

September 2006

Retrospective Assessment of Childhood Sexual and Physical Abuse: A Comparison of Scaled and Behaviorally Specific Approaches

David DiLillo

University of Nebraska - Lincoln, ddilillo@unl.edu

Michelle A. Fortier

University of Nebraska - Lincoln

Sarah A. Hayes

University of Nebraska - Lincoln

Emily Trask

University of Nebraska - Lincoln

Andrea R. Perry

University of Nebraska - Lincoln

See next page for additional authors

Follow this and additional works at: <http://digitalcommons.unl.edu/psychfacpub>



Part of the [Psychiatry and Psychology Commons](#)

DiLillo, David; Fortier, Michelle A.; Hayes, Sarah A.; Trask, Emily; Perry, Andrea R.; Messman-Moore, Terri; Fauchier, Angèle; and Nash, Cindy, "Retrospective Assessment of Childhood Sexual and Physical Abuse: A Comparison of Scaled and Behaviorally Specific Approaches" (2006). *Faculty Publications, Department of Psychology*. 125.

<http://digitalcommons.unl.edu/psychfacpub/125>

This Article is brought to you for free and open access by the Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Psychology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors

David DiLillo, Michelle A. Fortier, Sarah A. Hayes, Emily Trask, Andrea R. Perry, Terri Messman-Moore, Angèle Fauchier, and Cindy Nash

Retrospective Assessment of Childhood Sexual and Physical Abuse: A Comparison of Scaled and Behaviorally Specific Approaches

David DiLillo

Michelle A. Fortier

Sarah A. Hayes

Emily Trask

Andrea R. Perry

University of Nebraska–Lincoln

Terri Messman-Moore

Miami University, Ohio

Angèle Fauchier

University of Southern California

Cindy Nash

University of Nebraska–Lincoln

Abstract: This study compared retrospective reports of childhood sexual and physical abuse as assessed by two measures: the Childhood Trauma Questionnaire (CTQ), which uses a Likert-type scaling approach, and the Computer Assisted Maltreatment Inventory (CAMI), which employs a behaviorally specific means of assessment. Participants included 1,195 undergraduate students recruited from three geographically diverse universities. Agreement was high across the two measures in the classification of victim status (92% and 80% for sexual and physical abuse, respectively). However, the CTQ classified more participants as sexually abused than did the CAMI, whereas the opposite trend was found for physical abuse. For child physical abuse, many participants reporting abusive acts on the CAMI scored below the cut-point for physical abuse on the CTQ. Classification differences for both types of abuse were largely unrelated to demographic factors, socially desirable responding, or self-reported withholding of information. The implications of these results are discussed in light of future research using retrospective methods of assessing childhood abuse.

Keywords: sexual abuse; physical abuse; assessment; retrospective

During the past three decades, a multitude of studies have been instrumental in revealing the prevalence, nature, and have relied on adults' retrospective reports to assess a history long-term consequences of both forms of child maltreatment of childhood sexual and physical abuse.

These studies have been instrumental in revealing the prevalence, nature, and long-term consequences of both forms of child maltreatment (see Malinosky-Rummell & Hansen, 1993; Polusny & Follette, 1995, for reviews of this literature). Countless other investigations have included

retrospective assessments of abuse as secondary measures. Common formats for the retrospective assessment of sexual and physical abuse include interviews (in-person and telephone), paper-and-pencil questionnaires, chart reviews, and more recently, computer-assisted instruments (Abbey, BeShears, Clinton-Sherrod, & McAuslan, 2004; Nash, DiLillo, Messman-Moore, & Rinkol, 2002). Regardless of format, an important function of these measures is to identify individuals who have and have not experienced childhood abuse. Indeed, the ability to classify participants dichotomously according to the presence or absence of abuse is fundamental to research addressing issues such as the prevalence or long-term consequences of a particular form of maltreatment. Thus, the accuracy of retrospective measures of abuse history is critical to ensure that the results of studies using the methodology are not misleading.

Although prevalent in the literature, retrospective assessments of maltreatment may be limited by several factors, including errors in recall due to the passage of time, intentionally false responding (either false positives or false negatives), and the possible inaccessibility of memories for traumatic abuse (Widom & Morris, 1997). Abundant writings by memory and cognitive researchers, as well as trauma experts, have drawn attention to issues and limitations regarding the accuracy of self-reported childhood trauma (cf. Williams & Banyard, 1998). The primary question underlying the use of retrospective methods therefore becomes, to what degree do recollections of childhood victimization reflect the actual occurrence (or nonoccurrence) of abuse? As with any construct, the answer to this question depends to a large extent on evidence concerning the reliability and validity of the measurement techniques used. The reliability of retrospective reports can be evaluated in a fairly direct manner by examining the stability or fluctuation in abuse reports over time. Several investigations of concordance rates between reports of childhood sexual abuse (CSA) obtained on two different occasions have found consistently high temporal stability. For example, Friedrich, Talley, Panser, Fett, and Zinsmeister (1997); Meyer, Muenzenmaier, Cancienne, and Struening (1996); and Lesserman, Drossman, and Zhiming (1995) reported 85.6%, 91%, and 81% agreement, respectively, on reports of sexual abuse from Time 1 to Time 2 (intervals ranging from 2 weeks to 24 months). Based on this evidence, it appears that abuse-specific reports maintain a fairly high degree of stability throughout at least a 2-year time span.

The most telling measure of the validity of retrospective reporting involves the comparison of adult abuse reports to independent corroborative information obtained at the time the maltreatment occurred

(e.g., from hospitals, child protective services, or police records). Only a few studies have been equipped to evaluate the accuracy of adult recollections of abuse in this manner. Williams (1994) interviewed 129 women approximately 17 years after the occurrence of documented sexual abuse. Results revealed that 38% of these women failed to report the specific abuse incident for which they had been treated as a child (see Widom & Morris, 1997, for comparable findings). However, only 12% of women had no recollection of any sexual abuse whatsoever; that is, although some memories for specific episodes had been lost, 88% of participants retained generalized information regarding some type of CSA. In a similarly designed study, Goodman et al. (2003) reported that 81% of 175 adults with documented histories of sexual abuse recalled the target incidents during a follow-up interview approximately 13 years later. Finally, in examining retrospective reports of childhood physical abuse, Widom and Shepard (1996) found that depending on the definition used, between 60% and 92% of adults reported incidents of abuse that had been officially documented years earlier. Thus, when considering the important question of whether sexual and physical abuse status can be accurately assessed through retrospective methods, data from prospective studies suggest that although the majority—perhaps as many as 80% to 90%—of cases can be accurately identified, there remains a need to improve on the methods used.

Despite the importance of establishing measurement accuracy, with very few exceptions, instruments used in abuse history research have not been psychometrically evaluated (Hulme, 2004). One notable exception is the Childhood Trauma Questionnaire (CTQ; Bernstein & Fink, 1998; Bernstein, Fink, Handelsman, & Foote, 1994), which has been subjected to extensive psychometric testing and is commonly used in retrospective studies of childhood maltreatment. The most recent version of the CTQ (Bernstein & Fink, 1998) contains five subscales assessing different types of child maltreatment, including sexual abuse, physical abuse, emotional abuse, physical neglect, and emotional neglect; there are also three items assessing tendencies toward minimization and denial. The CTQ employs what can be called a scaling approach (Hulme, 2004) in that its Likert-type scales have been shown through exploratory and confirmatory factor analysis to represent latent constructs reflecting various subtypes of abuse (Bernstein & Fink, 1998). However, rather than adopting specific behavioral descriptors of abusive acts—and comparing self-reported experiences to these criteria—this instrument produces dimensional scales to which cutoff scores, supplied by the authors, are used to classify individuals as abused or not abused.

Unlike the scaled approach described above, the Computer Assisted Maltreatment Inventory (CAMI; Nash et al., 2002) is a newly developed, behaviorally specific instrument. Rather than using Likert-type scales, the CAMI employs a series of screener questions that conform to precise operational definitions of sexual and physical abuse. In assessing CSA, for example, the CAMI describes specific sexual activities ranging from kissing and fondling to intercourse and then asks whether, prior to the age of 18, the respondent experienced any of these activities under conditions that would be considered abusive. Physical abuse is assessed in a similar manner. For both abuse types, an affirmative response to one or more screener questions is followed by queries about the nature and circumstances of the reported activities (e.g., perpetrator identity, type and frequency of abusive incidents, age at onset and termination, use of verbal coercion or physical force). Although a positive screener is suggestive of abuse, answers to the more detailed follow-up questions are examined to make an ultimate determination of victimization status. Thus, in contrast to the CTQ, the CAMI uses activity-specific screeners followed by more detailed inquiries to distinguish participants who have and have not experienced abuse according to a specific operational definition. This approach intentionally minimizes the need for subjective interpretation of items by the respondent and avoids the use of potentially sensitive labels such as “abuse” and “victim.”

The contrasting approaches represented by the CTQ and the CAMI raise questions about the relative utility of each in detecting childhood sexual and physical abuse. Some authors have suggested that measures using multiple, behaviorally specific items to screen for abuse are superior to more general or single-item measures (Wyatt & Peters, 1986). This notion has been supported by several findings that multiple behaviorally specific questions are superior to omnibus screeners in eliciting reports of various traumatic events, including child maltreatment (DiLillo, Hayes, & Hope, 2006; Franklin, Sheeran, & Zimmerman, 2002; Fricker, Smith, Davis, & Hanson, 2003; Weaver, 1998). Lipschitz, Bernstein, Winegar, and Southwick (1999) compared adolescent inpatients' responses to the CTQ and the Traumatic Events Questionnaire–Adolescent version (TEQ-A; Winegar & Lipschitz, 1997), which uses multiple-choice questions to detect sexual and physical abuse history in a more behaviorally specific manner. These authors found that physical abuse status was reported consistently across both measures in 86% of cases, whereas sexual abuse was reported consistently in 71% of cases. Nevertheless, 35% of those classified as CSA victims on the CTQ were not identified as such on the TEQ-A, which suggests a fair level of inconsistency

in the identification of CSA using these instruments.

Although this work sheds light on similarities and differences in reporting patterns across measures, little is known about possible reasons for discrepant reporting on scaled versus behaviorally specific measures. One factor that may influence responding is social desirability, generally described as an individual's need for approval (Crowne & Marlowe, 1960). Social desirability is thought to be reflected through habitual response styles and personal expectations that are elicited when engaged in self-evaluation (Crowne & Marlowe, 1960; Leite & Beretvas, 2005). Past research has found that individuals high in social desirability may withhold sensitive information to appear more favorable to others (e.g., Dutton & Hemphill, 1992; Latkin, Vlahov, & Anthony, 1993). Similar response biases may affect the reporting of abuse history, particularly when assessments include the more graphic questions contained on a behaviorally specific instrument. To fully explore this possibility, the present study examines associations between self-reported abuse and social desirability as measured with two common instruments assessing separate but related dimensions of social desirability.

To further understand the types of maltreatment that are most likely to be detected (or missed) by the CTQ and CAMI, the current study also examines the abuse characteristics of participants reporting maltreatment on each measure. Factors such as participant demographics and the intentional withholding of information are explored as possible reasons for the differential reporting of abuse. Finally, this study appears to be the first to contrast responses to both types of measures using a large, geographically diverse sample of undergraduate students. This is an important gap in the literature, considering the popularity of the CTQ and the prevalence of abuse history studies with college-age samples.

METHOD

Participants

Participants were recruited from three geographically diverse universities: the University of Nebraska–Lincoln (UNL), Miami University (in Ohio, MU), and the University of Southern California (USC). The entire sample was composed of 1,195 undergraduate students ($n = 863$ women and 332 men), with an average age of 20.3 ($SD = 2.3$). Of the total sample, 402 participants were from UNL ($n = 108$ men and 294 women), 358 from MU ($n = 124$ men and 234 women), and 435 from USC ($n = 100$ men and 335 women). Approximately 4% of all the participants reported a family income while growing up of less than \$20,000, 23% reported between \$20,000 and \$50,000, and 73% report-

ed a family income of greater than \$50,000, suggesting that, overall, the group was composed of participants from mostly a middle to upper-middle income range. Approximately 63% of participants' fathers held at least a bachelor's degree, whereas 55% had mothers who had attained a similar level of education. Regarding ethnicity, 76.2% of participants were European American, 9.0% Asian American, 5.2 % Hispanic/Latin American, 3.4% African American, 0.2% Native American, 0.2% Hawaiian/Pacific Islander, and 5.9% Other (e.g., self-reported bi- or multiracial or bi- or multiethnic individuals).

Measures

Childhood Trauma Questionnaire. The CTQ (Bernstein et al., 1994; Bernstein & Fink, 1998) was originally developed as a 70-item instrument designed to assess a variety of child maltreatment experiences (Bernstein et al., 1994). In subsequent years, the authors have refined the item pool based on significant psychometric evaluation of the CTQ. The most recent version of the instrument consists of 28 items that were retained from the original 70-item version (Bernstein et al., 2003). It is these items that are recommended by the primary author of the CTQ (D. Bernstein, personal communication, November 2001) and thus were selected from the larger pool for inclusion in the present study. These items assess retrospective reports of childhood maltreatment in five areas: emotional abuse, physical abuse, sexual abuse, emotional neglect, and physical neglect. Each abuse subscale contains five items, with the remaining three assessing minimization and denial. Items are rated on a 5-point Likert-type scale (1 = *never true* to 5 = *very often true*). Some CTQ items are stated objectively, whereas others are more general and require subjective interpretation on the part of the respondent. Sample items from the sexual abuse subscale include the following: "Someone molested me" and "Someone tried to make me do sexual things or watch sexual things." Sample physical abuse items include the following: "I believe I was physically abused" and "People in my family hit me so hard that it left me with bruises or marks." Administration time for the CTQ is between 5 and 7 minutes.

The CTQ produces both dimensional and categorical levels for each form of maltreatment. Individuals must obtain a raw score of at least 6 on the CTQ item responses to be classified as sexually abused (Bernstein & Fink, 1998). Individuals who are considered sexually abused based on these guidelines can then be classified as experiencing low to moderate levels of abuse (score of 6-7), moderate to severe levels of abuse (score of 8-12), or severe to extreme levels of abuse (score of 13 or more). A raw score of at least 8 is required to be classi-

fied as physically abused. Individuals who are considered physically abused based on these guidelines can then be classified as experiencing low to moderate levels of physical abuse (score of 8-9), moderate to severe levels of physical abuse (score of 10-12), or severe to extreme levels of physical abuse (score of 13 or more).

Bernstein and Fink (1998) report mean internal consistency estimates of .92 for the sexual abuse subscale and .80 for the physical abuse subscale across eight samples (e.g., adult substance abusers, adolescent psychiatric inpatients, college undergraduates). The internal consistency in the current sample was .93 for the sexual abuse subscale and .75 for the physical abuse subscale. Test-retest reliabilities have been reported as .80 for physical abuse and .81 for sexual abuse throughout a 1.6- to 5.6 month time period (Bernstein & Fink, 1998). Confirmatory factor analyses indicate that the five-factor model utilized in the CTQ has been supported across three samples (i.e., adolescent psychiatric inpatients, adult substance abusers, and female health maintenance organization [HMO] members), which is suggestive of the measure's construct validity (Bernstein, Ahluvalia, Pogge, & Handelsman, 1997; Bernstein et al., 1994; Bernstein et al., 2003). Criterion-related validity of the CTQ also has been examined by Bernstein et al. (2003). Specifically, therapist ratings of childhood maltreatment were examined within an adolescent psychiatric inpatient sample. Therapist ratings were positively associated with participant scores on the CTQ scales ($r = .59$ for physical abuse and $r = .75$ for sexual abuse). These therapist ratings of trauma, along with other maltreatment data (obtained from the Evaluation of Lifetime Stressors questionnaire) available for the female HMO members, were used as validity criteria to develop cutoff scores for the CTQ using receiver operating characteristic (ROC) analysis (Hsiao, Bartko & Potter, 1989). ROC procedures seek to establish a balance between the false positive rates and false negative rates (i.e., sensitivity and specificity) for all possible cutoff scores, which can be subjected to significance testing. The lowest cutoff scores were established to detect a maximum number of low-severity abuse cases (at least 80% of cases reported in the criterion interview) while keeping the rate of false positives at less than 20%. Cutoff scores were derived from the female HMO sample because this sample was likely to include more cases of lesser severity abuse than did the adolescent psychiatric sample (Bernstein & Fink, 1998).

Computer Assisted Maltreatment Inventory. The CAMI (DiLillo, DeGue, Kras, Di Loreto-Colgan, & Nash, 2006; Nash et al., 2002) is a computer-based, self-report measure that is designed to screen and assess each major form of child maltreatment, including

sexual abuse, physical abuse, psychological abuse, neglect, and exposure to domestic violence. Questions that comprise each child maltreatment subscale are behaviorally specific and are based on acts commonly defined as abuse in the child maltreatment literature. Because the psychological abuse, neglect, and exposure to domestic violence subscales remain in the early stages of development, only the sexual abuse and physical abuse subscales are examined in the present study. Administration time for the full CAMI varies greatly but generally ranges between 7 and 20 minutes, depending on the type and extent of maltreatment history reported.

Sexual abuse is assessed on the CAMI by first presenting participants with a list of sexual activities ranging in intensity from sexual touching and kissing to oral, anal, or vaginal intercourse. Participants then indicate whether, prior to age 18, they experienced any of these activities under the following three conditions: (a) against their will or when they did not want the activities to happen, (b) with a close family member or relative, and (c) with someone more than 5 years older than themselves. Participants are instructed to exclude voluntary sexual activities with a dating partner and any consensual sexual play or exploration with a similar-age peer. Those who respond affirmatively to any of the screener questions are considered likely to have experienced some form of childhood or adolescent abuse and are therefore directed to a series of more detailed questions about the nature of these experiences. These follow-up questions ask about the identity of up to three individuals (i.e., perpetrators) with whom the sexual activities occurred. For each perpetrator, participants indicate the frequency of various sexual acts, including sexual kissing; fondling (by victim and perpetrator); masturbation (by victim and perpetrator); oral, anal, and/or vaginal penetration; and attempted or actual intercourse on a 5-point scale (1 = *never happened*, 2 = *1-2 times*, 3 = *3-5 times*, 4 = *6-10 times*, 5 = *more than 10 times*). Respondents also provide information about the age at which the abuse began and ended, reasons for termination, and whether physically or verbally coercive tactics were used by the perpetrator to obtain sexual contact.

Participants who endorse one or more screeners and whose follow-up responses are consistent with the following criteria are classified as victims of CSA. First, any non-consensual (forced or coerced) sexual experience involving hands-on physical contact (i.e., sexual touching; kissing; oral, anal, or vaginal intercourse) before the age of 14 was considered to be CSA. In addition, an age difference criterion that often has been employed in the literature was applied (cf. Rind, Tromovitch, & Bauserman, 1998). Specifically, for incidents

occurring prior to the age of 14, CSA constituted any hands-on activity with someone more than 5 years older than the respondent, regardless of consent. For adolescents age 14 to 17, CSA constituted any act that occurred with someone 10 or more years older, regardless of consent. Finally, any explicitly sexual act occurring with a close family member that was not characterized as sex play or exploration was considered to be CSA. In contrast to CSA, adolescent sexual assault (e.g., date or acquaintance rape) was defined as any nonconsensual, hands-on sexual activity occurring between the age of 14 to 17 with a perpetrator who was fewer than 10 years older than the respondent.

Similar to sexual abuse, child physical abuse (CPA) is measured by a series of screener questions that identify whether participants experienced physically abusive acts committed by a parent or caregiver before the age of 18. Those responding affirmatively to any physical abuse screener are then asked to identify up to three potential perpetrators and to provide information about the frequency (e.g., 1 = *never happened*, 5 = *more than 10 times*) of several acts, including shaking, slapping, pinching, severe spanking, punching, kicking, choking, burning, hitting with an object, and threatening with a weapon. Based on these responses, participants are classified as physically abused if they report the frequent occurrence (10 or more times) of lower severity acts such as being grabbed, spanked on the bottom with an object, or pinched hard. A classification of physical abuse also is made for participants reporting that moderately severe acts (being slapped, spanked so hard it left a mark or bruise, and/or hit on a body part other than the bottom by a parent or caregiver) occurred at least three times. Finally, those who report experiencing one or more instances of high-severity abusive acts also are considered to have been victims. These acts included being thrown or knocked down; hit with a hard object; grabbed around the neck or choked; kicked; punched with a fist; beat by slapping, hitting, and/or punching; burned or scalded on purpose; threatened with a weapon; or hit with an object that could cause major injury.

Because various means of physical punishment are commonly used by parents in the United States (Straus & Stewart, 1999), the above criteria were established based on the rationale that less severe acts occurring more frequently and more severe acts occurring less frequently can be delineated from more normative acts of physical discipline. This two-dimensional conceptualization, taking into account both the frequency and severity of acts, is consistent with the popular notion of child physical abuse as excessive discipline or inappropriate physical treatment of children that “crosses the line” into abuse because of its po-

tential to lead to harm or injury (Whipple & Richey, 1997). Here, emphasis is placed on the circumstances and nature of the acts, as opposed to the physical consequences per se, because even abusive behaviors do not always lead to visible harm (Wolfe, 1988). Furthermore, standard definitions of CPA (Hansen, Sedlar, & Warner-Rogers, 1999; Kolko, 2002) as well as other self-report instruments such as the Parent-Child Conflict Tactics Scales (Straus, Hamby, Finkelhor, Moore, & Runyan, 1998) focus on specific behaviors considered abusive (e.g., beating, burning, suffocating). The CAMI operationalization reflects a similar approach by assessing the occurrence of several types of behaviorally specific acts.

Because the CAMI is a newly developed instrument, data regarding its psychometric properties are not yet extensive. However, to glean test-retest reliability, 100 college students were administered the CAMI twice throughout a period of 1 to 2 weeks. Results indicated that test-retest reliability for the sexual and physical abuse subscales was .71 and .86, respectively (Nash et al., 2002). Initial indications of concurrent validity come from studies with various samples showing expected relationships between sexual and physical abuse reports and a variety of adverse outcomes, including psychological distress (Clemmons, 2005), adult revictimization (Messman-Moore, Denardi, Gaffey, Fauchier, & DiLillo, 2004), problematic dating and marital relationships (DiLillo, Di Loreto, & Perry, 2003), poor coping responses (Fortier et al., 2004), cruelty toward animals (DeGue, Perry, DiLillo, Messman-Moore, & Fauchier, 2004), and an expected overlap with other forms of maltreatment (Clemmons, Fortier, & DiLillo, 2003).

Marlowe-Crowne Social Desirability Scale (MCSDS). The MCSDS (Crowne & Marlowe, 1960) scale consists of 33 true and false items that reflect positive and negative statements (e.g., "I have never intensely disliked anyone" or "I am sometimes irritated by people who ask favors of me," respectively) that are highly improbable. Higher scores on the scale indicate a tendency to portray oneself in a positive light. A recent meta-analysis by Beretvas, Meyers, and Leite (2002) suggests that the test-retest reliability is variable (range = .38-.86) and internal consistency is acceptable (for men = .70, for women = .80). Internal consistency in this sample was .80. Furthermore, Paulhus (1991) found that the concurrent validity with the Balanced Inventory of Desirable Responding was good (.71).

Balanced Inventory of Desirable Responding, Version 6 (BIDR). The BIDR (Paulhus, 1994) is a 40-item questionnaire designed to measure socially desirable responding. Specifically, the BIDR encompasses two subscales—Self-Deceptive Enhancement and Impression

Management—with each consisting of 20 questions. According to the BIDR's author, self-deceptive enhancement refers to "the tendency to give self-reports that are honest but positively biased," whereas impression management is defined as "deliberate self-presentation to an audience" (Paulhus, 1991, p. 37). Both constructs are considered to be stable personality traits. Participants choose their responses by rating how much they agree with each of the 40 statements based on a 7-point Likert-type scale (ranging from 1 = *not true* to 7 = *very true*). Scores were obtained for the overall measure as well as for Impression Management and Self-Deception. Higher scores indicate increased social desirability. As reported by Paulhus (1991), the BIDR has excellent psychometric properties. Specifically, measures of reliability are as follows: internal consistency (self-deception, range = .68-.80; impression management, range = .75-.86; overall score = .83) and test-retest (for self-deception = .69; for impression management = .65). Measures of validity also were adequate (concurrent validity with the MCSDS = .71). In this study, reliability estimates were .72 for self-deception, .68 for impression management, and .76 for the total scale.

Feedback questionnaire. All participants received a feedback questionnaire following completion of the CAMI. Contained in this questionnaire were two items regarding participants' honesty in reporting: (a) "I purposely left out some information" and (b) "I answered all the questions as accurately and honestly as I could." These items were rated on a 5-point Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Procedure

Data for this study were collected as part of a multi-site project (including UNL, MU, and USC) examining relations between childhood maltreatment and adult adjustment. Internal Review Board (IRB) approval was obtained at all participating institutions. Participants were recruited from undergraduate psychology courses. In some cases, informed consent was obtained in person (at UNL and MU), and in other cases, it was obtained online via computer (USC). All measures were presented online. At UNL, the measures were completed onsite in a research laboratory. At the other two universities, the measures were completed offsite with instructions to complete the study in a private setting. The CAMI sub-scales were presented first, in a random order, followed by the feedback questionnaire; the remaining measures (BIDR, CTQ, MCSDS) were presented next, in a randomized order, along with other measures not used in the present study. Upon completion of the assessment, all participants were given a debriefing form and received course credit for their participation.

TABLE 1
CSA and CPA Victim Classification
on the CAMI and CTQ

CTQ	CAMI	
	Nonvictim	Victim
Child sexual abuse ($N = 1,195$)		
Nonvictim	1,021 (85.4%)	21 (1.8%)
Victim	69 (5.8%)	84 (7.0%)
Child physical abuse ($N = 1,195$)		
Nonvictim	786 (65.6%)	199 (16.8%)
Victim	38 (3.2%)	172 (14.4%)

NOTE: CSA = childhood sexual abuse; CPA = child physical abuse; CAMI = Computer Assisted Maltreatment Inventory; CTQ = Childhood Trauma Questionnaire.

RESULTS

Child Sexual Abuse

Prevalence of CSA as Assessed by the CAMI and CTQ

Participants were classified as being a CSA victim or nonvictim using the cutoff score of 6 recommended by the authors of the CTQ (Bernstein & Fink, 1998) and the behavioral criteria employed by the CAMI. Table 1 shows the number of individuals classified by each measure. Of the 1,195 individuals who responded to both CSA assessments, 1,021 (85.4%) were classified as nonvictims on both the CAMI and the CTQ, whereas 84 (7.0%) were identified as victims on both measures. Therefore, 92.4% were classified consistently across the two instruments. Cohen's Kappa for the relationship between the CAMI and the CTQ victim status was .61 ($p < .001$). Of the remaining 90 participants, 21 (1.8% of total sample) were classified as victims on the CAMI only and 69 (5.8% of total sample) were classified only by the CTQ. Of the total sample, the CAMI identified 105 individuals (8.8%) as victims of CSA, whereas the CTQ identified 153 (12.8%).

To better understand the nature of the inconsistencies between the CAMI and CTQ in the detection of CSA, several hypotheses were developed and explored. First, we examined associations between the objectively (i.e., behaviorally specific) and subjectively worded items on the CTQ and the objective abuse classifications on the CAMI under the assumption that items with similar formats would be more strongly related on the two measures. Second, response patterns of participants identified only by the CTQ and only by the CAMI as sexual abuse victims ($ns = 69$ and 21, respectively) were examined to shed light on discrepancies in reporting across the two measures. More specifically,

for CSA, hypotheses were tested that attributed inconsistent reporting to three major causes: low sensitivity (false negative reporting) on the CAMI, low sensitivity on the CTQ, or low specificity (false positive reporting) on the CTQ. Analyses exploring these possibilities are presented in the following sections. For each hypothesis, the percentage of discrepant cases reported is independent of other sources of inconsistency.

Comparisons Between CTQ Objective and Subjective Items and CAMI Classifications

As noted, it was expected that the behaviorally descriptive items on the CTQ would be more strongly associated with CAMI abuse classification than would be the CTQ subjective items. To test this, the objective and subjective items of the CSA subscale of the CTQ were each correlated with the dichotomous sexual abuse classification variable of the CAMI. For the objective CTQ items, correlations ranged from $r(1195) = .304$ to $r(1195) = .624$ ($M = .465$), whereas for the subjective CTQ items, correlations with the CAMI ranged from $r(1195) = .589$ to $r(1195) = .620$ ($M = .605$). All correlations were significant at the $p < .01$ level. When the objective and subjective items of the CTQ were summed to form subscales, correlations with the CAMI were $r(1195) = .539$ and $r(1195) = .628$, respectively ($p < .01$). Fisher's z results indicated that these correlations differed significantly ($z = 3.31$, $p < .01$), suggesting that the subjective items of the CTQ were more strongly correlated with the CAMI abuse classification variable for sexual abuse.

Low Sensitivity of the CAMI

Influence of social desirability. As noted earlier, aspects of social desirability such as impression management, self-deception, and a general need for approval may lead to particular response styles (e.g., withholding information) that could produce differential results for the two measures. In particular, we wondered whether the graphic nature of the behaviorally specific questions on the CAMI may be differentially related to social desirability, which could explain the lower reporting rate of CSA on this measure. To examine this possibility, mean social desirability scores on the BIDR (Impression Management, Self-Deception, and Total Scores) and the MCSDS were compared between the 21 participants classified as sexual abuse victims only by the CAMI (i.e., nonvictims on the CTQ) to the 69 respondents identified only by the CTQ (i.e., nonvictims on the CAMI). In this sample, the MCSDS was significantly correlated with the Impression Management BIDR subscore, $r(1195) = .13$, $p < .001$, Self-Deception BIDR subscore, $r(1195) = .12$, $p < .001$, and the Total BIDR score, $r(1195) = .14$, $p < .001$. Table 2

TABLE 2
Comparison of Mean Social Desirability Scores Between Participants Identified as Sexual and Physical Abuse Victims Only by the CAMI and Those Detected Only by the CTQ

	CAMI-Only Victims (CTQ Nonvictims)		CTQ-Only Victims (CAMI Nonvictims)		t Test		Effect Size
	M	SD	M	SD	t	p	d
Child sexual abuse							
BIDR Imp Mgt	4.14	3.17	3.54	2.78	-.85	.40	.21
BIDR Self-Dec	4.62	4.03	4.13	3.54	-.55	.59	.14
BIDR Total	8.76	6.39	7.67	5.32	-.79	.43	.20
MCSDS	14.86	6.59	13.58	4.20	-.84	.41	.26
Child physical abuse							
BIDR Imp Mgt	4.72	3.35	3.26	3.19	2.48	.01	.44
BIDR Self-Dec	5.07	3.77	3.71	2.97	2.11	.04	.37
BIDR Total	9.80	6.02	6.97	5.60	2.68	.01	.47
MCSDS	14.20	4.77	15.24	4.84	-1.22	.22	.22

NOTE: For child sexual abuse, CAMI-only victims (CTQ nonvictims), $n = 21$; CTQ-only victims (CAMI nonvictims), $n = 69$. For child physical abuse, CAMI-only victims (CTQ nonvictims), $n = 199$; CTQ-only victims (CAMI nonvictims), $n = 38$. CAMI = Computer Assisted Maltreatment Inventory; CTQ = Childhood Trauma Questionnaire; BIDR = Balanced Inventory of Desirable Responding; Imp Mgt = Impression Management; Self-Dec = Self-Deceptive Enhancement; MCSDS = Marlowe-Crowne Social Desirability Scale.

contains the results of the mean comparisons for these analyses. No significant differences were found for any of the social desirability subscales. Furthermore, effect sizes for these comparisons fell into the small range. Thus, social desirability appears to be an unlikely source of the differences in detection rates for sexual abuse between the CAMI and CTQ in this study.

Inaccurate reporting. A second hypothesis was that participants may have been less forthcoming about CSA history on the CAMI compared to the CTQ due to its more graphic nature. To explore this possibility, we examined responses to two questions contained on the feedback questionnaire administered after completion of the CAMI. These items were as follows: "I purposely left out information" and "I answered all questions accurately and honestly." Each item was recorded using a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). ANOVAs comparing the responses of the 21 participants identified as sexual abuse victims by the CAMI only to the 69 participants identified by the CTQ only were nonsignificant for both feedback questions, suggesting that individuals were similarly forthcoming on both measures. Examination of mean responses on these items revealed that neither the CAMI-only victims ($M = 1.57$, $SD = 1.12$) nor the CTQ-only victims ($M = 1.93$, $SD = 1.19$) reported purposely withholding information to any large degree. In addition, both groups ($M = 4.90$, $SD = 0.30$; $M = 4.72$, $SD = 0.54$; CAMI only and CTQ only, respectively) reported that they responded to items honestly. Thus, based on these self-report items, intentionally in-

accurate responding appears to be an unlikely source of the differences in detection rates for sexual abuse between the CAMI and CTQ in this study.

Underdetection of low-severity abuse by the CAMI. The final hypothesis examining sensitivity of the CAMI was that this measure failed to identify victims who were classified in the low to moderate category of sexual abuse on the CTQ. More specifically, as noted above, the CTQ includes three categories of child sexual abuse (low to moderate, moderate to severe, severe to extreme). Thus, analysis of the proportion of the 69 CTQ-only victims was conducted to determine the percentage of these individuals who were considered to have experienced a lower severity of abuse. This analysis indicated that 31 of these 69 participants (44.9%) were classified in the low-severity category on the CTQ, whereas 49 (31%) of the 154 total victims identified by the CTQ were classified as low severity. Chi-square analysis with these 154 individuals showed that the percentage of those in the low-severity category identified by the CTQ only was significantly greater than the entire sample identified by the CTQ, $\chi^2(1) = 6.23$, $p < .025$. Thus, a substantial proportion of the discrepancy between detection status of the CTQ and CAMI was due to the CAMI not identifying cases classified as low to moderate severity by the CTQ.

Low Specificity of the CTQ

Misclassification of adolescent sexual assault as child sexual abuse. Another potential explanation for the greater detection of CSA by the CTQ is that this

measure identified experiences that would better be classified as something other than childhood sexual abuse. To explore this possibility, the CAMI responses of the 69 sexual abuse victims who were identified by the CTQ only were examined. These responses were of interest because the CAMI contains behaviorally specific queries about all abuse-related sexual experiences prior to the age of 18 (CSA as well as adolescent sexual assault). Sixteen of the 69 child sexual abuse victims identified by the CTQ only (23%) reported experiences on the CAMI that occurred when victims were between the ages of 14 to 17, with perpetrators who ranged in age from 13 to 23 (a maximum age difference of 7 years). Thirteen of these 16 individuals (81.2%) described perpetrators who were friends or acquaintances, and several depicted clear date-rape experiences. Although these experiences may have been inappropriate or coercive, none of the 16 victims described experiences that met the common research definition of child sexual abuse employed by the CAMI. Rather, these activities were more consistent with common classifications of adolescent sexual assault (Koss & Cook, 1998; Marx, Van Wie, & Gross, 1996). Thus, a considerable proportion of the discrepancy in detection of victim status between the CAMI and CTQ (23%) may be attributable to the misclassification by the CTQ of peer sexual assault during mid- to late adolescence as childhood sexual abuse.

Misclassification of sexual exploration as child sexual abuse. A related hypothesis was that the CTQ detected activities that could be better described as voluntary sexual behaviors (i.e., sex play or exploration) between similarly aged peers or siblings. To examine this possibility, the CAMI responses of the 69 victims detected only by the CTQ were again examined. Four of these victims (6%) reported experiences on the CAMI that were described by participants as voluntary sex play. Although all four of these victims reported activities that occurred while younger than age 10, the largest age difference between themselves and a perpetrator was 2 years (victim age 9 and perpetrator age 11). In two of these cases, the victim and perpetrator were the same age. In addition, the majority of those labeled as perpetrators were identified as siblings (75%), and one was a friend. The most severe sexual act reported by these individuals was fondling; however, the most common act was kissing, and force was not reported in any of the responses. In the narrative section of the CAMI, one respondent described the activities as a result of curiosity and another described the behaviors as simulating activities seen on television, also the result of curiosity. All of the activities were reported to have been voluntary. Thus, a small proportion of the discrepancy (6%) between the CAMI and CTQ appear attrib-

utable to the misclassification of sexual exploration as child sexual abuse by the CTQ.

Classification of noncontact experiences as child sexual abuse. The final hypothesis exploring possible misclassification by the CTQ was the possibility that some CTQ-only victims had experienced only noncontact forms of abuse (e.g., exposure, witnessing sexual activity of others). To explore this hypothesis, the CAMI responses of the 69 victims detected only by the CTQ again were examined. Only one victim reported noncontact abuse in isolation (one-time witnessing of two male strangers: one who exposed himself and one who was masturbating in a public area). Thus, the reporting on noncontact events accounted for a very small proportion (1.4%) of the discrepancy between the CAMI and CTQ.

Low Sensitivity of the CTQ

Misclassification of child sexual abuse as no abuse. Although there were a number of CSA victims detected by the CTQ who were not identified by the CAMI ($n = 69$), there were 21 cases of CSA detected only by the CAMI. An exploration of the responses to the CSA screener questions of these individuals indicated that all met one or more screening criteria for sexual abuse (acts occurring against one's will, with an immediate family member, or with someone significantly older). The most common scenario, reported by 15 respondents (71.4%), involved reports of sexual activities against the participants' will with an adult or other individual meeting the "significantly older" age criteria. Examination of the abuse characteristics reported by these 21 participants on the CAMI indicated that the majority (95.2%) reported 1 perpetrator, whereas one respondent (4.8%) reported 3 perpetrators. Of the 23 total perpetrators reported, 9 (39.1%) were male strangers or nonfamily members; 7 (30.4%) were male friends, acquaintances, or neighbors; 2 (8.7%) were male babysitters or teachers; 2 (8.7%) were male relatives (uncle, grandfather); and 2 (8.7%) were female friends or cousins. Regarding nature of the abusive acts, one third (33%) experienced penetration (e.g., with a finger or foreign object) or intercourse as the most severe act committed. The remaining respondents indicated the most severe act included fondling (48%), kissing (14%), and oral sex (5%). Moreover, 38% reported that physical force or violence was used to obtain unwanted sexual acts, whereas a similar proportion (38%) reported that the sexual abuse involved the use of verbal coercion. Thus, although these participants were not identified as victims of sexual abuse by the CTQ, their responses on the CAMI clearly suggested that they met criteria for a history of childhood sexual abuse.

Differences on the CTQ and CAMI Due to Demographic Variables

A series of analyses were conducted to rule out the possibility that differences in the victim status on the CAMI versus the CTQ were related to demographic characteristics. Neither participant age, $F(1, 88) = 0.04$, $p = .84$, gender, $\chi^2(1) = 1.88$, $p = .17$, nor ethnicity, $\chi^2(3) = 0.71$, $p = .87$, were related to victim status on only the CAMI versus only the CTQ. Due to cell sizes, ethnicity analyses were limited to those who reported being European American, African American, Hispanic/Latino, or Asian American. Taken together, demographic variables do not appear to account for the discrepancy in victim detection status between the CAMI and the CTQ in this sample.

Child Physical Abuse

Prevalence of CPA as assessed by the CAMI and CTQ.

Participants were classified as being a CPA victim or nonvictim using the cutoff score of 8 recommended by the authors of the CTQ (Bernstein & Fink, 1998) and the behavioral criteria employed by the CAMI. Table 1 depicts the number of individuals classified by each measure. Of 1,195 total participants, 786 (65.8%) were classified as nonvictims on both the CAMI and the CTQ, whereas 172 (14.4%) were victims on both measures. Therefore, 80.2% were classified consistently on both measures. Cohen's Kappa for the relationship between the CAMI and the CTQ victim status was 0.47 ($p < .001$). Of the remaining 237 participants, 199 (16.7% of total sample) were classified as victims on the CAMI only and 38 (3.2% of total sample) were classified only by the CTQ. Of the total sample, the CAMI identified 371 individuals (31.0%) as victims, whereas the CTQ identified 210 (17.6%). Similar to the exploration of child sexual abuse, several sets of analyses were conducted to examine the inconsistencies noted above. More specifically, low sensitivity on either measure was evaluated as the source of inconsistent reporting. Again, for each hypothesis, the percentage of discrepant cases reported is independent of other sources of inconsistency.

Comparisons Between CTQ Objective and Subjective Items and CAMI Classifications

Similar to the CSA analyses, the objective and subjective items of the physical abuse subscale of the CTQ were correlated with the dichotomous physical abuse classification variable of the CAMI. For the objective CTQ items, correlations with the CAMI ranged from $r(1195) = .069$ to $r(1195) = .443$ ($M = .285$). All correla-

tions were significant at the $p < .01$ level, except $r = .069$, which was significant at the $p < .05$ level. The one item of the CTQ physical abuse scale that is considered subjective ("I believe I was physically abused") correlated with the CAMI at $r(1195) = .449$, $p < .01$. When the objective items of the CTQ were summed to form a subscale, its correlation with the CAMI was $r(1195) = .439$. A Fisher's z test indicated that this correlation was not significantly different from that obtained for the single objective CTQ item ($z = 0.30$).

Low Sensitivity of the CTQ

Influence of social desirability. Although our a priori assumption had been that the more graphic nature of the CAMI questions would elicit socially biased reporting, the many fewer CTQ-only victims ($n = 38$) in comparison to CAMI-only victims ($n = 199$) of physical abuse led us to hypothesize that the CTQ might actually be more susceptible to the influence of social desirability. If this were true, then we would expect that participants who went undetected by the CTQ (CAMI-only victims) would score higher than CTQ-only victims on the social desirability measures. To examine this possibility, mean social desirability scores on the BIDR (Impression Management, Self-Deception, and Total Scores) and the MCSDS were compared between these two groups. As Table 2 indicates, significant differences were found for the BIDR Impression Management and Self-Deceptive Enhancement subscales as well as the BIDR Total Score. As expected, those who were identified as victims of CPA on the CAMI but who failed to report such acts on the CTQ scored higher in the components of social desirability assessed by the BIDR. No such differences were found on the MCSDS.

Intentionally inaccurate reporting. A related hypothesis, again based on the lower detection rates of the CTQ, was that participants may have responded with less self-reported honesty and accuracy on this measure. To examine this possibility, the two feedback questions relating to the accuracy of reporting (i.e., "I purposely left out information" and "I answered all questions accurately and honestly") were compared for the 199 victims of physical abuse detected only by the CAMI to the 38 victims identified only by the CTQ using ANOVA. Results were nonsignificant for both feedback questions, suggesting that individuals were similarly forthcoming on both measures. Examination of mean responses on these items revealed that, overall, neither the CAMI-only victims ($M = 1.43$, $SD = 0.88$) nor the CTQ-only victims ($M = 1.41$, $SD = 0.73$) reported that they purposely withheld information. Furthermore, both groups ($M = 4.90$, $SD = 0.39$; $M = 4.86$, $SD = 0.47$; CAMI only and CTQ only, respectively) report-

ed that they responded to items honestly. Thus, intentionally inaccurate responding appears to be an unlikely source of the differences in detection rates for physical abuse between the CAMI and CTQ.

Abuse characteristics of victims identified only by the CAMI. The self-reported abuse characteristics of the 199 CAMI-only victims were examined to evaluate possible underreporting on the CTQ. Each of these respondents reported levels of severity and frequency of physically abusive acts that met criteria for child physical abuse employed by the CAMI. More specifically, 70 respondents (35.2%) reported experiencing at least one high-severity act (e.g., being punched, choked, thrown down, burned). Furthermore, 122 participants (61.3%) reported experiencing moderate-level severity acts on three or more occasions. These included being slapped on the face or head, receiving spankings that left a mark or bruise, or being hit on a part of the body other than the bottom with an object. Finally, 25 of the CAMI-only victims (12.6%) experienced low-level severity acts of being grabbed and shaken, spanked on the bottom with an object, or pinched hard by caregivers on more than 10 occasions. Thus, each of the 199 victims of child physical abuse identified only by the CAMI met criteria for physical abuse based on both the severity and frequency of acts reported.

To further explore discrepancies, the CTQ responses of the CAMI-only victims also were examined. Of these 199 individuals, 45 responded *rarely true* and 23 reported that it was *sometimes true* that they were “punished with a belt, board, or hard object”; 41 individuals responded *rarely true* and 5 reported that it was *sometimes true* that “people hit me so hard that it left bruises/marks”; 20 individuals responded *rarely true* and 2 reported that it was *sometimes true* that they believed that they were physically abused; 2 individuals responded *rarely true* that they were “hit so hard that [they] had to see a doctor or go to the hospital”; and no one reported being “beaten so badly that it was noticed by someone like a teacher, neighbor, or doctor.” Thus, although many respondents endorsed physical abuse items on the CTQ, these responses did not exceed the victim status cutoff score.

Low Sensitivity of the CAMI

Child physical abuse victims identified only by the CTQ. Even though detection rates of physical abuse were greater on the CAMI, 38 respondents were identified as victims of physical abuse by the CTQ only, suggesting that in some instances the CAMI may have underdetected abuse. Of these individuals, 1 participant responded that it was *often true* and 24 reported that it was *sometimes* or *rarely true* that they had been “hit so hard that it left a bruises or red marks.” In responding

to a question about “being punished with a belt, cord, or some other or hard object,” 22 reported that it was *rarely* or *sometimes true* and 12 reported that it was *often* or *very often true*. To a question asking if participants had been “hit so hard by someone in [their] family that [they] had to see a doctor or go to this hospital,” 12 reported that it was *rarely* or *sometimes true* and 5 reported that it was *often* or *very often true*. Three individuals reported that it was *true* or *very often true* and 11 reported that it was *sometimes* or *rarely true* that they had been “beaten so badly that it was noticed by someone like a teacher, neighbor, or doctor.” Finally, 17 reported that it was *rarely* or *sometimes true* and 6 reported that it was *often* or *very often true* that they believed they were physically abused.

The CAMI responses of these CTQ-only victims also were examined to assist in explaining the discrepancy in classification. Of these 38 individuals, 22 (57.9%) did not endorse any of the physical abuse screening questions included in the CAMI. The remaining 16 individuals reported levels of severity and frequency of physically abusive acts that were below the threshold of child physical abuse established by the CAMI. More specifically, 12 respondents (31.6%) reported experiencing low-level severity acts, including grabbing, shaking, and spanking that left a mark fewer than 10 times; 4 (10.5%) reported experiencing the low-severity act of pinching fewer than 5 times; and 1 (2.6%) participant reported experiencing the low-severity act of spanking on the bottom with an object fewer than 2 times. Furthermore, 13 respondents (34.2%) reported experiencing moderate-level severity acts (as defined by the CAMI physical abuse criteria), such as slapping on the face or head and hitting on a part of the body other than bottom with an object on one or two occasions. None of the 16 participants reported any of the acts considered severe in the CAMI criteria (e.g., punching, choking, threw down, burned). Thus, the majority (57.9%) of the CPA victims identified only by the CTQ endorsed no physical abuse on the CAMI, and the remaining (42.1%) reported experiences that did not meet the CAMI physical abuse criteria.

Differences on the CTQ and CAMI Due to Demographic Variables

A series of analyses was conducted to rule out the possibility that differences in the victim status on the CAMI versus the CTQ were related to demographic characteristics. Neither participant age, $F(1, 238) = 0.68, p = .41$, gender, $\chi^2(1) = 0.34, p = .56$, nor ethnicity, $\chi^2(3) = 4.48, p = .21$, were related to victim status on only the CAMI versus only the CTQ. Due to cell sizes, ethnicity was limited to those who reported being European American, African American, Hispanic/Latino, or Asian

American. Thus, demographic variables do not appear to account for the discrepancy in victim detection status between the CAMI and the CTQ in this sample.

DISCUSSION

This study compared retrospective reports of childhood sexual and physical abuse as assessed by two measures: the CTQ, which uses a Likert-type scaling approach, and the CAMI, which employs a behaviorally specific means of assessment. These instruments showed high percentage agreement in the classification of sexual abuse status (92.5%) and strong, although somewhat lower, consistency in the identification of physical abuse (80.2% agreement). These figures are slightly greater on average than the agreement rates found between the CTQ and an adolescent trauma questionnaire, the TEQ-A, in an analogous study with adolescent inpatients (Lipschitz et al., 1999). Despite the high rate of agreement in abuse status, the measures provided discrepant classifications in a number of cases. The CTQ, which applies cutoff scores to dimensional subscales, classified more participants as sexually abused than did the CAMI, which uses behaviorally specific screeners (12.8% vs. 8.8%, respectively). Conversely, the CAMI classified more participants as physically abused than did the CTQ (31.0% vs. 17.6%). By examining the responses of those identified as victims on one or the other measure, several possible explanations for these discrepancies were explored.

Prior comparisons between the CTQ and the TEQ-A suggested that the CTQ was more sensitive than its counterpart in detecting lower severity sexual abuse (Lipschitz et al., 1999). A similar pattern was found here. Of the 69 participants reporting sexual abuse on the CTQ only, 31 (44.9%) were classified as having experienced abuse that fell in the low-severity range. Thus, the CAMI may not be as sensitive as the CTQ in capturing less serious forms of CSA. In such cases, individuals may have experienced less severe yet abusive events that went undetected by the CAMI. One method for addressing this issue might be to increase the number and specificity of CAMI screening questions, which past research has found to increase the reporting of CSA (Hulme, 2004; Williams, Siegel, & Pomeroy, 2000). On the other hand, it is possible the CTQ produces false positive classifications of lower severity sexual abuse. Using the CTQ's cutoff score of 6, a participant need only to endorse a response of *rarely true* to a single item on the sexual abuse scale to be classified as a CSA victim. Although the present data are not definitive, this rather minimal threshold may overidentify cases of low-severity CSA, a possibility that has been raised by other authors as well (Lipschitz et al., 1999).

The current findings also highlight that the CTQ is not structured to distinguish between cases of child sexual abuse and incidents more accurately labeled date rape or adolescent sexual assault. Indeed, of the 69 CSA cases identified only by the CTQ, 16 (23%) reported experiences on the CAMI that were more consistent with accepted definitions of date or acquaintance rape (i.e., forced sexual activities occurring between teenage peers; Koss & Cook, 1998; Marx et al., 1996). Although some participants may have described entirely different encounters on each measure (e.g., CSA on the CTQ but date rape on the CAMI), this explanation is unlikely to account for all such discrepancies. The CTQ instructions, which ask participants to report quite generally about experiences "When I was growing up..." also may contribute to mislabeling by not allowing respondents to distinguish between childhood sexual abuse and other forms of sexual coercion. Such distinctions seem important, however, given the different etiologies and social contexts in which these experiences occur (Grauerholz & Koralewski, 1991; Russell & Bolen, 2000).

A related finding was that, contrary to expectations, the CTQ subjective items were more highly correlated than the objective items with CAMI sexual abuse classifications. One explanation for this finding may be that the behaviorally specific questions on the CTQ (e.g., "Someone tried to touch me in a sexual way or make me touch them") more often tapped into incidents that did not fit the CAMI criteria of sexual abuse (e.g., adolescent sexual assault, sex play, or very low-level experiences). In such instances, respondents who endorsed these items on the CTQ probably did not view themselves as having experienced childhood sexual abuse. If this is the case, then it is not surprising that the CAMI classifications were more strongly associated with experiences that respondents subjectively labeled as abusive rather than the objective CTQ items.

In addition to the 69 CTQ-only cases, there were also 21 individuals classified as victims of CSA only on the CAMI. The CAMI responses of these individuals suggested clear instances of sexual abuse, with more than 70% reporting that prior to age 18 they had experienced unwanted sexual activities with an adult or someone significantly older. One third of these participants reported penetration as a part of their abuse, and 72% indicated that verbal or physical coercion occurred during these interactions. Despite these rather clear instances of sexual abuse on the CAMI, all 21 of these individuals fell below the victim threshold on the CTQ.

In contrast to sexual abuse, the CAMI detected 199 cases of physical abuse that were not identified as such on the CTQ (a difference of 44%). All of these individuals disclosed incidents that met the CAMI's behavioral

criteria for CPA, with 35.2% disclosing the occurrence of seriously abusive events, such as being punched, choked, thrown down, or burned by a parent or caregiver. An examination of the physical abuse subscale of the CTQ may shed light on these apparently false negative classifications. Using the recommended cutoff score of 8, participants could report *often true* to items such as “I got hit so hard by someone in my family that I had to see a doctor or go to the hospital” or “I believe I was physically abused” yet still fall below the minimum CTQ criteria for having been physically abused. Indeed, many CAMI-only CPA victims endorsed physical abuse items on the CTQ but at a subthreshold level. In comparison, the multiple CAMI screeners, which cover a range of experiences, may have made this instrument more sensitive in detecting a range of CPA. Despite the apparent sensitivity, there were 38 participants identified as victims of physical abuse on the CTQ only. These cases fell into two categories: those who failed to endorse any CAMI screeners (58%) and those who reported harsh discipline by caregivers that fell below the CAMI abuse threshold (42%). Although the threshold for CPA used by the CAMI is conceptually consistent with prior work, some subjective judgment must inevitably be applied in determining the types and frequencies of acts constituting physical abuse. Thus, it is possible that some of the cases falling below the CAMI threshold used here could reasonably be defined as abusive.

In general, those who reported abuse on only one measure did not differ with respect to social desirability or a self-reported tendency to withhold abuse-related information. An exception to this trend was that CAMI-only CPA victims (i.e., those who failed to report CPA on the CTQ) had higher social desirability scores on the BIDR. This finding is somewhat inconsistent with Bernstein and Fink's (1998) report of low correlations between CPA and social desirability and suggests that impression management and self-deceptive enhancement may interfere with disclosure of physical abuse in a subset of individuals. Although it is unclear why such relations were found here, further work could explore these discrepant findings. Regarding the examination of demographic variables, results indicated that participants who reported abuse on only one measure did not differ from each other on any of the characteristics examined (age, gender, or ethnicity). These findings support the possibility that the differences in detection status between the CAMI and the CTQ are likely related to the approach of the two measures (i.e., behaviorally specific vs. scaled) rather than to demographic characteristics of the participants.

Several limitations of this study shed light on avenues for future research. First, although comparisons

across instruments can be informative, the present study did not evaluate the actual sensitivity or specificity of the abuse measures. It would be most revealing to evaluate responses to different measures in relation to documented reports obtained at the time of abuse. Only through this approach can definitive conclusions be drawn about the actual sensitivity and specificity of abuse measures. Comparison of different types self-report data to documented cases also could shed light on the reasons for certain discrepancies in classification that could not be explained here (e.g., why some reported CPA on the CTQ but not on the CAMI). Relatedly, future work is needed to understand the effects of different questionnaire formats on participant responses to abuse history measures. Controlled studies with random assignment to conditions are needed to examine disclosure rates and participant responses to different assessment tools using both behaviorally specific or scaled approaches. Because conceptualizations of abuse are to a large degree socially constructed, it also would be beneficial to examine whether agreement across measures varies as a function of definitions used. An additional avenue for future research might explore the possibility of participant suggestibility in the assessment of child maltreatment. Although less likely than underreporting in a research context, the possibility of false positive reporting increases in more applied forensic or clinical settings (Ornstein, Ceci, & Loftus, 1998). Finally, although the present sample was large and represented an ethnically and geographically diverse group of college students, future investigations will need to include a wider range of sample characteristics, including different age, income, and educational levels, to more broadly represent those who have experienced child maltreatment.

In considering the implications of this study, researchers should take into account their goals when employing retrospective assessments of abuse. When assessing sexual abuse history, for example, the low-end cutoff score of the CTQ may be best used as a screening device rather than for the strict classification of individuals into victim and nonvictim groups. As the authors of the measure acknowledge (Bernstein & Fink, 1998)—and the current findings support—CTQ cutoff scores were established to maximize sensitivity and, therefore, may yield some false positive classifications. On the other hand, the continuous scores produced by the CTQ sexual and physical abuse subscales offer researchers greater conceptual and statistical flexibility in measuring the gradations of severity associated with each form of abuse.

At a broader level, researchers also should consider whether a scaled or behaviorally specific approach best fits their theoretical assumptions. For instance, in-

investigators needing to know what actually happened may opt for behaviorally specific instruments, such as the CAMI, which are designed to assess the precise nature and frequency of abusive acts, the identity of perpetrators, types of coercion, and other abuse characteristics. Findings have repeatedly shown the importance of these factors by demonstrating their associations with short- and long-term developmental outcomes (e.g., Trickett, Reiffman, Horowitz, & Putnam, 1997). Of course, the value of descriptive approaches may be affected by the many factors that can distort the accurate recall and reporting of abuse and its specific characteristics. Although some authors have concluded that salient aspects of past abuse can be clearly reported (Brewin, Andrews, & Gotlib, 1993), additional work is needed to understand the accuracy with which more detailed aspects of abuse can be recalled. On the other hand, if knowing the specific features of abuse is not important, then the briefer scaled approach represented by the CTQ may be desirable. As Hulme (2004) notes, unlike descriptive measures, scaled instruments assess theoretical constructs and therefore require a shift from conceptualizing abuse as a set of observable behaviors to viewing it more in abstract theoretical terms. Hulme (2004) contends that using a psychometrically sound scale such as the CTQ allows researchers to avoid the response error pitfalls associated with descriptive measures. She states that “the shift to [viewing abuse as] a theoretical construct allows for the distortion that time, memory, emotions, denial, and other factors put on the events and their recall to remain part of the abstraction” (p. 211).

In summary, although comparable in some respects, the two instruments studied here each have advantages and shortcomings. Scaled measures can provide multiple abuse indices within a brief administration time. Although the present study suggests caution in using the CTQ subscales dichotomously except for screening purposes, this measure’s strong psychometrics, including its well-established factor structure, are a clear asset. On the other hand, behaviorally specific questionnaires provide an account of specific abuse events without requiring the somewhat contorted theoretical shift to viewing abuse as an abstract construct rather than a clearly defined set of behaviors. In addition, by assessing abuse characteristics, behaviorally specific measures such as the CAMI increase the number of possible research questions that can be asked. On the downside, the current findings indicate that both measures may under- or overidentify certain cases of maltreatment. For this reason, researchers may wish to consider a multi-method approach to the retrospective assessment of sexual and physical abuse. Specifically, looking for convergence across both measures, as well as at

the more severe or clear-cut cases mentioned on only one measure, may yield more comprehensive case identification than either method alone. In addition, investigators should consider conducting follow-up interviews in cases where victimization has been identified through only one source. Such interviews may shed light on reporting discrepancies and clarify whether abuse in fact occurred. Although more time-intensive, this multi-method approach could capitalize on the advantages of both instruments and is therefore a promising means of thoroughly assessing maltreatment history.

REFERENCES

- Abbey, A., BeShears, R., Clinton-Sherrod, A. M., & McAuslan, P. (2004). Similarities and differences in women’s sexual assault experiences based on tactics used by the perpetrator. *Psychology of Women Quarterly*, 28, 323-332.
- Beretvas, N. S., Meyers, J. L., & Leite, W. L. (2002). A reliability generalization study of the Marlowe-Crowne Social Desirability Scale. *Educational and Psychological Measurement*, 62, 570-589.
- Bernstein, D. P., Ahluvalia, T., Pogge, D., & Handelsman, L. (1997). Validity of the Childhood Trauma Questionnaire in an adolescent psychiatric population. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36, 340-348.
- Bernstein, D., & Fink, L. (1998). *Childhood Trauma Questionnaire: A retrospective self-report*. San Antonio, TX: The Psychological Corporation.
- Bernstein, D. P., Fink, L., Handelsman, L., & Foote, J. (1994). Initial reliability and validity of a new retrospective measure of child abuse and neglect. *American Journal of Psychiatry*, 151, 1132-1136.
- Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., et al. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse and Neglect*, 27, 169-190.
- Brewin, C. R., Andrews, B., & Gotlib, I. H. (1993). Psychopathology and early experience: A reappraisal of retrospective reports. *Psychological Bulletin*, 113, 82-98.
- Clemmons, J. C. (2005). Multiple forms of child maltreatment and abuse-specific characteristics: Relationships to psychological adjustment. *Dissertation Abstracts International: Section B. The Sciences & Engineering*, 65(8-B), 4278.
- Clemmons, J. C., Fortier, M., & DiLillo, D. (2003, November). *The impact of co-occurring forms of child maltreatment and abuse-specific characteristics on adult adjustment*. Poster presented at the 37th annual convention of the Association for Advancement of Behavior Therapy, Boston, MA.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- DeGue, S. A., Perry, A. R., DiLillo, D., Messman-Moore, T. L., & Fauchier, A. (2004, November). *Cruelty towards animals and family violence*. Poster presented at the 38th annual convention of the Association for the Advancement of Behavior Therapy, New Orleans, LA.

- DiLillo, D., DeGue, S., Kras, A., Di Loreto-Colgan, A., & Nash, C. L. (2006). Participant responses to retrospective surveys of child maltreatment: Does method of assessment matter? *Violence and Victims, 21*, 419-424.
- DiLillo, D., Di Loreto, A., & Perry, A. (2003, October). *Childhood exposure to abusive parenting: Long-term implications for marital functioning*. Paper presented at the Jacobs Foundation Conference titled "In What Ways Do Families Matter In Young People's Development?" Marbach Castle, Lake Constance, Germany.
- DiLillo, D., Hayes, S., & Hope, D. A. (2006). Sensitivity of the Anxiety Disorder Interview Schedule-IV (ADIS-IV) in detecting potentially traumatic childhood abuse. *Journal of Psychopathology and Behavioral Assessment, 20*, 131-135.
- Dutton, D. G., & Hemphill, K. J. (1992). Patterns of socially desirable responding among perpetrators and victims of wife assault. *Violence & Victims, 7*, 29-40.
- Fortier, M. A., Perry, A. R., DiLillo, D., Messman-Moore, T. L., Tremblay, G. C., & Fauchier, A. (2004, November). *Coping with child sexual abuse: Does gender matter?* Poster presented at the 38th annual convention of the Association for the Advancement of Behavior Therapy, New Orleans, LA.
- Franklin, A. L., Sheeran, T., & Zimmerman, M. (2002). Screening for trauma histories, Posttraumatic Stress Disorder (PTSD), and subthreshold PTSD in psychiatric outpatients. *Psychological Assessment, 14*, 467-472.
- Friedrich, W. N., Talley, N. J., Panser, L., Fett, S., & Zinsmeister, A. R. (1997). Concordance of reports of childhood abuse by adults. *Child Maltreatment, 2*(2), 164-171.
- Fricker, A. E., Smith, D. W., Davis, J. L., & Hanson, R. F. (2003). Effects of context and question type on endorsement of childhood sexual abuse. *Journal of Traumatic Stress, 16*, 265-268.
- Goodman, G. S., Ghetti, S., Quas, J. A., Edelstein, R. S., Alexander, K. W., Redlich, A. D., et al. (2003). A prospective study of memory for child sexual abuse: New findings relevant to the repressed/lost memory controversy. *Psychological Science, 14*, 113-118.
- Grauerholz, E., & Koralewski, M. A. (Eds.). (1991). *Sexual coercion: A sourcebook on its nature, causes, and prevention*. Lexington, MA: Lexington Books.
- Hansen, D. J., Sedlar, G., & Warner-Rogers, J. E. (1999). Child physical abuse. In R. T. Ammerman & M. Hersen (Eds.), *Assessment of family violence* (2nd ed., pp. 127-156). New York: John Wiley.
- Hsiao, J. K., Bartko, J. J., & Potter, W. Z. (1989). Diagnosing diagnoses: Receiver operating characteristic methods and psychiatry. *Archives of General Psychiatry, 46*, 664-667.
- Hulme, P. A. (2004). Retrospective measurement of childhood sexual abuse: A review of instruments. *Child Maltreatment, 9*, 201-217.
- Kolko, D. J. (2002). Child physical abuse. In J. E. Myers, L. Berliner, J. Briere, T. C. Hendrix, C. Jenny, & T. A. Reid (Eds.), *The APSAC handbook on child maltreatment* (2nd ed., pp. 21-54). Thousand Oaks, CA: Sage.
- Koss, M. P., & Cook, S. L. (1998). Facing the facts: Date and acquaintance rape are significant problems for women. In R. K. Bergen (Ed.), *Issues in intimate violence* (pp. 147-156). Thousand Oaks, CA: Sage.
- Latkin, C. A., Vlahov, D., & Anthony, J. C. (1993). Socially desirable responding and self-reported HIV infection risk behavior among intravenous drug users. *Addiction, 88*, 517-526.
- Leite, W. L., & Beretvas, S. N. (2005). Validation of scores on the Marlowe-Crowne Social Desirability Scale and the Balanced Inventory of Desirable Responding. *Educational and Psychological Measurement, 65*(1), 140-154.
- Lesserman, J., Drossman, D. A., & Zhiming, L. (1995). The reliability and validity of a sexual and physical abuse history questionnaire in female patients with gastrointestinal disorders. *Behavioral Medicine, 21*, 141-150.
- Lipschitz, D. S., Bernstein, D. P., Winegar, R. K., & Southwick, S. M. (1999). Hospitalized adolescents' reports of sexual and physical abuse: A comparison of two self-report measures. *Journal of Traumatic Stress, 12*, 641-654.
- Malinosky-Rummell, R., & Hansen, D. J. (1993). Long-term consequences of childhood physical abuse. *Psychological Bulletin, 114*, 68-79.
- Marx, B. P., Van Wie, V., & Gross, A. M. (1996). Date rape risk factors: A review and methodological critique of the literature. *Aggression & Violent Behavior, 1*, 27-45.
- Messman-Moore, T. L., DeNardi, K. A., Gaffey, K. J., Fauchier, A., & DiLillo, D. (2004, November). *Predicting psychological distress and revictimization among childhood sexual abuse and adolescent sexual assault survivors: The impact of various coping strategies*. Poster presented at the 38th annual convention of the Association for the Advancement of Behavior Therapy, New Orleans, LA.
- Meyer, I. H., Muenzenmaier, K., Cancienne, J., & Struening, E. (1996). Reliability and validity of a measure of sexual and physical abuse histories among women with serious mental illness. *Child Abuse and Neglect, 20*, 213-219.
- Nash, C. L., DiLillo, D., Messman-Moore, T. L., & Rinkol, S. (2002, November). *The Computer Assisted Maltreatment Inventory (CAMI): An investigation of the test-retest reliability and criterion validity*. Poster presented at the meeting of the Association for the Advancement of Behavior Therapy, Reno, NV.
- Ornstein, P. A., Ceci, S. J., & Loftus, E. F. (1998). Adult recollections of childhood abuse: Cognitive and developmental perspectives. *Psychology, Public Policy, and Law, 4*, 1025-1051.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes*. San Diego, CA: Academic Press.
- Paulhus, D. L. (1994). *Balanced Inventory of Desirable Responding: Reference manual for BIDR Version 6*. Unpublished manuscript, University of British Columbia, Vancouver, Canada.
- Polusny, M. M., & Follete, V. M. (1995). Long-term correlates of childhood sexual abuse: Theory and empirical findings. *Applied and Preventive Psychology, 4*, 143-166.
- 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.
- Russell, D. E. H., & Bolen, R. M. (2000). *The epidemic of rape and child sexual abuse in the United States*. Thousand Oaks, CA: Sage.
- Straus, M. A., Hamby, S. L., Finkelhor, D., Moore, D. W., & Runyan, D. (1998). Identification of child maltreatment with the parent-child conflict tactics scales: Development and psychometric data

- for a national sample of American parents. *Child Abuse and Neglect*, 22, 249-270.
- Straus, M. A., & Stewart, J. H. (1999). Corporal punishment by American parents: National data on prevalence, chronicity, severity, and duration, in relation to child and family characteristics. *Clinical Child & Family Psychology Review*, 2, 55-70.
- Trickett, P. K., Reiffman, A., Horowitz, L. A., & Putnam, F. W. (1997). Characteristics of sexual abuse trauma and the prediction of developmental outcomes. In D. Cicchetti & S. Toth (Eds.), *Rochester symposium on developmental psychopathology: Developmental perspectives on trauma. Theory, research, and intervention* (pp. 289-314). Rochester, NY: University of Rochester Press.
- Weaver, T. L. (1998). Method variance and sensitivity screening for traumatic stressors. *Journal of Traumatic Stress*, 11, 181-185.
- Whipple, E. E., & Richey, C. A. (1997). Crossing the line from physical discipline to child abuse: How much is too much? *Child Abuse & Neglect*, 21, 431-444.
- Widom, C. S., & Morris, S. (1997). Accuracy of adult recollections of childhood victimization: Part 2. Childhood sexual abuse. *Psychological Assessment*, 9, 34-46.
- Widom, C. S., & Shepard, R. L. (1996). Accuracy of adult recollections of childhood victimization: Part 1. Childhood physical abuse. *Psychological Assessment*, 8, 412-421.
- Williams, L. M. (1994). Recall of childhood trauma: A prospective study of women's memories of child sexual abuse. *Journal of Consulting and Clinical Psychology*, 62, 1167-1176.
- Williams, L., & Banyard, V. (1998). *Trauma and memory*. Newbury Park, CA: Sage.
- Williams, L. M., Siegel, J. A., & Pomeroy, J. J. (2000). Validity of women's self-reports of documented child sexual abuse. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *Science of self-report: Implications for research and practice* (pp. 211-226). Mahwah, NJ: Lawrence Erlbaum.
- Winegar, R. K., & Lipschitz, D. S. (1997). *The reliability of adolescents' reports of maltreatment and a comparison of inquiry formats*. Paper presented at the 5th International Family Violence Research Conference, Durham, NH.
- Wolfe, D. A. (1988). Child abuse and neglect. In E. J. Mash & L. G. Terdal (Eds.), *Behavioral assessment of childhood disorders* (2nd ed., pp. 626-666). New York: Guilford.
- Wyatt, G. E., & Peters, S. D. (1986). Methodological considerations in research on the prevalence of child sexual abuse. *Child Abuse & Neglect*, 10, 241-251.
- David DiLillo**, PhD, is an associate professor of clinical psychology at the University of Nebraska–Lincoln. His research interests lie in the areas of family violence and marital and couple relations. He is in the midst of a National Institutes of Mental Health–funded project exploring associations between childhood abuse history and adult marital and parent-child functioning. He also maintains an interest in methodological issues associated with the retrospective assessment of child maltreatment.
- Michelle A. Fortier** received her PhD in clinical psychology from the University of Nebraska–Lincoln in 2006. She is currently a postdoctoral fellow in pediatric psychology at the Mayo Clinic in Rochester Minnesota. Her research interests include coping processes associated with childhood maltreatment and chronic illness.
- Sarah A. Hayes** is a doctoral candidate in clinical psychology at the University of Nebraska-Lincoln. She is currently completing her predoctoral internship at the University of Washington School of Medicine in Seattle, Washington. Her research interests include the treatment of anxiety disorders and the consequences of childhood maltreatment.
- Emily Trask** is a doctoral student in clinical psychology at the University of Lincoln–Nebraska. She received her master's degree in psychology at San Diego State University in 2004. She has research and clinical interests in the treatment of childhood sexual abuse.
- Andrea R. Perry** received a bachelor of science degree in psychology from Middle Tennessee State University in 2002. She is currently a doctoral candidate in the Clinical Psychology Training Program at the University of Nebraska–Lincoln. Her research and clinical interests include childhood maltreatment, adult romantic relationships, and processes that mediate these constructs.
- Terri Messman-Moore** received her PhD from Oklahoma State University and is currently an associate professor of clinical psychology at Miami University. Her research interests include the impact of childhood maltreatment and sexual violence on interpersonal functioning and health-risk behavior among young women.
- Angèle Fauchier** completed her doctorate in clinical psychology at the University of Southern California. She is currently a postdoctoral fellow in the Family Research Laboratory, University of New Hampshire. Her research examines the interplay between family violence and family relationships in different dyads.
- Cindy Nash**, PhD, is currently the psychologist for the Lincoln Regional Center Sex Offender Program. Her research interests include looking at the relationship between perpetrators who were sexually abused as children, their motivation to change, distorted thinking patterns, and their victim characteristics. She also is involved in the development of a program for slow and nonresponders and program evaluation.