Avoiding False Claims of Child Sexual Abuse: Empty Promises

Kathy Pezdek


Stable URL: http://links.jstor.org/sici?sici=0197-6664%28199407%2943%3A3%3C258%3AAFCOCS%3E2.0.CO%3B2-9

*Family Relations* is currently published by National Council on Family Relations.

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/NCFR.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.
This article provides a commentary to the piece, “The Professional Response to Child Sexual Abuse: Whose Interests are Served?” by Fincham, Beach, Moore, and Diener (1994). In that article, it is argued that in their zeal to “do something” about child abuse, practitioners who work in child protective services may have adopted ineffective procedures that are unwittingly harmful to children. The major source of harm addressed is that resulting from pursuing claims of sexual abuse that are false. This commentary focuses on one point—the importance of recognizing that any attempts to reduce the probability of false claims of child abuse (Type II error) would necessarily result in increasing the probability of missing true claims of child abuse (Type I error). The hypothesis-testing framework is offered as a useful heuristic for conceptualizing the tradeoffs between these two types of errors.

The job of separating real from false instances of child sexual abuse is often a truly ambiguous task. It involves decision making in the face of uncertainty, in a highly emotional situation, with dire consequences for errors. Fincham et al. (1994) have identified numerous problems that might result from current practices in responding to child abuse allegations. It is hard to argue with their position that practitioners must be careful to avoid “merely replacing one form of denial (that sexual abuse of children does not occur) with another (that professionals always act in the best interests of children and do not harm them)” (p. 253). And, few professionals would disagree that there is room for improvement in the methods of (a) identifying child abuse when it occurs and (b) effectively dealing with the child and the perpetrator. However, the over-arching logic of the position of Fincham et al. is fallacious.

THE HYPOTHESIS-TESTING FRAMEWORK

In a perfect world, one could easily jump on the crusade proposed by Fincham et al. (1994) and join in the battle cry to reduce or eliminate false claims of child abuse. But the world is not perfect. Fincham et al. make an empty promise by failing to recognize the fact that any attempts to reduce the probability of false claims of child abuse would necessarily result in increasing the probability of missing true claims of child abuse. I suggest that the hypothesis-testing framework offers a useful heuristic for conceptualizing the tradeoffs between reducing false claims and missing true claims of child abuse. Figure 1 presents the hypothetical probability distributions for two populations: The distribution on the left represents the distribution of the amount of evidence of abuse for the population in which abuse did not occur; the distribution on the right represents the distribution of the amount of evidence of abuse for the population in which abuse did occur.

The point in presenting these two distributions in this way is that it captures several characteristics of the reality of the situation. First, the amount of evidence of abuse is a continuous variable that ranges from little evidence on the left to much evidence on the right. The evidence of abuse is not an all-or-none phenomenon; it is not useful to think of the evidence of abuse as either present or absent. Second, each population distribution spans the full range along the horizontal axis from the minimum to the maximum amounts of evidence of abuse. This reflects the ambiguity in the situation—even if there is minimal evidence of child abuse, it may be that the abuse really did occur, and even if there is a great deal of evidence of child abuse, it may be that the abuse did not really occur. However, because it is assumed that these are normal distributions, the probability of these extreme cases is quite small.

Third, the most important point in characterizing the problem of detecting child abuse according to the hypothesis-testing framework is that it clarifies the fact that there is an overlap between these distributions in terms of the amount of evidence for abuse that might exist. That is, one may not be able to discriminate between many abusive and nonabusive cases, as the amount of evidence for abuse could be similar for both types of cases. These more ambiguous cases exist in the overlap of the distributions in Figure 1.

Given the ambiguity in this situation, how does one make a decision for individual cases, regarding whether abuse occurred or not? Another way of framing this question is to ask, “Where does one place the decision criterion?” The criterion will be placed at some point along the horizontal axis, such that if in an individual case, the amount of evidence of abuse is greater than the criterion (to the right of the criterion in Figure 1), the case will be reported as one suspected of abuse, and if in an individual case, the amount of evidence of abuse is less than the criterion (to the left of the criterion in Figure 1), the case will be reported as one in which abuse is not suspected. The vertical line in the middle of Figure 1 represents one possible placement of the decision criterion. Generally speaking, the criterion in Figure 1 is positioned at a reasonable point; most of the cases where abuse did occur would correctly be reported as abuse, and most of the cases where abuse did not occur would correctly be reported as not abusive.

The above framework also elucidates two inevitable errors in making decisions about the occurrence of abuse. One error focused on by Fincham et al. (1994) is that a false accusation might be made. This would be a false positive error or Type I error. Referring to Figure 1, the probability of a false positive error corresponds to the shaded area to the right of the criterion, under the distribution representing the population in which abuse did not occur.

* I am grateful to Edward Teyber for his critical feedback on this article and thank Joel Smith for constructing Figure 1.

Kathy Pezdek is Professor of Psychology at the Center for Organizational and Behavioral Sciences, The Claremont Graduate School, Claremont CA 91711, email: PEZDEK@CGSVAX.CLAIREMONT.EDU.

Key Words: child abuse, false accusations, incest, sexual abuse, suggestibility.

(Family Relations, 1994, 43, 258-260.)

July 1994
Figure 1. Hypothetical Probability Distributions for Cases in Which Abuse Occurred and Cases in Which Abuse Did Not Occur.

<table>
<thead>
<tr>
<th>Probability of Abuse Did Not Occur</th>
<th>Probability of Abuse Did Occur</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abuse Not Reported</td>
<td>Abuse Reported</td>
</tr>
<tr>
<td>False Negatives</td>
<td>False Positives</td>
</tr>
<tr>
<td>Amount of Evidence of Abuse</td>
<td></td>
</tr>
</tbody>
</table>

Note: The distribution of the amount of evidence of abuse for the population in which abuse did not occur (on the left) and the population in which abuse did occur (on the right). The decision criterion (the vertical line) is positioned such that abuse is not reported in cases that fall to the left of the criterion, and abuse is reported in cases that fall to the right of the criterion. Accordingly, the probability of false positives (Type I error) and false negatives (Type II error) are indicated.

which abuse did not occur. That is, abuse did not occur, but it is erroneously decided that abuse did occur. The other type of error is a false negative or Type II error. Referring to Figure 1, the probability of a false negative error corresponds to the shaded area to the left of the criterion, under the distribution representing the population in which abuse did occur. That is, abuse did occur, but it is erroneously decided that abuse did not occur.

Now, to illustrate the thesis of this commentary, the inevitable tradeoff between false positive error and false negative error can be seen by observing what happens to the size of the area of the curve corresponding to the false positive error rate and the false negative error rate as the criterion is shifted to the right and then to the left. Referring to Figure 1, if one wants to reduce the incidence of false positives, the criterion should be raised (i.e., shifted to the right), but in so doing, the incidence of false negatives would increase. If one wants to reduce the incidence of false negatives, the criterion should be lowered (i.e., shifted to the left), but an increase in false positives would correspondingly result. The point not made in the article by Fincham et al. (1994) is that there is a price to be paid for reducing the number of false claims of child sexual abuse: it is impossible to reduce the incidence of false positives without a corresponding increase in false negatives.

**Evaluating the Costs of False Negative Errors and False Positive Errors**

So how do practitioners in the field decide where hypothetically to place the criterion in terms of how much evidence will be required prior to pursuing a claim of child sexual abuse? Ordinarily, the statistical standard is to place the criterion such that the rate of false positives (also called alpha error) is limited to .05. However, in any domain, the decision regarding where the criterion should be placed depends on weighing the costs of false positives against the costs of false negatives. Surely we must agree with Fincham et al. (1994), that it is incorrect to assume that “if there is no abuse ‘no harm will come to the family or child’” (Walker, Bonner, & Kaufman, 1988, p. 18). Some awkwardness, if not psychological harm, is surely caused by taking a child or a parent from the home for some period of time, by subjecting a child to an unnecessary physical examination, and by raising suspicions in the school and in the neighborhood. These would be the costs of a false positive error, that is, pursuing a claim of child sexual abuse when no abuse in fact occurred.

On the other hand, what would be the costs of a false negative error? If sexual abuse is occurring and nothing is done to prevent its reoccurrence, a child could be subjected to a childhood of frequent sexual abuse. Although Fincham et al. (1994) argue: “the widespread assertion that the impact of sexual abuse is serious” is “misleading” (p. 251), I think that most people would be appalled by this position. I recently heard someone argue at a public forum that what makes incest so bad for children is the unnecessary fuss made about it by society. This view does not seem a far cry from the position taken on this issue in the target piece. Fincham et al. propose that more research is necessary to determine how serious the consequences of child sexual abuse are. In particular, it is argued that studies are necessary to compare “abused children to children reported to CPS [Child Protective Services] but whose cases are unfounded,” (p. 250) to separate out the effects of the abuse from the effects of the professionals’ response to the abuse. Although from a logical positivist point of view I agree that such a study is necessary to isolate the effects of these two variables, I can only imagine Senator Proxmire’s delight in bestowing the “Golden Fleece Award” on such a research enterprise.

I have difficulty even imagining a convincing argument supporting the view that the severity of the consequences of raising false accusations of sexual abuse even approaches the severity of the consequences of leaving a child in a situation where he or she will be repeatedly abused. Fincham et al. (1994) demain the motivation of child protective services practitioners by implying that it is their zeal to simply do something that may be generating a very large false positive rate. It is more likely their astute awareness of the dire consequences of leaving a child in an abusive situation that has motivated them to set their decision criterion high to minimize false negative errors.

Fincham et al. (1994) refer to a very thorough review of the literature on the impact of sexual abuse on children by Kendall-Tackett, Williams, and Finkelhor (1993), in which it is reported that there does not appear to be a single pattern of symptoms associated with sexual abuse. Although Fincham et al. interpret this finding as minimizing the seriousness of the effects of child abuse (“just about any worrisome childhood behavior has been associated with child sexual abuse,” p. 250), this was not at all the...
of suggestively planting memories for events that did not occur.

Several other points reported by Fincham et al. (1994) in the section on children’s memory are not consistent with the conclusions reached from a thorough literature review. The authors are especially suspect, as no original source material was referenced for each. Rather than providing a detailed critique and review of the relevant research in this commentary, I simply want to point out that the following are among the conclusions reported by Fincham et al. that are inaccurate by virtue of being deceivingly global (the italics were added for emphasis):

“children may inaccurately report as remembered aspects of the event material communicated to them by others” (p. 252).

“children sometimes embellish the information that they incorporate into their memories, providing vivid and detailed accounts of events that never occurred” (p. 252).

“individual differences exist in children’s susceptibility to the influences mentioned; some children are extremely resistant to such influences” (p. 252).

“children sometimes lie when the motivational structure is tilted toward lying” (p. 252).

A review of the research literature on the suggestibility of children’s memory would lead to the qualified conclusion that, at present, there is no evidence that illusory memories for sexual abuse are a widespread phenomenon.

IMPLICATIONS OF THE HYPOTHESIS-TESTING FRAMEWORK FOR FUTURE RESEARCH

One interesting research suggestion follows from using the hypothesis-testing approach to frame the issue of separating real versus false claims of sexual abuse. There is one change that would simultaneously decrease the rate of false negative errors and false positive errors, and that is, decreasing the variance within the two distributions. Referring to Figure 1, if the means of the two distributions remained constant (there is no way to move these distributions farther apart), and the criterion remained fixed, reducing the variance within the two distributions (i.e., making both distributions taller and skinnier) would decrease both types of error. The variance of these two distributions would be decreased if we had less measurement error, that is, a more sensitive and precise technique for assessing the amount of evidence of abuse. If we had a better understanding of the critical indicators of child sexual abuse, there would be less overlap in the two distributions in Figure 1, indicating less uncertainty regarding whether sexual abuse is evident or not in any particular situation. This argues for the need for research focused on developing a battery of measures that more accurately predicts when abuse has occurred. Application of the hypothesis-testing framework to the issue of child sexual abuse thus has clear implications for conceptually framing the tradeoffs between false positive errors and false negative errors, and suggests productive avenues for future research.

REFERENCES


