CHAPTER 3

Standardized Self-Report Measures of Civilian Trauma and PTSD

FRAN H. NORRIS
JESSICA L. HAMBLEN

This chapter reviews 24 standardized self-report measures for traumatic stress that are suitable, with some modification, for use with adults by professional or lay interviewers or in paper-and-pencil questionnaires. Each scale is described in terms of its content, number of items, and response formats and is evaluated in terms of the available evidence regarding its reliability and validity. We also describe the population or populations on whom the scale was validated. We note strengths and weaknesses but stop short of recommending one scale for all situations. In fact, our assumption is that different scales may be more or less suitable for different purposes in different contexts involving traumatic stress responses.

Because measures of combat-related trauma are described elsewhere in this volume, this chapter focuses on scales that are suitable for studying civilian trauma in clinical or community populations. These populations may include veterans of military service but are not limited to them. The measures described herein are those that either have been significant to this field historically or appear quite promising for future research. In deciding which scales warranted inclusion in this chapter, we relied heavily on the published literature and, to a lesser extent, on information gained from networking with investigators working in this area. The selected scales make up a reasonable cross-section of standardized self-report measures available in the field today.

The scales reviewed here fall into two broad categories: seven that measure DSM-IV posttraumatic stress disorder (PTSD) criterion A, or trauma his-

tories, and 17 that measure DSM-IV PTSD criteria B–D, or symptom histories. The chapter is organized accordingly. Researchers and practitioners should plan on selecting one scale from each category to fully capture the phenomenon of trauma.

DSM-IV PTSD CRITERION A: ASSESSING TRAUMATIC EVENTS

Over the past two decades, the definition of a traumatic event has changed considerably. These changes in definition have a significant impact on what events qualify for PTSD and must be considered when determining what trauma exposure measure to use. In DSM-III, a trauma was defined as a "recognizable stressor that would evoke significant symptoms of distress in almost anyone" (American Psychiatric Association, 1980, p. 238). In 1987, the DSM-III-R definition of trauma was revised to mean an event that is "outside the range of usual human experience and that would be markedly distressing to almost anyone" (American Psychiatric Association, 1987, p. 250). These two definitions were intended to capture catastrophic events that happen with low frequency and to exclude more common events such as simple bereavement, chronic illness, business loss, and marital conflict. The current DSM-IV (American Psychiatric Association, 1994) defines a traumatic event as one in which both of the following were present: "(1) the person experienced, witnessed, or was confronted with an event or events that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others (criterion A1), and (2) the person's response involved intense fear, helplessness, or horror" (criterion A2; pp. 427–428). Thus the current definition has been expanded to include events that would not have been considered in earlier versions because of their frequency, such as personal illness. On the other hand, the definition has been made narrower by requiring a subjective response of fear, helplessness, and horror. Some controversy continues to exist among experts as to exactly which events should be characterized as traumatic.

In this section, we review seven scales in which criterion A is the sole or primary focus. Most scales do not assess criterion A2. The scales are the Traumatic Stress Schedule (TSS; Norris, 1990), the Traumatic Events Questionnaire (TEQ; Vrana & Lauterbach, 1994), the Trauma History Questionnaire (THQ; Green, 1996), the Stressful Life Events Screening Questionnaire (SLESQ; Goodman, Corcoran, Turner, Yuan, & Green, 1998), the Traumatic Life Events Questionnaire (TLEQ; Kubany et al., 2000), the Life Stressor Checklist—Revised (LSC-R; Wolfe, Kimerling, Brown, Chrestman, & Levin 1996), and the Brief Trauma Questionnaire (BTQ; Schnurr, Vielhauer, Weathers, and Findler 1999). Because of our focus on brief measures that can be self-administered, we excluded some measures that are seriously worthy of consideration in situations in which a more in-depth assessment (e.g., Poten-

tial Stressful Events Interview; Kilpatrick, Resnick, & Freedy, 1991) or clinician administration (e.g., Evaluation of Lifetime Stressors; Krinsley, Gallagher, Weathers, Kaloupek, & Vielhauer, 1997) is feasible.

Also excluded from this chapter were measures that detail the experiences of specific trauma populations, such as adult survivors of child abuse (e.g., Briere, 1992), refugees (Mollica et al., 1995), or victims of natural disasters (e.g., Norris & Kaniasty, 1992). Their exclusion should be taken neither as a criticism nor as a statement that such measures are unimportant. Rather, we excluded them because such instruments almost inevitably need to be tailored to the specific event, population, and context and thus are difficult to describe or evaluate in a standardized way. The scales described here screen for the occurrence of potentially traumatic events more broadly. They are best used to supplement more targeted assessments of a focal event or experience. In clinical practice, one of these measures could be used to identify experiences that might subsequently be probed for greater detail in a less structured way.

For each scale, we note which events are specifically assessed and provide evidence for the scale's reliability and validity where such data exist. For self-reported trauma histories, reliability evidence has typically taken the form of test–retest correlations. Internal consistency (e.g., Cronbach's alpha) is not applicable to event measures because the experience of one event does not necessarily imply the experience of another. It should be noted, however, that length restrictions prevent us from providing the exact wording of the events. For any scale of interest, we recommend obtaining the specific instrument to determine whether the wording is appropriate for the intended use.

Validity is difficult to establish unequivocally for these scales. To the extent that face validity may be counted, construct validity has been used most often; that is, checklists of events typically "seem" reasonable. Criterion validity is virtually impossible to establish because no external standard of accuracy exists. Concurrent validity is sometimes evidenced when similar estimates of trauma prevalence are yielded by different scales (see Resnick, Falsetti, Kilpatrick, & Freedy, 1996). In our opinion, content validity could receive much more attention than it has in the development of these scales. Any list of life events, traumatic or otherwise, is a sample representing a larger population of life events. Bruce Dohrenwend (e.g., Dohrenwend, Krasnoff, Askenasy, & Dohrenwend, 1978) must be credited with directing researchers' attention to the fact that decisions made in constructing the list will ultimately determine the kinds of inferences and generalizations that can be made. He raised two basic and related questions: How do we define the events to be sampled? And, what is the population of events from which the sample is to be drawn? Life-event-scale developers seldom have described *explicitly* the population of events that the items on their scales purportedly represent. Some consensus among researchers is implicit in these measures: If we exclude the contributions of open-ended or "catch-all" items, no trauma scales reviewed here are so broad as to include all events demanding readjustment (e.g., moving to a new place) or even all undesirable life events (e.g., losing a job). Yet consensus

still has not emerged with regard to just where to draw the line between traumatic events and other undesirable events. This is a critical issue for content validity, which, like construct validity, is often established more on conceptual than on empirical grounds (Wilson, 1994).

Traumatic Stress Schedule

Among the earliest published self-report measures was the Traumatic Stress Schedule (TSS), developed by Fran Norris (1990) as a short screening instrument for assessing traumatic stress in the general population. The format of the scale followed from two basic assumptions: first, that it was important to assess rates of impairment within specific event-defined populations (e.g., crime victims) in addition to assessing those rates within the population at large; and second, that it is important to quantify stressful experiences generically, using descriptors such as life threat, loss, and scope that are not unique to any one event.

In selecting the items for the scale, Norris relied on the DSM-III-R (American Psychiatric Association, 1987) definition of criterion A, in which the defining feature was that events should be beyond the realm of normal human experience. For research purposes, she proposed a more restricted definition of the relevant event population as that involving "violent encounters with nature, technology, or humankind" (p. 1706). She defined a violent event as one that (1) is marked by extreme and/or sudden force, (2) involves an external agent, and (3) is typically capable of arousing intense fear or aversion. The events were selected to provide a reasonable cross-section of this population of events. The scale, as initially published, assessed eight potentially traumatic events: (1) robbery, a theft involving force or threat of force; (2) physical assault; (3) sexual assault, that is, forced unwanted sexual activity of any kind; (4) loss of a loved one through accident, homicide, or suicide; (5) personal injury or property loss as a result of fire, severe weather, or disaster; (6) being forced to evacuate or otherwise learning of an imminent danger or hazard in the environment; (7) having a motor vehicle accident serious enough to cause injury to one or more passengers; and (8) "some other terrifying or shocking experience." The current version has 10 items; fire was separated from disaster, and serving in combat was added. For each stressor, six dimensions are assessed: loss (the tangible loss of persons or property), scope (the extent to which persons other than the respondent were affected by the incident), threat to life and physical integrity (including actual physical injury), blame, familiarity, and four probes assessing posttraumatic stress reactions. This last dimension of posttraumatic stress shifted the focus from assessing the characteristics of the stressor to assessing the response to that stressor, and it can be used as a brief stress measure for each endorsed event.

The event portion of this scale has performed well in research (see Norris, 1992). Norris and Perilla (1996) reported a test–retest correlation of .88 between English and Spanish versions completed by 53 bilingual volunteers 1

week apart. Estimates of exposure to trauma have been strikingly stable across purposive and random community samples. Excluding events that were the focus of these studies, such as Hurricanes Hugo and Andrew, sample frequencies of exposure to one or more traumatic events (using an "ever" time frame) have ranged from 62 to 75%, with an average of 69%. Quite reasonably, higher frequencies (82%) emerged in a study of family members of homicide victims in inner-city Atlanta (M. Thompson, personal communication, March 24, 1995). The symptom portion of this scale is moderately reliable (alpha =.76) and may be useful as a quick screen for posttraumatic stress, but we do not recommend its use as a measure of PTSD. It does not assess all 17 criterion symptoms and assesses neither duration of distress nor functional impairment. The scale does not include an assessment of A2.

The strength of this scale is that is a brief measure that assesses criterion A1 events only. It is alone in establishing equivalence between English and Spanish versions. The probes provide information on experiences that cross particular events (e.g., number of life-threatening events). In addition, the symptom portion of this measure can be used as an indication of posttraumatic stress. The scale does not provide information on age at time of trauma, does not ask specifically about childhood events, and does not query about fear, helplessness, or horror.

Traumatic Events Questionnaire

The Traumatic Events Questionnaire (TEQ), developed by Scott Vrana and Dean Lauterbach (1994), assesses 11 specific traumatic events: (1) combat, (2) large fires/explosions, (3) serious industrial/farm accidents, (4) sexual assault/rape (forced unwanted sexual activity), (5) natural disasters, (6) violent crime, (7) adult abusive relationships, (8) physical/sexual child abuse, (9) witnessing someone being mutilated, seriously injured, or violently killed, (10) other life-threatening situations, and (11) violent or unexpected death of a loved one. Two nonspecific questions, "other event" and "can't tell," complete the scale. Probes assess dimensions such as life threat and injury after any affirmative response.

Over a 2-week test–retest interval, very high reliability for the total scale was observed (.91) in a sample of 51 students (Lauterbach & Vrana, 1996). In another student sample (N = 440), 84% reported at least one event, which is higher than other rates that have been reported in the literature. Endorsement of "catch-all" events was especially high: 30% had some other life threatening experience, 23% had some other event, and 9% endorsed "can't tell." Specific events also showed high prevalence rates. A particularly striking statistic was that almost half (49%) of Vrana and Lauterbach's (1994) sample reported having experienced a violent or unexpected death of a loved one. This scale defined the event population to include unexpected natural deaths, as well as those due to violence from technology or humankind. This expansion is consistent with the present wording of criterion A1 in DSM-IV.

Like the TSS, this measure provides a good, quick screen for traumatic events. Criterion A1 is asked about specifically for each item. The inclusion of events about which the respondent "can't tell" is interesting and constitutes both a strength (comprehensiveness) and shortcoming (the researcher can't tell if the event meets criterion). The scale does not inquire about age at time of trauma or assess for criterion A2.

Trauma History Questionnaire

The Trauma History Quesionnaire (THQ) was developed by Bonnie Green and her associates at Georgetown University (Green, 1996). The THQ aims to provide a comprehensive assessment of exposure and to be suitable for both research and clinical populations. The scale has 24 items: (1) mugging, (2) robbery—a theft by force, (3) break-in with respondent present, (4) break-in with respondent absent, (5) serious accident at work, in a car, or somewhere else, (6) natural disaster with respondent or loved ones in danger, (7) disaster of human origin with respondent or loved ones in danger, (8) toxin exposure, (9) other serious injury, (10) other situation in which respondent feared being killed or injured, (11) witnessed serious injury or death, (12) handled/seen dead bodies, (13) close friend or family member murdered or killed by a drunk driver, (14) spouse, romantic partner, or child died, (15) respondent had serious or life-threatening illness, (16) someone close experienced serious or life-threatening illness, injury, or unexpected death, (17) combat, (18) forced intercourse, oral, or anal sex, (19) forced touching of private parts, (20) other unwanted sexual contact, (21) aggravated assault, (22) simple assault, (23) beaten, spanked, or pushed hard enough to cause injury, and (24) any other extraordinarily stressful situation or event. Each event is followed by probes assessing the number of times that event has occurred and the respondent's age at the time.

Green (1996) provided reliability data collected from 25 female participants that were tested twice over a 2–3 month interval. Excluding the total severe-threat index that received a stability coefficient of only .14, test–retest correlations ranged from .54 for total bereavement to .92 for total crime. Scale means were higher in an outpatient sample than in a university sample, which provides some additional evidence of validity. Green's take on the population of relevant events is the broadest of all those reviewed here, as this scale includes deaths and illnesses of significant others, even if expected and due to natural causes, which many would argue should not qualify as meeting criterion A1. This strategy was chosen because, in the research, respondents who provided affirmative responses were interviewed in more detail about their experiences. Two-thirds of the students in Green's pilot study ($N = 423$) reported that someone close to them had become seriously ill at some time, making the frequency for this one event as high as the total frequency across events obtained using the TSS.

In the past few years, the scale has been used in a variety of populations, including cancer, epilepsy, and chronic pain patients, battered women, per-

sons with serious mental illness, and adult offspring of Holocaust survivors. The strengths of this measure include its comprehensiveness and careful wording. This measure includes a range of both traumatic and stressful life events. Additional information is available about the frequency of the event and the age at time of trauma. There is no assessment of fear, helplessness, or horror.

Traumatic Life Events Questionnaire

The Traumatic Life Events Questionnaire (TLEQ) was described by Edward Kubany and colleagues (2000). The scale was designed for both clinical and research purposes. The present version, expanded from the experimental version described by Norris and Riad (1997), assesses the occurrence of 23 events: (1) natural disaster, (2) motor vehicle accident involving injury or death, (3) other accident involving injury or death, (4) combat, (5) sudden and unexpected death of a close friend or loved one due to accident, illness, suicide, or murder, (6) loved one surviving life-threatening illness, accident, assault, (7) life-threatening illness, (8) mugging or robbing by someone with a weapon, (9) physical assault by an acquaintance or stranger, (10) witnessing someone being attacked or assaulted, (11) being threatened with death or bodily harm, (12) childhood physical abuse, (13) witnessing severe family violence, (14) physical abuse from intimate partner, (15) childhood sexual touching by someone at least 5 years older (probes for force, penetration), (16) childhood sexual touching by someone less than 5 years older, (17) adolescent unwanted sexual activity (probes for force, penetration), (18) adulthood unwanted sexual activity (probes for force, penetration), (19) sexual harassment, (20) stalking, (21) miscarriage, self or partner, (22) abortion, self or partner, and (23) other extremely disturbing or distressing experience. As with the THQ, many would argue that some of the events included would not qualify as meeting criterion A1, such as sexual harassment or abortion. This scale provides a good match to criterion A2 by including a probe after each experienced event that reads, "Did you experience intense fear, helplessness, or horror when it happened?" Additional questions ask about frequency, injury, whether any of the events occurred within the past 2 months or 12 months, and which event caused the most distress. There is also a brief version of the scale that assesses for the same events but with fewer probes.

Kubany et al. (2000) described the results in a series of five studies. The first four were conducted using the earlier 16-item version; the last study was conducted using the expanded 21-item version. In the first study, the authors generated a preliminary version of the measure and sent it to seven published experts in the area. On average, the reviewers believed that the items were worded "very well" and sampled the range of events "very well." In Study 2, 49 patients completed the TLEQ twice, over a 60-day interval. When assessed item by item, test–retest percent agreements averaged 83%. Kappas varied widely because some of the events assessed were extremely rare. The stability was lowest for items assessing "other" accidents and childhood sexual abuse by someone less than 5 years older. In Study 3, 51 veterans completed the

TLEQ two times. The length of the interval varied from 5 to 45 days, with a median of 13 days. The results were quite similar to those of the second study; percent agreements averaged 84%. In Study 4, 62 undergraduate students completed the standard self-report version of the TLEQ. One week later, they were interviewed using a structured measure with similar content. Percent agreements were again high for most items. There were no significant differences in proportions disclosing traumatic experiences across the two modalities. In Study 5, 42 members of a support group for battered women completed the 21-item TLEQ 2 weeks apart. Overall percent agreement was 86%.

This scale provides information on a range of potentially traumatic events and is unusual in that it assesses for both criteria A1 and A2. The authors have done an exceptional job of researching the scale's psychometric qualities. The scale has some novel inclusions, such as sexual harassment, abortion, and miscarriage, although there is some debate as to whether these events should be included as criterion A1 events.

Stressful Life Events Screening Questionnaire

The Stressful Life Events Screening Questionnaire (SLESQ) is a 13-item self-report screening measure designed to assess lifetime exposure to potentially traumatic events (Goodman et al., 1998). Trauma was defined according to DSM-IV as an event that involves actual or threatened death or serious injury or a threat to physical integrity of self or others. The SLESQ was designed to be brief. The measure provides more detail on interpersonal trauma than does the TSS but less information about exposure to hazards such as fire or disasters. The events are: (1) life-threatening illness, (2) life-threatening accident, (3) robbery/mugging (4) loss of loved one because of accident, homicide, suicide, (5) forced intercourse, oral, or anal sex, (6) attempted forced intercourse, oral, or anal sex, (7) unwanted sexual touching, (8) childhood physical abuse, (9) domestic violence, (10) threats with weapons, (11) being present when another person was killed, injured, or assaulted, (12) other injury or life threat, and (13) other extremely frightening or horrifying event. For each experienced event, the questionnaire asks for the respondent's age at the time of the trauma. Probes vary across events to provide more detail on the nature of the event.

Psychometric data were collected from a sample of 140 college students assessed twice 2 weeks apart (Goodman et al., 1998). At least one event was reported by 72% of the respondents, a rate similar to others reported in the literature. The correlation between the total number of events reported at Time 1 and the total number reported at Time 2 was .89. Kappas for specific events averaged .73. The least reliably assessed events (i.e., kappa < .60) were attempted rape, witnessing injuries or trauma to others, other injury/life threat, and other extreme event.

The SLEQ is a brief, carefully researched measure of trauma that would be useful in many situations. This scale provides information on age at time of

trauma, frequency, and life threat. The measure includes an "other" category. However, there is no assessment of criterion A2.

Life Stressor Checklist—Revised

The Life Stressor Checklist—Revised (LSC-R) was designed to screen for events that would meet DSM-IV criterion A, as well as for some events that are stressful but unlikely traumatic. The measure has a special focus on events that may be relevant to women, such as abortion, but can also be used with men. Developed by Wolfe et al. (1996), it assesses 30 events, including: (1) serious disaster, (2) serious accident, (3), witnessing a serious accident, (4), close family member being sent to jail, (5), being sent to jail, (6) being in foster care or put up for adoption, (7) parents separating or divorcing, (8) separation or divorce, (9) serious financial problems, (10) serious physical or mental illness, (11) emotional neglect, (12) physical neglect, (13) abortion/miscarriage, (14) separation from child against one's will, (15) severe physical or mental handicap of one's child, (16) primary responsibility for someone with severe mental or physical handicap, (17) sudden or unexpected death of someone close, (18) death of someone close, (19) witnessing family violence, (20) seeing a robbery, mugging, or attack, (21) being robbed, mugged, or attacked, (22) physical abuse, (23) physical assault, (24) sexual harassment, (25) forced genital touching before age 16, (26) forced genital touching after age 16, (27) forced intercourse before age 16, (28) forced intercourse after age 16, (29) "other," (30) being seriously upset by any of these events happening to someone close, even though the respondent did not see it. For each endorsed event, respondents are asked between two and five follow-up questions, depending on the event, including: How old were you when it happened/started; How old were you when it ended; Did you believe that you/someone else could be killed or seriously harmed; At the time, did you experience feelings of fear, helplessness, or horror, and how much has it affected your life in the past year? Thus this measure explicitly assesses for both criteria A1 and A2.

McHugo et al. (2004) provided psychometric data for an adapted version of the measure, collected as part of the Women, Co-Occurring Disorders, and Violence Study. Primary differences between the LSC-R and the adapted version used in the study include: the omission of the A2 probe, the addition of several stressors (homelessness and unwanted sex for money or goods), and the rewording of a few items (such as combining abortion and miscarriage). Data were collected on 2,729 women, who were recruited into the study if they had a diagnosis of both mental and substance use disorders and if they reported experiencing physical or sexual abuse during their lifetimes. A test–retest sample was completed on a subset of 186 women who completed the measure on average 7 days later. Kappas ranged from a low of .52 for physical abuse to a high of .97 for miscarriage and averaged .70. Percent agreement ranged from a low of 79% for serious physical or mental illness to a high of 98% for miscarriage.

The Life Stressor Checklist is the longest measure reviewed in this chapter, because it encompasses both potentially traumatic and other seriously stressful life events. This scale asks about age at time of event, assesses for criteria A1 and A2, and asks about how much the event has affected the person in the past year. It is particularly sensitive to the stressors of women and has been shown to be well tolerated in consumer samples.

Brief Trauma Questionnaire

The Brief Trauma Questionnaire (BTQ), developed by Paula Schnurr and colleagues (1999), assesses 10 traumatic events: (1) combat, (2) serious car accident, (3) major natural or technological disaster, (4) life-threatening illness, (5) physical punishment as child, (6) physical assault, (7) unwanted sexual contact, (8) other situation in which respondent was seriously injured or feared being seriously injured or killed, (9) violent death of close friend or family member, and (10) witnessing a situation in which someone was seriously injured or killed or in which respondent feared someone might be seriously injured or killed.

Although the psychometrics of the BTQ are only currently being established, it is included here as a promising new measure. One of its strengths is that it includes explicit assessment of criterion A1. For all endorsed traumatic events, respondents are asked if they thought their lives were in danger or if they thought they might be seriously injured or were in fact injured. Perceived life threat as measured by the BTQ has been shown to be related to higher dissociation scores (Morgan, Hazelett, Wang, Richardson, Schnurr, & Southwick, 2001). In a study of more than 400 military veterans from World War II and the Korean conflict, interrater reliability was established on a subset of interviews (Schnurr, Spiro, Vielhauer, Findler, & Hamblen, 2002). Kappa coefficients for the presence of trauma that met DSM criterion A1 were above .70 (range .74–1.00) for all events except for illness (.69) and "other life-threatening events" (.60).

The BTQ is a brief measure of trauma that explicitly assesses for criterion A1. It does not inquire about criterion A2, age at time of trauma, or childhood events. Preliminary data look promising.

Summary

Table 3.1 summarizes the descriptions of these seven measures of potentially traumatic events. For each scale, the table lists the number of event items included in the scale, the type of data provided, evidence of stability, the population that the measure was developed on, and whether or not the scale assesses for criteria A1 and A2. Evidence of validity was not included because none of these scales are especially well validated, nor are any apparently especially weak in this regard. General guidelines are provided for considering which scales may be the most useful in different settings.

TABLE 3.1. Summary Descriptions of Seven Standardized Self-Report Measures of Trauma Exposure

Scale	Number of event items	Evidence of stability	Population developed on	Assesses for criterion A1	Assesses for criterion A2
Traumatic Stress Schedule	10	TR, total no. = .88	Multicultural	Yes	No
Traumatic Events Questionnaire	11	TR, total no. = .91	Undergraduates	Yes	No
Trauma History Questionnaire	24	TR, by type = .54–.92	Female undergraduates	Partial	No
Traumatic Life Events Questionnaire	17	*M* agreement, by type = 83%	Variety	Partial	Yes
Stressful Life Events Screening Questionnaire	13	TR, total no. = .89 *M* kappa, by type = .73	Undergraduates	Yes	No
Brief Trauma Questionnaire	10	Kappa, by type = .60–1.0	Veterans	Yes	No
Life Stressor Checklist—Revised	30	Agreement, by type = 79–98%	Women	Partial	Yes

Note. Tr, test–retest; na, data not available.

Perhaps the most basic issue to consider in comparing the measures is how well they satisfy criteria A1 and A2. In terms of criterion A1, all of the scales include a range of traumatic events, but they differ in the definitional boundaries of the relevant population of events. The TSS appears to use the most objective and restricted definition. Other measures, such as the THQ, TLEQ, and LSC-R, use a broader definition and include events that are arguably not traumatic and would not satisfy criterion A1. Decisions regarding how broadly or narrowly to define the relevant domain of events depend on the assessor's intent. For example, clinicians may find that the broader measures can better inform their clinical work, whereas researchers may prefer a more restricted range of included events.

A second important issue is assessment of criterion A2. Only the TLEQ and the LSC-R explicitly ask respondents about whether their subjective reactions to the event included fear, helplessness, or horror. In some cases, the inclusion of A2 may not be an issue. For example, event checklists are often used to screen for trauma exposure by identifying a single or most upsetting experience, and then additional questions are used to determine if the event meets criteria A1 and A2. However, if the assessor needs to know how many criterion A traumas the respondent has experienced, it is essential that both criteria A1 and A2 be assessed for each event. It should be noted that many of these measures were structured so that the event is a gate question, followed by additional probes if answered affirmatively. In such cases, it would not be difficult to add a probe that explicitly assesses A2.

A third issue to consider is whether the range of traumatic events being assessed is sufficient. Clearly, scales with more items have a greater likelihood of identifying traumatic events. However, they take longer to administer, and some include items that would not qualify as traumatic. Multipurpose surveys may find it difficult to include one of the longer scales. Also, some measures give more attention to certain types of events. For example, for studying long-term consequences of childhood trauma, measures that explicitly differentiate child abuse from adulthood assault are recommended. A review of the items in the longer scales shows that it is usually exposure to sexual assault and domestic violence that receives additional explicit attention, and this may be very useful in many contexts.

One issue that requires additional consideration is the use of "catch-all" events that compromise specificity. All of the reviewed scales used this technique. The reasoning behind the inclusion of these items is clear. It would be too difficult, costly, and unacceptable to researchers to enumerate every traumatic event that might conceivably occur. Such items also give respondents the chance to report experiences that were important to them, which can be informative, as well as helpful in building rapport. On the other hand, these items may be tapping into personal crises and failures that are not truly in the domain of traumatic life events.

Norris and Riad (1997) reviewed the responses to this open-ended question provided by persons who participated in their study of Hurricane An-

drew. Of 404 respondents, 36 (9%) reported some other "shocking or terrifying experience" on the TSS. Fourteen people told of events that clearly qualified as traumatic, according to DSM-III-R definitions, but these events were not asked about directly (or specifically enough) in other TSS questions (e.g., being at the scene of a bank robbery, in a train accident, threatened with a gun). Five told of events that should have been picked up by other items but for some reason were not. For these 19 respondents (roughly half), the item served its purpose: catching other trauma histories not elsewhere recorded. An additional six people described life-threatening or very serious illnesses experienced by themselves (going blind temporarily, surgery) or loved ones (grandfather's cancer) that qualify under DSM-IV PTSD criteria, although they were not in the domain of experiences Norris was initially attempting to capture with the TSS. Three people mentioned deaths due to natural causes of loved ones, and seven told of other unfortunate (husband convicted of murder, son in prison) or unusual (paranormal) experiences. Thus roughly 28–44% of the events captured by this item (2–3% of the total sample) would not qualify under more restrictive definitions. When reviewing the literature across these scales, it is striking that catch-all items seem to have even higher rates of endorsement on longer measures. Compared with 9% of Norris and Riad's (1997) sample, 23% of Vrana and Lauterbach's (1994) sample reported some other event. Even in Green's (1996) sample, 14% reported some other event, although the THQ asked about 23 specific events, including serious illnesses of respondent and others and deaths of close family members, regardless of cause. When affirmative answers were explored in subsequent interviews, Green found that few of the events qualified as criterion A events (personal communication, April 12, 1995).

A related issue with these items was highlighted by Vrana and Lauterbach's (1994) finding that a high percentage of TEQ respondents rated the "other event" as their very worst. This finding may reflect an intrinsic bias wherein participants primarily note another event *if* it was their subjective worst. Conceivably, all respondents have experienced undesirable changes in their lives, but they do not always bring these to mind. The issue here is again one of content validity. If "traumatic events" and "undesirable events" are synonymous terms, these scales need to be expanded to capture the range of undesirable events that have been important in life events research more generally (e.g., Dohrenwend et al., 1978). In our opinion, it is better for measures of PTSD criterion A to focus more specifically on a clearly defined population of events. This is not to say that other events are not important in the lives of individuals but simply that they are beyond the domain of concern for these measures. In many studies, it is advisable to include a scale of normative life events, in addition to a scale of traumatic life events. Perhaps this is analogous to developing a scale for anxiety rather than or in addition to a scale of depression or generalized distress. To summarize, these catch-all questions seem necessary, but the responses they elicit may be seriously compromising the content validity of all of the measures that were reviewed here. Regardless of

which scale is selected, the researcher or clinician should probe for content of these events.

Finally, although we were unable to detail the exact wording of each question on the reviewed measures, such wording is nonetheless quite important. Overall, we believed the included measures were careful and clear in their wording. One lesson from earlier research that has clearly been learned is that items must include behavioral descriptions of events. For example, not one of these measures used the term "rape." Instead, each referred to unwanted or forced sexual activity. Wording of prospective measures should be reviewed prior to use.

Thus determining what measure is best really depends on the intended purpose. The TSS, TEQ, and BTQ are brief screens for traumatic stress, which may increase their appeal to researchers focusing on many constructs in addition to trauma, whereas the THQ and TLEQ aim to provide comprehensive trauma histories and may be more suitable for research in which length of the instrument is not an issue. The SLESQ provides an exceptional amount of information about sexual trauma and interpersonal violence. Still other measures, such as the LSC-R, may be most suitable for occasions on which the researcher or clinician does not wish to confine the domain of concern to trauma per se but seeks to include other seriously stressful events, such as divorce. Sources for obtaining these seven measures are provided in Appendix 3.1.

DSM-IV PTSD CRITERIA B–D: INTRUSTION, AVOIDANCE, AND AROUSAL

In this section, we review 17 scales that purport to measure symptomatic criteria for PTSD. We describe each scale in terms of its length and format, provide some background regarding its development, and evaluate its psychometric properties. Rules for establishing reliability and validity are developed much better for symptom measures than for event measures, which raise the standards by which these symptom scales are judged. Regarding reliability, it is usually important for symptom measures to establish both internal consistency and stability over time. Validity data for symptom scales usually takes the form of criterion validity or construct validity. Sometimes, criterion validity is established in terms of a scale's correlations with more established measures in the field. A PTSD scale should correlate highly—but not too highly—with measures of general psychopathology and should correlate most highly with other measures of posttraumatic stress. Most highly regarded is evidence that the scale can correctly classify subjects into diagnostic groups, determined by some independent criterion. Statistics are usually provided regarding the measure's sensitivity (the proportion of cases correctly classified) and specificity (the proportion of noncases correctly classified). Construct validity is important as well. In this case, validity is usually established by showing that

scale scores differ across groups having different objective trauma histories. Sometimes, construct validity is examined by exploring how well the observed factor structure of the scale conforms to theoretical predictions.

The difficulty of creating a measure that is both sensitive and specific to PTSD should not be taken lightly, because the disorder is composed of a broad, if unique, constellation of psychological symptoms. In the tradition of the American Psychiatric Association's DSM, these symptoms are grouped into three clusters. DSM-IV PTSD criterion B is the reexperiencing of the trauma. Intrusive symptoms, such as thinking about the event when the individual does not intend to, having nightmares or flashbacks, or being suddenly reminded of the event by environmental stimuli are extremely common experiences following traumatic life events. Criterion C encompasses avoidance and a numbing of responsiveness to the external world. Often, trauma victims avoid people and places that remind them of the event, feel estranged from other people, or lose interest in things they formerly enjoyed. Criterion D refers to a varied collection of symptoms indicative of increased arousal. Being jumpy, easily startled, or hyperalert, having trouble sleeping or concentrating, or feeling easily angered characterize criterion D. To satisfy DSM-IV criteria for PTSD, the person must show at least one intrusion symptom, three avoidance symptoms, and two arousal symptoms.

The measures that are included here are not the only self-report measures of PTSD but are those that appear to be the most commonly used in recent research. These measures reflect varying strategies for assessing PTSD. Perhaps now the most common strategy is to create measures that map directly onto the 17 criterion symptoms included in DSM-IV. Such measures include the National Women's Study PTSD Module (Kilpatrick, Resnick, Saunders, & Best, 1989); the Posttraumatic Stress Diagnostic Scale (Foa, Cashman, Jaycox, & Perry, 1997), which evolved from Foa's earlier PTSD Symptom Scale (Foa, Riggs, Dancu, & Rothbaum, 1993); the PTSD Checklist (PCL) developed by Weathers, Litz, Herman, Huska, and Keane (1993); the Davidson Trauma Scale (Davidson et al., 1997); the Purdue PTSD Scale (Lauterbach & Vrana, 1996); the PTSD Interview (Watson, Juba, Manifold, Kucala, & Anderson, 1991); the Screen for Posttraumatic Stress Symptoms (Carlson, 2001); and the Self-Rating Interview for PTSD (Hovens, Bramsen, & van der Ploeg, 2002). The second strategy has been to develop scales that assess symptoms of posttraumatic stress continuously and in a manner less rigidly tied to DSM guidelines. The Posttraumatic Symptom Scale (Holen, 1990), the Penn Inventory (Hammarberg, 1992), the Trauma Symptom Checklist—40 (Briere & Runtz, 1989), and the Trauma Symptom Inventory (Briere, 1995) are examples here. The Impact of Event Scale (IES) is described by Weiss (Chapter 7, this volume) so will not be included, though this is the group of scales with which it would belong. The third strategy has been to derive PTSD subscales from larger symptom inventories that are commonly used in clinical practice and research. Examples here are the Minnesota Multiphasic Personality Inventory (MMPI; Keane, Malloy, & Fairbank, 1984) and the Symptom Checklist—90 (SCL-90;

Saunders, Arata, & Kilpatrick, 1990; Ursano, Fullerton, Kao, & Bhartiya, 1995) PTSD scales. This strategy is most useful in settings in which MMPI and SCL-90 data are being collected and in which it would be difficult to add a measure specifically focused on PTSD. A fourth strategy has been to develop measures that are tailored to assess culturally relevant outcomes. The Harvard Trauma Questionnaire (Mollica et al., 1992) is the premiere example of this approach. We also included the Revised Civilian Mississippi Scale (Norris & Perilla, 1996) under this strategy rather than the second because it was established so as to be equivalent in English and Spanish. The remainder of this chapter is organized according to this scheme.

PTSD SCALES THAT CLOSELY FOLLOW DSM SYMPTOM CRITERIA

National Women's Study PTSD Module

The National Women's Study (NWS) PTSD Module developed by Dean Kilpatrick and colleagues (1989) was revised from the version of the Diagnostic Interview Schedule (DIS) used in the National Vietnam Veterans Readjustment Survey. Designed for use by lay interviewers, the measure begins with 20 symptom items that span the range of symptoms associated with PTSD. Questions are first answered yes or no. Then, dates of first and last experiences of that symptom are recorded for all affirmative responses. None of the items is anchored to the specific event or events experienced. This characteristic of the scale makes it easy to administer to people with multiple or complex trauma histories. Another advantage of this assessment approach is that the respondent is not required to attribute the symptom to a specific experience, a characteristic for which the original DIS was criticized (Solomon & Canino, 1990). However, open-ended probes are used to assess symptom content in specific instances. For example, if an individual reports nightmares, he or she is asked what the nightmares are about. After the symptom questions, the scale assesses amnesic experiences, timing and co-occurrence of symptoms, and functional impairment. The scale has typically been scored to yield dichotomous measures of lifetime and current PTSD rather than to yield a continuous measure of PTSD symptomatology.

Because of the dichotomous nature of the scoring algorithms, data regarding the scale's reliability and validity have taken the form of kappa coefficients. Resnick, Kilpatrick, Dansky, Saunders, and Best (1993) reported that stability over a 1-year interval for lifetime PTSD was adequate (kappa = .45). Data collected from clinical cases as part of the DSM-IV field trials provided evidence of concurrent validity. Kappa coefficients of agreement between a PTSD diagnosis made on the basis of this module and the Structured Clinical Interview for DSM-III-R (SCID) were .71 for current PTSD and .77 for lifetime PTSD. These analyses also indicated that the NWS Module had high sensitivity for lifetime (.99) and current (.96) PTSD. Specificity was somewhat lower: .79 for lifetime and .80 for current PTSD.

Posttraumatic Stress Diagnostic Scale

The Posttraumatic Stress Diagnostic Scale (PTDS) was developed by Edna Foa and her colleagues (1997) to address the various shortcomings of preexisting self-report measures. It follows DSM-IV closely. The PTSD Symptom Scale (Foa et al., 1993) that was described in Norris and Riad (1997) was the precursor to the PTDS. The PTDS is a measure of current (previous month) PTSD anchored to the single event that "bothers" the respondent the most. Thus the PTDS actually begins with a 12-item checklist of traumatic events followed by a question that asks the respondent to identify the single event that has disturbed him or her the most in the previous month. Criteria A1 and A2 are then assessed by 4 dichotomous questions regarding physical injury, threat, terror, and helplessness. This section is followed by 17 symptom items answered on a 4-point Likert scale of frequency during the previous month. The scale concludes with 9 questions that address functional impairment. An excellent feature of the scale is that it yields both dichotomous (diagnostic) and continuous scores.

Foa et al. (1997) presented impressive validation data, derived from a sample of 248 men and women, of whom 110 composed a retest sample. Participants were excluded from the retest sample if they selected a different event that bothered them the most, which was not uncommon. Internal consistency was high for each symptom cluster B–D (alphas = 78–.84) and for the total scale (alpha = .92). Test–retest reliability coefficients over 2–3 weeks were likewise high for each cluster (r's = .77–.85) and for the total scale (r = .83). When scored continuously, the PTDS correlates highly with other symptom measures, such as the Beck Depression Inventory (BDI; Beck, Ward, Mendelsohn, Mock, & Erbaugh, 1961) (.79) and the IES—Revised, Intrusion subscale (Weiss & Marmar, 1997) (.78). Respondents in the initial sample were classified as meeting diagnostic criteria for PTSD using the SCID-PTSD module. The PTDS and SCID yielded the same diagnosis 82% of the time (kappa = .65). The sensitivity of the PTDS was .89, specificity .75.

Foa's scale has a number of excellent features. It provides both a DSM-IV diagnosis and a severity scale. It is both internally consistent and stable when used to study the aftermath of the same trauma over a 2- to 3-week interval. Moreover, it showed good agreement with SCID diagnosis. The high correlations with depression raise questions regarding discriminant validity, but this shortcoming reflects a general issue with the PTSD diagnosis rather than an issue specific to this scale.

PTSD Checklist

The PTSD Checklist, Civilian Version (PCL-C) was developed by Frank Weathers and his colleagues at the National Center for PTSD (1993). The scale consists of 17 questions that now correspond to DSM-IV. Respondents are asked how often they have been bothered by each symptom in the previous month on a 5-point severity scale. According to the authors, the ques-

tions may be worded generically to refer to "stressful experiences in the past" (PCL-C) or to describe reactions to a specific event (PCL-S). Initial psychometric data were derived by using a military version of the PCL (PCL-M) in a sample of Vietnam veterans, in which the prevalence of PTSD was high. Internal consistency coefficients were very high for the total scale (.97) and for each subscale (.92–.93). Test–retest reliability over 2–3 days was .96. The PCL-M correlated highly with the Mississippi Scale for Combat-Related PTSD (Keane, Caddell, & Taylor, 1988) (.93), the PK scale of the MMPI (.77), and the Impact of Event Scale (.90). In this sample, the PCL-M was quite predictive of PTSD as assessed with the SCID; a cutoff score of 50 had a sensitivity of .82, a specificity of .83, and a kappa of .64. (The reader should note that cutoff scores may vary depending on the prevalence of disorder in a sample.)

Other researchers have also presented evidence supporting the reliability and validity of the PCL-C or PCL-S. In a sample of 40 motor vehicle accident and sexual assault victims, of whom 18 had PTSD on the Clinician-Administered PTSD Scale (CAPS), Blanchard, Jones-Alexander, Buckley, and Forneris (1996) found an alpha of .94 and an overall correlation between total PCL-S and CAPS scores of .93. They found that a score of 44 (rather than 50) maximized diagnostic efficiency (sensitivity of .94, specificity of .86, overall efficiency of .90). In a sample of individuals in France who had experienced a variety of events, Ventureya, Yao, Cottraux, Note, and De May-Guillard (2002) reported excellent internal consistency (.86) and test–retest reliability (.80) for the total PCL-S score. Using the cutpoint of 44 recommended by Blanchard et al. (1996), the PCL-S showed a sensitivity of .97, a specificity of .87, and an overall diagnostic efficacy of .94.

The PCL appears to have much to recommend it. Because it was developed by the National Center for PTSD, it is in the public domain. It is reliable, and the M and S versions map directly onto DSM criteria. The M and S versions have been shown to correlate highly with clinician-administered measures. Less information is available about version C—the civilian version that does not identify a specific event—and the reader should be cautious about generalizing psychometric findings from one version of the scale to another. Also, the published cutpoints should be used with caution, as they were derived from samples with high prevalence rates of current PTSD and may not be appropriate for samples with lower rates.

Davidson Trauma Scale

The Davidson Trauma Scale (DTS) was developed by Jonathan Davidson and his colleagues (1997) as a self-rating scale for PTSD that is reliable, valid, and sensitive to treatment effects in a variety of trauma survivors. The scale assesses 17 symptoms that correspond to DSM-IV, and each is rated for both frequency and severity on 5-point scales using a past-week time frame. The response formats vary somewhat across questions, making the format for the scale longer than similar 17-item PTSD scales.

Davidson et al. (1997) showed that the scale was quite reliable. In a large sample, composed of participants in various studies, alpha coefficients for internal consistency were very high (.97–.99) for the frequency, severity, and total scales. The test–retest correlation over a 2-week interval was .86 in a small clinical sample that had been rated as showing no change on an independent measure of clinical improvement. In a sample of 129 participants, of whom 67 met SCID criteria for PTSD, a total score of 40 most accurately predicted diagnosis, having a sensitivity of .69, a specificity of .95, and an overall efficiency of .83. Among 102 participants who were administered the CAPS, the DTS correlated .78 with the total CAPS score and .64 with the Impact of Event Scale. An interesting feature of Davidson's analysis was his consideration of whether scores on the scale changed given clinical treatment and improvement. Those who improved during treatment had pre- and postscores of 74 and 40, respectively, whereas those who did not improve had pre- and postscores of 87 and 86, respectively.

Purdue PTSD Scale—Revised

The Purdue PTSD Scale was developed a number of years ago by Don Hartsough and his students at Purdue University (e.g., Wojcik, 1988). Dean Lauterbach and Scott Vrana (1996) revised and regenerated this scale for use in heterogeneous event populations. The Revised Purdue Scale (PPTSD-R) corresponds to DSM-III-R criteria. Like Foa's measure, the PPTSD-R anchors reporting of symptoms to a single worst event identified by a screen for traumatic experience. Respondents report how often they have experienced each symptom in the previous month on a 5-point scale, from *not at all* to *often*. The scale can be scored either continuously or dichotomously.

Lauterbach and Vrana (1996) described three studies undertaken to assess the reliability and validity of the PPTSD-R. Both women and men were well represented in all studies. In the first, 440 undergraduates who had experienced a variety of traumatic events were tested once. All subscales appeared internally consistent. Alphas were .91, .84, .79, and .81 for the total, reexperiencing, avoidance, and arousal scales, respectively. In the second study, 51 undergraduates were tested twice over a 2-week interval. Test–retest correlations were .72, .48, .67, and .71, respectively.

As for validity data, the Purdue Scale correlated highly with both the IES (.66) and the Civilian Mississippi Scale (.50) in the larger sample. These correlations were stronger than those between the scale and general measures of distress, such as the BDI (.37–.39), providing preliminary support for convergent and discriminant validity. These relations were examined further by adding a third group of 35 students receiving psychology services to the sample. Reexperiencing and arousal scores were significantly higher (1) among persons reporting a traumatic event on the TEQ than among persons not reporting an event, (2) among patients than among nonpatients, and (3) among patients seeking treatment because of a traumatic event than among patients

seeking treatment for other reasons. However, whereas avoidance scores differed between patient and nonpatient groups, they did not differ between trauma and no-trauma groups.

In summary, this scale has a number of good features. It was developed for use in heterogeneous samples. As the authors correctly noted at the time of its publication, very few scales had been developed and validated on a broad cross-section of trauma survivors. (This is less true today.) In addition to this strength, the scale is internally consistent and correlates with other measures of trauma exposure and outcome in meaningful ways. However, before the scale can be recommended without reservations, two issues must be resolved. One is the lack of stability in the reexperiencing subscale. It is not altogether clear that respondents were thinking about the same event on the two testing occasions, which could deflate test–retest coefficients. The second issue is the sensitivity of the avoidance measure. Scores on this subscale did not differ between respondents reporting a traumatic event and respondents who did not. In traumatized populations, criterion C is satisfied less often than criteria B or D and therefore has a strong impact on classification (e.g., Solomon & Canino, 1990; Norris, 1992).

PTSD-Interview

Charles Watson and his colleagues (1991) developed the PTSD—Interview (PTSD-I) for use with veteran populations, but the scale could easily be applied to other groups. Seventeen items were generated that reflect PTSD symptoms as outlined in DSM-III-R. Each question is answered on a 7-point scale, from *no* to *extremely* or *never* to *always*. The scale can be scored continuously or dichotomously. The authors recommend that any symptom receiving a score of 4 or higher be counted toward PTSD diagnosis but note that users could substitute higher or lower values, depending on the purpose of the assessment. It was designed to be suitable for use by lay interviewers.

Watson et al. reported that the scale has a test–retest reliability coefficient, over 1 week, of .95. This was tested in a sample of 31 veterans, 30 of whom had been in combat. The scale was also internally consistent (alpha = .92).

The scale appears to have substantial validity in veteran populations. Watson et al. (1991) administered the PTSD-I and the Modified DIS-PTSD module (a structured interview) to 53 patients and 8 staff members at a VA medical center. Although the DIS-PTSD measure has been criticized (Weiss, 1993), the authors noted that the issues pertain to its utility with the general population rather than with clinical samples. The correlations between PTSD-I items and their DIS counterparts averaged .77. Using the DIS as the standard, the kappa was .84, which is quite high. The PTSD-I showed a sensitivity of .89, a specificity of .94, and an overall hit rate of 92%. Watson et al. (1994) examined the convergent validity of the scale in a sample of 80 help-seeking veterans. Scored continuously, the PTSD-I correlated .84 with the

Mississippi Scale for Combat-Related PTSD and .79 with the MMPI-PTSD scale; validity coefficients were equal to the Mississippi and superior to the MMPI-PTSD scale. Scored dichotomously, kappa coefficients were .59 and .60. There was about 80% agreement between the PTSD-I and each of the other two scales regarding who did or did not qualify as a case; the three scales' concordances with one another did not differ significantly in this case.

Watson et al.'s scale originally had 20 items. The first question asked whether the interviewee had experienced an unusual or extremely distressful event. By current standards in the field, a single item would not provide an adequate assessment of PTSD criterion A; thus users of this scale would be wise to supplement the PTSD-I with a trauma history or screener. (In later research, e.g., Watson et al., 1994, it appears that the authors may have revised this aspect of the scale so as to provide a list of catastrophic experiences, but this list was not detailed or published.) Two final questions determine whether symptoms have been present for at least 1 month.

All in all, this scale has many good features. It is flexible in scoring and appears to be reliable and valid. Although developed initially for veterans, it was subsequently used with a variety of trauma populations, including medical trauma victims, auto accident victims, and women who have been sexually or physically assaulted (Watson, personal communication, April 19, 1995). The scale also has been translated into French and Spanish.

Screen for Posttraumatic Stress Symptoms

The Screen for Posttraumatic Stress Symptoms (SPTSS) was developed by Eve Carlson (2001) to provide a measure that does not require the respondent to focus on a single event—or any event, for that matter. Thus the scale may be particularly useful when respondents are likely to have experienced multiple traumas, a situation that is not at all uncommon. Although the SPTSS was not intended to provide a diagnosis of PTSD, its items match the 17 DSM-IV criteria except that the symptom is not linked to a particular traumatic stressor. Participants rate their experience of each symptom on an 11-point scale using a past-2-week time frame. The scale is scored as the mean of all items, and thus scores have a potential range of 0 to 10.

In a study of 136 adult psychiatric inpatients, Carlson obtained an alpha of .91 for the total scale, which is indicative of high internal consistency. She also presented considerable evidence of validity. Scores on the SPTSS were higher among participants who had experienced a traumatic event than among participants who had not, and, within the subset of participants who had experienced trauma, scores were far higher for those who met criteria for PTSD on a structured interview than for those who did not. A total SPTSS score of 4 had high sensitivity (.94) though lower specificity (.60). Specificity may have been difficult to establish because of the psychiatric status of participants, and thus further research with community populations is needed.

Self-Rating Inventory for PTSD

The 22-item Self-Rating Inventory for PTSD (SRIP) was developed by J. E. Hovens and colleagues (2002) as a shortened version of an earlier 52-item measure. Like Carlson's SPTSS, the SRIP was developed to assess current symptoms without identifying specific traumatic experiences. Some questions refer to "past events," and others make no reference to events at all. Each item on the SRIP assesses distress over the previous 4 weeks using a 4-point scale from *not at all* to *extremely*. Psychometric data for the 22-item version were collected from several samples of trauma survivors, older adults, peacekeepers, and medical students in the Netherlands. The total scale is highly internally consistent, with alphas ranging from .90 to .94 across samples. Test–retest correlations were also high, ranging from .60 to .97, depending on the length of the interval between tests (the shorter the interval, the higher the correlation). In the trauma survivor sample that was administered other scales measuring PTSD, the SRIP correlated highly with the Civilian Mississippi Scale and MMPI-PTSD scales. In this same sample of survivors, 41 of 76 had PTSD according to the CAPS. Using the CAPS as the criterion, a SRIP cutoff score of 52 had a sensitivity of 86%, specificity of 71%, and efficiency of 78%. However, in a sample of older adults (van Zelst et al., 2003), which had a very low rate of current PTSD, a score of 52 was not at all sensitive (23%) to PTSD as assessed by the Composite International Diagnostic Interview (CIDI). In this case, a score of 39 was superior (sensitivity = 74%, specificity = 81%). This research illustrates our earlier point quite well—that cutpoints developed in clinical samples may not work well in community samples and should be applied with caution. The scale may also be scored according to DSM criteria. Both Dutch and English versions of the SRIP are available from the authors.

OTHER SCALES OF POSTTRAUMATIC STRESS

Post-Traumatic Symptom Scale

One of the earliest measures developed in the field was the Post-Traumatic Symptom Scale (PTSS) developed by Are Holen (1990) for use in studying survivors of the 1980 North Sea oil rig disaster. The scale has both 10- and 12-item versions and has been administered by using a dichotomous yes–no response format, as well as by using a 7-point frequency scale. The scale uses a past-week time frame. The scale does not map onto DSM criteria precisely but does provide a brief assessment of a variety of posttraumatic stress symptoms, including depressed mood, unstable mood, guilt, and tension, as well as selected criterion symptoms, such as sleep difficulties, nightmares, startle, and fears of reminders.

When used with the dichotomous response format, the 10-item and 12-item versions both have alphas of .85, which is good for a scale of this length. When used with the 7-point response format, the alpha increases to .90. The

scale correlates more highly with the Global Severity Index of the SCL-90 (.83, .84) than it does with the Impact of Event Scale (.70 , .69). This finding is appropriate given that the scale does not purport to assess PTSD alone.

Penn Inventory for PTSD

The Penn Inventory was developed by Melvyn Hammarberg (1992). The scale has 26 items. Each item is composed of four sentences, scored 0–3, that represent different levels (severity or frequency) of a feeling or thought. The respondent selects the sentence that best describes himself or herself. Although developed for veterans initially, the wording of the scale is not specific to the military.

Hammarberg examined the reliability and validity of the instrument in three phases. The first employed a sample of 83 participants: 28 inpatient combat veterans diagnosed with PTSD, 24 combat veterans who had previously been diagnosed with PTSD but were now at least 6 months into posttreatment, 15 age-matched veterans without PTSD, and 16 age-matched nonveterans without PTSD. The scale was found to be quite reliable, in terms of both internal consistency (alpha of .94) and stability over a 5-day interval (r = .96). Mean scale scores differed between groups who had PTSD at the time of testing or previously and the groups who did not have PTSD. However, inpatient and posttreatment groups did not differ. Using a score of 35 as the cutpoint, the scale demonstrated a sensitivity of .90 and a specificity of 1.0.

In the second phase, 98 new participants were selected and assigned to the same four categories: 39 inpatient combat veterans diagnosed with PTSD, 26 combat veterans who had previously been diagnosed with PTSD but were now at least 6 months into posttreatment, 17 age-matched veterans without PTSD, and 16 age-matched nonveterans without PTSD. The scale again demonstrated high internal consistency with an alpha of .94. Results of between-group tests replicated the findings of Phase 1: PTSD participants differed significantly from participants without PTSD, but inpatients did not differ from former patients. Again using a cutoff of 35, sensitivity was .98 and specificity was .94, for an overall hit rate of .97.

Hammarberg's (1992) third phase involved a wider range of psychiatric cases, including 39 veteran patients with PTSD, 18 veteran inpatients with a diagnosis other than PTSD, and 19 survivors of an oil rig disaster, of whom 16 were diagnosed as having PTSD. The groups without PTSD showed significantly lower means on the Penn Inventory than did groups with PTSD. With respect to the veterans in the sample, the Penn again showed excellent sensitivity (.97), although specificity (.61) was lower this time. The Mississippi Scale was also included in this phase of the study and performed similarly. The overall hit rates of the Penn and Mississippi were .86 and .88, respectively. Both performances seem excellent when it is recalled that the scales were discriminating between different groups of psychiatric patients. With respect to the disaster victims, sensitivity was .94 and specificity was 1.0. The high prevalence of PTSD in this group needs to be kept in mind when interpreting these

results. Yet they provide evidence that the scale could function effectively with trauma populations other than combat veterans.

Kutcher, Tremont, Burda, and Mellman (1994) administered the Penn Inventory, Combat Mississippi, and MMPI-2-PTSD Scale to 109 inpatient veterans, of whom 54 had been diagnosed as having PTSD. Correlations of the Penn with the other two measures were .78 and .72, respectively, showing good convergent validity. However, as did Hammarberg (1992), these investigators found the Penn to correlate more highly with depressive symptomatology than would be ideal for showing good divergent validity. The BDI's correlation with the Penn (.82) was higher than its correlations with the Mississippi (.65) or the MMPI-2-PTSD scale (.68). Showing a specificity of only .33 in this study, the Penn Inventory was less successful than the Mississippi at discriminating PTSD patients from veterans with other psychiatric diagnoses.

Trauma Symptom Checklist—40

John Briere and Marsha Runtz (1989) created the Trauma Symptom Checklist (TSC) for use in clinical research with adult survivors of childhood sexual abuse. The TSC originally had 33 items divided into five subscales: anxiety, depression, dissociation, post-sexual-abuse trauma, and sleep disturbance. Briere and Runtz established that the original scale was adequately reliable, with the exception of the sleep disturbance scale. The scale was then expanded to improve this subscale and to add a subscale for sexual problems. This version has 40 items. Subjects rate the relevance of each item to their own experience on a 5-point scale from *not at all true* to *very often true*. The reporting period is 2 months.

Using data collected from a large sample (N = 2,963) of professional women, Elliott and Briere (1992) determined that the TSC-40 has high internal consistency (alpha = .90). The revision was effective in improving the internal consistency of the Sleep Disturbance subscale (alpha = .77). The scale related to sexual problems also performed well (alpha = .73). Elliott and Briere also showed that the scale discriminates between women who have and have not been abused as children. This difference held strongly for all subscales, as well as for the total scale. Similarly, Gold, Milan, Mayall, and Johnson (1994) administered the TSC-40 to 669 female college students, divided into groups with no sexual assault or abuse (N = 438), childhood sexual assault/abuse (N = 96), adulthood sexual assault/abuse (N = 89), and both childhood and adulthood sexual assault/abuse (N = 31). Groups differed in meaningful ways except on the sleep disturbance subscale.

Trauma Symptom Inventory

For clinical purposes, or for whenever a longer measure is acceptable, Briere (1995) developed the Trauma Symptom Inventory (TSI). The TSI is not a measure of PTSD per se, but rather a global measure of trauma sequelae. It is

unique among the measures reviewed here in using a time frame for reporting symptoms of 6 months. The TSI has a total of 100 items, scored on a 4-point scale, and contains 10 clinical scales: Anxious Arousal (AA; 8 items, alpha = .86), Depression (D; 8 items, alpha = .91), Anger/Irritability (AI; 9 items, alpha = .90), Intrusive Experiences (IE; 8 items, alpha = .89), Defensive Avoidance (DA; 8 items, alpha = .90), Dissociation (DIS; 9 items, alpha = .82), Sexual Concerns (SC; 9 items, alpha = .87), Dysfunctional Sexual Behavior (DSB; 9 items, alpha = .85), Impaired Self-Reference (ISR; 9 items, alpha = .88), and Tension Reduction Behavior (TRB; 8 items, alpha = .74). In addition, the inventory includes three validity scales. The scale can be self-administered by anyone with a fifth-grade reading level or higher. Norms and *T*-scores were derived on the basis of a large mail-survey sample (n = 836) that was approximately representative of the U.S. population in terms of sex, ethnicity, and state of residence.

Briere (1995) provided confirmatory factor analyses as evidence of the inventory's construct validity. Although the factors were highly interrelated, these analyses justify conceptualizing the scale in terms of three higher order constructs. Four of the scales—IE, DA, DIS, and ISR (34 items total)—may be considered as manifestations of traumatic stress, whereas three of the scales—AI, D, and AA (25 items total)—are best viewed as manifestations of generalized dysphoria. The remaining subscales appear to reflect a third factor, Self, that may be more specific to the experience of sexual trauma and dysfunction. Also to assess construct validity, respondents in the national survey were categorized as having experienced childhood or adulthood disaster or interpersonal violence and compared with respondents who had not experienced trauma. All four trauma types were significantly associated with elevated TSI scores. Studies that have been conducted with clinical samples have yielded similar results (Briere, Elliott, Harris, & Cotman, 1995). However, it should be noted that the TSI does not tie the experience of symptoms to any specific stressor.

PTSD SCALES DERIVED FROM ESTABLISHED SYMPTOM INVENTORIES

MMPI-PTSD (PK) Scale

A different approach to developing measures of PTSD has been to derive new subscales for symptom inventories that are commonly used in clinical practice. The best known among these empirically (as opposed to rationally) derived measures is the MMPI-PTSD (PK) Scale developed by Terry Keane and colleagues (1984). The scale was modified slightly when the MMPI-2 was released. The original PK scale had 49 items, whereas the MMPI-2 version has 46 (see Lyons & Keane, 1992). The items were selected because they discriminated between veterans who did and did not have diagnoses of PTSD. Items are dichotomous, but the scale provides a continuous measure of symptomatology.

Herman, Weathers, Litz, Joaquim, and Keane (1993) provided strong evidence of scale reliability. In their studies, the alpha was .95, and test–retest reliability, over 2–3 days, was .94. Notwithstanding its excellent reliability, the validity of the scale has been challenged. Because it draws from available items in the MMPI, the PK scale does not explicitly measure all PTSD symptoms as defined in DSM-IV. However, Watson, Juba, Anderson, and Manifold (1990) found that the scale correlates highly and equally well with various diagnosed symptoms such as intrusive memories, flashbacks, detachment, arousal, and cognitive interference. These were important data for establishing the scale's validity, because otherwise high scores may have indicated the presence of some, but not necessarily all, criterion symptoms. An area of much debate in the literature has been the determination of the scale value that provides the optimal cutpoint for discriminating cases from noncases. Keane et al. (1984) originally suggested a cutpoint of 30, but other investigators subsequently suggested using much lower values (see Watson et al., 1990). Based on a series of psychometric studies (Herman et al., 1993), a score of 23 was recommended. This value yielded a sensitivity of .79 and a specificity of .71 in veteran samples.

The PK scale has been used primarily with veterans. Reliability and validity data derived from veteran populations need to be viewed with caution when the scale is used with other populations. Nonetheless, there is nothing specific to combat or military experience in the wording of the MMPI items and thus no reason why the scale could not be equally applicable to other groups. Data from several studies support this conclusion. Koretsky and Peck (1990) administered the original 49-item MMPI-PTSD scale to 18 adults diagnosed as having civilian trauma and 27 controls who had a variety of psychiatric conditions. Using a cutoff score of 19, the scale correctly classified 89% of the PTSD cases and 85% of the other cases. The scale performed equally well in a second sample of 15 PTSD patients and 9 other psychiatric cases. Dutton, Hohnecker, Halle, and Burghardt (1994) compared scores obtained from forensic and clinical samples of battered women. Quite reasonably, both groups were very distressed, as measured by the PK scale: The mean of 22 in the clinical sample approached the currently recommended cutpoint, and the mean of 28 in the forensic sample exceeded it. However, the two groups' means were not different significantly, whereas their IES and CR-PTSD means were. Neal, Busuttil, Rollins, Herepath, Strike, and Turnbull (1994) examined the convergent validity of the scale in a heterogeneous sample of 70 trauma victims; many participants had service-related trauma, but others were victims of assaults, accidents, or childhood abuse. The MMPI-PTSD scale correlated highly with CAPS measures of endorsed symptoms (r = .84) and symptom intensity (r = .85) and with the IES (.79). On the other hand, correlations were equally high with a measure of general distress (.82). On the basis of CAPS diagnoses of PTSD, a cutoff score of 21 successfully classified 80% of the cases (sensitivity .83, specificity .79.) The IES performed slightly better in this same study. In Hovens and van der Ploeg's (1993) study of 53 psychiatric inpatients

in the Netherlands, trauma victims and patients with no trauma differed significantly from one another on their MMPI-PK scores. These differences were of comparable strength to those found for the Civilian Mississippi Scale and greater than those found for the SCL-90. These two scales were highly correlated (r = .89), suggesting high concurrent validity.

All in all, the MMPI-PK scale has performed reasonably well in both veteran and civilian samples, although the shifting cutpoints should be noted. However, there is little evidence that the measure is superior to shorter scales presently available. Using the PK scale may therefore make the most sense in settings in which the MMPI is administered routinely.

Symptom Checklist—90 PTSD Scales

CR-PTSD

A similar approach was taken by Ben Saunders and his colleagues at the Crime Victims Research and Treatment Center (1990). The Symptom Checklist—90 (SCL-90; Derogatis, 1977) is a commonly used 90-item self-report symptom inventory. The 90 items are categorized into nine subscales measuring somatization, depression, anxiety, phobic anxiety, hostility, obsessive–compulsive behavior, paranoid ideation, interpersonal insensitivity, and psychoticism. All items are scored on a 5-point scale (0 = *not at all*, 4 = *extremely*). Using items on the SCL-90-R, Saunders et al. derived a 28-item scale that discriminated between crime victims with and without PTSD. Originally named the SCL-PTSD, it later became known as the CR-PTSD scale.

The CR-PTSD has high internal consistency, as evidenced in its alpha of .93. Arata, Saunders, and Kilpatrick (1991) compared the CR-PTSD scale with the IES in a sample of 266 women with a history of criminal victimization. The rate of PTSD was 7.5%. Victims with and without PTSD differed greatly on both the CR-PTSD and IES scales. The CR-PTSD scale was only moderately correlated with the IES (.44), suggesting that the two measures might be tapping different aspects of the same phenomenon. Regression analyses confirmed this impression: The SCL scale made a unique contribution to the prediction of caseness over and above the contribution of the IES. The unique contribution of the CR-PTSD scale was actually somewhat greater than the unique contribution of the IES. Of the 20 cases, the IES correctly classified 17, compared to 15 for the CR-PTSD. This difference in sensitivity was not statistically significant. Of the 246 noncases, the IES correctly classified 207, compared with 223 for the CR-PTSD. This difference in specificity was significant, with the SCL appearing superior. These results need to be viewed with some caution because the validation sample was not completely independent of the derivation sample. Dutton et al. (1994) found forensic and clinical samples of battered women to differ significantly on the CR-PTSD scale, but the difference was no greater than that obtained for the Global Severity Index (GSI) of the SCL-90. The difference was equivalent to that found

for the IES Avoidance subscale but smaller than the two groups' difference on the IES Intrusion subscale.

Like the MMPI-PK scale, an advantage of the CR-PTSD scale is that it can be administered, and often is, without knowledge of trauma history. Also like the MMPI, the SCL-90 is used in many settings anyway, so the PTSD subscale can be scored at no additional cost. However, its precision as a measure of posttraumatic stress is uncertain.

SCL-Supplemented PTSD

Robert Ursano and his colleagues (1995; see also Fullerton et al., 2000) also created a PTSD measure for the SCL-90. Theirs was rationally rather than empirically derived, that is, 31 items were selected on their apparent relevance and then assigned to categories B, C, and D. To provide coverage of criterion symptoms that were not well measured, they added 12 items, such as nightmares, feelings of reliving something unpleasant, avoidance, and hyperalertness. An advantage of this scale over Saunders et al.'s SCL-90 PTSD scale is that DSM guidelines, rather than a cutpoint, can be used to classify respondents as "probable PTSD" or not.

The scale alpha was .77 in a sample of motor vehicle accident survivors (Fullerton et al., 2000). Validity was assessed by comparing results obtained using this scale with results obtained using the MMPI-PTSD scale and a score of 19 as the cutpoint. In four community samples of disaster victims, sensitivity averaged 67% and specificity 91%. Overall, 88% were classified correctly. The scale was also related highly to the IES. Given its similar measurement strategy and controversy over optimum cutpoint, the MMPI PTSD may not have been the best choice as a criterion measure for the purpose of documenting the precision of this scale as a measure of PTSD. Fullerton et al. (2000) reported correlations in the range of .19 and .50 between the SCID and their measure. Importantly, Fullerton and colleagues also demonstrated that the sensitivity and specificity of the measure varied considerably depending on the scoring rule used, the percentage of PTSD in the sample, and whether the PTSD being assessed was of an acute or chronic nature.

SCALES DEVELOPED FOR CULTURALLY SPECIFIC OR CROSS-CULTURAL RESEARCH

Revised Civilian Mississippi Scale

The Civilian Mississippi Scale for PTSD was one of the earliest self-report scales for assessing posttraumatic stress. The Mississippi Scale for Combat-Related PTSD measured self-reported symptoms of posttraumatic stress in veteran populations. Because of its excellent psychometric characteristics, Terry Keane and other researchers associated with the Veterans Administration subsequently developed a civilian form of the scale. The scale had 35

items when used in the National Vietnam Veterans Readjustment Survey (NVVRS; Kukla et al., 1990). Four items were subsequently added. The original 35 items fall into four categories, three that align with criteria for PTSD and a fourth that taps self-persecution (guilt and suicidality). Whereas the Mississippi Scale for Combat-Related PTSD elicited information about symptoms experienced "since I was in the military," the civilian form elicits frequency of symptoms "in the past." Vreven, Gudanowski, King, and King (1995) presented psychometric data from a sample of 668 civilians who participated in the NVVRS. They found the civilian form of the Mississippi Scale to have high internal consistency (.86) but questionable discriminant validity (see also Lauterbach, Vrana, King, and King, 1995). In an analysis of the factor structure of the original Civilian Mississippi Scale, Inkelas, Loux, Bourque, Widawski, and Nguyen (2000) found that the positively worded items grouped together into a single factor, regardless of the criterion they might be assumed to reflect. The total scale was more internally consistent when these items were removed.

Fran Norris and Julia Perilla (1996) revised the Civilian Mississippi in a number of ways, partly to shorten the scale but also to sharpen its focus on posttraumatic stress. The Revised Civilian Mississippi Scale (RCMS) has 30 items. Twenty-eight were selected from the 39-item form. Two intrusion items were selected from the TSS because they had received high endorsement in previous research with victims of traumatic events (Norris, 1992). Other changes concerned question formats. As noted, Kukla et al.'s (1990) civilian form elicits frequency of symptoms "in the past." Another reason this scale may act more as a general measure of distress than as a scale of posttraumatic stress is that this wording is not tied very closely to specific trauma experiences. Norris and Perilla (1996) therefore argued (see also the discussion of Vreven et al., 1995) that it would be better to elicit feelings surrounding a specific stressful event rather than to refer vaguely to feelings "in the past." They also divided the 30 items into two parts: The first 18 items "anchored" the symptom to a specific event (e.g., "Since the event, unexpected noises make me jump"); the last 12 items did not ("I am able to get emotionally close to others"). Another change they made was to score all items on the same 5-point scale (1 = *not at all true*; 5 = *extremely true*). This eases administration considerably when data are being collected by lay interviewers or by self-administration.

Norris and Perilla (1996) developed equivalent Spanish and English versions of the RCMS, using back translation and centering (Brislin, Lonner, & Thorndike, 1973), and conducted a study to assess the instruments' cross-language stability. Participants were 53 bilingual volunteers who completed paper-and-pencil instruments twice, with a 1-week interval between tests. The total scale was reasonably consistent internally, with alphas in the bilingual sample of .86 and .88 for the English and Spanish versions, respectively. Norris and Perilla also presented data from a study involving 404 victims of Hurricane Andrew. This time, the data for the English (n = 299) and Spanish

(n = 94) versions of the RCMS were provided by different respondents, assigned according to their own language preference. Both versions of the scale again were found to have good internal consistency. Alphas were .92 and .88 for the Spanish and English versions, respectively. Norris, Perilla, and Murphy (2001) also used the RCMS to compare the structure of PTSD across samples of disaster victims from the United States (Hurricane Andrew) and Mexico (Hurricane Paulina). In an analysis that excluded the noncriterion symptoms, a 4-factor measurement model (Intrusion, Avoidance, Numbing, Arousal) fit the data of the U.S. and Mexican samples equally well. Norris et al. (2001) also administered the RCMS to a subset of respondents in a larger epidemiological study of trauma in Mexico, in which PTSD was assessed by using the CIDI (version 2.1). When RCMS symptoms were dichotomized as absent (*not* or *slightly true*) or present (*somewhat*, *very*, or *extremely true*) and counted according to DSM-IV criteria, this measure yielded the same diagnosis as the CIDI 84% of the time. Given that the RCMS was not intended for use in clinical settings, this amount of agreement is sufficient to suggest that the scale is valid as a measure of posttraumatic stress.

Altogether, the RCMS has some shortcomings relative to other, more recently developed diagnostic scales, but it performs well as a continuous measure of posttraumatic stress and stands out in terms of its validation for use with Spanish-speaking populations.

Harvard Trauma Questionnaire

The Harvard Trauma Questionnaire (HTQ) was developed by Richard Mollica and his colleagues (1992). Both traumatic events and symptoms are included in the questionnaire. In the first section, 17 items describe a range of stressors experienced by refugees, such as torture, rape, murder, and lack of food or water. For each item, the respondent notes whether he or she has (1) not experienced, (2) heard about, (3) witnessed, or (4) personally experienced that stressor. The symptom portion consists of 30 items, 16 of which correspond to DSM-IV criteria and 14 of which tap other aspects of distress as it is expressed in Indochinese culture. Items are scored on a 4-point scale from *not at all* = 1 to *extremely* = 4, and the investigators now recommend scoring the scale as the mean item value (Mollica, personal communication, April 18, 1995). The HTQ is available in Khmer, Lao, and Vietnamese, in addition to English. Linguistic equivalence was established using back translation and centering.

The HTQ is important to review here because it illustrates an approach to the cross-cultural assessment of trauma and PTSD. The investigators (Mollica et al., 1995) note that it is important to *adapt* rather than merely translate the questionnaire for each trauma population and culture. According to Mollica, the "core" PTSD section should be kept equivalent across languages, but the remaining symptom questions should vary so that they are specific and relevant to the culture of respondents. These items should be

identified by ethnographic studies, clinical experience, key informants, and healers in the setting of interest (Mollica et al., 1995).

Mollica et al. (1992) examined the reliability and validity of the Cambodian, Lao, and Vietnamese versions of the instrument in a sample of 91 Indochinese refugees, of whom 34 were men and 57 were women, and of whom 55 were Cambodian, 20 Laotian, and 16 Vietnamese. Reliability was very high: The symptom portion of the HTQ yielded an alpha of .96 and a test–retest correlation of .92, with a 1-week interval between tests. To assess criterion validity, research participants were divided into groups on the basis of independent diagnoses. The PTSD group (n = 65) showed significantly higher symptom scores than the non-PTSD group (n = 26). A cutpoint of 75 (mean item value of 2.5) was found to maximize classification accuracy. Sensitivity was .78, specificity was .65, and the overall hit rate was .75. These initial studies provided the tools used in a large-scale study (Mollica, Poole, & Tor, 1998) involving a random sample of nearly 1,000 Cambodian refugees living in camps along the Thai–Cambodian border. Approximately one-third of the sample had PTSD scores in the clinical range (2.5+), and two-thirds had depression scores in the clinical range. Most relevant to the purpose of this chapter were the exceptionally strong relations between traumatic experiences and symptom scores. Rates of PTSD varied from 14% among refugees reporting four or fewer trauma events to 81% among refugees reporting 25 or more trauma events. The relative odds ratio was 38.9 in the most traumatized group. Rates of depression varied from 45 to 93%. In this case, the relative odds ratio was 21.8 in the most traumatized group. These data are instructive in showing that posttraumatic stress symptoms were more specifically associated with the cumulative amount of trauma, whereas depressive symptoms were more pervasive among the refugees.

SUMMARY

Table 3.2 summarizes the information available on these 17 scales. Sources for obtaining these measures are shown in Appendix 3.1. All of the scales reviewed here show acceptable reliability and validity, although some test creators have documented these attributes more completely than have others. Undoubtedly, clinician-administered interviews will remain the "gold standard" in the field. Yet, as a group, these self-report measures performed well when contrasted directly with them. In Table 3.2, we have reserved the descriptor of "strong" validity for those scales that have shown sensitivity and specificity in clinical samples within studies that have been subjected to peer review. This crude summary may give undue weight to criterion validity at the expense of construct validity, which is excellent among many of the scales whose validity is described only as moderate in Table 3.2. Even more important, these data on clinical validity need to be interpreted most cautiously. Much of it resulted from researchers identifying a scale score cutpoint in a sin-

TABLE 3.2. Summary Descriptions of 17 Standardized Self-Report Measures of Posttraumatic Stress

Scale	Number of items	Evidence of stability	Evidence of consistency	Evidence of validity	Reporting period	Anchored to identified event
NWS Module	20+	Kappa = .45	na	Strong	Lifetime	No
PTDS	17	r = .83	.92	Strong	Past month	Yes
PCL	17	r = .96	.97	Strong	Past month	Varies
Davidson TS	34	r = .86	.97	Strong	Past week	Yes
Purdue PTSD-R	17	r = .71	.91	Moderate	Past month	Yes
PTSD-Interview	20	r = .95	.92	Strong	Lifetime	Yes
SPTSS	17	na	.91	Moderate	Past 2 weeks	No
SRIP	22	.60–.97	.90–.94	Strong	Past 4 weeks	No
PTSS	10–12	na	.85–.90	Moderate	Past week	No
Penn Inventory	26	r = .96	.94	Moderate–strong	Past week	No
TSC-40	40	na	.90–.92	Moderate	2 months	No
TSI	100	na	.74–.90	Moderate–strong	6 months	No
MMPI-PTSD	46	r = .94	.95	Moderate–strong	Not explicit	No
CR-PTSD	28	na	.93	Moderate	Past 2 weeks	No
SCL-Supplemented PTSD	43	na	na	Moderate	Past 2 weeks	No
Revised Civilian Mississippi	30	r = .84	.86–.92	Moderate–strong	Varies	Partially
HTQ	16 + 14	r = .92	.96	Moderate–strong	na	Partially

Note. na, data not available.

gle sample that was sensitive and specific to PTSD. Such data establish only that the measure in question *can* predict caseness; they do not establish that the identified cutpoint is appropriate for other samples and populations.

With so many adequate symptom measures available, how should the reader decide which measure to use? The answer to this question is found by considering the wide array of choices these scales offer for measuring posttraumatic stress. Some, but not all, are in the public domain. Some scales adhere closely to DSM-IV criteria; others take a broader sweep. Some are relatively short, whereas others are relatively long. Some take advantage of available clinical data, such as the MMPI or SCL-90; most require additional assessment materials. Most assess current symptoms, whereas a few assess symptoms over the lifetime or since a specific event. Some require the specification of a single or most stressful event, but some refer broadly to past events. Thus a person who is studying reactions to a variety of traumatic events that cannot easily be distinguished but who has few constraints in terms of cost or time of administration may make one choice of measure, whereas a person who is studying a specific event but who needs a brief measure that is available at no cost may make a different choice; yet both choices are equally valid and defensible.

Before leaving this point, we need to acknowledge that the extent to which a PTSD measure must be anchored to a specific traumatic experience is among the points of most controversy in trauma assessment. On the one hand, when symptoms are not tied to a specific stressor, it is difficult to establish for certain that the respondent met criterion A (see the first part of this chapter) or even that the various symptoms pertain to an event at all. (For example, trauma is certainly not the only source of irritability.) On the other hand, epidemiological research has shown quite clearly that it is not uncommon for people to experience multiple events, and victims may not be cognizant of the reason they feel a certain way. There are experts who advocate for each point of view quite strongly, and whichever approach the researcher decides on, it is reasonable to expect at least some criticism from proponents of the other perspective. The best way to manage this dilemma is to acknowledge the issue and to be clear about the reasons for deciding on one measurement approach or the other.

In the first edition of this volume, Norris and Riad (1997) noted that progress in the measurement of civilian trauma had lagged behind that related to military trauma. That statement may no longer be true. That there are now numerous reliable and valid self-report measures of PTSD should aid epidemiological and community-based studies immensely over the next few years. Notwithstanding the quality of these measures, we believe there is still room for improvement in the methods used to validate them. Systematic research comparing various self-report measures in representative community samples, as well as clinical and survivor samples, is needed. Perhaps it is our own bias, but we were disappointed in the lack of attention to diversity in validation samples in this literature. The SRIP was unusual in having been evaluated in a

sample of older adults. Excluding the Revised Civilian Mississippi Scale and HTQ, little attention was given to potential ethnic or cross-cultural differences in symptom expression. We concluded this chapter with the HTQ because it illustrates a forward-thinking approach that balances cross-cultural standardization with cultural specificity in developing assessment tools. In our increasingly global and mobile society, cross-cultural equivalence and relevance are extremely important issues for psychometricians to address in future research.

In summary, we believe that future progress in this area would be served best by efforts to refine and cross-validate the existing measures of PTSD. Can we, if only for awhile, forego the temptation of generating new, but largely similar, scales? We hope the answer to this question is yes. As measurement becomes more standardized, we can build a database that elucidates the prevalence and nature of PTSD across different populations and events.

APPENDIX 3.1. Sources for Obtaining Standardized Self-Report Scales

Scale and contact person	Affiliation and e-mail	Telephone
Brief Trauma Questionnaire Paula Schnurr	National Center for PTSD (NCPTSD) White River Junction, VT *Paula.P.Schnurr@Dartmouth.edu*	(802) 296-5132
Civilian Mississippi—Revised Fran Norris	Dartmouth Medical School/ NCPTSD *Fran.Norris@Dartmouth.edu*	(802) 296-5132
Davidson Trauma Scale Jonathan Davidson	Duke University Medical Center *tolme@acpub.duke.edu*	(919) 684-2880
Harvard Trauma Questionnaire Richard Mollica	Harvard Program in Refugee Trauma *rmollica@partners.org*	(617) 876-7879
Life Stressor Checklist—Revised Rachel Kimerling	NCPTSD/ Palo Alto VAMC *Rachel.Kimerling@med.va.gov*	(650) 493-5000 X 23218
MMPI-PTSD Terence Keane	Boston University/ NCPTSD Boston VAMC *Terry.Keane@med.va.gov*	(617) 278-4551
NWS PTSD Module Heidi Resnick	Crime Victims Research and Treatment Center Medical University of South Carolina *Resnickh@musc.edu*	(843) 792-2945
Penn Inventory Melvyn Hammarberg	University of Pennsylvania *mhammarb@ccat.sas.upenn.edu*	(215) 898-0981

Scale and contact person	Affiliation and e-mail	Telephone
PTSD Checklist Frank Weathers	Auburn University *weathfw@auburn.edu*	(334) 844-6495
PTSD-Interview Charles Watson	St. Cloud MN, DVAMC	(570) 824-3521 X 7818
Posttraumatic Stress Diagnostic Scale Edna Foa	University of Pennsylvania, Department of Psychiatry *Foa@mail.med.upenn.edu*	(215) 746-3327
Post-Traumatic Symptom Scale Are Holen	Norwegian University of Science and Technology *are.holen@ntnu.no*	47-7-355-1513
Purdue PTSD Scale—Revised Scott Vrana	Virginia Commonwealth University *srvrana@saturn.vcu.edu*	(804) 828-6273
Screen for Posttraumatic Stress Symptoms Eve Carlson	NCPTSD/Menlo Park, CA *Eve.Carlson@med.va.go*	
Self-Rating Inventory for Posttraumatic Stress Disorder J. E. Hovens	Delta Psychiatric Teaching Hospital *hans.hovens@deltabouman.nl*	31-10-503-1512
SCL-PTSD (CR-PTSD) Ben Saunders	National Crime Victims Research and Treatment Center Medical University of South Carolina *Saunders@musc.edu*	(843) 792-2945
SCL-Supplemented PTSD Robert Ursano	Uniformed Services University School of Medicine *rursano@usuhs.mil*	(301) 295-3293
Stressful Life Events Screening Questionnaire Lisa Goodman	Boston College *goodmalc@bc.edu*	(617) 552-1725
Trauma History Questionnaire Bonnie Green	Georgetown University Medical School *Bgreen01@georgetown.edu*	(202) 687-6529
Traumatic Events Questionnaire Dean Lauterbach	Eastern Michigan University *dlauterba@emich.edu*	(734) 487-0785
Traumatic Life Events Questionnaire Edward Kubany	NCPTSD/ Honolulu DVAMC *edward.kubany@med.va.gov* or *kubany@hawaii.rr.com*	(808) 284-4497

Scale and contact person	Affiliation and e-mail	Telephone
Traumatic Stress Schedule Fran Norris	Dartmouth Medical School and NCPTSD *Fran.Norris@Dartmouth.edu*	(802) 296-5132
Trauma Symptom Checklist–40 John Briere	Department of Psychiatry, Keck School of Medicine, University of South Carolina *www.johnbriere.com/psych_tests.htm*	1-800-331-Test
Trauma Symptom Inventory John Briere	Department of Psychiatry, Keck School of Medicine, University of South Carolina *www.johnbriere.com/psych_tests.htm*	1-800-331-Test

REFERENCES

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

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

Arata, C., Saunders, B., & Kilpatrick, D. (1991). Concurrent validity of a crime-related posttraumatic stress disorder scale for women with the Symptom Checklist—90—Revised. *Violence and Victims*, *6*, 191–199.

Beck, A. T., Ward, C. H., Mendelsohn, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561–571.

Blanchard, E., Jones-Alexander, J., Buckley, T., & Forneris, C. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy*, *34*, 669–673.

Briere, J. (1992). *Child abuse trauma: Theory and treatment of the lasting effects.* Newbury Park, CA: Sage.

Briere, J. (1995). *Trauma Symptom Inventory (TSI): Professional manual.* Odessa, FL: Psychological Assessment Resources.

Briere, J., Elliott, D., Harris, K., & Cotman, A. (1995). Trauma Symptom Inventory: Psychometrics and association with childhood and adult victimization in clinical samples. *Journal of Interpersonal Violence*, *10*, 387–401.

Briere, J., & Runtz, M. (1989). The Trauma Symptom Checklist (TSC-33): Early data on a new scale. *Journal of Interpersonal Violence*, *4*, 151–163.

Brislin, R., Lonner, W., & Thorndike, R. (1973). *Cross-cultural research methods.* New York: Wiley.

Carlson, E. (2001). Psychometric study of a brief screen for PTSD: Assessing the impact of multiple traumatic events. *Psychological Assessment*, *8*, 431–441.

Davidson, J., Book, S., Colket, L., Tupler, L., Roth, S., David, D., et al. (1997). Assessment of a new self-rating scale for posttraumatic stress disorder. *Psychological Medicine*, *27*, 153–160.

Derogatis, L. (1977). *SCL-90: Administration, scoring, and procedure manual for the revised version.* Baltimore: John Hopkins University School of Medicine.

Dohrenwend, B. S., Krasnoff, B., Askenasy, A., & Dohrenwend, B. P. (1978). Exemplification of a method for scaling life events: The PERI Life Events Scale. *Journal of Health and Social Behavior, 19*, 205–229.

Dutton, M., Hohnecker, L., Halle, P., & Burghardt, K. (1994). Traumatic responses among battered women who kill. *Journal of Traumatic Stress, 7*, 549–564.

Elliott, D., & Briere, J. (1992). Sexual abuse trauma among professional women: Validating the Trauma Symptom Checklist—40 (TSC-40). *Child Abuse and Neglect, 16*, 391–398.

Foa, E., Cashman, L., Jaycox, L, & Perry, K. (1997). The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. *Psychological Assessment, 9*, 445–451.

Foa, E. B., Riggs, D. S., Dancu, C. V., & Rothbaum, B. O. (1993). Reliability and validity of a brief instrument for assessing posttraumatic stress disorder. *Journal of Traumatic Stress, 6*, 459–473.

Fullerton, C., Ursano, R., Epstein, R., Crowley, B., Vance, K., Craig, K., & Baum, A. (2000). Measurement of posttraumatic stress disorder in community samples. *Nordic Journal of Psychiatry, 54*, 5–12.

Gold, S., Milan, L., Mayall, A., & Johnson, A. (1994). A cross-validation of the Trauma Symptom Checklist: The role of mediating variables. *Journal of Interpersonal Violence, 9*, 12–25.

Goodman, L., Corcoran, C., Turner, K., Yuan, N., & Green, B. (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. (1996). Trauma History Questionnaire. In B. H. Stamm & E. M. Varra (Eds.), *Measurement of stress, trauma, and adaptation* (pp. 366–368). Lutherville, MD: Sidran Press.

Hammarberg, M. (1992). Penn Inventory for Posttraumatic Stress Disorder: Psychometric properties. *Psychological Assessment, 4*, 67–76.

Herman, D., Weathers, F., Litz, B., Joaquim, S., & Keane, T. (1993, October). *The PK Scale of the MMPI-2: Reliability and validity of the embedded and stand-alone versions.* Paper presented at the annual meeting of the International Society for Traumatic Stress Studies, San Antonio, TX.

Holen, A. (1990). *A long-term outcome study of survivors from a disaster.* Oslo, Norway: University of Oslo Press.

Hovens, J., Bramsen, I., & van der Ploeg, H. (2002). Self-rating inventory for posttraumatic stress disorder: Review of the psychometric properties of a new brief Dutch screening instrument. *Perceptual and Motor Skills, 94*, 996–1008.

Hovens, J., & van der Ploeg, H. M. (1993). Posttraumatic stress disorder in Dutch psychiatric in-patients. *Journal of Traumatic Stress, 6*, 91–102.

Inkelas, M., Loux, L., Bourque, L., Widawski, M., & Nguyen, L. (2000). Dimensionality and reliability of the Civilian Mississippi Scale for PTSD in a postearthquake community. *Journal of Traumatic Stress, 13*, 149–167.

Keane, T. M., Caddell, J. M., & Taylor, K. L. (1988). Mississippi scale for combat-related posttraumatic stress disorder: Three studies in reliability and validity. *Journal of Consulting and Clinical Psychology, 56*, 85–90.

Keane, T., Malloy, P., & Fairbank, J. (1984). Empirical development of an MMPI sub-

scale for the assessment of combat-related posttraumatic stress disorders. *Journal of Consulting and Clinical Psychology, 52*, 888–891.

Kilpatrick, D., Resnick, H., & Freedy, J. (1991). *The Potential Stressful Events Interview*. Charleston, SC: Medical University of South Carolina, Department of Psychiatry, Crime Victim Research and Treatment Center.

Kilpatrick, D., Resnick, H., Saunders, B., & Best, C. (1989). *The National Women's Study PTSD module*. Charleston, SC: Medical University of South Carolina, Department of Psychiatry, Crime Victim Research and Treatment Center.

Koretsky, M., & Peck, A. (1990). Validation and cross-validation of the PTSD subscale of the MMPI with civilian trauma victims. *Journal of Clinical Psychology, 46*, 296–300.

Krinsley, K., Gallagher, J., Weathers, F., Kaloupek, D., & Vielhauer, M. (2003). Reliability and validity of the Evaluation of Lifetime Stressors questionnaire. *Journal of Traumatic Stress, 16*, 399–409.

Kubany, E., Haynes, S., Leisen, M., Owens, J., Kaplan, A., Watson, S., & Burns, K. (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., Schlenger, W., Fairbank, J., Hough, R., Jordan, B. K., Marmar, C., & Weiss, D. (1990). *Trauma and the Vietnam war generation: Report of findings from the National Vietnam Veterans Readjustment Study*. New York: Brunner/Mazel.

Kutcher, G., Tremont, M., Burda, P., & Mellman, T. (1994). *The effectiveness of PTSD self-report measures with an inpatient veteran population*. Paper presented at the annual meeting of the International Society for Traumatic Stress Studies, Chicago, IL.

Lauterbach, D., & Vrana, S. (1996). Three studies on the reliability and validity of a self-report measure of posttraumatic stress disorder. *Assessment, 3*, 17–25.

Lauterbach, D., Vrana, S., King, D., & King, L. (1995, May). *Psychometric properties of the Civilian Version of the Mississippi PTSD Scale*. Paper presented at the Midwestern Psychological Association, Chicago, IL.

Lyons, J., & Keane, T. (1992). Keane PTSD Scale: MMPI and MMPI-2 update. *Journal of Traumatic Stress, 5*, 111–117.

McHugo, G., Caspi, Y., Kammerer, N., Mazelis, R., Jackson, E., Russell, L., et al. (2004). *The assessment of trauma history in women with co-occurring mental health and substance use disorders and a history of interpersonal violence*. Unpublished manuscript.

Mollica, R., 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., Caspi-Yavin, Y., Lavelle, J., Tor, S., Yang, T., Chan, S., et al. (1995). *Manual for the Harvard Trauma Questionnaire*. Brighton, MA: Indochinese Psychiatry Clinic.

Mollica, R., Poole, C., & Tor, S. (1998). Symptoms, functioning, and health functioning in a massively traumatized population: The legacy of the Cambodian tragedy. In B. P. Dohrenwend (Ed.), *Adversity, stress, and psychopathology* (pp. 34–51). New York: American Psychiatric Association Press.

Morgan, C.A., Hazelett, M.G., Wang, S., Richardson, E.G., Schnurr, P., & Southwick, S. (2001). Symptoms of dissociation in humans experiencing acute, uncontrollable stress: A prospective investigation. *American Journal of Psychiatry, 158*, 1239–1247.

Neal, L., Busuttil, W., Rollins, J., Herepath, R., Strike, P., & Turnbull, G. (1994). Convergent validity of measures of posttraumatic stress disorder in a mixed military and civilian population. *Journal of Traumatic Stress, 7*, 447–456.

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

Norris, F. (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., & Kaniasty, K. (1992). Reliability of delayed self-reports in disaster research. *Journal of Traumatic Stress, 5*, 575–588.

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

Norris, F., Perilla, J., & Murphy, A. (2001). Postdisaster stress in the United States and Mexico: A cross-cultural test of the multicriterion conceptual model of posttraumatic stress disorder. *Journal of Abnormal Psychology, 110*, 553–563.

Norris, F., & Riad, J. (1997). Standardized self-report measures of civilian trauma and posttraumatic stress disorder. In J. P. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 7–42). New York: Guilford Press.

Resnick, H., Falsetti, S., Kilpatrick, D., & Freedy, J. (1996). Assessment of rape and other civilian trauma-related posttraumatic stress disorder: Emphasis on assessment of potentially traumatic events. In T. Miller (Ed.), *Theory and assessment of stressful life events* (pp. 235–271). Madison, CT: International Universities Press.

Resnick, H., Kilpatrick, D., Dansky, B., Saunders, B., & Best, C. (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.

Saunders, B., Arata, C., & Kilpatrick, D. (1990). Development of a crime-related posttraumatic stress disorder scale for women with the Symptom Checklist—90—Revised. *Journal of Traumatic Stress, 3*, 439–448.

Schnurr, P., Spiro, A., Vielhauer, M., Findler, M., & Hamblen, J. (2002). Trauma in the lives of older men: Findings from the normative aging study. *Journal of Clinical Geropsychology, 8*, 175–187.

Schnurr, P., Vielhauer, M., Weathers, F., & Findler, M. (1999). *The Brief Trauma Questionnaire.* White River Junction, VT: National Center for PTSD.

Solomon, S., & Canino, G. (1990). The appropriateness of DSM-III-R criteria for posttraumatic stress disorder. *Comprehensive Psychiatry, 31*, 227–237.

Ursano, R., Fullerton, C., Kao, T., & Bhartiya, V. (1995). Longitudinal assessment of posttraumatic stress disorder and depression after exposure to traumatic death. *Journal of Nervous and Mental Disease, 183*, 36–42.

van Zelst, W., de Beurs, E., Beekman, A., Deeg, D., Bramsen, I., & van Dyck, R. (2003). Criterion validity of the self-rating inventory for posttraumatic stress disorder (SRIP) in the community of older adults. *Journal of Affective Disorders, 76*, 229–235.

Ventureya, V., Yao, S., Cottraux, J., Note, I., & De Mey-Guillard, C. (2002). The vali-

dation of the Posttraumatic Stress Disorder Checklist Scale in posttraumatic stress disorder and nonclinical subjects. *Psychotherapy and Psychosomatics, 71*(1), 47–53.

Vrana, S., & Lauterbach, D. (1994). Prevalence of traumatic events and posttraumatic psychological symptoms in a nonclinical sample of college students. *Journal of Traumatic Stress, 7*, 289–302.

Vreven, D., Gudanowski, D., King, L., & King, D. (1995). The Civilian Version of the Mississippi PTSD Scale: A psychometric evaluation. *Journal of Traumatic Stress, 8*, 91–110.

Watson, C., Juba, M., Anderson, P., & Manifold, V. (1990). What does the Keane et al. PTSD Scale for the MMPI measure? *Journal of Clinical Psychology, 46*, 600–606.

Watson, C., Juba, M., Manifold, V., Kucala, T., & Anderson, P. (1991). The PTSD Interview: Rationale, description, reliability and concurrent validity of a DSM-III-based technique. *Journal of Clinical Psychology, 47*, 179–185.

Watson, C., Plemel, D., DeMotts, J., Howard, M., Tuorilla, J., Moog, R., et al. (1994). A comparison of four PTSD measures' convergent validities in Vietnam veterans. *Journal of Traumatic Stress, 7*, 75–82.

Weathers, F., Litz, B., Herman, D., Huska, J., & Keane, T. (1993, October). *The PTSD Checklist (PCL): Reliability, validity, and diagnostic utility.* Paper presented at the annual meeting of the International Society for Traumatic Stress Studies, San Antonio, TX.

Weiss, D. (1993). Structured clinical interview techniques. In J. Wilson & B. Raphael (Eds.), *International handbook of traumatic stress syndromes* (pp. 179–189). New York: Plenum.

Weiss, D. S., & Marmar, C. R. (1997). The Impact of Event Scale—Revised. In J. P. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD: A handbook for practitioners* (pp. 399–411). New York: Guilford Press.

Wilson, J. (1994). The historical evolution of PTSD diagnostic criteria: From Freud to DSM-IV. *Journal of Traumatic Stress, 7*, 681–689.

Wojcik, E. (1988). Disruption to pretrauma social support networks considered as a factor in posttraumatic stress reaction (Doctoral dissertation, Purdue University, 1987). *Dissertation Abstracts International, 49*, 246B.

Wolfe, J., Kimerling, R., Brown, P. J., Chrestman, K. R., & Levin, K. (1996). Psychometric review of the Life Stressor Checklist—Revised. In B. H. Stamm (Ed.), *Measurement of stress, trauma, and adaptation* (pp. 198–201). Lutherville, MD: Sidran Press.