Moreover, he noted that national governments and parliaments in these jurisdictions have frequently appointed expert commissions to scrutinize age of consent laws and make recommendations. Most of these commissions recommended 14, but some went as low as 12. The Canadian age of consent is also 14. These findings call into question claims that no distinctions should be made between adolescents and children, and that even 14 or 15 year olds can never consent and are invariably abused and harmed in sexual involvements with older persons.

*Summary.* The complaint that our study was flawed because it assumed consent was seriously misguided from the start. We never stated or implied that children can give informed consent. Our use of the consent construct was derived from the primary studies and pertained to simple, not informed, consent. Its use had predictive utility and thus was empirically valid, which is the critical issue from a scientific point of view, because the purpose of science is to uncover what is, not to endorse what ought to be. Other empirical research has even more forcefully demonstrated the predictive utility of consent as a moderator. Even though informed consent was irrelevant to our use of the consent construct, it is important to note that assertions that adolescents cannot provide informed consent are not well supported scientifically or cross-culturally.

### **Response to the Congressional Resolution**

Claiming among other things that “all credible studies in this area...condemn child sexual abuse as criminal and harmful to children,” and citing a 1982 Supreme Court opinion that expressed this view, *H. Con. Res. 107* proclaimed our study to be “severely flawed.” It condemned and denounced “all suggestions in the article...that indicate that sexual relationships between adults and ‘willing’ children are less harmful than believed and might be positive for ‘willing’ children.” It urged President Clinton likewise to reject and condemn any suggestion that sexual relations between children and adults—regardless of the child’s frame of mind—are anything but abusive and destructive. The resolution concluded by encouraging “competent investigations to continue to research the effects of child sexual abuse using the best methodology, so that the public, and public policymakers, may act upon accurate information.”

This resolution was built on the criticisms made by NARTH, “Dr. Laura,” the FRC, and the Leadership Council. Its specific comments and wording were heavily drawn from the Alaska resolution, which was itself based on some of these criticisms. Given that we have just shown these criticisms to be without merit, the characterization of our study as “severely flawed” must be seen as in-

valid. It is not the case that all credible studies have concluded harm, as our review amply demonstrated, unless concluding harm is viewed as a prerequisite to judging a study “credible.” The 1982 Supreme Court opinion was made when CSA research was relatively scant and unsophisticated; in our review, almost all the studies (57 of 59) were published after 1982. Willingness is in fact relevant to outcome. We neither stated nor implied that CSA might be positive for willing children, as “positive” in this context connotes beneficial. Instead, we accurately summarized what the students themselves reported in terms of perceived reactions and effects, some of which were positive. Finally, using the best methodology involves being strongly grounded in methodological logic (specifically regarding issues of generalizability and causality) and statistical precision, qualities that characterized our review above most others.

### **Threat to Science**

We believe that we have shown our critics to be consistently in error in their attacks on our research. However, even if many or most of the criticisms had been true, the process and practice of science would still have been ill-served by the events that unfolded. The collaboration between organizations that claim to be devoted to science, such as NARTH and the Leadership Council, and social conservatives with an overarching moral agenda, is a recipe for trouble. Add to this congressional condemnation, and we have a clear assault on the integrity of science and academic freedom. This assault is aggravated by the fact that most if not all of the criticisms were specious. The threat to science posed by the forces behind the attacks needs careful scrutiny.

*NARTH.* On its web page, NARTH claims to be “devoted to scientific inquiry.” NARTH notes that it is “dedicated to the...prevention of homosexuality” and that it welcomes the support of religious organizations that “turn to us for scientific evidence which may support their traditional doctrines.” NARTH claims that “many researchers have been intimidated into trading the truth for silence”—the “truth” being that homosexual attractions are a

“sign that deep emotional wounding has occurred” and arise from conscious and unconscious conflicts that include problems in early family relationships, sexual seduction, and self-labeling. Moreover, NARTH says, homosexuality “distorts the natural bond of friendship” between members of the same sex and “works against society’s essential male/female design and the all-important family unit.” Consistent with this, it claims that “human beings are governed by certain immutable laws of nature.”

NARTH president and co-founder, psychoanalyst Charles Socarides, once wrote in the *Journal of the American Medical Association* that “homosexuality is a dread dysfunction, malignant in character, which has risen to epidemic proportions” (Socarides, 1970). More recently, in 1995, he asserted that the removal of homosexuality from DSM was a “Trojan Horse which, once admitted into the gates of the heterosexual world, has led to a sexual and social dementia,” and that homosexuality is “a freedom that cannot be given.” NARTH’s executive director and co-founder, psychologist Joseph Nicolosi, has written that gay sex is neurotically driven with an addictive dimension that is not about sex itself, but instead about stabilizing the homosexual’s “fragmented personality structure.” Writing for the FRC, Nicolosi proclaimed that “the gay adolescent” is a myth and that no one can decide his or her sexual orientation until they are in their twenties.

In support of its views on mental disturbance, NARTH cites decades-old psychoanalytic research (e.g., Bieber, 1962) and personal experience: “As clinicians, we have witnessed the intense suffering caused by homosexuality.” In making these claims, NARTH ignores the vast array of more sophisticated non-clinical research done since the 1960s that concludes that homosexuality is not associated with psychopathology and does not have roots in family and social environment disturbance (e.g., Bell et al., 1981). Its assertions about the essential male/female design, the all-important family unit, and immutable laws of nature are drawn from theology, not nature herself. Since the 1960s, a growing body of historical, cross-cultural, and cross-species research (e.g., Greenberg, 1988; Vasey, 1995) has confirmed what Ford and Beach (1951) concluded a half century ago: “The cross-cultural and cross-spe-

cies comparisons...combine to suggest that a biological tendency for [homosexual] behavior is inherent in most if not all mammals including the human species" (p. 143). Claims that "the gay adolescent" is a myth ignore the extensive social science research showing otherwise and have the potential to harm, not help, gay youth (cf. Savin-Williams, 1998).

NARTH's approach to homosexuality explains its approach to critiquing our study: start with a presumption of pathology based upon moral conviction, look only at clinical data obtained by clinicians already holding this presumption, interpret these data as firm evidence for causation and view them as generalizable to all cases, irrespective of population or context. But this is *not* science: it is circular to start with an assumption and allow it to influence all inferences used as evidence for the assumption; it is a fallacy to draw causal inferences from clinical case studies; it is faulty generalization to assume that the general population can be accurately described by clinical samples; and it is bias to selectively ignore research based on more representative samples that contradict basic assumptions.

*The Leadership Council.* In its press release attacking our study (May 24, 1999), the Leadership Council claimed that its "membership includes many of the nation's most prominent mental health leaders," with a mission "to insure the public receives accurate information about mental health issues." It noted the ethical duty of researchers to report research "in accordance with recognized standards of scientific competence and to minimize the possibility that results will be misleading." In his correspondence with the APA (June 3, 1999), Fink noted that his organization was concerned about a "misuse of the media...and the use of sound bites [sic] rather than real science." He noted that they "are committed to a high level of scientific discourse."

Given this self-characterization, one might expect that the Leadership Council's criticisms would be sound, competent, and not misleading; the organization would avoid providing sound bites of its own to the media; and it would not assist a governmental body in undermining a rigorously peer-reviewed article in a premiere scientific journal. Such goals could be achieved by submitting a

critique to a journal for peer-review and publication, allowing us to respond so that interested researchers could draw their own conclusions. Instead, their criticisms were flawed and misleading; they provided sound bites to media critics eager to attack our study; and they offered *non-peer reviewed* criticisms to Congress as if they were established fact. Science was ill-served.

The Leadership Council presents itself as a group of prominent mental health leaders, implying that its views are authoritative and representative of the field. However, its membership largely represents an extremely controversial segment of the mental health field, whose views are *not* widely accepted as valid, particularly among academic researchers. Its membership consists of predominantly psychoanalytically- or psychodynamically-oriented practitioners or advocates, who appear to accept repressed memories as undisputed fact and believe in recovered memory therapy as well as the validity of MPD. Fink's correspondence with the APA (June 3, 1999) regarding the background and scientific basis of the Leadership Council's work makes the group's goals clear. Fink complained about the "destructive theories that justify a trivialization of sexual abuse" and an "effort to reduce and destroy psychotherapy by undermining some of the basic principles" by which they conduct their work. He noted that his group was committed to a reversal of trends in which the courts and media were defending those "who would impugn all of psychiatry" because of personal distress. Additionally, he noted that his group wanted to "protect GOOD psychotherapists from attack and from financial ruin as a result of suits that are costly both financially and emotionally" (capitalization in the original).

Recovered memory therapists, enjoying widespread acceptance in the late 1980s and early 1990s, have more recently been on the defensive. Early on, recovered memories formed the basis for thousands of suits against alleged perpetrators (e.g., fathers) of CSA and often SRA as well. But then former patients began retracting these memories, holding their therapists responsible for implantation of false memories and subsequent injury to them and their families, and then suing their therapists for damages. As Fink indicated in his correspondence with the APA, these suits substantially con-

cern his group. Some of his group's members have been subjected to such suits. Ubiquitous and severe harm is a premise that underlies belief in widespread repression of memories of CSA and in the occurrence of MPD. A challenge to this premise adds to the growing trend questioning the validity of these therapeutic beliefs. As such, the Leadership Council might be expected to challenge a study such as ours. The problem from our view is not the challenge itself, but the manner in which it was conducted. Peer-review of criticisms in scientific forums reduces bias or misrepresentation by not accepting such criticisms for publication, or by publishing them but permitting those challenged to respond. Circumventing scientific protocol by providing one-sided and severely flawed critiques to conservative social commentators and political entities, however, is a challenge that is destructive to science.

The following example shows clearly the Leadership Council's commitment to advocacy rather than scientific accuracy. As discussed previously, day care sexual and satanic abuse cases have now been all but discredited. In the original case (the McMartin preschool in California) in 1984, hundreds of children were interviewed by social workers, resulting in accusations of being sexually assaulted in hot-air balloons, faraway farms, the shoulders of busy highways, cemeteries, and tunnels under the schoolyard; witnessing animals being tortured and ritually sacrificed and staff members dressing up as witches and flying in the air; and being forced to drink the blood of a baby who had just been murdered. These accusations were almost certainly all false, the result of coercive interrogations premised on faulty assumptions about CSA effects on mental health and memory (Ceci & Bruck, 1993; Garven et al., 1998; Nathan, 1990; Nathan & Snedeker, 1995). Later that year in Malden, Massachusetts a similar case occurred at the Fells Acres day care center. One prosecutor flew to California to consult with the McMartin trial team. The Fells Acres children's allegations subsequently took on many of the same ritualistic and bizarre aspects as the McMartin case. Under interrogation by a pediatric nurse committed to belief in SRA, the children complained of such things as being sadistically molested in "secret" or "magic" rooms by clowns and lobsters, assaulted with a

butcher knife, stomped on by elephants, threatened by robots if they disclosed the abuse, and exposed to a dog being killed and buried. The Amiraults—a mother and her adult son and daughter who ran the center—were convicted and sentenced to lengthy prison terms (Harris, 1995; Pendergrast, 1996).

In this case, the mother and daughter spent a decade in prison until a judge ordered a new trial for them based on a technicality in 1995. The mother died of cancer soon afterwards. In 1998, another judge dismissed all charges against the daughter, ruling that the children in the case had been so manipulated in interviews that their testimony was forever tainted. To illustrate, the judge read a transcript showing how investigators insistently refused to take “no” for an answer to questions about abuse (*New York Times*, June 13, 1998, A6). The judge was influenced by the new thinking on children’s suggestibility, which resulted from the empirical research of Ceci and others, which had already played a key role in causing similar cases around the country to fall apart in the 1990s. In 1999, the Supreme Judicial Court of Massachusetts ruled on whether to reinstate the conviction. The Leadership Council submitted an amicus curiae brief, that stated:

It is the position of the undersigned, as an organization comprised of well-respected, oft-published, eminent professionals in the field of psychiatry and developmental science with regard to trauma including child sexual abuse, that no new evidence has emerged in the area of suggestibility with respect to preschool-aged victims of child sexual abuse since the trial in this matter.

Thus, incredibly, the Leadership Council maintained that between 1987 and 1999 *no* relevant research on CSA and suggestibility had been conducted. This remarkable position reflects strong advocacy rather than dispassionate science, and perhaps gives some context to the Leadership Council’s seriously flawed critique of our review.

The Supreme Judicial Court reinstated the conviction, ruling that the law was properly carried out in the 1987 trial. A vice president of the Leadership Council reacted by commenting to the media, “Stick a fork in them. They’re done” (*Boston Herald*, August 19, 1999, p. 5). The Amiraults are most likely completely innocent,



given the striking similarity between their case and the discredited ones all around the country (Rosenthal, 1999). As such, it does not appear that "justice" (part of the Leadership Council's full title) was well served.

*"Dr. Laura"*

In an opinion piece defending our review, psychologist Romy Cawood wrote that she scoured our article for misstatements, overstatements, and anything that could be construed as dramatic enough to make "Dr. Laura's" web page. "I came up empty," she wrote, noting by contrast that "Dr. Laura's" web page "offered all sorts of misstatements and drama." She concluded with a suggestion: "Investigate how Dr. Laura, in good conscience, can present herself as 'Dr.' in an advice-giving profession when her Ph.D. is in a different field. Now that is scandal material" (*Greensboro News & Record*, July 4, 1999, H7). This prompted an angry letter from Bellows, the general manager of the "Dr. Laura" program. Bellows defended "Dr. Laura" by writing that, although she is well qualified to analyze research findings, she did not rely on her own analysis but instead "solicited opinions from a noted clinical psychologist in Holland and has consistently referenced an analysis refuting the study by the Leadership Council" (*Dayton Daily News*, July 1, 1999, 11A).

But the "noted" clinical psychologist from Holland, van den Aardweg, is a NARTH member. His opinions, provided in a transcript on "Dr. Laura's" web page, included no substantive criticism of our methods except a claim that our measures were bad, embedded in personal attacks that, among other things, labeled us ideologues who cannot be stopped by reason or reasonable arguments and that compared us to the Nazi doctors of the 1930s and their belief in racial superiority. Despite NARTH's claims of devotion to scientific inquiry, it relies on outdated psychoanalytic theorizing to label homosexuality pathological. As for the Leadership Council, the reality is that their analysis was a flawed assortment of false assertions, inaccurate speculations, and selective criticisms. Thus, the basis for "Dr. Laura's" criticisms was invalid. "Dr. Laura" her-

self denounced our study as putting a bunch of meaningless findings together and stirring them up with mathematics, adding that she had never heard of a real scientist who uses such procedures. Meta-analysis, however, is an important tool for synthesis, widely used not just in psychology but in other disciplines such as medicine. Whether “Dr. Laura” could have provided an informed critique is a moot point, because she did not. Instead, she provided her listeners with distortions combined with moral denunciations. It likely was this sort of attack that AAAS was referring to in its rebuke of our most vocal critics. Science is ill-served by those who use their credentials to influence people on scientific issues when they themselves are inadequately informed.

### *The Family Research Council*

The FRC’s mission is to promote traditional Christian morality, not scientific knowledge. As has been the case throughout history, if science appears to challenge moral dogma, then the science itself must be challenged. The FRC’s reaction to research on homosexuality is a prime example. The FRC sees homosexuality as “unhealthy, immoral, and destructive to individuals, families, and society.” To combat homosexuality, the FRC publishes “critiques” of important research that disputes the premise that homosexuality is anything but pathological. It is instructive to review one example.

Thomas Landass, writing for the FRC, authored an article entitled, “The Evelyn Hooker Study and the Normalization of Homosexuality” (<http://www.frc.org>). Hooker’s (1957) study was a breakthrough at the time. She argued that prevailing beliefs that homosexuals were necessarily maladjusted were suspect, because all the research was based on clinical or prison samples. Instead, she examined a non-clinical sample of homosexuals, comparing their adjustment with a matched heterosexual sample, finding no differences. Landass ignored the logic behind this study, instead searching for weaknesses to discredit it. What he produced was a series of trivialities, from which he argued that the entire study was invalid. For example, he complained that Hooker wrote in the text of her article that her subjects’ IQs ranged from 90 to 135, but a

table showed the lowest IQ to be 91. He noted that the average education for the homosexual subjects was 14.0, not the 13.9 that Hooker reported. He complained that the highest educational level of her subjects was one year beyond a master's program, not the master's level itself as Hooker reported in the text of her article. Although minor in degree, he argued, these discrepancies suggest that we cannot trust Hooker as a researcher. Elsewhere, he repeatedly called her a "rat-runner," implying that we could not trust her measurements of humans. He criticized her administration of tests of adjustment because she was not blind to her subjects' orientation and impugned her motives by claiming she had a vested interest in showing that homosexuals are not necessarily pathological. Noteworthy is that these criticisms apply even more strongly to clinical research, whose findings Landass accepted without challenge. Simply put, Landass' critique is not scientific; the criticisms were selectively applied and the strategy of attack was fishing for minor points to create an illusion of invalidity. Such a critique would fail to be published in a mainstream peer-reviewed journal for these reasons. It clearly reflects an effort to promote moral condemnation of homosexuality, not advance scientific research. We find much in common between the attack on Hooker's study and the attacks on ours from the FRC, the Leadership Council, "Dr. Laura," and so on. It is a threat to the integrity of sexual science when an organization focused on a moral agenda, such as the FRC, is granted a powerful voice in debates on sexual science methodology.

*Congress.* Political attacks on scientific sex research are not new, as Hunt (1998) recently described. He noted that Kinsey and his colleagues wrote in the introduction to their male report that they were "repeatedly warned of the dangers involved in the undertaking and were threatened with specific trouble." The 1953 female report offended and outraged even more people, producing claims that the Kinsey reports were contributing to depravity of a whole generation and the spread of juvenile delinquency. Conservative Congressman B. Carroll Reece from Tennessee headed a committee investigating the Kinsey reports and their financial supporter, the Rockefeller Foundation. The president of this foundation, Dean Rusk, made a strong defense of the Kinsey funding. Soon after,

however, the Foundation's support stopped, and it gave a large grant to the Union Theological Seminary. As a consequence, Kinsey's team was left without funding, and much of its data went unpublished for decades. There are important parallels in our own case. Conservative congressmen attacked the paper and the APA as well. Dr. Fowler of the APA initially defended our review as a good study. Shortly afterward, however, the APA made historic concessions such as agreeing to ask for an outside review, to consider the social policy implications of future articles, and to write an *amicus curiae* brief stating that our study was invalid for legal use.

But our case is also different from Kinsey's, because our research was condemned in a congressional resolution, bringing matters to a new level. Critics of Congress and APA's concessions have analogized this situation to that of Galileo. At the APA's annual convention in August 1999 in Boston, an anonymous source, unknown to us, posted stickers around the convention that satirically read, "If the US Congress says that 'The sun revolves around the earth,' then that fact will be given most careful consideration in all articles published by the APA." Berry and Berry (2000), writing for the *Skeptical Inquirer*, recounted the long history of church/state attacks on scientific inquiry, leading up to the trial of Galileo. More recently, they noted, the former Soviet Union created governmental boards that reviewed all research for its acceptability to the state, and the behavioral sciences floundered in the USSR. To avoid punishment, Galileo recanted his position. Berry and Berry noted that, similarly, Fowler "recanted" in his letter to the House of Representatives during the "congressional inquisition," which brought the APA praise in *H. Con. Res. 107*. They noted the historical pattern: religions or states exert pressure upon researchers to self-censor findings disliked by those in power—most recently by threatening to withdraw financial support for research deemed offensive. *H. Con. Res. 107*, Berry and Berry commented, represented a joint effort on behalf of the state and church, which is inimical to the health of science. They argued that moral conservatives were now likely to target behavioral science research on gender, sexuality, and parenting roles. Writing in the *National Journal*

(August 7, 1999), Rauch also noted that conservatives would be emboldened by how easy it was to smear research on a sensitive subject. As a result, he asked, "would you be surprised if this happened again? And again?"

*Synthesis.* In the foregoing discussion, we considered each of the major forces central to the condemnation of our review. Generally speaking, these forces consisted of two varieties of psychoanalysts (anti-homosexual and repressed memory) and various types of social conservatives. The psychoanalysts preached science in their attacks on our article but practiced defense of morality and certain psychoanalytic assumptions. These assumptions, which are integral to their therapy, have recently been under intense attack. It seems that our article may have added to their problem. The social conservatives unambiguously attacked our article in defense of morality. To them, it seems, science is true when its conclusions are "acceptable," but bogus otherwise, which helps to explain their eagerness to uncritically accept and promulgate the flawed critiques of the psychoanalysts. The collaboration between these psychoanalysts and social conservatives is not surprising or novel. Indeed, it has historical roots (cf. Szasz, 1987, 1990).

Psychiatrist Thomas Szasz has been a fierce critic of the alliance between psychiatry and religion, citing its negative influence on science and therapy. Although Szasz has taken what many regard as extreme positions, and we ourselves do not intend to impugn all therapists, his comments regarding the synergy between certain psychoanalysts and certain religious conservatives in the area of sexual behavior seem quite apt to our review and the collaborative attacks on it. Szasz (1990) noted that for millennia clerics were the recognized experts on sexual sin and were accorded special privileges and powers to punish the sinners. But, since the birth of modern medicine, it has been the clinicians who have filled this role. The practice is the same, but with different terms (e.g., "disease" instead of "sin"). Szasz traced the religious-to-medical transformation from female sexual desire as mental disease to masturbatory insanity to homosexuality qua disease. Acerbically, he characterized the early but influential psychiatric opinion of the late nineteenth century on sexual disease as pseudoscientific fraud, full of

falsehoods pretentiously presented as if they were the fruits of hard-won scientific research. Rather than truly helping individuals or advancing science, he argued, the new thinking was instead intent on supplanting the waning power of the church with the waxing power of medicine. Likewise, many modern psychoanalysts, he continued, think of themselves as experts on sex and daringly progressive in their positions, but instead have continued the disease model for all unconventional sex, which they believe themselves to be especially adept at diagnosing and treating.

Szasz (1990) opined that many psychoanalysts tend to treat the disease concept of sexual “perversions” defensively, as if it were a valuable possession of which they feared being robbed. Szasz’s comments apply strikingly well to NARTH and the Leadership Council, who view themselves as progressive experts, but who seem to fear being robbed of their precepts. Both APAs (psychological and psychiatric) have expressed the view that it is bad therapy and harmful to patients to attempt to change homosexuals to heterosexuals. Such therapy is precisely what NARTH clinicians practice. Recovered memory therapy has not only been subjected to intense scientific challenge recently but costly lawsuits as well. Such therapy is what the Leadership Council appears to advocate. NARTH’s position rests on the intrinsic pathology of homosexuality (which, they claim, can result from childhood seduction); the Leadership Council’s position rests on the pervasive, traumatic nature of CSA. Both groups, it would appear, see our review as increasing the threat to “rob” them of their foundations. Further, the foundational premises of these groups have deep ties with traditional morality, which abhors sexual deviance. Before modern medicine, traditional moralists needed only to label sex as sinful to suppress it. But they now depend on medical pronouncements of pathology. The collaboration among these groups is thus unsurprising. Because their attacks revolved around what “must be” rather than what is, they were defensive and biased rather than objective and dispassionate in their attacks. Because of the vehemence of these attacks, conducted through the media and in conjunction with politicians, science was harmed in this area and is at risk in the future.

## Concluding Remarks

It is important that the central message in our review not be lost because of the distraction created by our critics. Various commentators have summarized this message in their replies to these critics. Psychologist Carol Tavis asked: "Shouldn't this be good news? Shouldn't we be glad to know which experiences are in fact traumatic for children, and which are not upsetting to them? Shouldn't we be pleased to get more evidence of the heartening resilience of children?" Her answers to all these questions were strongly in the affirmative (*Los Angeles Times*, July 19, 1999). Psychologist Sharon Lamb, offering a different feminist view from that which supports repressed memories, called our findings "great news for victims." She continued: "If it is true that many victims do not suffer lifelong consequences from abuse, and that many victims are not traumatized, permanently damaged, or wounded for life, we should be happy." Lamb argued that thirty years ago, when society was in denial about incest, advocacy was needed, but a different strategy is needed today, one that acknowledges that abuse comes in all shapes and sizes and that it is flawed to characterize all instances by the most extreme ones (*Boston Globe*, August 1, 1999, E1). Commentator Andrew Sullivan, in his article entitled "The Assault on Good News," noted that, although our findings should have been seen as a reason for relief, they were distorted by social conservatives, who seem to make a habit of undermining various types of "good news" that do not fit their moral agenda (*The New York Times Magazine*, November 7, 1999, 38-40).

Commentators also pointed out that stubborn refusal to acknowledge the findings in our review can be harmful. Lamb argued that the way to support victims is not to call each one a "survivor," which does not serve them but instead serves our culture's requirement to see all victims as damaged for life. She was critical of the APA for distancing itself from our review, which she argued does a disservice to victims who have "a right to get over it" (*Boston Globe*, August 1, 1999, E1). Writer Jonathan Rauch commented that the actions of the "Kulturkampf conservatives" serve to stifle our knowledge about sexual abuse, which will hurt children (*Na-*

*tional Journal*, August 7, 1999). These sentiments are consistent with problems of secondary harm, which we discussed earlier in the current article (cf. Baurmann, 1983; Constantine, 1983).

*Need to separate morality and science.* In closing, we want to elaborate on the continuing struggle that scientific research on socially sensitive issues, especially sexual ones, faces. In a television documentary entitled, "The power of belief" (aired on ABC in October 1998), John Stossel interviewed Emory University psychologist Claire Coles, who was one of the first to argue that the conventional wisdom that "crack babies" are permanently damaged is false. She noted that the conventional wisdom was based on flawed research, in which cocaine use was confounded with alcohol use and poverty, but everyone ignored the confounding and attributed all negative correlates to the cocaine. Both conservatives and liberals had a stake in cocaine's pathogenicity—conservatives wanted to demonize cocaine users and liberals wanted more money for their programs. When Coles argued her position, Stossel noted, she was "attacked viciously by politicians, called incompetent, accused of making data up, or believing in drug abuse." Stossel commented to Coles, "People confuse morality and science." Coles answered, "Well, they did. Cocaine is bad, therefore the effects must be bad."

Coles' plight has obvious parallels to our own. Wrongfulness was confused with harmfulness by those espousing the conventional wisdom, despite our own warning about this error in our review, and questioning this association elicited intense reaction. Sexual science has often been concerned with behaviors that some or many see as wrongful and thus has been especially vulnerable to attacks from those who stubbornly refuse to separate morality and science. This problem was so acute a half century ago that the authors of two of the most important works in the history of sexual science explicitly eschewed all moral judgments in the analyses of their data (Ford & Beach, 1951; Kinsey et al., 1948). Ford and Beach firmly asserted in the introduction to their seminal review of the cross-species and cross-cultural data on sexual behavior that "[m]oral evaluations form no part of this book" (p. 14). In distancing themselves from the moral perspective, Kinsey et al. commented



that “[t]oo often the study of [sexual] behavior has been little more than a rationalization of the mores masquerading under the guise of objective science” (p. 203). These two sets of researchers viewed avoidance of moral judgments and interpretations as essential to striving for scientific objectivity.

Today, the conflation of morality and science in the study of sex is most intense in behaviors that are currently classified as CSA. We followed the model advanced by Kinsey et al. and Ford and Beach in our review and were roundly condemned by moralists for doing so. But it is the scientifically correct path to pursue, which future researchers examining this important issue, or other sexual issues entwined in taboo, will hopefully follow.

## References

- Ageton, S. S. (1988). Vulnerability to sexual assault. In A. W. Burgess (Ed.), *Rape and sexual assault II* (pp. 221-243). New York: Garland.
- Baker, A. W. & Duncan, S. P. (1985). Child sexual abuse: A study of prevalence in Great Britain. *Child Abuse & Neglect*, 9, 457-467.
- Bass, E. & Davis, L. (1988). *The Courage to Heal: A guide for women survivors of child sexual abuse*. New York: HarperPerennial.
- Baurmann, M. C. (1983). *Sexualität, Gewalt und psychische Folgen* [Sexuality, force, and psychological consequences]. Wiesbaden, Germany: Bundeskriminalamt.
- Bauserman, R. & Rind, B. (1997). Psychological correlates of male child and adolescent sexual experiences with adults: A review of the nonclinical literature. *Archives of Sexual Behavior*, 26, 105-141.
- Bell, A. P., Weinberg, M. S., & Hammersmith, S. K. (1981). *Sexual preference: Its development in men and women*. Bloomington, IN: Indiana University Press.
- Bergdahl, J. (1983). Differences in demographic features and present anxiety level among college women who report type, extent and rated severity of childhood and adolescent sexual contacts with adults (Doctoral dissertation, Ohio University, 1982). *Dissertation Abstracts International*, 44, 296.
- Berry, K. K. & Berry, J. (2000). Commentary: The congressional censure of a research paper: Return of the Inquisition? *Skeptical Inquirer*, Jan/Feb, p. 20.
- Best, J. (1997). Victimization and the victim industry. *Society*, May/June, 9-17.
- Bieber, I. (1962). *Homosexuality: A psychoanalytic study*. New York: Basic Books.
- Bogan, Y. K. H. (1992). Child sexual abuse: The effects of mislabeling sexual experiences on long-term adjustment. Doctoral dissertation, University of Georgia.

- Brandon, S., Boakes, J., Glaser, D., & Green, R. (1998). Recovered memories of childhood sexual abuse. *British Journal of Psychiatry*, 172, 296-307.
- Briere, J. (1992). Methodological issues in the study of sexual abuse effects. *Journal of Consulting and Clinical Psychology*, 60, 196-203.
- Browne, A. & Finkelhor, D. (1986). Initial and long-term effects: A review of the research. In D. Finkelhor (Ed.), *A Sourcebook on Child Sexual Abuse* (pp. 143-179). London: Sage.
- Burgess, A. W., Hartman, C. R., McCausland, M. P., & Powers, P. (1984). Response patterns in children and adolescents exploited through sex rings and pornography. *American Journal of Psychiatry*, 141, 656-662.
- Ceci, S. J. & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, 113, 403-439.
- Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cole, S. (2000). Unpopular psychology. *Lingua Franca*, February, 12-14.
- Condy, S., Templer, D., Brown, R., & Veaco, L. (1987). Parameters of sexual contact of boys with women. *Archives of Sexual Behavior*, 16, 379-394.
- Constantine, L. (1983). Child sexuality: Recent developments and implications for treatment, prevention, and social policy. *Medicine and Law*, 2, 55-67.
- Coxell, A., King, M., Mezey, G., & Gordon, D. (1999). Lifetime prevalence, characteristics, and associated problems of non-consensual sex in men: Cross sectional survey. *BMJ*, 318, 846-850.
- Crews, F. (1998). *Unauthorized Freud: Doubters confront a legend*. New York: Penguin Books.
- Dallam, S., Gleaves, D., Spiegel, D., & Kraemer, H. (1999). An analysis of Rind et al.'s meta-analysis of the long-term effects of child sexual abuse. Preliminary analysis in preparation for full review article rebutting the Rind et al. study.
- Esman, A. H. (1994). "Sexual abuse," pathogenesis, and enlightened skepticism. *American Journal of Psychiatry*, 151, 1101-1103.
- Fink, P., Silberg, J., & Dallam, S. (1999). Scientific organization investigates controversial sex abuse study. Unpublished letter sent to the *Los Angeles Times* in response to Carol Tavris's July 19, 1999 op-ed piece, July 28, 1999.
- Finkelhor, D. (1979). *Sexually victimized children*. New York: Free Press.
- Finkelhor, D. (1984). *Childhood sexual abuse: New theory and research*. New York: Free Press.
- Fishman, J. (1991). Prevalence, impact, and meaning attribution of childhood sexual experiences of undergraduate males (Doctoral dissertation, University of Massachusetts, 1990). *Dissertation Abstracts International*, 52, 114.
- Ford, C. S. & Beach, F. A. (1951). *Patterns of Sexual Behavior*. New York: Harper & Row, Pub.
- Fromuth, M. & Burkhart, B. (1987). Childhood sexual victimization among college men: Definitional and methodological issues. *Violence & Victims*, 2, 241-253.

- Fromuth, M. & Burkhart, B. (1989). Long-term psychological correlates of childhood sexual abuse in two samples of college men. *Child Abuse & Neglect*, 13, 533-542.
- Frontline (1991). "Innocence lost," TV documentary produced by Ofra Bikel, broadcast on PBS May 7, 1991.
- Frontline (1993). "Innocence lost: The verdict," TV documentary produced by Ofra Bikel, broadcast on PBS July 20 and 21, 1993.
- Frontline (1995a). "Divided memories," TV documentary produced by Ofra Bikel, broadcast on PBS April 4 and 11, 1995.
- Frontline (1995b). "Search for Satan," TV documentary produced by Ofra Bikel, broadcast on PBS October 24, 1995.
- Frontline (1998). "The child terror," TV documentary produced by Michael Kirk and Rick Young, broadcast on PBS October 27, 1998.
- Gardner, R. (1993). Modern Witch Hunt—Child abuse charges. *Wall Street Journal*, Feb 22.
- Garven, S., Wood, J. M., Malpass, R. S., & Shaw, J. S. III (1998). More than suggestions: The effect of interviewing techniques from the McMartin Preschool case. *Journal of Applied Psychology*, 83, 347-359.
- Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. New York: Free Press.
- Graupner, H. (1997). Sexuelle Mündigkeit: Die Strafgesetzgebung in europäischen Ländern. *Zeitschrift für Sexualforschung*, 10, 281-378.
- Graupner, H. (1999). Love versus abuse: Crossgenerational sexual relations of minors: A gay rights issue? *Journal of Homosexuality*, 37, 23-56.
- Graupner, H. (in press). Sexual consent: The criminal law in Europe and overseas. *Archives of Sexual Behavior*.
- Green, R. (1992). *Sexual science and the law*. Cambridge, MA: Harvard University Press.
- Greenberg, D. (1988). *The Construction of Homosexuality*. Chicago: University of Chicago Press.
- Harris, J. (1995). The Fells Acres Daycare Case. Unpublished manuscript.
- Hooker, E. (1957). The adjustment of the male overt homosexual. *Journal of Projective Techniques*, 21, 18-31.
- Hunt, M. (1998). *The New Know-Nothings: The political foes of the scientific study of human nature*. New Brunswick (U.S.): Transaction Publishers.
- Jaroff, L. (Nov. 29, 1993). Lies of the Mind. *Time*, 52-59.
- Jumper, S. (1995). A meta-analysis of the relationship of child sexual abuse to adult psychological adjustment. *Child Abuse & Neglect*, 19, 715-728.
- Kelly, R. J. (1993). Overall level of distress. In J. Waterman, R. J. Kelly, J. McCord, & M. K. Oliveri (eds.), *Behind the Playground Walls: Sexual abuse in preschools*. New York: Guilford Press.
- Kendall-Tackett, K. A., Williams, L. M., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin*, 113, 164-180.
- Kihlstrom, J. (1997). Memory, abuse, and science. *American Psychologist*, 52, 994-995.

- Kilpatrick, A. C. (1987). Childhood sexual experiences: Problems and issues in studying long-range effects. *The Journal of Sex Research*, 23, 173-196.
- Kinsey, A., Pomeroy, W. B., & Martin, C. E. (1948). *Sexual Behavior in the Human Male*. Philadelphia: Saunders.
- Landis, J. (1956). Experiences of 500 children with adult sexual deviation. *Psychiatric Quarterly Supplement*, 30, 91-109.
- Laumann, E. O., Gagnon, J. H., Michael, R. T., & Michaels, S. (1994). *The Social Organization of Sexuality: Sexual practices in the United States*. Chicago: The University of Chicago Press.
- Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist*, 48, 518-537.
- Loftus, E. F. (1997). Creating childhood memories. *Applied Cognitive Psychology*, 11, S75-S86.
- Loftus, E. F. & Polage, D. C. (1999). Repressed memories: When are they real? How are they false? *Psychiatric Clinics of North America*, 22, 61-70.
- Long, P. & Jackson, J. (1993). Initial emotional response to childhood sexual abuse: Emotion profiles of victims and relationship to later adjustment. *Journal of Family Violence*, 8, 167-181.
- López, F., Carpintero, E., Hernández, M., & Fuertes, A. (1995). Prevalencia y consecuencias del abuso sexual al menor en España. [Prevalence and consequences of child sexual abuse in Spain.] *Child Abuse & Neglect*, 19, 1039-1050.
- McCarty, R. (1999). A brief comment by APA executive director for science. *Psychological Science Agenda*, 12, no. 6, Nov/Dec, 2-3.
- Money, J. & Weinrich, J. D. (1983). Juvenile, pedophile, heterophile: Hermeneutics of science, medicine and law in two outcome studies. *International Journal of Medicine and Law*, 2, 39-54.
- Myers, M. F. (1989). Men sexually assaulted as adults and sexually abused as boys. *Archives of Sexual Behavior*, 18, 203-215.
- Nash, M. R., Hulse, T. L., Sexton, M. C., Harralson, T. L., & Lambert, W. (1993). Reply to comments by Briere and Elliot. *Journal of Consulting and Clinical Psychology*, 61, 289-290.
- Nathan, D. (June 12, 1990). What McMartin started: The ritual sex abuse hoax. *The Village Voice*, 36-44.
- Nathan, D. & Snedeker, M. (1995). *Satan's silence: Ritual abuse and the making of a modern American witch hunt*. New York: Basic Books.
- Nelson, J. A. (1989). Intergenerational sexual contact: A continuum model of participants and experiences. *Journal of Sex Education & Therapy*, 15, 3-12.
- Neumann, D. A., Houskamp, B. M., Pollock, V. E., & Briere, J. (1996). The long-term sequelae of childhood sexual abuse in women: A meta-analytic review. *Child Maltreatment*, 1, 6-16.
- Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Ofshe, R. & Watters, E. (1993). Making monsters. *Society*, 30, 4-16.
- Okami, P. (1990). Sociopolitical biases in the contemporary scientific literature on adult human sexual behavior with children and adolescents. In J.

- R. Feierman (Ed.), *Pedophilia: Biosocial dimensions* (pp. 91-121). New York: Springer-Verlag.
- Okami, P. (1994). "Slippage" in research of child sexual abuse: Science as social advocacy. In J. J. Krivacska & J. Money (Eds.), *The handbook of forensic sexology: Biomedical and criminological perspectives* (pp. 559-575). Amhurst, NY: Prometheus.
- Pendergrast, M. (1996). *Victims of Memory: Sex abuse accusations and shattered lives*. Hinesburg, Vermont: Upper Access.
- Pope, H. G. & Hudson, J. I. (1995). Does childhood sexual abuse cause adult psychiatric disorders? Essentials of methodology. *Journal of Psychiatry and Law*, 23, 363-381.
- Rind, B. & Bauserman, R. (1993). Biased terminology effects and biased information processing in research on adult-nonadult sexual interactions: An empirical investigation. *Journal of Sex Research*, 30, 260-269.
- Rind, B. & Tromovitch, P. (1997). A meta-analytic review of findings from national samples on psychological correlates of child sexual abuse. *Journal of Sex Research*, 34, 237-255.
- Rind, B., Tromovitch, P., & Bauserman, R. (1998). A meta-analytic examination of assumed properties of child sexual abuse using college samples. *Psychological Bulletin*, 124, 22-53.
- Risin, L. & Koss, M. (1987). The sexual abuse of boys: Prevalence and descriptive characteristics of childhood victimizations. *Journal of Interpersonal Violence*, 2, 309-323.
- Rosenthal, B. (1999). No finality in Fells Acres. Unpublished manuscript.
- Rosenthal, R. (1984). *Meta-analytic Procedures for Social Research*. London: Sage.
- Rosenthal, R. (1994). Science and ethics in conducting, analyzing, and reporting psychological research. *Psychological Science*, 5, 127-133.
- Sandfort, T. G. (1992). The argument for adult-child sexual contact: A critical appraisal and new data. In O'Donohue, James, et al. (eds), *The sexual abuse of children (vol. 1): Theory and research* (pp. 38-48). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Savin-Williams, R. C. (1998). *And then I became gay: Young men's stories*. New York: Routledge.
- Schlessinger, L. (1999). Article on pedophilia is just "junk science." *The Times-Picayune*, April 18, E6.
- Seligman, M. E. P. (1994). *What you can change and what you can't*. New York: Knopf.
- Socarides, C. W. (1970). Homosexuality and medicine. *Journal of the American Medical Association*, 212, 1199-1202.
- Szasz, T. (1987). Justifying coercion through religion and psychiatry. *Journal of Humanistic Psychology*, 27, 158-174.
- Szasz, T. (1990). *Sex by Prescription: The startling truth about today's sex therapy*. New York: Syracuse University Press.
- Tavris, C. (1993). Beware the incest-survivor machine. *New York Times Book Review*, January 3, p.1 and pp. 16-17.

**62     Sexuality & Culture**

- Vasey, P. L. (1995). Homosexual behavior in primates: A review of evidence and theory. *International Journal of Primatology*, 16, 173-203.
- West, D. J. (1998). Boys and sexual abuse: An English opinion. *Archives of Sexual Behavior*, 27, 539-559.
- West, D. J. & Woodhouse, T. (1993). Sexual encounters between boys and adults. In C. Li, D. West, & T. Woodhouse (eds.), *Children's Sexual Encounters with Adults* (pp. 3-137). New York: Prometheus.