

# Assessing War Trauma in Refugees: Properties of the Comprehensive Trauma Inventory-104

Michael Hollifield

*Department of Psychiatry and Behavioral Sciences, University of Louisville,  
School of Medicine, Louisville, KY*

Teddy D. Warner

*Department of Family and Community Medicine, University of New Mexico,  
Albuquerque, NM*

Janis Jenkins

*Departments of Anthropology and Psychiatry, Case Western Reserve University, Cleveland, OH*

Nityamo Sinclair-Lian

*Department of Family and Community Medicine, University of New Mexico,  
Albuquerque, NM*

Barry Krakow

*Sleep and Human Health Institute, Albuquerque, NM*

Valorie Eckert

*State of California Department of Health, Sacramento, CA*

Pary Karadaghi

*Kurdish Human Rights Watch, San Diego, CA*

Joseph Westermeyer

*Department of Psychiatry, University of Minnesota, Minneapolis, MN and the Minneapolis  
Veterans Administration Medical Center, Minneapolis, MN*

*In this article, the authors describe the properties of the Comprehensive Trauma Inventory-104 (CTI-104), developed and designed empirically to improve assessment of traumatic war-related events. The mean number of events reported by 252 community dwelling Kurdish and Vietnamese refugees was 32 (SD = 27) out of the 104 items. Internal and test-retest reliability was excellent, and the validity of the CTI-104 as a measure of war trauma was supported by its high correlation with standard measures of known outcomes of trauma. The CTI-104 is reliable and valid, and assesses a broader range of traumatic war-related events in a broader range of refugees than currently available instruments.*

---

The research was conducted while Michael Hollifield was with the Departments of Family and Community Medicine and Psychiatry, University of New Mexico, Albuquerque, NM.

This research was supported by the National Institute of Mental Health, Grant # RO1 59574. We are deeply indebted to our advisors and assistants, Nguyen Yen, Bui Toan, Dinh Ky, Bobbie Nobles, Don Aaker, Jiyun Gozoh, Johnson Yuam, Quach Hoa, Nguyen Dien, Nguyen Phong, Nazar Gozoh, Parwaz Khaswrah, Diary Mohammad, Mohammad Dowlati, Chinar Hussein, Alan Beker, Corbaan Raza, Rzgar Abdula, and Faridoun Veriyani.

Correspondence concerning this article should be addressed to: Michael Hollifield, Department of Psychiatry and Behavioral Sciences, University of Louisville School of Medicine, Med Center One, 501 E. Broadway, Suite 340, Louisville, KY 40202. E-mail: m.hollifield@louisville.edu.

© 2006 International Society for Traumatic Stress Studies. Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/jts.20137

. . .there was shooting around the city and we had a lot of fears because we knew that they had orders to rape women and . . . kill them, or destroy that house while the people are in it . . . they destroyed our house, and the things we saw . . . I pray to God that nobody should see that . . . (A Kurdish woman)

Although the few current instruments that assess war-related events in refugees have contributed substantially to a once nascent field of inquiry (Clarke, Sack, & Goff, 1993; Mollica, n.d.; Mollica et al., 1992), the surge of global attention to the prevalence and impact of war-related trauma (World Health Organization, 2001) requires more precision in assessing such events. Specifically, extant measures were not empirically developed from community samples and do not assess the broad range of events experienced by both men and women and by both soldiers and civilians. An instrument that includes a more comprehensive range of events will improve the documenting of human rights abuses and the validity of measuring the relationship between war trauma and various kinds of health outcomes (Hollifield et al., 2002).

The recently reported Comprehensive Trauma Inventory-104 (CTI-104) was designed to be such an instrument (Hollifield et al., 2005). We developed the CTI-104 using recommended methods (Weathers, Keane, King, & King, 1997), including a multimethod design to gather information from Kurdish and Vietnamese refugees. In-depth interviews and focus groups were conducted to obtain as many war-related traumatic events and types of events as possible. These qualitative data were combined with data from a preliminary version of the CTI to determine the range and relative prevalence of events reported by refugees. An iterative process was then employed to distill over 250 items to a more efficient set of 104 items that represented the sample. This article is about the psychometric properties of the CTI-104.

## METHOD

### Study Design

The CTI-104 was evaluated in Phase II of the New Mexico Refugee Project (NMRP), a cross-sectional, retrospective study designed to improve measurement of trauma and health outcomes in refugees. As part of a larger battery of tests, eight instruments were selected a priori for assessing validity of the CTI-104. The Harvard Trauma Questionnaire (Mollica et al., 1992), the current standard for assessing refugee war trauma, was chosen to test concurrent validity. Seven instruments were chosen to test the ability of the CTI-104 to predict various health outcomes, including symptoms, disorders, impairment, and coping. Eighty (48 Kurdish and 32 Vietnamese) of the total 252 study participants were consecutively sampled for a second administration of the CTI-104 4 to 6 weeks after the first administration to determine test-retest reliability of the CTI-104.

### Participants and Sampling Method

The sample was drawn from communities of Kurdish refugees in Colorado Springs, CO and San Diego, CA, and Vietnamese refugees in Albuquerque, NM. The estimated populations of refugees in these communities were 250, 5,000, and 6,000, respectively. Kurds primarily originated from northern Iraq and resettled in Colorado after 1990 and in California since the late 1970s. Vietnamese have resettled in Albuquerque since 1975. Sampling was designed to include a broad range of people who experienced a broad range of events to develop and test the CTI-104. Purposive sampling is a good choice to obtain such a representative sample (Waters & Biernask, 1989), and was conducted using a chain-referral method. The Trauma Experiences Questionnaire (TEQ) developed in our research (Hollifield et al., 2005) by partly adapting criteria used by Thompson and McGorry (1995), and a screening grid

were used to obtain the sample by two strata: (1) TEQ category and (2) gender and ethnicity.

The TEQ was used to sample participants into the first strata of having experienced either torture, non-torture war-related trauma, or no war-related trauma. The United Nations' Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment (United Nations, 1985) definition of torture was operationalized with a checklist for identifying torture survivors. Non-torture war-related trauma survivors met criteria for having experienced war-related events but not torture, and no war-related trauma participants identified themselves as not having experienced torture or significant war-related events.

Power analyses to determine sample size was conducted by using expected health outcome scores of participants in the three TEQ groups. Literature exists about expected levels of depression and posttraumatic stress disorder (PTSD) in refugees with differing levels of trauma; however, no data exist for the CTI-104. Thus, we were conservative in estimating sample size by using the Hopkins Symptom Checklist-25 depression measure (Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974), which would likely have a smaller difference by TEQ group than PTSD, to determine that a sample of 240 participants (three TEQ groups of 80, stratified by ethnicity and gender) provided power of  $\geq .80$  at Cronbach's alpha .05 to detect moderate effect-size differences between TEQ groups. Because it was difficult to find Kurdish people who had no war-related trauma and women who were tortured, we over-sampled in other categories. All participants provided verbal and written informed consent, and were reimbursed \$20 for their participation. The study was approved by the University of New Mexico's Institutional Review Board.

## Instruments

**Instrument translation and administration.** The translation of language is complex and must be adapted for specific purposes (Westermeyer & Janca, 1997). We trans-

lated all questionnaires in the NMRP into Kurdish and Vietnamese using standard, back-and-forth blinded techniques and consensus approaches. Instruments were administered individually during one session and in the same order for each participant as part of the larger battery of assessments, which took approximately 3 hours to complete.

**The Comprehensive Trauma Inventory-104.** The Comprehensive Trauma Inventory-104 (CTI-104) is a self-administered questionnaire that assesses war-related events in community dwelling refugees (Hollifield et al., 2005). It has English, Kurdish, and Vietnamese versions, and has 104 event items divided among 12 event-type scales. The response format has participants check whether or not they experienced the event and, if they did, how much impact the event had in terms of fear or threat to their life or safety. This response format was meant to be relevant to Criterion A of the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* PTSD diagnosis (DSM-IV; American Psychiatric Association, 1994). Thus, each item has five possible responses: 0 = *did not happen*, 1 = *a little fear or threat*, 2 = *moderate fear or threat*, 3 = *a lot of fear or threat*, and 4 = *extreme fear or threat*. Scoring may be either a sum of dichotomous values (number of events) or a sum of 5-point severity scores, which were both modeled as continuous data. Further information about the CTI-104 and its 12 scales is available from the authors.

**The Harvard Trauma Questionnaire.** The Harvard Trauma Questionnaire (HTQ) was developed by expert consensus methods from treatment studies with refugees in clinical settings (Mollica, n.d.; Mollica et al., 1992). It has been widely used as the current research standard in the field for assessing war-related trauma in refugees (Hollifield et al., 2002). Excellent interrater reliability for all events ( $r = .93$ ); test-retest reliability ( $r = .89$ ); and internal consistency (Cronbach's  $\alpha = .90$ ) was shown in a convenience sample of 91 Southeast Asian psychiatric outpatients (Mollica et al., 1992). Each of the 17 trauma

items has four response options: experienced, witnessed, heard about, or no experience of the event. All response options that apply to a participant for each item are to be checked. We scored the HTQ as recommended by adding the number of experienced, witnessed, and heard about events for all 17 items.

**Health outcomes: Symptoms and disorders.** The Hopkins Symptom Checklist-25 (HSCL-25) assesses symptoms of anxiety and depression and is valid and reliable in the general U.S. population and in Indochinese refugee groups (Derogatis et al., 1974; Mollica, Wyshak, deMarneffe, Khuon, & Lavelle, 1987; Winokour, Winokur, Rickels, & Cox, 1984). Reviews in the cultural psychiatry literature note the transcultural validity of the HSCL-25 (Butcher, 1991; Kinzie & Manson, 1987). Continuous or dichotomous scoring may be used for analyses; we used both. Cutoff scores using item-averages of  $\geq 1.75$  for each scale are valid as diagnostic proxies for clinically significant anxiety and depression in general U.S. and refugee samples (Derogatis et al., 1974; Winokur et al., 1984).

The PTSD Symptom Scale-Self Report (PSS-SR) has not been used in refugee research, but has been shown to be a valid predictor of the PTSD diagnosis in U.S. populations (Foa, Riggs, Dancu, & Rothbaum, 1993). Cronbach's alpha for internal consistency is .91 for the total score, and one-month test-retest reliability is .74. The 17 items on the scale, each scored from 0 to 3 for frequency of symptoms, are essentially *DSM-IV* PTSD diagnostic items. The PSS-SR may be scored as (a) continuous; (b) ordered, based on severity; or (c) dichotomous based on case versus noncase (i.e., PTSD vs. no PTSD).

The New Mexico Refugee Symptom Checklist (NMRSCCL) has been developed by our research group and will be the topic of future articles. It has 121 items in 12 physical and psychological symptom scales. Each item is scored from 0 (*not at all*) to 4 (*extremely*) as a reflection of the persistent and bothersome nature of the symptom over the past year. The NMRSCCL is a reliable and valid indicator of symptoms in refugees (Hollifield & Warner, 1999).

**Health outcome measures: Impairment and coping.** The Medical Outcomes Study SF-36 is brief, psychometrically well tested, appropriate for administration to persons over the age of 14 years, and may be self- or interviewer-administered (McHorney, Ware, & Raczek, 1993; Ware, Snow, & Kosinski, 1993). The SF-36 measures eight domains of daily functioning and generates continuous data. The SF-36 has been shown to be useful for international comparisons (Aaronson et al., 1992).

The Sheehan Disability Inventory (SDI) has three 10-point Likert scales that measure impairment due to the disruption of daily activities in the areas of work, social, home and family life, and it has a 5-point global disability rating scale (Sheehan, 1983). Cronbach's alpha coefficients for the four scales range from .56 to .86. The SDI discriminates between patients with and without panic disorder and is sensitive to change with treatment (Leon, Shear, Portera, & Klerman, 1992).

The Sense of Coherence-13 (SOC-13), developed in work with Jewish holocaust survivors, measures a global coping resource in the face of overwhelming stress. Data from the SOC are continuous. The SOC has been supported cross-culturally (Antonovsky, 1987); its reliability and validity have been established in Hebrew and in English (Antonovsky, 1983, 1984). In addition, acceptable test-retest reliability has been shown in a study of Cambodian refugees (Cheung, 1995).

## Data Analyses

**Descriptive.** The prevalence and range of event number and severity scores for the CTI-104 total scale and 12 trauma type scales in our sample are described. We used analyses of variance (ANOVA), and multivariate analysis of variance (MANOVA) to analyze the CTI-104 total score and 12 scales (as repeated-measures dependent variables) with a factorial between group design of ethnicity by gender, enabling group comparisons for the CTI-104 total scale and 12 trauma-type scales.

**Reliability.** Cronbach's alphas for the CTI-104 total scale and the 12 trauma-type scales were calculated for the

complete sample ( $N = 252$ ) and for the retest subsample ( $n = 80$ ). Test-retest correlation coefficients for the CTI-104 total scale and the 12 trauma-type scales were determined from the subsample ( $n = 80$ ).

**Validity.** Validity was examined by correlating the CTI-104 with (a) the HTQ, (b) the health outcomes of symptoms and disorders, and (c) the health outcomes of impairment and coping. We used  $t$  tests of paired correlation coefficients to compare the CTI-104 with the HTQ on strength of association with health outcome measures.

## RESULTS

### Descriptive

**Participants.** Of the 252 participants, 135 (54%) were Vietnamese (V) and 117 (46%) were Kurdish (K); 27% lived in Colorado Springs, Colorado, and 19% lived in San Diego, California. There were 135 (54%) men and 117 (46%) women. Ages ranged from 19 to 77 years with a mean age of 44 ( $SD = 14$ ), with the Vietnamese being older (47 vs. 39 years),  $t(250) = 4.79$ ,  $p < .01$ . One-hundred sixteen were classified as survivors of torture (62 K, 54 V; 77 men, 39 women), 84 were survivors of non-torture war related trauma (46 K, 38 V; 36 men, 48 women), and 48 had no war-related trauma (8 K, 40 V; 21 men, 27 women). Four people were not classified by the TEQ. Because sampling was conducted by chain-referral methods, and those who may not have been interested did not contact the research team, nonparticipation rates are not available.

### CTI-104 scores and quartiles for the total sample.

The mean number of CTI-104 events reported was 32 ( $SD = 27$ ), range: 0–104. Quartiles (25th, 50th, and 75th percentiles) were 9, 28, and 46 events, respectively. The average CTI-104 severity score reported was 91 ( $SD = 95$ ), range: 0–414. In contrast, the mean number of HTQ experienced events reported was 5.2 ( $SD = 4.9$ ); the mean number of total HTQ events reported was 16.8 ( $SD = 12.8$ ). Pearson's correlation between total CTI-104 events and total severity score was .95, and correlations between events

and severity on the 12 scales ranged from .84 to .96, with 10 of 12 scales  $> .90$ . Given the high correlation between events and severity, the remainder of the analyses was with events only (Table 1).

**CTI-104 events by ethnicity and gender.** A Gender by ethnic group ANOVA on CTI-104 total events showed that Kurds reported more events than Vietnamese,  $M = 41$  ( $SD = 29$ ) versus  $M = 23$  ( $SD = 23$ ), Cohen's  $d = .69$ , Ethnic main effect,  $F(1, 250) = 30.9$ ,  $p < .001$ . Men reported more events than women,  $M = 38$  ( $SD = 27$ ), versus  $M = 25$  ( $SD = 25$ ), Cohen's  $d = .50$ , gender main effect,  $F(1, 250) = 16.63$ ,  $p < .001$ . Ethnicity and gender together accounted for 17% of the variance in total CTI-104 scores with a larger effect from ethnicity (12%) than gender (5%).

A factorial ethnicity and gender MANOVA conducted on the 12-scale scores showed both multivariate ethnicity,  $F(1, 237) = 18.37$ ,  $p < .01$ , and gender,  $F(1, 237) = 5.99$ ,  $p < .01$ , main effects on the reported number of events with no ethnicity by gender interaction. Follow-up univariate ANOVA conducted on the 12-scale scores showed ethnic group differences were significant for 10 of the 12 scales (all but the Physical Injury and the Separation and Isolation Scales) with Kurds reporting more events for each scale than Vietnamese. Gender differences reached statistical significance for 7 of the 12 scales (all but Sexual Trauma and Abuse, Domestic Discord, Displacement, Separation and Isolation, and Difficulties During Migration scales) with men reporting more events than women.

### Reliability

**Internal reliability.** The CTI-104 was factor analyzed, resulting in 12 coherent factors that corresponded closely to a priori conceptual groupings of items. Details of the factor analysis and all reliability analyses are available from the authors. For the total sample ( $N = 252$ ), Cronbach's alpha coefficient for the CTI-104 was .99, and for the 12 scales ranged from .68 to .98, with 10 scales being larger than .86. For the subsample ( $n = 80$ ), Cronbach's alpha coefficient for the CTI-104 was .99, and for the

**Table 1.** CTTI-104 Scale Descriptive Statistics, Total Sample (N = 252)

CTTI-104 Scale (number of items)	Score range	M	SD	M % of total scale items endorsed	Number (n) endorsing ≥ one scale item	% endorsing ≥ one scale item	Item-average severity score (endorsed by respondents)
CTTI-104 total (104)	E 0-104 S 0-414	31.74 90.85	27.0 94.61	31	233	92.5	2.86
Psychological Injury (15)	E 0-15 S 0-59	5.48 15.42	4.28 14.44	37	212	84.1	2.78
Physical Injury (3)	E 0-3 S 0-12	0.54 1.63	0.93 3.06	18	79	31.3	3.00
Detention and Intentional Abuse (32)	E 0-32 S 0-128	7.52 22.06	9.72 33.58	24	173	68.7	2.88
Sexual Trauma or Abuse (6)	E 0-6 S 0-24	0.40 1.27	1.31 4.50	7	29	11.5	3.00
Witnessing Abuse, Injury, or Death (14)	E 0-14 S 0-56	5.90 16.61	4.58 16.12	42	204	81.0	2.83
Hearing about Injury and Death (3)	E 0-3 S 0-12	2.09 5.80	1.28 4.48	70	189	75.0	2.78
Deprivation and Discrimination (6)	E 0-6 S 0-24	3.07 8.33	2.33 7.81	51	191	75.8	2.73
Betrayal (8)	E 0-8 S 0-32	1.86 5.42	2.33 7.73	23	144	57.1	2.96
Domestic Discord and Violence (2)	E 0-2 S 0-8	0.60 1.64	0.88 2.73	30	83	32.9	2.73
Displacement (3)	E 0-3 S 0-12	1.05 2.94	1.11 3.36	35	145	57.5	2.80
Separation and Isolation (7)	E 0-7 S 0-28	2.05 6.22	2.26 7.74	29	158	62.7	3.07
Difficulties during Migration (5)	E 0-5 S 0-20	1.18 3.54	1.66 5.57	24	111	44.0	2.96

Note. E = Events reported; S = severity scale.

<sup>a</sup>Severity was rated on a scale indicating how frightening the event was in terms of it being a threat to life or safety: 1 = a little fear or threat, 2 = moderate fear or threat, 3 = a lot of fear or threat, and 4 = extreme fear or threat. Zero indicated the event did not happen.

12 scales ranged from .75 to .99, with 11 scales being larger than .90.

**Test-retest reliability.** Total CTI-104 events retest correlation was .83. The Hearing About Abuse Scale had a low retest correlation of .29; three scales had acceptable retest reliability ranging from .66 to .69, and eight scales had good retest reliability ranging from .74 to .83. The mean number of events reported on the first administration for the subsample ( $n = 80$ ) was 46.2, which was higher than the mean of 32 for the overall sample ( $N = 252$ ). The mean number of events reported on the second administration increased significantly from the first administration for the total CTI-104,  $M = 52.9$  versus  $M = 46.2$ ,  $t(78) = 3.10$ ,  $p < .01$ , and for eight scales.

## Validity

**Concurrent validity.** The CTI-104 total event score was correlated with the HTQ total score,  $r = .65$ ,  $p < .01$ , and the HTQ torture item,  $r = .48$ ,  $p < .01$ .

**Construct validity: Discriminating by entry criteria on the TEQ.** The CTI-104 and its 12 scales discriminated between the three TEQ groups (torture vs. non-torture war-related trauma vs. no war-related trauma) in a one-way ANOVA, e.g., total CTI-104,  $r = .54$ ,  $F(2, 245) = 49.21$ ,  $p < .001$ . Mean scores for the total CTI-104 and its scales were higher for the torture group than for the non-torture war-related trauma group, Cohen's  $d$  range = 0.12 to 0.87,  $p < .05$ , which in turn were higher than scores for the no war-related trauma group, Cohen's  $d$  range = 0.19 to 1.17,  $p < .01$ , for total CTI-104 and 11 of 12 scales, although 2 scales did not reach statistical significance (Table 2).

**Construct validity: Correlation with health outcomes—symptoms and disorders.** The CTI-104 event number was strongly correlated with the PSS-SR continuous PTSD symptom score,  $r = .50$ ,  $p < .001$ , and the dichotomous PTSD proxy diagnosis,  $r = .43$ ,  $p < .001$ . CTI-104 event number was also strongly correlated with HSCL-25 Continuous Anxiety,  $r = .48$ ,  $p < .001$ , and Depression,

$r = .47$ ,  $p < .001$ , scores as well as with dichotomous measures of both (representing positive vs. negative cases), and with the NMRSCS Symptom Number,  $r = .43$ ,  $p < .001$ , and Severity,  $r = .47$ ,  $p < .001$ . In contrast, the corresponding symptom and disorder correlations with the HTQ were about 60% in magnitude of those with the CTI-104,  $r$  range = .24 to .32. When compared directly to the correlations of the HTQ on symptoms and diagnoses, the CTI-104 correlated significantly better with all outcomes (all  $ps < .01$ ) than did the HTQ (Table 3). Not shown in the table, there were few differences between ethnic-gender groups regarding correlation of the CTI-104 to health outcomes. For example, CTI-104 event number was not as strongly correlated with HSCL-25 depression score in Vietnamese women as it was in the other three ethnic-gender groups. Complete details of the ethnic-gender by health outcome analyses are available from the authors.

**Construct validity: Correlation with health outcomes—impairment and coping.** The CTI-104 event number was correlated with all four SDI scales,  $r = .34$  to .47, all  $ps < .001$ , all eight SF-36 scales,  $r = .19$  to .39, all  $ps < .01$ , and the SOC scale,  $r = .21$ ,  $p < .01$  (Table 3). A  $t$  test comparing paired correlations showed that the CTI-104 was more significantly correlated with 10 of the 13 impairment and coping measures than the HTQ.

**Construct validity: The CTI-104 as a predictor of PTSD, depression, and anxiety.** The CTI-104 was dichotomized at the median to determine the risk of having clinically significant (vs. nonsignificant) anxiety and depression and having PTSD or no PTSD, dependent on low (0 to 27 events,  $n = 124$ ) or high (28 to 104 events,  $n = 128$ ) CTI-104 scores. Refugees with high CTI-104 scores were 7.03 (95% CI: 3.45–14.33) times more likely to score above the cutoff for clinically significant anxiety, 4.06 (95% CI: 2.27–7.29) times more likely to score above the cutoff for clinically significant depression, and 4.85 (95% CI: 2.78–8.45) times more likely to have the proxy PTSD diagnosis than refugees with low CTI-104 scores. When graphs crossing exposure (CTI-104 scores) and PTSD, depression, or anxiety were viewed,

**Table 2.** Means and Standard Deviations of CTI-104 Scale Event Scores by Study Entry Criteria on the Trauma Experiences Questionnaire (TEQ)

CTI-104 Scale	TEQ Torture ( <i>n</i> = 116)		TEQ Non-torture war-related trauma ( <i>n</i> = 48)		TEQ No war-related trauma ( <i>n</i> = 48)		Cohen's <i>d</i> <sup>a</sup>	Cohen's <i>d</i> <sup>b</sup>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
CTI-104 total	45.60	26.50	27.80	22.85	7.58	9.38	49.21**	0.71
Psychological Injury	7.30	4.01	5.48	3.93	1.50	2.17	41.85**	0.46
Physical Injury	0.85	1.07	0.43	0.83	0.04	0.20	13.60**	0.43
Detention and Intentional Abuse	12.65	10.63	4.49	7.38	1.02	1.90	40.48**	0.87
Sexual Trauma or Abuse	0.68	1.63	0.24	1.14	0.06	0.25	4.96*	0.30
Witnessing Abuse, Injury, or Death	8.09	4.09	5.62	4.20	1.58	2.52	48.40**	0.60
Hearing about Injury and Death	2.46	1.04	2.26	1.17	1.00	1.31	28.81**	0.18
Deprivation and Discrimination	4.16	1.94	2.88	2.27	1.02	1.64	42.28**	0.61
Betrayal	2.69	2.53	1.65	2.16	0.35	0.84	20.34**	0.44
Domestic Discord and Violence	0.78	0.94	0.67	0.91	0.08	0.35	11.64**	0.12
Displacement	1.44	1.14	1.01	1.05	0.27	0.61	22.30**	0.39
Separation and Isolation	2.91	2.31	1.93	2.21	0.35	0.70	26.36**	0.43
Difficulties during Migration	1.61	1.83	1.14	1.58	0.29	0.87	11.65**	0.27

Note. *N* = 248. Four cases for TEQ were missing or uncategorizable.

<sup>a</sup>Cohen's *d* comparing torture group with non-torture war-related trauma group. <sup>b</sup>Cohen's *d* comparing non-torture war-related trauma group with no war-related trauma group.

\**p* < .01. \*\**p* < .001.



**Table 3.** Comparison of the CTI-104 Scale and the Harvard Trauma Questionnaire (HTQ) With Health Outcome Measures ( $N = 252$ )

	CTI-104 Total severity score <i>r</i>	CTI-104 Total events <i>r</i>	HTQ Total events <i>r</i>	Comparison of paired correlation coefficients (CTI-104 vs. HTQ) <i>r - r</i>
<i>Symptoms/diagnoses<sup>a</sup></i>				
HSCL Anxiety score	.48**	.48**	.26**	.22**
HSCL Anxiety diagnosis	.40**	.42**	.25**	.17**
HSCL Depression score	.48**	.47**	.30**	.17**
HSCL Depression diagnosis	.40**	.41**	.26**	.15*
PSS-SR Total score	.49**	.50**	.32**	.18**
PSS-SR PTSD diagnosis	.39**	.43**	.26**	.17**
NMRSCL Total events	.35**	.43**	.24**	.19**
NMRSCL Total severity	.44**	.47**	.27**	.20**
<i>Impairment/Coping<sup>b</sup></i>				
SDI Work Impairment	.37**	.37**	.24**	.13*
SDI Social/Life Impairment	.42**	.44**	.24**	.20**
SDI Family Life Impairment	.47**	.47**	.28**	.19**
SDI Global Disability	.32**	.34**	.23*	.11*
SF36 Physical Functioning	.14*	.19**	.01	.18**
SF36 Role-Physical	.25**	.28**	.22**	.06
SF36 Bodily Pain	.40**	.39**	.33**	.06
SF36 General Health	.26**	.30**	.14*	.16**
SF36 Vitality	.21**	.23**	.18**	.05
SF36 Social Functioning	.31**	.33**	.11	.22**
SF36 Role-Emotional	.27**	.28**	.14*	.14*
SF36 Mental Health	.34**	.35**	.22**	.13*
SOC Total score	.20**	.21**	.08	.13*

*Note.* HSCL = Hopkins Symptom Checklist-25; PSS-SR = Posttraumatic Symptom Scale–Self Report; NMRSCL = New Mexico Refugee Symptom Checklist; SDI = Sheehan Disability Inventory; SF36 = Medical Outcomes Study Short-Form 36; SOC = sense of coherence.

<sup>a</sup>All but three health outcome measures were continuous data. HSCL Anxiety Diagnosis, HSCL Depression Diagnosis, and PSS-SR PTSD Diagnosis each had two categories (yes or no), and thus reduced the power for analysis by reducing the number of possible responses. <sup>b</sup>Actual negative correlations of SF-36 and SOC are not shown. Lower SF-36 scores imply worse functioning, and, as number of events increase, SF-36 scores decrease. Lower SOC implies worse coping, and, as the number of trauma events increase, the SOC scores decrease.

\* $p < .05$ . \*\* $p < .01$ .

it appeared that the risk increased at 20 reported events and then again at 60 reported events. As shown in Table 4, the risk for meeting criteria for proxy PTSD diagnosis was 5% for no CTI events, 17% for 1 to 20 CTI events, 50% for 21 to 60 CTI events, and 76% for more than 60 CTI events. The event-related risk for having significant depression or anxiety followed a similar pattern.

The 12 CTI-104 scales varied modestly in their strength of correlation with health outcome measures. For example, the range of correlations between each of the 12 scales and

the HSCL-25 anxiety scores was .20 to .44, between each of the 12 scales and the depression scores was .24 to .41, and between each of the 12 scales and the PTSD scores was .16 to .42 (not shown in tables).

## DISCUSSION

The CTI-104 is a reliable and valid self-report instrument that assesses a wide range of traumatic war-related events in a broad range of community-dwelling refugees who experienced events at different times and in different

**Table 4.** Relationship of CTI-104 Scale Events to Mean Scores and Number (%) Above Diagnostic Cutoff Scores for PTSD, Depression, and Anxiety

Health outcome measure	Number of CTI-104 total events											
	0 (n = 19)			1-20 (n = 88)			21-60 (n = 107)			61-104 (n = 38)		
	M	SD	n (%) Above diagnostic cutoff	M	SD	n (%) Above diagnostic cutoff	M	SD	n (%) Above diagnostic cutoff	M	SD	n (%) Above diagnostic cutoff
PSS-SR PTSD <sup>a</sup> (range 0-51)	2.79	3.95	1 (5.3%)	5.10	7.67	15 (17.0%)	13.18	10.59	53 (49.5%)	21.48	11.21	29 (76.3%)
HSCL-25 Depression <sup>b</sup> (range 1-4)	1.24	0.19	0 (0.0%)	1.29	0.31	10 (11.4%)	1.77	0.64	42 (39.3%)	2.14	0.61	27 (71.1%)
HSCL-25 Anxiety <sup>b</sup> (range 1-4)	1.22	0.24	1 (5.3%)	1.22	0.24	2 (<1%)	1.62	0.57	38 (35.5%)	2.04	0.67	22 (57.9%)

<sup>a</sup>PSS-SR Above diagnostic cutoff based on dichotomous scoring for caseness for PTSD. <sup>b</sup>HSCL-25 Depression and anxiety scores are item averages; above diagnostic cutoff is  $\geq 1.75$ , considered clinically significant.

geopolitical contexts. The development of the CTI-104 was systematic to ensure that the events, the types of events, and the language used in the items were relevant to both soldiers and civilians, both men and women, and both young and old adults.

The CTI-104 is a more comprehensive measure of events than is currently available. Because the CTI-104 was developed using multiple methods in community samples, it assesses more relevant events than current instruments that were developed by rational, expert methods. Thus, the mean number of events reported using the CTI-104 is larger than in other research (e.g., Mollica, McInnes, Poole, & Tor, 1998; Mollica et al., 1990, 1999) and has good variability across the sample. Although it discriminates by ethnic and gender groups, the variance in total CTI-104 scores accounted for by these group variables was a modest 17%, most of which was due to ethnic group. This implies that the CTI-104 may be generalizable to various refugee populations, and it may discriminate between different war contexts. These features allow the CTI-104 to be a standard used to document a broad range of war-related experiences and human rights abuses in refugees.

The CTI-104 is a good measure of traumatic events because it has good-to-excellent construct validity. In this study, the CTI-104 was a significant predictor of clinically significant symptoms of and diagnostic proxies for anxiety, depression, and PTSD, physical and psychological symptoms, and of impairment and coping. The CTI-104 was more strongly correlated with these health outcome measures than the Harvard Trauma Questionnaire (Table 3). Establishing the dose-effect response is an essential component for inferring the etiology of PTSD (True, Goldberg, & Eisen, 1988), and the CTI-104 seems to be an advance in this regard. Other research with the HTQ has shown modest associations of war-related events with health outcomes (Crescenzi et al., 2002; Mollica et al., 1998, 1999; Saab, Chaaya, Doumit, & Farhood, 2002). We did not directly compare the CTI-104 to the War Trauma Scale (WTS) developed by Clarke and colleagues (Clarke et al., 1993). However, in one study the WTS accounted for 15% of PTSD score variance and 7% of depression score variance (Clarke et al., 1993), whereas the CTI-104 accounted

for 25% of PTSD score variance and 22% of depression score variance in this study. The validity of the CTI-104 allows it to be a standard for research about the traumatic effects of war experiences of refugees. Replication of our findings will be another important step in the process of establishing evidence for the validity of the CTI-104.

We believe that CTI-104 validity was enhanced because of attention to important translation and cultural syntax in developing the language and the format for the CTI-104. Many CTI-104 items retain words and syntax from interviews with participants. During the developmental phase, key words and phrases in both languages were agreed on by scientific and refugee advisors. For example, there are multiple Vietnamese words and phrases that might mean "torture." We agreed to use a term phrased such that most Vietnamese people would likely understand it to mean "torture" when read in text. Likewise, the questionnaire format was assessed for and altered to promote broad acceptance.

The CTI-104 has good-to-excellent internal and test-retest reliability, which has generally not been well tested in other measures developed or adapted for refugee research (Hollifield et al., 2002). The high reliability for the CTI-104 total score (.99) and the Detention and Intentional Abuse scale (.98) is no doubt partly an artifact of a high number of scale items because a Cronbach's alpha of  $>.90$  can be achieved with 100 items with inter-item correlations of only .10. However, 9 of the 12 CTI-104 scales have between two and eight items, and 7 of these 9 scales have alpha  $\geq .86$ , the other 2 scales are .68 and .78. The Hearing About Abuse and Violence scale was the single exception to overall good reliability, with a test-retest correlation of .29, although its internal reliability was excellent, Cronbach's  $\alpha = .94$ . We do not know why this scale had retest properties that were different from other scales, but it may be related to being the only scale about events that came to the awareness of respondents through another person or media.

The 80 participants in the retest sample reported a significantly higher number of events at the second administration for 9 of the 12 CTI-104 scales. This phenomenon is consistent with other research (Mollica & Caspin-Yavin, 1991). Less recall at Time 1 relative to Time 2 may be due to

long-standing repression followed by priming of memory by the first administration. Repressing specific horrors of war may be one mechanism used to defend ones' psychological integrity. We don't know whether participants had more intrusions about events after administration of the CTI-104. However, during Phase I of the NMRP the majority of participants said they felt better in the few weeks after their interview. All (15%) who said they had more symptoms after the interviews were not regretful that they participated in the study. This phenomenon of sensitization warrants more investigation from a scientific and an ethical perspective.

In this sample, the CTI-104 total severity score was not an appreciably different index than simply measuring whether or not an event occurred. This was unexpected. It may be that CTI-104 events are sufficiently traumatic that the severity scale is rendered less meaningful. The CTI-104 was partly constructed from interview data where participants were asked to disclose war-related events that were specifically traumatic. Thus, whether a participant circled 1—*little fear or threat*, or 4—*extreme fear or threat*, the event may have involved enough fear, helplessness, or horror to be a traumatic Criterion A event for PTSD. This is supported by the fact that the correlations between CTI-104 events and health outcomes are at least as strong as between CTI-104 severity and health outcomes (Table 3). It may also be that the severity scale wording was not clear. Further study about how to measure severity or impact of war-related events in refugees is needed.

On the other hand, the 12 subscales or types of events are important. In addition to these 12 types being identified by refugees during interviews (Hollifield et al., 2005), quantitative data demonstrate modest differences between them on demographic, reliability, and validity measures. Previous research has often assumed that the categories of physical, psychological, and sexual trauma were adequate descriptors of the refugee experience, but our data demonstrates the importance of other categories. This is not surprising because other research with refugee (Momartin, Silove, Manicavasagar, & Steel, 2003) and nonrefugee (Fullilove, 1996) populations have determined specific effects of categories such as threats to life, traumatic loss, and displace-

ment. These 12 types, in addition to the breadth of items, help validate experiences not only of soldiers or those imprisoned, but also of women who raised children at home alone while their husbands were away or being tortured in camps, and of people who witnessed death and destruction as they fled the horrors of war.

The primary limitation to the scope, validity, and generalizability of the CTI-104 is that our samples were refugees settled in three U.S. locales who had experienced events from two conflicts out of many. Survivors of other conflicts and those who are not displaced may have different experiences and perceptions. Another potential limitation is that the symptom and illness health outcome measures in our study were not diagnostic assessments. However, we used the most valid self-report measures available. It is not clear whether currently defined psychiatric diagnoses are the most appropriate constructs for traumatized refugees because culture and language complicate diagnosis (deGirolamo & McFarlane, 1996), and polytrauma causes symptom complexes that may be different from those seen in PTSD caused by single events (Basoglu, Paker, Ozmen, Tasdemir, & Sahin, 1994; Basoglu et al., 1994; Punakami, 1989). The instruments used for this analysis were part of a larger battery administered in the same order to all participants in Phase II of the NMRP. There may be order effects and participant fatigue that we have not accounted for. However, the large number of instruments precluded randomization of questionnaire ordering, and order effects would likely produce only main effects and would not change the relationship between dependent and independent variables. Furthermore, reliability of the CTI-104 was demonstrated. The CTI-104 was not developed to be all-inclusive for all traumatic events experienced by refugees, who have traumatic experiences that predate war and extend beyond migration. This is evidenced by the 5% PTSD risk in our sample associated with not having experienced any CTI-104 events.

We recommend the CTI-104 in its current form to investigators of refugee war-related trauma. It was intentionally developed to contain a broad range of events that represent a broad range of refugees. The 104 items are an efficient set of items from more than 250 events described

by refugees. If administration time is an issue, altering the CTI-104 to have only two responses, “yes, it happened to me,” and “no, it did not happen to me” should not compromise validity. The use of individual CTI-104 scales for efficiency is also appropriate because we have now reported community data for each scale. However, except for the scales that highly correlate with the total CTI-104, these scales could not be used for CTI-104 proxies with the same assumptions about validity.

## REFERENCES

- Aaronson, N. K., Acquadro, C., Alonso, J., Apolone, G., Bucquet, D., Bullinger, M., et al. (1992). International quality of life assessment (IQOLA) project. *Quality of Life Research*, 1, 349–351.
- Antonovsky, A. (1983). The sense of coherence: Development of a research instrument. *Tel Aviv University Newsletter and Research Report*, 1, 11–22. (Available from Tel Aviv University, P.O. Box 49040, Tel Aviv 68978, Israel).
- Antonovsky, A. (1984). A call for a new question—Salutogenesis—and a proposed answer—The sense of coherence. *Journal of Preventive Psychiatry*, 2, 1–13.
- Antonovsky, A. (1987). *Unraveling the mystery of health*. San Francisco: Jossey-Bass.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Basoglu, M., Paker, M., Ozmen, E., Tasdemir, O., & Sahin, D. (1994). Factors related to long-term traumatic stress responses in survivors of torture in Turkey. *Journal of the American Medical Association*, 272, 357–363.
- Basoglu, M., Paker, M., Paker, O., Ozmen, E., Marks, I., Incesu, C., et al. (1994). Psychological effects of torture: A comparison of tortured with nontortured political activists in Turkey. *American Journal of Psychiatry*, 151, 76–81.
- Butcher, J. N. (1991). Psychological evaluation. In J. Westermeyer, C. Williams, & A. Nguyen (Eds.), *Mental health services for refugees* (pp. 111–122). Washington, DC: Government Printing Office.
- Cheung, P. (1995). Acculturation and psychiatric morbidity among Cambodian refugees in New Zealand. *International Journal of Social Psychiatry*, 41, 108–119.
- Clarke, G., Sack, W. H., & Goff, B. (1993). Three forms of stress in Cambodian adolescent refugees. *Journal of Abnormal Child Psychology*, 21, 65–77.
- Crescenzi, A., Ketzer, E., Van Ommeren, M., Phuntsok, K., Komproe, I., & de Jong, J. (2002). Effect of political imprisonment and trauma history on recent Tibetan refugees in India. *Journal of Traumatic Stress*, 15, 369–375.
- deGirolamo, G. D., & McFarlane, A. C. (1996). The epidemiology of PTSD: A comprehensive review of the international literature. In A. Marsella, M. Friedman, E. Gerrity, & R. Scurfield (Eds.), *Ethnocultural aspects of posttraumatic stress disorder* (pp. 33–86). Washington, DC: American Psychological Association.
- Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Science*, 19, 1–15.
- Foa, E. B., Riggs, D. S., Dancu, C. V., & Rothbaum, B. O. (1993). Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *Journal of Traumatic Stress*, 6, 459–473.
- Fullilove, M. (1996). Psychiatric implications of displacement: Contributions for the psychology of place. *American Journal of Psychiatry*, 153, 1516–1523.
- Hollifield, M., Eckert, V., Warner, T. D., Jenkins, J., Krakow, B., Ruiz, J., et al. (2005). Development of an inventory for measuring war-related events in refugees. *Comprehensive Psychiatry*, 46, 67–80.
- Hollifield, M., & Warner, T. D. (1999). The New Mexico Refugee Symptom Checklist (NMRSLC). Unpublished manuscript, University of New Mexico, Albuquerque.
- Hollifield, M., Warner, T. D., Lian, N., Krakow, B., Jenkins, J. H., Kesler, J., et al. (2002). Measuring trauma and health status in refugees: A critical review. *Journal of the American Medical Association*, 288, 611–621.
- Kinzie, J. D., & Manson, S. M. (1987). The use of self-rating scales in cross-cultural psychiatry. *Hospital and Community Psychiatry*, 38, 190–196.
- Leon, A. C., Shear, M. K., Portera, L., & Klerman, G. L. (1992). Assessing impairment in patients with panic disorder: The Sheehan Disability Scale. *Social Psychiatry and Psychiatric Epidemiology*, 27, 78–82.
- Levenson, H. (1981). *Differentiation among internality, powerful others and chance* (Vol. 1). New York: Academic Press.
- McHorney, C. A., Ware, J. E., & Raczek, A. E. (1993). The MOS 36-item Short Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care*, 31, 247–263.

- Mollica, R. F. (n.d.). The Harvard Trauma Questionnaire manual: Indochinese version. Unpublished manuscript, Harvard University, Cambridge, MA.
- Mollica, R. F., & Caspi-Yavin, Y. (1991). Measuring torture and torture-related symptoms. *Psychological Assessment*, 3, 581–587.
- Mollica, R. F., Caspi-Yavin, Y., Bollini, P., Truong, T., Tor, S., & Lavelle, J. (1992). The Harvard Trauma Questionnaire: Validating a cross-cultural instrument for measuring torture, trauma and posttraumatic stress disorder in Indochinese refugees. *Journal of Nervous and Mental Disease*, 180, 111–116.
- Mollica, R. F., McInnes, K., Poole, C., & Tor, S. (1998). Dose-effect relationships of trauma to symptoms of depression and posttraumatic stress disorder among Cambodian survivors of mass violence. *British Journal of Psychiatry*, 173, 482–488.
- Mollica, R. F., McInnes, K., Sarajlic, N., Lavelle, J., Sarajlic, I., Massagli, M. P., et al. (1999). Disability associated with psychiatric comorbidity and health status in Bosnian refugees living in Croatia. *Journal of the American Medical Association*, 282, 433–439.
- Mollica, R. F., Wyshak, G., de-Marneffe, D., Khuon, F., & Lavelle, J. (1987). Indochinese versions of the Hopkins Symptom Checklist-25: A screening instrument for the psychiatric care of refugees. *American Journal of Psychiatry*, 144, 497–500.
- Mollica, R. F., Wyshak, G., Lavelle, J., Truong, T., Tor, S., & Yang, T. (1990). Assessing symptom change in Southeast Asian refugee survivors of mass violence and torture. *American Journal of Psychiatry*, 147, 83–88.
- Momartin, S., Silove, D., Manicavasagar, V., & Steel, Z. (2003). Dimensions of trauma associated with posttraumatic stress disorder (PTSD) caseness, severity and functional impairment: A study of Bosnian refugees resettled in Australia. *Social Science and Medicine*, 57, 775–781.
- Punakami, R. I. (1989). Political violence and mental health. *International Journal of Mental Health*, 17, 3–15.
- Saab, B., Chaaya, M., Doumit, M., & Farhood, L. (2003). Predictors of psychological distress in Lebanese hostages of war. *Social Science and Medicine*, 57, 1249–1257.
- Sheehan, D. (1983). *The anxiety disease*. New York: Scribner.
- Thompson, M., & McGorry, P. (1995). Psychological sequelae of torture and trauma in Chilean and Salvadorean migrants: A pilot study. *Australian and New Zealand Journal of Psychiatry*, 29, 84–95.
- True, W. R., Goldberg, J., & Eisen, S. A. (1988). Stress symptomatology among Vietnam veterans. Analysis of the Veterans Administration Survey of Veterans. II. *American Journal of Epidemiology*, 128, 85–92.
- United Nations. (1985). *United Nations Declaration Against Torture: Article 1*. New York: Author.
- Ware, J. E., Snow, K. K., & Kosinski, M. (1993). *SF-36 health survey manual and interpretation guide*. Boston: New England Health Center, The Health Institute.
- Waters, J. D., & Biernask, P. (1989). Targeted sampling: Options for the study of hidden populations. *Social Problems*, 36, 416–430.
- Weathers, F., Keane, T., King, L., & King, D. (1997). Psychometric theory in the development of posttraumatic stress disorder assessment tools. In J. Wilson & T. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 98–135). New York: Guilford Press.
- Westermeyer, J., & Janca, A. (1997). Language, culture and psychopathology: Conceptual and methodological issues. *Transcultural Psychiatry*, 34, 291–311.
- Winokur, A., Winokur, D. F., Rickels, K., & Cox, D. (1984). Symptoms of emotional distress in a family planning service: Stability over a four-week period. *British Journal of Psychiatry*, 144, 395–399.
- World Health Organization. (2001). *The world health report. Mental health: New understandings, new hope*. Geneva, Switzerland: Author.

Copyright of *Journal of Traumatic Stress* is the property of John Wiley & Sons, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.