Journal of Traumatic Stress, Vol. 16, No. 4, August 2003, pp. 399–409 (© 2003)

Consistency of Retrospective Reporting About Exposure to Traumatic Events[1]

Karen E. Krinsley,[2] James G. Gallagher,[3,4] Frank W. Weathers,[3,5] Catherine J. Kutter,[3] and Danny G. Kaloupek[2,6]

Lifetime exposure to traumatic events was assessed by means of a multimethod protocol applied to 76 male military veterans. Consistency of retrospective reporting was determined for physical and sexual assault and abuse, accidents, disasters, combat and warzone experiences, serious illness or injury, and hazardous duty. Findings demonstrate that respondents are generally consistent in reporting traumatic events, although the majority report more events upon reevaluation. Reporting about traumatic events shows some variation as a function of the life epoch in which events occurred, whether they were directly or indirectly experienced, and the type of trauma involved. Discussion addresses memory-related processes triggered by trauma evaluation or tied to characteristics of events themselves as potential sources of inconsistency.

KEY WORDS: trauma assessment; consistency of reporting; Evaluation of Lifetime Stressors; traumatic event characteristics; memory for traumatic events.

Research increasingly demonstrates that exposure to psychologically traumatic events is widespread (Breslau et al., 1998; Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995; Resnick, Kilpatrick, Dansky, Saunders, & Best, 1993) and potentially detrimental to mental and physical health, as well as to economic welfare (Solomon & Davidson, 1997). Among the prominent sequelae of trauma is posttraumatic stress disorder (PTSD), an often severe and disabling condition with a lifetime prevalence rate of approximately 8% in the United States (Kessler et al., 1995), making it one of the most common psychological disorders. Sectors of the population with elevated trauma exposure, such as military personnel in warzones or women exposed to interpersonal violence, have higher than average rates of PTSD, ranging from 35 to 69% (Kulka et al., 1990; Norris, 1992; Resnick et al., 1993). Compared to people without a trauma history, trauma survivors are at higher risk for multiple problems including depression, substance abuse, suicidal and homicidal behavior, physical illness, and poverty (e.g., Burnam et al., 1988; Byrne, Resnick, Kilpatrick, Best, & Saunders, 1999; Keane & Wolfe, 1990; Kessler et al., 1995; Schnurr, Spiro, Aldwin, & Stukel, 1998). Thus, both the pervasiveness and consequences of exposure indicate the importance of securing a comprehensive trauma history from individuals in healthcare settings.

The assessment literature has documented a set of obstacles to the collection of valid trauma histories.

[1]Portions of this paper have been presented at the Second World Conference of the International Society for Traumatic Stress Studies in 1996 and at the Annual Meeting of the International Society for Traumatic Stress Studies in 1998.

[2]VA Boston Healthcare System and Department of Psychiatry, National Center for PTSD, Boston University School of Medicine, Boston, Massachusetts.

[3]VA Boston Healthcare System, National Center for PTSD, Boston, Massachusetts.

[4]Now in private practice, Denver, Colorado.

[5]Department of Psychology, Auburn University, Auburn, Alabama.

[6]To whom correspondence should be addressed at VA Boston Healthcare System, National Center for PTSD (116B-2), 150 S. Huntington Avenue, Boston, Massachusetts 02130; e-mail: danny.kaloupek@med.va.gov.

0894-9867/03/0800-0399/1 © 2003 International Society for Traumatic Stress Studies

These challenges include (a) labeling and self-definition issues (e.g., a woman not calling a forced sexual experience "rape" because she knew her assailant); (b) reporting impediments tied to stigma and shame (Della Femina, Yeager, & Lewis, 1990; Kilpatrick, 1983; Mollica & Caspi-Yavin, 1991); and (c) memory deficits ranging from incomplete recall to psychogenic amnesia (Briere & Conte, 1993; Green & Grace, 1988). There is evidence of inconsistent reporting, in particular a trend toward increased reporting of traumatic events over repeated assessments (Arnow et al., 1999; Goodman, Corcoran, Turner, Yuan, & Green, 1998; King et al., 2000; Leserman, Drossman, & Li, 1995; Martin, Anderson, Romans, Mullen, & O'Shea, 1993; Roemer, Litz, Orsillo, Ehlich, & Friedman, 1998; Southwick, Morgan, Nicolaou, & Charney, 1997; Wyshak, 1994). Explanations for such increases have focused on both underreporting at initial assessment (Leserman et al., 1995) and the reporting of additional low magnitude stressor events at later assessments (Goodman et al., 1998; Martin et al., 1993). Two studies (Roemer et al., 1998; Southwick et al., 1997) provide evidence that increased reporting of traumatic events by combat veterans is paralleled by increased reporting of PTSD symptoms. This association raises questions about the potential impact of ambient emotional states on recall and reporting of trauma. King et al. (2000) have addressed the issue with complex time-based statistical techniques and concluded that event reporting is relatively stable, although there is a slight tendency toward enhanced remembering tied to increases in PTSD symptom severity. This effect may resemble state-dependent memory retrieval (Bower, 1981).

The primary purpose of the current investigation was to test three hypotheses concerning the potential impact of event characteristics on consistency of reporting: First, we predicted that change in the number of traumas reported from a first to a second interview would be predominantly in the direction of more traumas reported at the second interview. This prediction is consistent with existing evidence and with the theoretical expectation that initial assessment provides cues that can stimulate memory and aid recall. Second, we predicted that consistency of reporting would be inversely related to the specificity and detail used to describe the traumatic events. This prediction reflects the fact that consistency may become more difficult to achieve as the measurement target becomes more complex or more precise. Third, we predicted that consistency of reporting would be greatest for the most clearly traumatic events. This prediction is based on laboratory evidence that emotional arousal may enhance memory for stressful events (e.g., Cahill & McGaugh, 1995). The vehicle for collecting information about lifetime trauma exposure was a multimethod protocol called the Evaluation of Lifetime Stressors (ELS).[7]

Method

Instrument Development

Content Identification

Initial content for the ELS was drawn from extensive literature reviews and consultation with experts on both clinical and research aspects of trauma exposure across a variety of populations. These consultants provided guidance regarding the scope, importance, and order of items. Feedback on content also was solicited from trauma-exposed individuals who were evaluated during pilot testing of the instrument.

The ELS attempts to be comprehensive with respect to (a) range of trauma types; (b) lifespan coverage; and (c) critical dimensions of traumatic events, such as severity, perceived life threat, helplessness, and fear. It addresses both objective and subjective dimensions of trauma events as defined by the two-part PTSD diagnostic Criterion A outlined in the *DSM-IV* (Diagnostic and statistical manual of mental disorders, 4th ed.; American Psychiatric Association, 1994). Events meeting only Criterion A1 are characterized as potentially traumatic while those meeting both Criteria A1 and A2 are characterized as traumatic.

Self-report and interview formats each have strengths when it comes to eliciting trauma-related information (e.g., Dill, Chu, Grob & Eisen, 1991; Koss & Gidcyz, 1985; Leserman et al., 1995; Martin et al., 1993; Steele, Henderson & Duncan-Jones, 1980; Stinson & Hendrick, 1992). Accordingly, the ELS uses a multimethod format consisting of an initial self-report questionnaire (ELS-Q), followed by a semistructured interview (ELS-I). This multimethod format allows varied but overlapping opportunities for reporting.

The ELS-Q is a screening questionnaire that addresses behavioral correlates of trauma exposure such as problems with trust, as well as a wide variety of specific potentially traumatic events (e.g., physical and sexual abuse and assault, disasters, warzone exposure, serious illnesses, and accidents), similar to other comprehensive trauma measures (e.g., Goodman et al., 1998; Norris, 1990). It contains a total of 53 questions and takes 10–20 min to complete. Key features of the ELS-Q are (a) open response options that address situations not easily

[7]The complete ELS and training manual are available from D. G. Kaloupek.

categorized in terms of the primary responses of yes or no (i.e., "I'm not sure" and "It happened to someone I knew"); (b) separate endorsement for childhood and adulthood experiences; (c) hierarchical arrangement of questions that progresses from least to most emotionally evocative; and (d) item phrasing that ranges from general (e.g., "Growing up, were there people you did not trust or did not feel safe with?") to explicit (e.g., "Have you ever been burned on purpose, with a cigarette, with hot water, or with something else?"). Items on the ELS-Q do not have a one-to-one correspondence with events reported through the ELS-I; a single item on the ELS-Q may lead to probing for several different events on the interview.

ELS-I administration takes one to three hours and involves interviewer queries of all nonnegative responses on the ELS-Q. For example, if a participant endorsed "not sure" for a question on the ELS-Q, the ELS-I interviewer would ask him to expand on this answer. Specific probe questions are provided for this follow-up, and the interviewer is prompted to ask about important aspects of the event (e.g., identification of the perpetrator, whether a weapon was used) if the respondent has not revealed this information spontaneously.

The ELS-I provides a structure for summarizing all potentially traumatic events, categorizing relatively objective characteristics of each event (the perpetrator, frequency of the event, whether a weapon was used, if physical injury occurred, etc.), as well as noting presence or absence of the subjective emotional responses of fear, helplessness or horror. At the end, interviewees are asked to select their worst traumatic event in childhood and in adulthood, guided by the interviewer as necessary.

Present Study

Participants

The study enrolled 86 male Vietnam-era veterans who served in the military between August 1964 and May 1973. Both combat and noncombat veterans were recruited from a Veterans Affairs Medical Center and from a large New England community via newspaper advertisements and flyers. The sample was limited to Vietnam-era veterans to reduce the potential for demographic confounds related to age and military experiences. It included both treatment-seeking and community-dwelling individuals in an effort to increase dispersion with respect to both trauma exposure and health status. Potential participants were excluded if they were actively psychotic, suicidal, homicidal, or unable to refrain from substance use for 24 hr prior to and during the study. A written consent form informed participants that the purpose of the study was to develop more accurate and reliable psychological evaluations for lifetime stressors. It assured them that all information provided for the study would be kept confidential and would not be recorded in their medical record or shared with any provider without a signed release, unless there was imminent risk of harm to self or others (i.e., the standard exception to confidentiality). Participants who completed the study were reimbursed $150 for their time and expenses.

Table 1. Participant Demographics ($N = 76$)

Variable	n	Percentage of respondents
Race		
Caucasian	67	88.2
African American	6	7.9
Native American	1	1.3
Other	2	2.6
Married/living with a partner	38	50.0
Served in a warzone	57	75.0
Branch of service		
Army	39	51.3
Marines	16	21.1
Navy	11	14.5
Air Force	10	13.2
Currently employed	31	41.3
Income below $20,000	37	48.7
Religious preference		
Roman Catholic	33	43.4
Protestant	19	25.0
Jewish	2	2.6
Other	7	9.2
No religion	15	19.7

Ten participants did not complete the protocol, failing to return for either the second or third session. The 76 participants in the final sample averaged 49 years of age (range = 37–65 years; $SD = 5.4$) and 13 years of education (range = 10–15 years; $SD = 0.9$). Additional demographic characteristics shown in Table 1 indicate a largely Caucasian sample comprised of individuals with considerable warzone military service, but relatively low current income and employment rates.

Procedure and Design

All participants received two independent administrations of the full ELS protocol. Interviewers were randomly assigned to an equal number of participants, and their assignment to first or second ELS-I administrations was counterbalanced. There was a minimum of two and a maximum of seven days between each administration of the ELS-I. In addition, participants were randomly assigned to one of two administration conditions to address the potential reactive effects of undergoing trauma

assessment: ELS-Q, ELS-I, ELS-Q, ELS-I was the standard order and ELS-Q, ELS-Q, ELS-I, ELS-I was the nonstandard order for which both interviews were referenced to the second questionnaire administration.

Measures

Participants completed three measures of PTSD symptoms, the Clinician- Administered PTSD Scale, a psychometrically robust interview (CAPS; Blake et al., 1995; Weathers, Keane, & Davidson, 2001); the Mississippi Scale for PTSD, a widely used self-report measure (M-PTSD; Kulka et al., 1990; Vreven, Gudanowski, King, & King, 1995); and the PTSD Checklist, a self-report version of *DSM-IV* PTSD symptoms (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993). This set of measures was selected to provide both diagnostic and psychometric indicators of PTSD and, because the latter two instruments were administered in their general civilian versions, all were suitable for measuring the impact of both combat and noncombat trauma. Additional measures were administered to obtain information about other diagnoses and symptoms, intelligence, and family functioning, but these are not reported because they do not address the aims of this study.

Interviewer Training

Interviews were conducted by seven advanced masters and doctoral level psychologists, all of whom had clinical experience with traumatized individuals. Training for the ELS was extensive and ongoing, with all interviewers trained to a consensus standard developed by the first and second authors. The manualized procedures for ELS administration and coding were explained and discussed in a series of training meetings incorporating lecture, role-play, and video formats, followed by written tests. After training, all new interviewers observed at least two interviews and were observed for a minimum of two interviews themselves. Both interviewers and observers completed and independently coded an ELS-I form during these sessions, after which the forms were exhaustively compared and discussed by the two people involved. In addition, each training interview was reviewed in a group meeting so that issues could be identified and resolved collectively. Weekly meetings of all interviewers occurred throughout the study to address rating drift, review the coding manual, and generally ensure that the consensus standard was maintained.

Training for CAPS interviewing followed the standard protocol used at the National Center for PTSD including videotape, role-play, and direct observation. The training was supplemented by direct observation by senior clinicians who commented on interview style and checked coding decisions. Coding questions raised by interviewers were resolved by consensus reached through group discussion at weekly interviewer meetings.

Event Coding

For this study, all potentially traumatic events were categorized into the event classifications, Criterion A status, exposure type, and epoch shown in Fig. 1. The coding manual provided a brief definition for each event category, followed by multiple examples of matches and exceptions for each. Summary definitions for the 10 event categories are (1) physical abuse or assault is a physical attack or fight involving a perpetrator; (2) sexual abuse or assault is unwanted or inappropriate sexual contact, sexual exploitation, or sexual coercion; (3) disasters are natural and human-created events that occur in public places and have the potential to adversely affect groups of people; (4) hazardous duty is performance of an occupation that routinely involves threat to one's life or physical well-being; (5) warzone exposure applies to events that occur as part of military combat, including injuries received; (6) serious illness or injury applies to potentially serious or life-threatening events with medical consequences, including witnessing the aftermath of such an event; (7) accidents are potentially life-threatening events not attributable to an identifiable perpetrator, including vehicular, work-related, and household types; (8) intentional perpetration is direct participation in purposeful sexual or physical violence toward another person; (9) unintentional perpetration is direct participation in accidental physical violence toward another person; and (10) captivity or torture is the systematic use of physical deprivation, medical procedures, or imprisonment for punishment or coercion. Criterion A status followed the *DSM-IV* definition. The distinction between experienced and witnessed events was based on whether the individual played a direct, active role or indirect, passive role during an event, and whether or not the event represented an immediate threat to their physical safety. Childhood was defined as the period prior to age 18, with the exception that events during military service were always assigned to adulthood.

Recall of Worst Traumas

Participants identified their worst childhood and adulthood traumas at the end of each ELS-I administration. Interviewers guided participants by reading them a

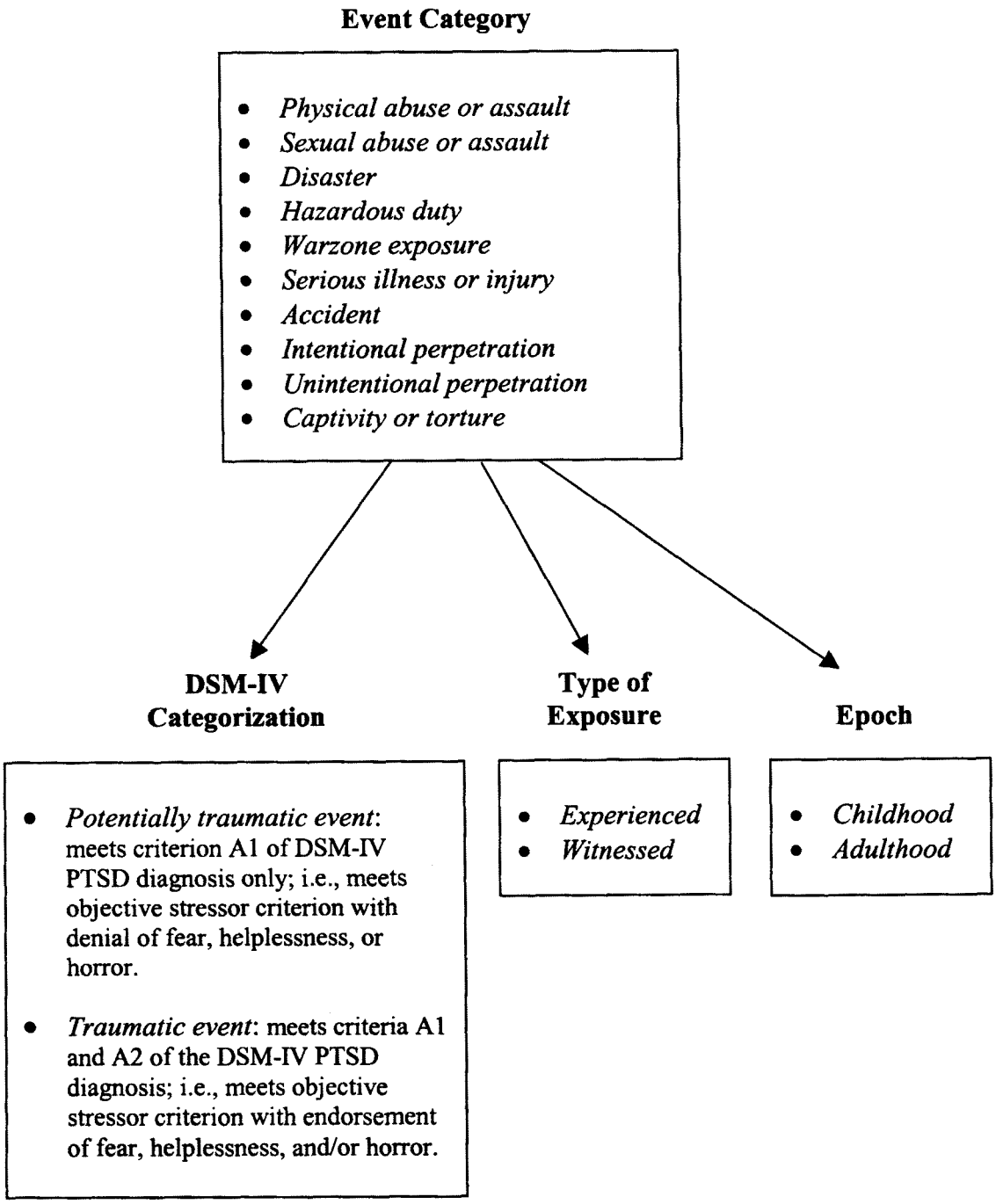

Fig. 1. Four basic category dimensions used for trauma event coding with the Evaluation of Lifetime Stressors. For the dimension of exposure type, the *DSM-IV* categories of "confronted with" and "threatened" also were applied but were not subjected to separate analysis in this investigation because of low frequency.

list of all the potentially traumatic events they had reported. Worst events selected in the first interview were listed and sealed in an envelope that was opened when the second interview concluded. The interviewer presented events identified during the first interview and asked whether they were the same as or different from those identified during the second interview. If different, the participant was asked why he had changed his report. These

answers were recorded verbatim and assigned to descriptive categories developed for the study.

Collateral Confirmation

An attempt was made to obtain corroborating information for each participant's worst childhood and adulthood traumas reported in the second interview. Skepticism about retrospective trauma reporting makes it necessary to attempt such external validation (Brewin, Andrews & Gotlib, 1993; Sutker, Uddo-Crane & Allain, 1991), despite numerous obstacles to obtaining the information and the fact that some events are inherently difficult or impossible to confirm because the only witness may be the perpetrator. The aim was to obtain independent information that an event took place, not to verify details.

At the end of the second interview, each participant was asked to name any people or institutions (e.g., court, hospital, school, military, police records) that they thought might be able and willing to corroborate their worst traumatic events. Participation in this procedure was optional for reasons of subject protection. We did not attempt to confirm an event if there were no known witnesses or records unless the participant had talked about it with a confidante who might confirm the disclosure. If an individual person was identified as the corroborating source, the participant was given the opportunity to contact them first, before contact by a research team member. Whether or not the participant chose to make initial contact, a member of the research team eventually attempted phone contact with the identified person. During this call, the reason for contact was explained, including the $15.00 payment that would be provided in exchange for speaking with the staff member, whether or not information was provided. If the collateral agreed, they were then asked for any recollection of the target event. Questioning was open-ended and nondirective, and any statements relevant to the index event were recorded and later reviewed by the first author. Releases were requested from participants and records were obtained when the potential corroborating source was an institution of some sort. Irrespective of source, the information obtained had to clearly match the reported trauma in order to be coded as an instance of positive corroboration.

Analytic Strategy

Initial descriptive analysis examined the prevalence of each event category. We then examined associations between trauma exposure and measures of PTSD in order to determine whether potentially traumatic events identified by the ELS demonstrate a moderate positive relationship with PTSD symptoms similar to that obtained with other trauma measures. The third step involved comparison of event frequency within individuals across interviews. This was followed by the primary analysis that examined consistency of reporting on the ELS-I by means of percent agreement and standard kappa statistics (Cohen, 1960). Consistency was examined relative to event category, Criterion A categorization, and exposure type, with childhood and adulthood events examined separately under each classification. In addition, consistency of reporting for the worst traumas from first to second interview was examined in terms of percent agreement. Finally, we examined the degree to which participants' worst traumas were corroborated by collateral data sources.

Results

Event Prevalence

Table 2 lists percentages of participants who reported being exposed to each of the 10 classes of potentially traumatic events. Most frequent endorsements were for physical abuse and accidents in childhood, and for serious illness or injury, warzone exposure, and physical assault in adulthood. Of these, only warzone exposure is uniquely tied to the military background of the sample.

Relationship With PTSD Symptoms

The trauma exposure index was operationally defined as the total number of reported events that met at least *DSM-IV* Criterion A1. Test–retest reliability for this score across interviews was high, $r(75) = .83$, $p < .001$; therefore, we averaged scores across the two interviews to create a single index for analysis. Measures of PTSD were

Table 2. Male Military Veterans' Reports of Exposure to a Traumatic or Potentially Traumatic Event During Childhood and Adulthood ($N = 76$)

	Childhood		Adulthood	
Event type	*n*	%	*n*	%
Physical abuse or assault	53	69.7	52	68.4
Sexual abuse or assault	16	21.1	9	11.8
Disaster	36	47.4	28	36.8
Hazardous duty	0	0.0	4	5.3
Warzone exposure	0	0.0	57	75.0
Serious illness or injury	34	44.7	63	82.9
Accident	42	55.3	37	48.7
Intentional perpetration	1	1.3	7	9.2
Unintentional perpetration	0	0.0	3	3.9
Captivity or torture	2	2.6	0	0.0

(1) lifetime severity calculated by summing frequency and intensity for all 17 primary symptoms on the CAPS interview (see Weathers, Ruscio, & Keane, 1999); (2) total score for the M-PTSD; and (3) total score for the PCL. As expected, the index of total exposure to potential trauma was moderately associated with each PTSD measure: lifetime CAPS total score, $r(75) = .35$, $p < .01$; M-PTSD, $r(7) = .41$, $p < .001$; and PCL, $r(75) = .42$, $p < .001$.

Event Frequency Across Interviews

The mean number of lifetime potential traumas increased from 10.4 at Interview 1 to 11.2 at Interview 2. This difference was in the predicted direction and statistically significant, $t(75) = 2.12$, $p < .05$. Separate analyses by epoch revealed that events from adulthood showed differences in the expected direction, but only childhood events yielded statistically significant differences based on averages of 4.4 traumas at Interview 1 and 4.8 traumas at Interview 2, $t(75) = 2.11$, $p < .05$. The maximum increase in event reports by an individual across the interviews was 12, whereas the maximum decrease was 6.

Examination of reporting in terms of individuals rather than events shows that 51% of participants reported more lifetime events at the second interview, 38% reported fewer events, and 11% reported the same number of events. As follow-up, we compared the group of participants who reported more events at the second interview with those who reported fewer events in terms of eight demographic, symptom, and trauma exposure variables: Marital and employment status, education, ethnic background, income, living situation, income received for psychiatric disability, warzone military service, PTSD diagnosis, PTSD symptom severity, and the index of total exposure to potential trauma. There were no differences on any of these variables, apart from a trend toward more lifetime trauma in adulthood (averaged across two interviews) for people who reported fewer events at the second interview, $t(60) = 1.86$, $p = .068$.

The magnitude and nature of decreased reporting was then examined in detail in hopes of gaining insight into its occurrence. Examination of childhood events revealed that 15 of 21 participants who showed decreases did so on the basis of one event, while 5 of the 21 decreased by two events. Of 27 childhood events that were not repeated at Interview 2, 11 were lost due to subtle changes in reporting that affected coding decisions. A similar pattern was found for adulthood events, with 13 of 25 participants who showed decreases doing so based on one event, while an additional 5 decreased by two events. Of 58 adulthood events that were not repeated at Interview 2, 10 were lost due to subtle changes in reporting that affected coding decisions.

General Consistency Analyses

Consistency was quantified via both percent agreement and kappa coefficients. This dual approach was used because very low or high base rates produce low kappas despite moderate or high percent agreement. Despite this shortcoming of the kappa statistic, no alternative statistical tests are available that are less affected by base rates (Langenbucher, Labouvie, & Morgenstern, 1996).

Kappa coefficients were generated separately for the two orders of administration and were compared to determine if there were order-related differences. This is a subjective determination because no acceptable method exists for the statistical comparison. No systematic differences were apparent across orders; therefore, data were collapsed across this variable for all succeeding analyses.

We first examined consistency of reported events across the two interviews using kappa matrices computed for the 10 categories of traumatic events listed in Fig. 1. Overall kappas for event categories collapsed across *DSM* Criterion A categorization and exposure type for childhood ranged from .64 to .94, with associated percent agreement values ranging from 83 to 97%. For adulthood, kappas ranged from .51 to .97, with associated percent agreement from 80 to 99%. The highest kappa values were obtained for physical abuse during childhood and for warzone exposure during adulthood.

The next set of analyses examined the consistency of reported events according to their correspondence with *DSM-IV* Criterion A categorization. Table 3 reveals moderate-to-high kappa values for most event categories irrespective of associated endorsement of fear, helplessness, and/or horror. However, consistency of reporting is almost always higher for full-fledged traumatic events than for events that lack the emotional component reflected by Criterion A2.

Table 4 displays the percent agreement and kappa coefficients that result when events are distinguished according to exposure classification. Moderate-to-high kappa values are found overall, but coefficients among adulthood events are relatively higher for those that were directly experienced as compared to those that were witnessed. Childhood events show a similar pattern for three categories, but there is a reversal for both serious illness or injury and accidents, with larger coefficients for witnessed events relative to experienced events.

Table 3. Consistency of Reporting Traumatic and Potentially Traumatic Events Across Interview 1 and 2

	Traumatic events				Potentially traumatic events			
	Childhood		Adulthood		Childhood		Adulthood	
Trauma type	% Agree	κ	% Agree	κ	% Agree	κ	% Agree	κ
Physical abuse or assault	86.8	0.73	82.9	0.66	81.6	0.63	72.4	0.45
Sexual abuse or assault	90.8	0.62	94.7	0.64	92.1	0.37	96.1	−0.02
Disaster	88.2	0.54	86.8	0.64	78.9	0.55	76.3	0.41
Hazardous duty	—	—	97.4	0.65	—	—	94.7	0.47
Warzone exposure	—	—	96.1	0.91	—	—	89.5	0.28
Serious illness or injury	82.9	0.61	80.3	0.54	78.9	0.37	65.8	0.31
Accident	84.2	0.66	77.6	0.50	75.0	0.39	78.9	0.42

Note. Reliability statistics on the trauma type categories of Intentional Perpetration, Unintentional Perpetration, and Captivity or Torture are not presented due to low frequencies and low base rates.

Consistency of Worst Traumas

As can be seen in Table 5, physical abuse was the most frequent category for events identified as the worst in childhood, and warzone exposure was the most frequent category for events identified as worst during adulthood. If different worst events were reported across interviews, we used the event selected in the second interview as the index event. The same worst event was identified at both interviews for 68% of childhood events and for 78% of adulthood events. The two most frequent explanations offered for discrepancies were reconsideration of an event that had been identified during Interview 1 but not designated as the worst (10%) and recall of a new event that superceded the original worst event (12%). The worst event identified at the first interview was not reported at all during the second interview in only three instances.

Corroboration of Worst Traumas

Sources were not available for 18% of the 152 index events (17 childhood events; 11 adulthood events). Participants denied permission to contact collaterals more often for childhood ($n = 25$) than for adulthood events ($n = 5$), $\chi^2(1, N = 121) = 22.03$, $p < .001$. Participants who gave permission to contact at least one collateral were compared with those who did not give permission for contact in terms of a set of eight demographic variables, PTSD diagnosis and severity, and number of reported traumatic events. No differences were found between the groups on any of the variables examined.

A potential source was identified and permission to contact was granted for 91 (60%) of the potential 152 index events. Sixty-five (71%) of the 91 events were corroborated, reflecting 19 (61%) childhood events and 46 (77%) adulthood events. Twenty-three of the 26 events that were not corroborated were due to our inability to reach the collateral contact. There were only three instances (two childhood; one adulthood) in which a collateral was reached and an event was explicitly not corroborated. Finally, comparison of percent agreement values revealed no differences in the consistency of reporting for corroborated versus noncorroborated events in childhood or adulthood. For childhood events, 78.9% of respondents whose worst events were corroborated and 64.9% of

Table 4. Consistency of Reporting Experienced and Witnessed Traumatic and Potentially Traumatic Events Across Interview 1 and 2

	Experienced events				Witnessed traumatic events			
	Childhood		Adulthood		Childhood		Adulthood	
Trauma type	% Agree	κ	% Agree	κ	% Agree	κ	% Agree	κ
Physical abuse or assault	98.7	0.97	93.4	0.86	90.8	0.82	86.8	0.73
Sexual abuse or assault	90.8	0.70	98.7	0.88	98.7	0.66	96.1	−0.02
Disaster	88.2	0.76	85.5	0.70	92.1	0.53	93.4	0.51
Hazardous duty	—	—	94.7	0.57	—	—	—	—
Warzone exposure	—	—	100.0	1.00	—	—	—	—
Serious illness or injury	90.8	0.58	85.5	0.70	89.5	0.70	81.6	0.63
Accident	78.9	0.57	85.5	0.69	93.4	0.79	86.8	0.56

Note. See *Note* to Table 3. Hazardous duty and warzone exposure are, by definition, trauma types that can only be experienced; no instances of witnessing either of these trauma types were recorded.

Table 5. Traumatic Events Identified as the "Worst"

Trauma type	Childhood (%)	Adulthood (%)
Physical abuse or assault	47.4	6.6
Sexual Abuse or assault	7.9	0.0
Disaster	7.9	2.6
Hazardous duty	0.0	2.6
Warzone exposure	0.0	51.3
Serious illness or injury	19.7	30.3
Accident	13.2	6.6
None reported	3.9	0.0

those whose worst events were not corroborated identified the same worst event during both interviews. This difference is not statistically significant, $\chi^2(1, N = 76) = 1.30$, *ns*. For adulthood events, 80.4% of respondents whose worst events were corroborated and 73.3% of those whose worst events were not corroborated identified the same worst event during both interviews. This difference also was not significant, $\chi^2(1, N = 76) < 1$.

Discussion

This investigation examined the relationship between traumatic event characteristics and the consistency with which they are reported across two occasions. The first prediction for the study was that change in the number of reported events across interviews would be predominantly in the increasing direction. This pattern was observed with respect to (a) average number of reported events, especially those from childhood, (b) the proportion of participants who reported more events at the second interview, and (c) the range of increases relative to the range of decreases in number of events reported. These findings are consistent with the idea that assessment of trauma history is a reactive process that can affect memory retrieval and/or reporting.

Our second hypothesis specified that consistency is easier to achieve when target categories are broad and inclusive rather than narrow and precise. Analyses based on kappa matrices computed for seven classes of potentially traumatic events demonstrate moderate-to-high consistency at the most global level, indexed by whether or not a participant reports the same categories of events at two assessments. This level of analysis is similar to typical psychometric examinations of trauma questionnaires (e.g., Bernstein et al., 1994; Goodman et al., 1998; Kubany et al., 2000; Norris & Perilla, 1996). As predicted, consistency was lower, although still substantial, when evaluated in relation to the narrower target of self-identified "worst" events from childhood and adulthood. This comparison provides a simple demonstration of how consistency is dependent upon the criterion by which it is evaluated. It remains to be determined what level of specificity we should be using to evaluate consistency in the context of trauma reporting, and it is likely that the answer will vary depending on the aims served by collecting the information.

Consistent with our third prediction, traumatic events that meet both parts of the *DSM-IV* definition for PTSD Criterion A generally were reported with greater consistency than were the larger class of potentially traumatic events that meet only the objective stressor criterion (A1). Examination of Table 3 reveals a kappa range of .50 to .91 for traumatic events, with a median of .64, compared to a range of −.02 to .63 for potentially traumatic events, with a median of .40. These results are compatible with the *DSM-IV* definition of traumatic event, which requires a pronounced emotional reaction in conjunction with stressor exposure. They also fit with laboratory findings that demonstrate the ability of emotional arousal to enhance memory for stressful events (Cahill, 1997; Cahill & McGaugh, 1995).

Other event characteristics also are related to the consistency of reporting. Adulthood events that were experienced directly are reported more consistently than events that were experienced from the perspective of a witness. By contrast, witnessed accidents and serious illness or injury in childhood have higher kappa values than the same types of events that were experienced directly. Our interviewers noted that major accidents and very serious illnesses or injuries witnessed during childhood were often particularly horrifying to the individual. In addition, when the same categories of childhood events were experienced directly, they sometimes produced clouded consciousness or confusion at the time that made recall more difficult. The results also indicate that some categories of events are reported consistently, independent of whether they were experienced directly or whether they meet both components of PTSD Criterion A. The prime example is physical abuse in childhood, which has high kappa coefficients irrespective of categorical qualifiers.

Independent corroboration was sought for self-identified worst trauma experiences from both childhood and adulthood. Among events for which a potential corroborating source was identified and able to be contacted, 71% were corroborated and only 3% were explicitly disconfirmed. Despite this positive evidence, 57% of the 152 identified worst events were not corroborated for a variety of reasons that included failure to obtain permission from the participant, inability to contact a collateral or information source, and absence of a potential source for the event in question. It is somewhat reassuring that uncorroborated events do not appear to be reported any less consistently

than corroborated events. Nonetheless, it is evident that validation of retrospective reports of traumatic events remains a challenge because sources of information about specific traumatic experiences are fallible, often difficult to contact, and sometimes simply do not exist.

Generalizability of findings from this study may be limited by the exclusively male sample and the high prevalence of warzone trauma not found in most civilian populations. In particular, the absence of female participants may be responsible for the low base rate and resulting low kappa coefficients for the event category of sexual abuse or assault. The age range of the sample also is a potential limitation, although it does have the advantage of ensuring that the span of adulthood is at least as long as the 18 years of childhood. Ultimately, the only way to determine the extent to which findings based on this sample generalize is to examine consistency of trauma reporting by individuals of both genders, drawn from backgrounds not marked by warzone military service, who span both younger and older ages.

It is difficult to compare across studies with regard to relative levels of trauma exposure because methods for quantifying and collecting trauma information vary widely. That said, the amount of trauma reported in this sample appears to be similar to quantities reported by both Vietnam veterans and other clinical samples (e.g., Kubany et al., 2000), but it is generally higher than epidemiological and college student samples (Breslau et al., 1998; Goodman et al., 1998; Kessler et al., 1995; Norris, 1992). This is understandable given the subset of treatment-seeking individuals in the current sample.

The value of obtaining trauma histories is indicated by an ever-growing collection of studies showing that exposure to potentially traumatic events is common among the general population (e.g., Kessler et al., 1995) and even more common among individuals seeking mental and physical health services (e.g., Arnow et al., 1999). It is clear that documenting lifetime trauma exposure is a challenge that necessitates thoughtful and well-developed assessment methodologies. This study demonstrates that retrospective reporting about exposure to potentially traumatic events is relatively consistent when the information is obtained in a comprehensive and well-structured manner. Nonetheless, the act of undergoing assessment for trauma exposure may influence recall and/or reporting in ways that are potentially relevant in research, clinical, and forensic contexts. In addition, characteristics of events such as the intensity of accompanying emotion, the age at which it occurs, and whether it is directly experienced or witnessed, appear to be related to variations in consistency of reporting and may offer clues to the processes that are involved.

Acknowledgments

This research was supported by a Department of Veterans Affairs Research Advisory Group grant to the first author, and by the National Center for PTSD, Behavioral Science Division.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arnow, B. A., Hart, S., Scott, C., Dea, R., O'Connell, L., & Taylor, C. B. (1999). Childhood sexual abuse, psychological distress, and medical use among women. *Psychosomatic Medicine, 61,* 762–770.

Bernstein, D. P., Fink, L., Handelsman, L., Foote, J., Lovejoy, M., Wenzel, K., et al. (1994). Initial reliability and validity of a new retrospective measure of child abuse and neglect. *American Journal of Psychiatry, 151,* 1132–1136.

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., et al. (1995). The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress, 8,* 75–90.

Bower, G. B. (1981). Mood and memory. *American Psychologist, 36,* 129–148.

Breslau, N., Kessler, R. C., Chilcoat, H. D., Schultz, L. R., Davis, G. C., & Andreski, P. (1998). Trauma and posttraumatic stress disorder in the community: The 1996 Detroit Area Survey of Trauma. *Archives of General Psychiatry, 55,* 626–631.

Brewin, C. R., Andrews, B., & Gotlib, I. H. (1993). Psychopathology and early experience: A reappraisal of retrospective reports. *Psychological Bulletin, 113,* 82–98.

Briere, J., & Conte, J. (1993). Self-reported amnesia for abuse in adults molested as children. *Journal of Traumatic Stress, 6,* 21–31.

Burnam, M. A., Stein, J. A., Golding, J. M., Siegel, J., Sorenson, S. B., Forsythe, A. B., et al. (1988). Sexual assault and mental disorders in community population. *Journal of Consulting and Clinical Psychology, 56,* 843–851.

Byrne, C. A., Resnick, H. S., Kilpatrick, D. G., Best, C. L., & Saunders, B. E. (1999). The socioeconomic impact of interpersonal violence on women. *Journal of Consulting and Clinical Psychology, 67,* 362–366.

Cahill, L. (1997). The neurobiology of emotionally influenced memory. Implications for understanding traumatic memory. In R. Yehuda & A.C. McFarlane (Eds.), *Psychobiology of posttraumatic stress disorder* (pp. 238–246). New York: New York Academy of Sciences.

Cahill, L., & McGaugh, J. L. (1995). A novel demonstration of enhanced memory associated with emotional arousal. *Consciousness and Cognition: An International Journal, 4,* 410–421.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46.

Della Femina, D., Yeager, C. A., & Lewis, D. O. (1990). Child abuse: Adolescent records vs. adult recall. *Child Abuse and Neglect, 14,* 227–231.

Dill, D. L., Chu, J. A., Grob, M. C., & Eisen, S. V. (1991). The reliability of abuse history reports: A comparison of two inquiry formats. *Comprehensive Psychiatry, 32,* 166–169.

Goodman, L. A., Corcoran, C., Turner, K., Yuan, N., & Green, B. L. (1998). Assessing traumatic event exposure: General issues and preliminary findings for the Stressful Life Events Screening Questionnaire. *Journal of Traumatic Stress, 11,* 521–542.

Green, B. L., & Grace, M. C. (1988). Conceptual issues in research with survivors and illustrations from a follow-up study. In J. P. Wilson, Z. Harel, & B. Kahana (Eds.), *Human adaptation to extreme stress: From the Holocaust to Vietnam* (pp. 105–124). New York: Plenum.

Keane,T. M., & Wolfe, J. (1990). Comorbidity in post-traumatic stress disorder: An analysis of community and clinical studies. *Journal of Applied Social Psychology, 20,* 1776–1788.

Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry, 52,* 1048–1060.

Kilpatrick, D. M. (1983). Rape victims: Detection, assessment, and treatment. *The Clinical Psychologist, 36,* 92–95.

King, D. W., King, L. A., Erickson, D. J., Huang, M. T., Sharkansky, E. J., & Wolfe, J. (2000). Posttraumatic stress disorder and retrospectively reported stressor exposure: A longitudinal prediction model. *Journal of Abnormal Psychology.*

Koss, M. P., & Gidcyz, C. A. (1985). Sexual experiences survey: Reliability and validity. *Journal of Consulting and Clinical Psychology, 53,* 422–423.

Kubany, E. S., Haynes, S. N., Leisen, M. B., Owens, J. A., Kaplan, A. S., Watson, S. B., et al. (2000). Development and preliminary validation of a brief broad-spectrum measure of trauma exposure: The Traumatic Life Events Questionnaire. *Psychological Assessment, 12,* 210–224.

Kulka, R. A., Schlenger, W. E., Fairbank, J. A., Hough, R. L., Jordan, B. K., Marmar, C. R., et al. (1990). *Trauma and the Vietnam war generation: Report of findings from the National Vietnam Veterans Readjustment Study* . New York: Brunner/Mazel.

Langenbucher, J., Labouvie, E., & Morgenstern, J. (1996). Measuring diagnostic agreement. *Journal of Consulting and Clinical Psychology, 64,* 1285–1289.

Leserman, J., Drossman, D. A., & Li, Z. (1995). The reliability and validity of a sexual and physical abuse history questionnaire in female patients with gastrointestinal disorders. *Behavioral Medicine, 21,* 141–150.

Martin, J., Anderson, J., Romans, S., Mullen, P., & O'Shea, M. (1993). Asking about child sexual abuse: Methodological implications of a two stage survey. *Child Abuse and Neglect, 17,* 383–392.

Mollica, R. F., & Caspi-Yavin, Y. (1991). Measuring torture and torture-related symptoms. *Psychological Assessment, 3,* 581–587.

Norris, F. H. (1990). Screening for traumatic stress: A scale for use in the general population. *Journal of Applied Social Psychology, 20,* 1704–1718.

Norris, F. H. (1992). Epidemiology of trauma: Frequency and impact of different potentially traumatic events on different demographic groups. *Journal of Consulting and Clinical Psychology, 60,* 409–418.

Norris, F. H., & Perilla, J. L. (1996). The revised Civilian Mississippi Scale for PTSD: Reliability, validity, and cross-language stability. *Journal of Traumatic Stress, 9,* 285–298.

Resnick, H. S., Kilpatrick, D. G., Dansky, B. G., Saunders, B. E., & Best, C. L. (1993). Prevalence of civilian trauma and posttraumatic stress disorder in a representative national sample of women. *Journal of Consulting and Clinical Psychology, 61,* 984–991.

Roemer, L., Litz, B. T., Orsillo, S. M., Ehlich, P. J., & Friedman, M. J. (1998). Increases in retrospective accounts of war-zone exposure over time: The role of PTSD symptom severity. *Journal of Traumatic Stress, 11,* 597–605.

Schnurr, P. P., Spiro, A., III, Aldwin, C. M., & Stukel, T. A. (1998). Physical symptom trajectories following trauma exposure: Longitudinal findings from the Normative Aging Study. *Journal of Nervous and Mental Disease, 186,* 522–528.

Solomon, S. D., & Davidson, J. R. T. (1997). Trauma: prevalence, impairment, service use, and cost. *Journal of Clinical Psychiatry, 58,* 5–11.

Southwick, S. M., Morgan, A., Nicolaou, A. L., & Charney, D. S. (1997). Consistency of memory for combat-related traumatic events in veterans of Operation Desert Storm. *American Journal of Psychiatry, 154,* 173–177.

Steele, G. P., Henderson, S., & Duncan-Jones, P. (1980). The reliability of reporting adverse experiences. *Psychological Medicine, 10,* 301–306.

Stinson, M. H., & Hendrick, S. S. (1992). Reported childhood sexual abuse in university counseling center clients. *Journal of Counseling Psychology, 39,* 370–374.

Sutker, P. B., Uddo-Crane, M., & Allain, A. N., Jr. (1991). Clinical and research assessment of posttraumatic stress disorder: A conceptual overview. *Psychological Assessment, 3,* 520–530.

Vreven, D. L, Gudanowski, D. M., King, L. A., & King, D. W. (1995). The civilian version of the Mississippi PTSD Scale: A psychometric evaluation. *Journal of Traumatic Stress, 8,* 91–109.

Weathers, F. W., Keane, T. M., & Davidson, J. R. T. (2001). The clinician administered PTSD scale (CAPS): A review of the first ten years of research. *Depression and Anxiety, 13,* 132–156.

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., & Keane, T. M. (1993, October). *The PTSD checklist (PCL): Reliability, validity and diagnostic utility.* Paper presented at the Ninth Annual Meeting of the International Society for Traumatic Stress Studies, San Antonio, TX.

Weathers, F. W., Ruscio, A. M., & Keane, T. M. (1999). Psychometric properties of nine scoring rules for the Clinician-Administered Posttraumatic Stress Disorder Scale. *Psychological Assessment, 11,* 124–133.

Wyshak, G. (1994). The relation between change in reports of traumatic events and symptoms of psychiatric distress. *General Hospital Psychiatry, 16,* 290–297.