Reports of childhood sexual abuse by adolescents and young adults: Stability over time

Article *in* The Journal of Sex Research · December 2002

DOI: 10.1080/00224490209552149 · Source: PubMed

CITATIONS
41

READS
307

5 authors, including:

Matthew Aalsma
Indiana University School of Medicine
241 PUBLICATIONS **2,742** CITATIONS
SEE PROFILE

Gregory Zimet
Indiana University School of Medicine
537 PUBLICATIONS **32,093** CITATIONS
SEE PROFILE

Margaret J Blythe
Indiana University – Purdue University Indianapolis
74 PUBLICATIONS **2,260** CITATIONS
SEE PROFILE

Donald P Orr
Indiana University – Purdue University Indianapolis
195 PUBLICATIONS **7,031** CITATIONS
SEE PROFILE

Reports of Childhood Sexual Abuse by Adolescents and Young Adults: Stability Over Time

Matthew C. Aalsma, Gregory D. Zimet, J. Dennis Fortenberry, Margaret Blythe, and Donald P. Orr
Indiana University School of Medicine

The consistency of childhood sexual abuse (CSA) reporting was explored in this study. Two-hundred seventeen adolescents and young adults (ages 14-24) enrolled in urban health care clinics completed self-report questionnaires assessing CSA and other problem behaviors at enrollment and at 7 months. Results indicated that the stability of CSA self-report at two time points was poor (58% consistent nonreporters of CSA, 20% consistent reporters, 22% inconsistent reporters). Consistent and inconsistent reporters were differentiated on risk measures. Adolescents who endorsed more items from the CSA scale were five times more likely to be consistent reporters. In sum, adolescent CSA reporting was quite inconsistent over time. Using multi-item scales and assessing CSA at two time points enhances accuracy of reporting.

Surveys of adolescent behavior often rely upon single-item, self-report measures. When this method is used, it is quite possible to misclassify experiences, particularly when assessing sensitive subjects such as adolescent sexual experiences (Newcomer & Udry, 1988). Given this reality, researchers have attempted to assess the reliability of certain behaviors in adolescent sexuality research, such as timing of initial sexual experiences and lifetime sexual intercourse (Alexander, Somerfield, Ensminger, Johnson, & Kim, 1993; Capaldi, 1996). The consistency of adolescent self-report of initial sexual experience timing, for example, appears poor (Alexander et al., 1993; Capaldi, 1996). In light of this fact, it seems possible that self-report of other sensitive experiences, such as childhood sexual abuse (CSA), may also be inconsistently reported in adolescent populations. Hence, the goal of this study is to assess the consistency of CSA reporting by adolescents.

CSA is linked to wide-ranging emotional, behavioral, and adjustment problems including adolescent pregnancy, aggression, anxiety, depression, risky sexual behaviors, low self-esteem, school problems, and withdrawn behaviors (Kendall-Tackett, Williams, & Finkelhor, 1997; Luster & Small, 1997). The relationship between CSA and future behavioral problems has previously been demonstrated; however, the specific psychosocial mechanisms leading to these outcomes are not clearly understood. One reason for this lack of clarity is methodological in nature, given that CSA measurement varies greatly across studies (Briere, 1992; Goldman & Padayachi, 2000; Roosa, Reyes, Reinholtz, & Angelini, 1998). For example, in many self-report questionnaire studies, CSA is measured by a single item (Bensley, Van Eenwyk, Spieker, & Schoder, 1999; Luster & Small, 1997; Thompson, Potter, Sanderson, & Maibach, 1997), which eliminates the possibility of assessing internal reliability. Other studies have used more extensive measurement methods, such as clinical interviews or medical/psychosocial evaluations (Brown, Kessel, Lourie, Ford, & Lipsitt, 1997; Meyer, Muenzenmaier, Cancienne, & Struening, 1996). Using an interview to assess CSA allows for information concerning the duration and nature of the abuse to be gained. However, these methods are time intensive and the reliability of interview methods has been poorly evaluated. Even relatively brief scales assessing childhood sexual abuse, such as the Early Sexual Experiences Checklist (Miller & Johnson, 1998) and the Unwanted Childhood Sexual Experiences Questionnaire (Stevenson, 1998), are lengthy: 20 items and 13 items respectively. In sum, studies using brief scales to assess CSA are noticeably absent. Using a brief self-report assessment of CSA is important when assessing a wide range of behaviors and psychosocial attitudes in one study.

Methodological research concerning the stability of CSA reporting over time has not been conducted even though researchers have called for data to be collected at more than one time point to assess for developmental differences in symptomatology (Kendall-Tackett, Williams, & Finkelhor, 1997). In an extensive literature search of studies related to CSA reporting, only one recent study (Costello, Angold, March, & Fairbank, 1998) measured the reliability of CSA reporting across two time points. In this study—the focus of which was to assess the utility of a measure of post-traumatic stress disorder (PTSD) with children—10 participants indicated they had been sexually abused. Two weeks later, 9 participants reported that they had been sexually abused, resulting in a Cohen's Kappa correlation of .81. Hence, a significant gap in the CSA research literature exists in terms of stability of reporting.

Self-report consistency at different time points is one approach to assessment of reporting stability (i.e., correctly identifying as non-CSA or CSA). This method has proven

This work was supported in part by grants NIH/NIAID (U19 AI43924) and MCHB (HRSA T71-00008). The authors greatly appreciate the contributions of Cathy Roberts and Patricia Brooks to this research project.

Address correspondence to Matthew C. Aalsma, Ph.D., Section of Adolescent Medicine, Indiana University School of Medicine, 575 N. West Dr., Rm. 070, Indianapolis, IN 46202; e-mail: maalsma@iupui.edu.

effective in past studies of adolescent self-report consistency (Alexander et al., 1993; Capaldi, 1996). Using two time points is expected to create three distinct groups of CSA reporting: consistent nonreporters (deny CSA at both time points), inconsistent reporters (endorse CSA at one point and deny at the other time point), and consistent reporters (endorse CSA at both time points). If the two-time-point approach reduces misclassification, we expect the following: the consistent nonreporters group will include a lower proportion of persons with a true experience of abuse; the inconsistent reporters group will include persons with true positive and false positive reports; and the consistent reporters group will include a larger proportion of true positive reports and a smaller proportion of false positive reports.

If the preceding premise is true, we would then expect problem behaviors associated with CSA to be differentially distributed across the three groups. This is expected given that CSA is associated with specific problem behaviors (see previous review). As a result, we expect consistent nonreporters to have lower levels of problem behaviors, inconsistent reporters to have intermediate levels, and consistent reporters to have higher levels. We expect intermediate levels of problem behavior among inconsistent reporters because their identification of CSA experiences may be influenced by temporal stress, such as relationship difficulty or other life stressors. The demonstration of differential problem behavior involvement will provide another measure of validity. Additionally, endorsement of two or more CSA items should be related to an increased likelihood of consistent reporting. Hence, the prediction of consistent and inconsistent CSA reporting will be explored via the number of CSA items endorsed.

Given the preceding literature review, we sought to explore a number of issues regarding the measurement of CSA in this study. These questions include the following:

1. How stable is self-reporting of CSA over time?
2. Do behavior differences exist that distinguish consistent versus inconsistent reporters of CSA?
3. How useful is a multi-item scale of CSA in comparison to a single-item scale in predicting consistent CSA reporting?

Methods

Adolescent and young adult subjects (14 to 24 years of age) were recruited from urban health clinics (which provide general health care to urban populations) and a sexually transmitted disease (STD) clinic in a large Midwestern city. These data were collected through the Mid-America Adolescent STD Research Center, which conducted a longitudinal study of the behavioral factors involved in STD acquisition. As a result, at the time of enrollment all subjects were diagnosed with or had sexual contact with an individual with an STD. Subjects were paid $15 for their participation.

Procedures

Data pertinent to this study were collected via written self-administered questionnaires (SAQ) at baseline and again at a 7-month follow-up visit. The SAQ included the following measures: a childhood sexual abuse scale, a sexual coercion scale, a sexual behavior assessment, a depressed mood assessment, and an alcohol and marijuana use measure.

Measures

The Childhood Sexual Abuse Scale. The 4-item Childhood Sexual Abuse Scale was created specifically for this research project with the intent to develop a brief, multiple-item tool to assess for CSA. Adolescent participants were instructed that the items referred to events that may have occurred before they were 12 years old. The use of this particular age cutoff was based in part on focus groups with adolescents in which the participants reached a consensus that the term CSA involved events occurring up to 12 years of age. We also wanted the CSA scale to address an age range during which consensual sexual experiences were less likely. To maintain brevity, the CSA scale did not include items regarding the specific nature of the abuse (e.g., whether penetration was involved) or the adolescent's relationship with the perpetrator. The CSA scale demonstrated excellent internal reliability at baseline (alpha = .81) and 7-month follow-up (alpha = .84). The four CSA items are displayed in Table 1.

The CSA scale was used in two ways. First, total scale scores were calculated by summing across items, such that a score of 0 indicated no abuse items had been endorsed and a score of 4 indicated that all items had been endorsed. Next, CSA was treated as a dichotomous variable, in which the first group was defined as those who endorsed a single item on the CSA scale and the second group was defined as those who endorsed two to four items on the CSA scale. The CSA scale was used in this manner to examine the utility of using a single-item measure of CSA versus a multiple-item scale.

The Sexual Coercion Scale. The 4-item Sexual Coercion Scale was also developed specifically for this study. The scale used a 3-point response format (1 = *never*, 3 = *two times or more*; alpha = .67). Sexual coercion was distinguished from CSA as being a more recent

Table 1. CSA Item Endorsement

CSA item (prior to age 12)	Baseline	7 months
Someone tried to touch me in a sexual way against my will.	52 (24%)	40 (18%)
Someone tried to make me touch them in a sexual way against my will.	66 (30%)	45 (21%)
I believe that I have been sexually abused by someone.	45 (21%)	29 (13%)
Someone threatened to tell lies about me or hurt me unless I did something sexual with them.	27 (12%)	13 (6%)

Note. $N = 217$.

experience, occurring after age 12, and it was described to include a wider variety of unwanted sexual experiences, including undesired sex with a boyfriend or girlfriend. The items include the following: "How often has someone used physical force (like punching you or holding you down) to make you have sex?"; "How often has someone used a weapon (like a gun or a knife) to make you have sex?"; "How often have you had sex with someone because you were afraid of them?"; "How often have you had sex when you really didn't want to?"

Sexual behavior. Assessed sexual behaviors included lifetime number of sexual partners, age at first sexual intercourse (focus group interviews indicated that adolescents could distinguish abuse from first consensual intercourse), and condom use (3-item scale). This scale examined the participants' reasons for condom use over the past 2 months with a 4-point response format (1 = *never*, 4 = *always*). Internal consistency was excellent (alpha = .94).

Depressed mood. Depressed mood was evaluated with a 4-item scale that focused on feelings of hopelessness and dysphoria (e.g., "In the past 6 months, have you felt pretty hopeless about the future?"). A 4-point response format was used (1 = *not at all*, 5 = *a lot*; alpha = .89).

Alcohol and marijuana use. Alcohol use over the past 2 months was measured with two items in a 5-point response format assessing frequency (0 = *not at all*, 4 = *daily use*) and amount (4 = *six or more cans of beer, glasses of wine, or drinks of liquor*, 0 = *I didn't drink at all*; alpha = .84) of alcohol use. The marijuana use scale used two items in a 5-point response format assessing lifetime use (1 = *never*, 5 = *very often*) and use during the past 2 months (1 = *never*, 5 = *about every day*; alpha = .84).

Of the 444 adolescents enrolled, complete data at both time points were available for 217 subjects. A comparison was made between those who completed the baseline assessment and those who completed both baseline and 7-month assessments. Participants who did not complete the 7-month assessment were more likely to be males ($p < .001$), to be older ($p < .001$), to drink alcohol ($p < .01$), to use marijuana ($p < .05$), and to have more sex partners in the last 2 months ($p < .01$). However, on other scales that are more salient to this exploration of the stability of CSA reporting, there were no statistically significant differences between the groups (i.e., CSA scale, sexual coercion scale, condom use scale, depressed mood scale).

In sum, 217 adolescents had complete data at baseline and at the 7-month follow-up. Subjects ranged in age from 14 to 24 years (mean = 17.0, standard deviation = 2.0) and 83% were female. The majority of the subjects were African American (82%). The gender, age, and race proportions of the subject cohort were similar to those of the urban health care clinic population. Written informed consent was obtained from all participants, and the university's institutional review board approved the research protocol. The requirement for parental permission was waived.

Results

Statistical Analyses

The statistical analysis was carried out in sequential steps. First, the stability of CSA reporting over time was evaluated by calculating Pearson product-moment correlations between baseline and 7-month CSA scale scores. A second approach to assess the stability of CSA reporting was to calculate the number of subjects who endorsed at least one item at one time point. We then determined the proportion of these individuals who endorsed at least one item at both time points (consistent reporters) and the proportion that endorsed at least one item at only one time point (inconsistent reporters).

The second analysis entailed distinguishing consistent reporters from inconsistent reporters in terms of demographic information, recent coercive sexual experiences, depression, substance abuse, and sexual behaviors. One-way Analysis of Variance (ANOVA) was subsequently used.

Third, we wanted to compare the groups that were formed on their overall CSA scale scores. Therefore, we conducted a *t*-test analysis on the individuals who endorsed at least one item from the CSA scale. For the purpose of this analysis, individuals who reported CSA at both time points (consistent reporters) were compared to inconsistent reporters on their CSA scale scores.

Last, we conducted a logistic regression to predict membership in the consistent or inconsistent reporting group. This analysis was conducted by using the cutoff of endorsing two or more items from the CSA scale as the dependent variable to predict the independent variable of group membership (i.e., consistent versus inconsistent CSA reporters).

Sample Characteristics

At baseline, 81 of the 217 adolescents (37%) endorsed at least one item from the CSA scale. At 7 months, 54 of the 217 adolescents (25%) endorsed at least one item from the CSA scale. Ninety-one adolescents (42%) reported CSA at either baseline or 7 months. Given the disparity in reporting of CSA over time (81 adolescents at baseline versus 54 at seven months), a Pearson product-moment correlation was conducted of baseline and 7-month reporting of CSA to assess the consistency of reporting at both time points. The resulting correlation was significant ($r = .64$, $p < .001$) but not very strong given that CSA should be a static entity as measured in this study. A Pearson product-moment correlation was then conducted with baseline and 7-month reporting of CSA only among subjects who reported CSA at some time point (91 participants). The resulting correlation was significant ($r = .34$, $p < .001$) and much lower than the previous correlation, which indicates that there is substantial variability in the number of CSA items endorsed.

Table 2 presents CSA reporting over time. The majority of adolescents (126 participants; 58%) were consistent

Table 2. CSA Reporting Over Baseline and 7 Months

Baseline CSA scale endorsement	7-month CSA scale endorsement	
	No	Yes
No	126 (58%)	10 (5%)
Yes	37 (17%)	44 (20%)

nonreporters of CSA across time. Forty-four participants (20%) were consistent reporters of CSA, in that they endorsed at least one item at both time points. Forty-seven respondents (22%) were inconsistent reporters of CSA, in that they either endorsed at least one item at baseline but not at 7 months (37 participants; 17%), or they did not endorse an item at baseline but did at 7 months (10 participants; 5%). These three groups (consistent nonreporters, inconsistent reporters, and consistent reporters) were then compared on relevant demographic, behavioral, and mental health variables from enrollment data using one-way ANOVA with posthoc Scheffe contrasts.

Table 3 presents significant means and standard deviations by the CSA groups. Significant univariate effects were found for the coercive sexual experiences scale ($F = 34.49$, $p < .001$) as well as for depression ($F = 7.16$, $p < .001$). The lifetime number of sexual partners variable evidenced outliers (0 to 100 partners); hence, the variable was standardized for the analyses. Subsequently, the lifetime number of sexual partners was significant ($F = 7.20$, $p < .001$). The remaining variables were nonsignificant.

To assess the ability of the CSA scale to predict stable reporting, the consistent reporters of CSA (44 participants) were compared to inconsistent reporters (47 participants) in terms of their overall score on the CSA scale at baseline. The group differences were statistically significant (t [89] $= 3.05$, $p < .01$) with consistent CSA reporters evidencing higher scale scores (mean = 2.64, $SD = .99$) than inconsistent reporters (mean = 2.00, $SD = 1.00$).

Last, we conducted a logistic regression to predict membership of consistent CSA (1) or inconsistent CSA (0) groups based on CSA scale scores. The scale scores on the CSA measures were divided into those who endorsed one item on the CSA scale and those who endorsed two to four items on the CSA scale. Hence, group membership acted as the dependant variable (consistent versus inconsistent reporting) and scale score on the CSA scale was the independent variable (those who endorsed one item versus those who endorsed two to four items). The logistic regression revealed that the model was significant ($\chi^2 = 10.42$; $df = 1$; $p < .001$). Based on this model, 89% of the consistent CSA reporters were correctly classified into the consistent reporting group and 40% of the inconsistent reporters were correctly classified. In sum, adolescents who endorsed two or more items on the CSA scale were over five times more likely to be consistent reporters of CSA.

Discussion

There was considerable instability in CSA scale scores over a 7-month time period. This instability was predominantly a function of inconsistent reports of abuse. Reports of no abuse were generally quite stable, as reflected in the fact that 93% of those reporting no abuse at baseline continued to report no abuse at the 7-month follow-up. The pattern of instability was similar across all four items of the CSA Scale. As noted in Table 1, a decrease in reporting of CSA occurred from baseline to 7-month reporting. There could be a number of reasons for this decline. For instance, although the CSA scale was a self-report measure, structured interviews were conducted at several time points in the course of this research project. Although this was not an intervention project, the subjects may have been less willing to report CSA as occurring as they felt more comfortable with the research assistant. Additionally, the subjects' level of stress may have been higher at baseline measurement given that enrollment in the study required an STD diagnosis. This may have resulted in subjective life interpretations, in turn resulting in higher endorsement of CSA at baseline. Regardless of the reasons, the CSA rates were quite inconsistent.

Significant differences existed between consistent, inconsistent, and nonreporters of CSA. Consistent CSA reporters endorsed the highest rates of sexual coercion experiences, depression, and lifetime number of sexual partners. The utility of assessing CSA at two time points is evident since consistent CSA reporters endorsed marked increases in measures of pathology (i.e., depression) and health-compromising behavior. Moreover, a linear trend is evident with the sexual coercion and depression scales. Although not statistically significant, nonreporters endorsed higher rates of sexual coercion and depression, while inconsistent reporters endorsed moderate levels on the sexual coercion and depression scales. This is again indicative of the utility and importance of assessing CSA at two time points given that CSA endorsement is unstable and that certain problem behavior rates are significantly different between the reporting groups.

Table 3. Means and Standard Deviations by CSA Reporting Group

Variable	Non-CSA	Inconsistent	Consistent
Sexual coercion	$9.20\ (2.06)_a$	$10.53\ (2.41)$	$13.07\ (3.99)_a$
Depression scale	$17.08\ (5.92)_a$	$18.98\ (6.01)_b$	$21.00\ (6.31)_{ab}$
Lifetime sexual partners	$9.27\ (11.96)_a$	$8.47\ (9.31)_b$	$15.61\ (19.79)_{ab}$

Note. Matching subscripts indicate significant differences at the $p < .01$ level.

We found that consistent reporters endorsed more CSA items than did inconsistent reporters. In light of this finding, a final analysis was conducted to predict membership into either the consistent or inconsistent reporters group based on participants' baseline CSA scale scores. The subjects were divided into those who endorsed two to four items from the baseline CSA scale and those who endorsed a single item. The results indicated that adolescents who endorsed at least two items on the CSA scale were over five times more likely to be consistent CSA reporters. The results of this analysis demonstrate the utility of using a scale to measure CSA rather than a single item. By using a scale to assess CSA, it was possible to predict group membership based on the number of items endorsed 89% of the time for consistent reporters and 40% of the time for inconsistent reporters. These findings argue for the use of a multi-item scale to measure CSA in self-report survey studies.

Through this research project, we have shown that reports of CSA are inconsistent. Moreover, clear health-compromising behavior differences exist between the consistent, inconsistent, and nonreporting CSA groups. Finally, a logistic regression revealed that adolescents endorsing two or more CSA scale items were five times more likely to be consistent reporters of CSA.

Based on the above research findings it is recommended that additional studies be conducted to assess the consistency of CSA reporting with other normative and at-risk populations. Given that this sample is a high-risk group of adolescents, the inherent stress of being diagnosed with an STD may have influenced CSA reporting. Nonetheless, data assessing the consistency of behavior-reporting from this and other samples is important given that behavioral research is often dependent on self-report data. Additionally, the clearly demonstrated advantages of our brief but multi-item scale over a single item suggest that future self-report survey research avoid the use of a single item to assess CSA. It should be highlighted that this study does not negate the clinical utility of CSA reporting. The inconsistency of CSA reporting in our sample has ramifications for measurement issues related to CSA, but implications for the consistency of CSA reporting in applied, clinical settings was not explored and the current research project does not address that research question.

References

Alexander, C. S., Somerfield, M. R., Ensminger, M. E., Johnson, K. E., & Kim, Y. J. (1993). Consistency of adolescents' self-report of sexual behavior in a longitudinal study. *Journal of Youth and Adolescence, 22,* 455–471.

Bensley, L. S., Van Eenwyk, J., Spieker, S. J., & Schoder, J. (1999). Self-reported abuse history and adolescent problem behaviors: I. Antisocial and suicidal behaviors. *Journal of Adolescent Health, 24,* 163–172.

Briere, J. (1992). Methodological issues in the study of sexual abuse effects. *Journal of Consulting and Clinical Psychology, 60,* 196–203.

Brown, L. K., Kessel, S. M., Lourie, K. J., Ford, H. H., & Lipsitt, L. P. (1997). Influence of sexual abuse on HIV-related attitudes and behaviors in adolescent psychiatric inpatients. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 316–322.

Capaldi, D. M. (1996). The reliability of retrospective report for timing first sexual intercourse in adolescent males. *Journal of Adolescent Research, 11,* 375–387.

Costello, E. J., Angold, A., March, J., & Fairbank, J. (1998). Life events and post-traumatic stress: The development of a new measure for children and adolescents. *Psychological Medicine, 28,* 1275–1288.

Goldman, J. D. G., & Padayachi, U. K. (2000). Some methodological problems in estimating incidence and prevalence in child sexual abuse research. *The Journal of Sex Research, 37,* 305–314.

Kendall-Tackett, K. A., Williams, L. M., & Finkelhor, D. (1997). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin, 113,* 164–180.

Luster, T., & Small, S. A. (1997). Sexual abuse history and number of sex partners among female adolescents. *Family Planning Perspectives, 29,* 204–211.

Meyer, I. H., Muenzenmaier, K., Cancienne, J., & Struening, E. (1996). Reliability and validity of sexual and physical abuse histories among women with serious mental illness. *Child Abuse & Neglect, 20,* 213–219.

Miller, R. S., & Johnson, J. A. (1998). Early sexual experiences checklist. In C. M. Davis, W. L. Yarber, R. Bauserman, G. Schreer, & S. L. Davis (Eds.), *Handbook of sexuality-related measures* (pp. 23–25). Thousand Oaks, NJ: Sage Publications.

Newcomer, S., & Udry, J. R. (1988). Adolescents' honesty in a survey of sexual behavior. *Journal of Adolescent Research, 3,* 419–423.

Roosa, M. W., Reyes, L., Reinholtz, C., & Angelini, P. J. (1998). Measurement of women's child sexual abuse experiences: An empirical demonstration of the impact of choice of measure on estimates of incidence rates and of relationships with pathology. *The Journal of Sex Research, 35,* 225–233.

Stevenson, M. R. (1998). Unwanted childhood sexual experiences questionnaire. In C. M. Davis, W. L. Yarber, R. Bauserman, G. Schreer, & S. L. Davis (Eds.), *Handbook of sexuality-related measures* (pp. 28–29). Thousand Oaks, NJ: Sage Publications.

Thompson, N. J., Potter, J. S., Sanderson, C. A., & Maibach, E. W. (1997). The relationship of sexual abuse and HIV risk behaviors among heterosexual adult female STD patients. *Child Abuse & Neglect, 21,* 149–156.

Manuscript accepted April 8, 2002