4-16-2010

Diagnostic efficiency of the computerized PTSD scale – multimedia version (CPS-M) in assessing posttraumatic stress disorder

Jennifer Boss Mainka
jbmainka@gmail.com

Diagnostic Efficiency of the Computerized PTSD Scale – Multimedia Version (CPS-M) in Assessing Posttraumatic Stress Disorder

by

Jennifer Boss Mainka, M.S.

Dissertation

Submitted to the Department of Psychology

Eastern Michigan University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Clinical Psychology

Dissertation Committee:

Dean Lauterbach, Ph.D., Chair

John Knapp, Ph.D.

Ellen Koch, Ph.D.

David C. S. Richard, Ph.D.

April 16, 2010

Ypsilanti, Michigan

ACKNOWLEDGEMENTS

I would like to thank my thesis committee, Dr. Dean Lauterbach, Dr. John Knapp, Dr. Ellen Koch, and Dr. David Richard. I am especially grateful to Dr. Lauterbach, who spent countless hours providing statistical assistance and meeting with me through the process. I am also very grateful to Dr. Sheila Rauch, my principal investigator at the VA hospital. Her persistence and guidance kept me going even when data collection seemed slow or impossible. I would like to thank research assistants Natalie Nugent, Lauren McSweeney, and Valentina Ivejaz for their assistance with data collection and entry. Special thanks also to Erin DeFever for assisting with collection of data at the VA. I would also like to thank my husband, Ben, for his support and patience throughout this project and my kids for understanding when mom needed to disappear for a while to write.

Abstract

The most commonly used interview for posttraumatic stress disorder (PTSD) is the Clinician-Administered PTSD Scale (CAPS), a semi-structured interview patterned after the DSM-IV criteria (Blake et al., 1990). The Computerized PTSD Scale – Multimedia Version (CPS-M: Richard, Mayo, Bohn, Haynes, & Kolman, 1997) is a computerized interview that is modeled after the CAPS. This study examined how well the CPS-M agreed with the CAPS diagnostically in a clinical sample. Ninety veterans completed the test protocol consisting of paper-and-pencil measures, the CPS-M, and the CAPS interview. Correlations between the CAPS and CPS-M were high at the item, subscale, and full-scale levels. Confidence interval analysis revealed that the CPS-M scales were not significantly different from their CAPS counterparts but failed to establish equivalence. Alpha scores for the scales indicated good internal consistency on both the CAPS and CPS-M. Difference scores between the two instruments were normally distributed, and scale effect sizes were negligible. ROC curve analysis for the CPS-M revealed high diagnostic accuracy. These results present a strong case for more widespread use of the CPS-M in the assessment of PTSD.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

TABLE OF ABBREVIATIONS

Abbreviation/ Symbol	Definition
AUC	Area under receiver operating characteristic (ROC) curve
CAPS	Clinician-Administered PTSD Scale
CPS – M	Computerized PTSD Scale – Multimedia Version
DSM	Diagnostic and Statistical Manual of Mental Disorders
n	Absence of a disorder
P	Probability
PCL	PTSD Checklist
PTSD	Posttraumatic stress disorder
ROC	Receiver operating characteristic
s	Presence of a disorder
S	Diagnosis of a disorder
SDT	Signal detection theory
TEQ	Traumatic Events Questionnaire
r_{pb}	Spearman's rank point-biserial correlation

Diagnostic Efficiency of the Computerized PTSD Scale – Multimedia Version (CPS-M) in Assessing Posttraumatic Stress Disorder

Psychological assessment instruments provide clinicians with information about the topography and/or function of behavior. Within a medical model, psychological assessment instruments also serve to augment diagnostic decision-making. Thus, it is critical to understand how accurately a new instrument classifies both positive and negative diagnostic cases. In order to understand how accurate an instrument is with regard to diagnostic classification, a criterion against which to evaluate the performance of the new instrument is required, an incontrovertible index or so-called "gold standard." The criterion can include results from another test, a behavioral criterion, or self-report results. The present study used the Clinician Administered PTSD Scale (CAPS), the parent instrument of the Computerized PTSD Scale – Multimedia version (CPS – M), as the criterion instrument. We examined the criterion-related validity of the CPS – M using signal detection theory. Signal detection theory calculates the percentage agreement between a new assessment instrument and a criterion with regard to the presence or absence of a particular diagnosis. We also determined equivalence of the two instruments using confidence interval analysis.

In the literature review that follows, psychometric properties, as well as limitations and advantages of the CAPS, will be discussed. The review will illustrate how the CPS-M avoids some of the shortcomings of the CAPS. Confidence interval analysis will be described, and signal detection theory will be introduced. Prior signal detection work involving the CAPS and other PTSD measures will be highlighted. In

addition, signal detection with non-PTSD psychological instruments and signal detection outside the field of psychology will be discussed.

The Clinician Administered PTSD Scale

The Clinician Administered PTSD Scale is a structured interview developed by the National Center for PTSD to diagnose posttraumatic stress symptoms (CAPS; Blake et al., 1990). The CAPS consists of three subscales, which correspond with Criteria B, C, and D for PTSD based on the DSM-IV-TR. Subscale B assesses reexperiencing, and this subscale is composed of five items. Subscale C measures avoidance and numbing symptoms in seven items, and Subscale D is composed of five items that assess hyperarousal. The CAPS provides a structured assessment of the frequency and intensity of each of the 17 PTSD symptoms. By combining the frequency and intensity scores for each reported symptom, a severity score is generated for that symptom. The sum of the 17 symptom severity scores is called the Total Severity Score. Weathers, Keane, and Davidson (2001) reported that the CAPS has been used in more than 200 studies and is the most commonly used interview for posttraumatic stress symptoms.

Several scoring rules are used to arrive at a diagnosis of PTSD on the CAPS (Weathers, Ruscio, & Keane, 1999). The 1-2 symptom presence rule classifies a symptom as being present when a rating of one or higher is provided for frequency and a two or higher is provided for intensity. The 2-2 symptom presence rule classifies a symptom as being present when ratings of two or higher are provided for both the frequency and intensity dimensions. The Sum 4 symptom presence rule classifies a symptom as being present when the sum of the frequency and intensity dimensions for a symptom is four or higher. The Total 65 rule provides a PTSD diagnosis if the sum of

the symptom severity scores is 65 or higher. Individuals are classified as PTSD-positive or PTSD-negative based on the dichotomous decision provided by these scoring rules. The CAPS also yields continuous scores for frequency and intensity of symptoms, but those values are not the focus of this paper.

Psychometric properties of the CAPS. Blake et al. (1990) performed a pilot study after the initial development of the CAPS that compared it to the Combat Exposure Scale (CES; Keane et al., 1989), the Mississippi Scale for Combat-Related PTSD (Mississippi Scale; Keane, Cadell, & Taylor, 1988), and the Keane PTSD Scale of the MMPI (PK Scale; Keane et al., 1984) in a group of 25 male combat veterans. Two researchers independently rated seven interviews, and correlations ranged from .92 to .99 across scales (Blake et al., 1990). Blake et al. also found internal consistency alphas between .73 and .85 for the three subscales. The CAPS correlated .70 with the Mississippi Scale, .84 with the PK Scale, and .42 with the Combat Exposure Scale (Blake et al., 1990).

In another study of the psychometric properties of the CAPS, Weathers, Keane, King, and King (2001) administered the instrument twice to participants, two to three days apart. Test-retest reliability correlations ranged from .90 to .98 across three separate rater pairs (Weathers et al., 2001). Internal consistency alphas ranged from .85 to .87 for the three subscales, with an alpha of .94 for the full scale (Weathers et al., 2001).

Palmieri, Weathers, Difede, and King (2007) examined the factor structure of the CAPS and PTSD Checklist in a sample of 2,960 utility workers exposed to the World Trade Center Ground Zero site. The relative fit of five previously supported measurement models were tested. The first was a single-factor model reflecting PTSD as

a single unitary construct. The second was a two-factor model composed of reexperiencing/avoidance and numbing/hyperarousal, based on the models in studies by Taylor, Kuch, Koch, Crockett, and Passey (1998), Buckley, Blanchard, and Hickling (1998), Simms, Watson, and Doebbeling (2002), and Asmundson, Wright, McCreary, and Pedlan (2003). The third was a three-factor model that parallels the DSM-IV diagnostic criteria. The fourth was a four-factor model composed of reexperiencing, avoidance, numbing, and hyperarousal (King et al., 1998). The fifth was a four-factor model composed of reexperiencing, avoidance, dysphoria, and hyperarousal, supported in Simms et al. (2002) and Baschnagel, O'Connor, Colder, and Hawk (2005). For each of the models tested, items were specified to load on a single construct, and error terms were uncorrelated. Palmieri et al. (2007) found that a four-factor model with distinct reexperiencing, avoidance, numbing, and hyperarousal factors best fit the data. These findings parallel those obtained by King et al.

Charney and Keane (2007) examined the factor structure and psychometric properties of a Bosnian translation of the CAPS with 115 help-seeking Bosnian refugees resettled in the Boston area. They found evidence for a two-factor model, with the first composed of symptoms of intrusion and avoidance and the second composed of hyperarousal and numbing (Charney & Keane, 2007). Charney and Keane calculated partial correlations to determine effect sizes for items and to determine which factors were most important in predicting the items. They first obtained the critical ratio (*t* statistic), indicating which items have significant loadings on each factor, and then they used this critical ratio to calculate the partial correlation for each loading. Intrusion and avoidance partial correlations based on the two-factor model ranged from .33 (D2:

irritability) to .82 (B4: psychological distress). Partial correlations for hyperarousal and numbing ranged from .24 (C5: detachment) to .59 (C7: foreshortened future; Charney & Keane, 2007).

Coefficient alpha in the Charney and Keane (2007) study was .92 for the CAPS, and the average item-total correlation was .66 ($p < .01$). Coefficient alphas for the two-factor model were .88 for intrusion/avoidance and .85 for hyperarousal/numbing (Charney & Keane, 2007). Charney and Keane reported correlations with other measures as follows: Beck Depression Inventory (BDI): .65, $p < .01$; Structured Clinical Interview for the DSM-IV (SCID): .50, $p < .01$; General Health Questionnaire (GHQ): .60, $p < .01$; Semistructured Interview for Survivors of War (SISOW): .20, $p < .05$.

Renner, Salem, and Ottomeyer (2006) administered the CAPS and several others measures to 150 asylum seekers from Chechnya, Afghanistan, and West Africa. Other measures used by Renner et al. (2006) included the Hopkins Symptom Checklist-25 (HSCL-25; Mollica et al., 1987), the Harvard Trauma Questionnaire (HTQ; Mollica et al., 1992), the Impact of Event Scale - Revised (IES-R; Weiss & Marmar, 1997), the Bradford Somatic Inventory (BSI; Mumford et al., 1991), and the Social Adaptation Self-Evaluation Scale (SASS; Bosc, Dubini, & Polin, 1997). Cronbach's alpha for the full scale was .90 for participants from Chechnya, .91 for those from Afghanistan, and .91 for those from West Africa (Renner et al., 2006). According to Renner et al., logistic regression revealed that the CAPS was superior in its ability to predict traumatization, relative to the other measures, in all three ethnic groups.

Hinton et al. (2006) administered the CAPS and the SCID to a sample of 179 Cambodian refugees exposed to genocide. Internal consistency alphas for the CAPS

ranged from .86 to .91 across the subscales, with a coefficient alpha of .92 for the full scale (Hinton et al., 2006). Hinton et al. reported corrected item-total correlations ranging from .48 to .85, and the CAPS demonstrated good interrater (intraclass correlation coefficient [ICC] = .92) and test-retest reliability (ICC = .84).

Summary of CAPS psychometric studies. Overall, the research suggests that the CAPS is a highly reliable diagnostic instrument. Internal consistency of the CAPS ranged from .90 to .94 for the full scale (Renner et al., 2006; Weathers et al., 2001) and from .73 to .91 (Blake et al., 1990; Hinton et al., 2006) for the subscales. Test-retest reliability was reported by Hinton et al. (2006) at one week as .84 and by Weathers et al. (2001) at two to three days as ranging from .90 to .98 across three separate rater pairs.

Although the CAPS was developed to represent the three-factor structure of PTSD as outlined in the diagnostic criteria, the extant empirical data suggest that the underlying factor structure is somewhat different. When performing confirmatory factor analysis, Palmieri et al. (2007) found evidence for a four-factor model composed of reexperiencing, avoidance, numbing, and hyperarousal, which paralleled results obtained by King et al. (1998). Using a Bosnian translation of the CAPS, Charney and Keane (2007) found evidence for a two-factor model composed of intrusion/avoidance and hyperarousal/numbing. Although this is not a comprehensive review of the literature on the factor structure of the CAPS, it illustrates the common finding that it is not best conceptualized as a three factor scale.

Studies examining the convergent and discriminant validity of the CAPS are largely supportive of the instrument. Several studies reported correlations with other PTSD measures, including the Combat Exposure Scale, the Mississippi Scale for

Combat-Related PTSD, the Keane PTSD Scale of the MMPI, the Semistructured Interview for Survivors of War, the Harvard Trauma Questionnaire, and the Impact of Events Scale. These correlations ranged from .20 with the Semistructured Interview for Survivors of War to .84 with the Mississippi Scale for Combat-Related PTSD. Correlations were also reported with other psychological measures of constructs similar to PTSD, but not isomorphic, including the Beck Depression Inventory, the Structured Clinical Interview for the DSM-IV, the General Health Questionnaire, the Hopkins Symptom Checklist-25, the Bradford Somatic Inventory, and the Social Adaptation Self-Evaluation Scale. Correlations between the CAPS and these measures ranged from .50 with the Structured Clinical Interview for the DSM-IV to .65 with the Beck Depression Inventory. As expected, in general the CAPS was better correlated with measures of PTSD than with measures of other psychological constructs or disorders.

Limitations of the CAPS. Weathers et al. (2001) listed three main criticisms of the CAPS. The authors point out that the CAPS is cumbersome because of its length. It takes almost an hour to administer and must be administered by a trained clinician. Although many CAPS questions are optional probes, individuals showing symptoms require the administration of these optional questions (Weathers et al., 2001). Second, critics of the CAPS assert that it takes too long to learn. A final criticism includes the overlap of the frequency and intensity prompts, which examiners have complained are redundant (Weathers et al., 2001). Additionally, although many individuals avoid disclosing sensitive information until they feel comfortable with a clinician, the CAPS requires individuals to report highly sensitive information to an examiner upon the first

meeting. Of course, this last shortcoming is not a limitation of the CAPS alone but is a characteristic of all interview-based assessments.

Computerized PTSD Scale – Multimedia Version

An alternative to the CAPS is a computerized adaptation that would address several of these limitations. Previous attempts have been made to computerize the CAPS. Neal, Busuttil, Herepath, and Strike (1994) administered a computerized version and the original CAPS to a group of 40 military members with varying trauma histories. They found that 95% of the cases were diagnosed accurately and reported a kappa of .90 between the computerized version and the CAPS with regard to PTSD presence (Neal et al., 1994). Alpha coefficients for the intensity scores were over .90 for both the computerized version and the CAPS, and the correlation for intensity ratings between the two versions was .95 (Neal et al., 1994).

More recently, the Computerized PTSD Scale-Multimedia Version, or CPS-M, was developed by Richard et al. (1997). Because it is computerized, the CPS-M requires no training to administer. It also requires no clinician time to administer, providing an additional time savings. In addition, participants can complete the administration in half the time required to administer the CAPS. Mainka (2005) administered the CAPS and CPS-M to a sample of 40 students with a trauma history and reported item-level correlations between the instruments ranging from .42 to .89 and scale-level correlations of .86, .89, and .84 for subscales B, C, and D, respectively. The Total Severity Score correlation between the two measures was .92, indicating strong associations between the instruments (Mainka, 2005). Using the CAPS as the criterion, Mainka found sensitivity of .63, specificity of .80, and overall diagnostic efficiency of .78 using a cut score of 45.

Aside from these two adaptations, no other attempts to computerize the CAPS are published in the literature. Furthermore, Neal et al. (1994) have not reported any further validation attempts for their instrument. This study aims to validate the CPS-M in a clinical sample using confidence interval analysis to determine equivalence and signal detection theory to determine diagnostic accuracy, relative to the CAPS.

Confidence Interval Analysis

Confidence interval analysis is a statistical procedure used to establish equivalency of two methods or instruments. Rogers, Howard, and Vessey (1993) argued that two instruments are equivalent if the confidence interval of the test instrument is contained entirely within a prespecified equivalence interval of the criterion. The equivalence interval has customarily been defined as ± 10% of the criterion mean and reflects the legacy of the confidence interval approach in evaluating the equivalence of newly developed pharmaceuticals (Rogers et al., 1993). To conduct the analysis, Rogers et al. recommend using the 90% confidence interval to determine the equivalency of the instrument while the 95% confidence interval is used to assess whether the test and criterion means are significantly different. If the 90% confidence interval for the CPS-M is entirely encompassed by the CAPS equivalence range (Mean of CAPS ± 10%), one would conclude that the two instruments are equivalent. If the 95% confidence interval of the test instrument includes zero (i.e., the difference between the means is zero), one may conclude that the two instruments are not statistically different.

Confidence intervals are determined using the following formula: $(M_1\text{-}M_2) \pm (z_\alpha)(S_{M1\text{-}M2})$. In this equation, M_1 represents the criterion mean and M_2 represents the test instrument mean. To establish a 90% confidence interval, $z_\alpha = 1.645$, and $z_\alpha = 1.96$ when

determining a 95% confidence interval. This value is multiplied by the standard error of the difference between means ($S_{M1\text{-}M2}$) and then added to or subtracted from the difference between means to produce the upper and lower confidence limits, respectively. The upper and lower limits of the confidence interval are then divided by the criterion mean, turning them into percentages of the criterion mean. This standardizes the values and allows for direct comparison with the ±10% equivalence interval.

Four classifications are possible using confidence interval analysis. "Statistically Different and Not Equivalent" applies when the 90% confidence interval is not contained entirely within the equivalence interval and the 95% confidence interval excludes zero. Two examples of this case are provided in the first panel of Figure 1. Figure 1 expresses the equivalence and confidence intervals as percentages of the criterion mean, which standardizes the values. "Not Different and Statistically Equivalent" applies when the 90% confidence interval falls within the equivalence interval and the 95% confidence interval includes zero. See the second panel of Figure 1. "Statistically Different and Statistically Equivalent" refers to cases in which the 90% confidence interval falls within the equivalence interval, but the 95% confidence interval excludes zero. See the third panel of Figure 1. Finally, "Not Different and Not Equivalent" refers to cases in which the 90% confidence interval is not contained entirely within the equivalence interval and the 95% confidence interval includes zero. See the fourth panel of Figure 1.

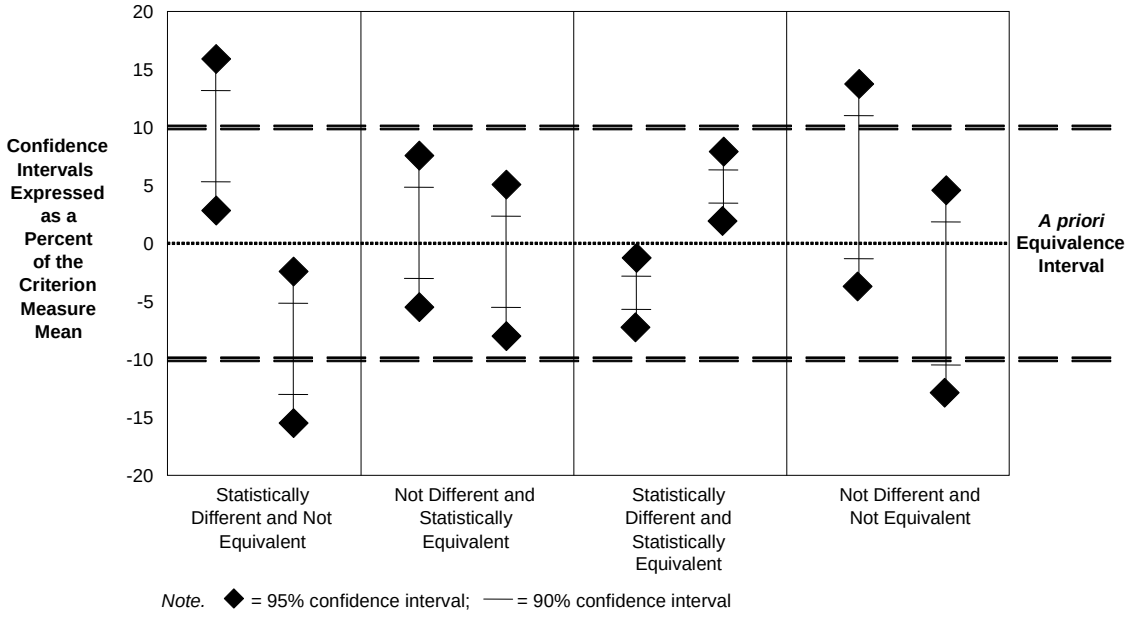

Figure 1. Rogers et al. (1993) confidence interval classifications.

Signal Detection Theory

When validating a new assessment instrument, it is necessary to determine its ability to accurately predict or its association with scores on a relevant outcome variable, also known as criterion related validity (Weathers, Keane, King, and King, 2001). Signal detection theory has been used in psychological and medical studies to compare new instruments, treatments, or diagnostic tools with already established "gold standards" within the field. Briefly, signal detection techniques provide a way to evaluate the effect of varying a cutoff score on a diagnostic test while holding constant a criterion against which the test is measured. As a result, a variety of indices regarding diagnostic efficiency may be calculated as a function of selected test cutoff points.

Signal detection theory was originally used in electrical engineering to determine the presence of a signal in background noise (DeCarlo, 1998; Siegel, Vukicevic, Elliott,

& Kraemer, 1989; Youngstrom, Findling, Danielson, & Calabrese, 2001). The same methods can also describe and predict the performance of a receiver, such as a diagnostic interview, in detecting the presence of a psychological stimulus of psychiatric diagnosis (Greig, 1990). Tanner and Swets (1954) were among the first to apply signal detection theory to the field of psychology when they examined the threshold at which humans were capable of detecting and recognizing sensory input. The technique was later expanded to allow researchers to determine the cut-point at which instruments made optimally accurate classifications (Swets, 1998).

Several signal detection terms are commonly employed. *Sensitivity* is the probability of a positive test, given a positive diagnosis on the criterion measure. *Specificity* is the probability of a negative test given a negative diagnosis on the criterion measure. The probability of a diagnosis on the criterion measure, given a positive test result, is called *positive predictive power*. The probability of non-diagnosis on the criterion measure, given a negative test result, is called *negative predictive power*.

Overall diagnostic efficiency, or *diagnostic utility*, is the rate of agreement between the test and criterion measure across all cases. In the present study, the term *actual diagnosis* refers to the diagnosis provided via ratings on the CAPS criterion. *True positives* or *true negatives* occur when the positive or negative diagnosis provided by the CPS – M agrees with the criterion. When a test assessment instrument gives a diagnosis that does not correspond with the actual diagnosis, the result is a *false positive* or *false negative*. *False positives* occur when the test instrument provides a positive diagnosis when the disorder is not present, and a *false negative* occurs when the test instrument fails to provide a positive diagnosis when the disorder is present. The cases where the

new instrument agrees with the actual diagnosis (*true positives* or *true negatives*) are *hits*, while cases where the new instrument does not agree with the actual diagnosis (*false positives* and *false negatives*) are termed *misses*.

Figure 2 shows the 2 X 2 contingency table that results from diagnostic efficiency calculations. According to Swets (1988), the decision criterion is the amount of favorable evidence necessary to issue a positive diagnosis. This decision criterion is determined by plotting the false positive proportion against the true positive proportion (Swets, 1988). Since there are as many potential diagnostic cutoff points on the test as there are discrete points on its scale of measurement, an underlying distribution of diagnostic efficiency can be created. This underlying probability distribution is represented by the receiver operating characteristic, or the ROC curve (DeCarlo, 1998; Swets, 1988).

		Test Result: Positive	Test Result: Negative	
Criterion	Present	a True Positive (Hit)	b False Negative (Miss)	a + b
	Absent	c False Positive (False Alarm)	d True Negative (Correct Rejection)	c + d
		a + c	b + d	a + b + c + d = N

$$\textit{Sensitivity} = \frac{a}{a+b} \qquad \textit{Specificity} = \frac{d}{c+d}$$

Figure 2. Two-by-two contingency table used for calculation of diagnostic efficiency statistics.

The ROC curve is obtained by plotting sensitivity, P(S/s) against false alarm rate, P(S/n), where P represents probability, S represents diagnosis of a disorder, and s and n represent presence or absence of a disorder, respectively (Treisman & Faulkner, 1984). In other words, the researcher plots sensitivity as a function of (1 – specificity). This produces a curve that allows for the determination of the optimal cut point for diagnosis. According to Dobie et al. (2002), a ROC curve will have an "initial steep section where sensitivity increases while the false positive rate (1 - specificity) changes only minimally. This is followed by a bend in the curve, then a flattened section where the false positive rate increases rapidly with little improvement in sensitivity (p. 369)." The ROC curve is used to determine the optimal cut point score that maximizes the detection of true positives and true negatives. This cut score is the *sensitivity/(1 - specificity)* value found at upper left corner of the curve (Dobie et al., 2002).

According to Youngstrom et al. (2001), the accuracy of the ROC can be determined by calculating the area under the curve or AUC. The area under the curve, referred to by Swets (1988) as *A* or A_z, denotes the amount of discrimination between true and false positives. This value also provides a natural summary of the test's overall diagnostic efficiency (Blume, 2009). When true and false positives are equal, the area under the curve will be 0.50 and is represented by a diagonal line dividing the total area in half, indicating no discrimination (Swets, 1988). In other words, the test measure is performing no better than chance. An AUC of 1.0 would identify perfect discrimination, so the line would trace the left vertical axis and the upper horizontal axis (Swets, 1988). Swets concluded that AUCs of .50 - .70 show low accuracy, AUCs of .70 - .90 show medium accuracy, while AUCs of .90 - 1.00 show high accuracy. Figure 3 shows a

sample ROC curve for two PTSD measures, with the first substantially outperforming the second.

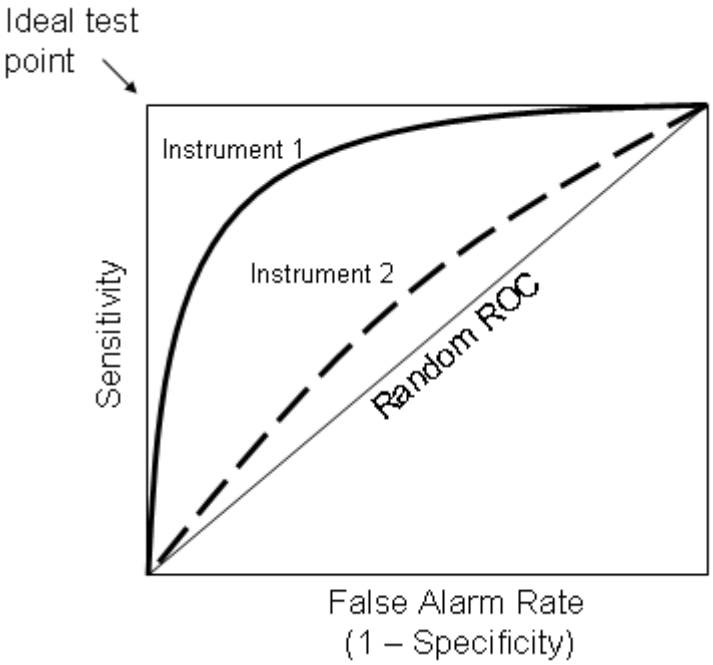

Figure 3. Sample Receiver Operating Characteristic (ROC) curves for two measures of PTSD.

A decision criterion for an instrument will depend on prior probabilities of true and false positives (Swets, 1988). A bias exists in diagnostic efficiency measures because of an instrument's tendency to favor a certain diagnostic alternative (Swets, 1988). For example, a diagnostic instrument such as the CPS – M is designed to test for the presence of PTSD symptoms, so it is biased toward true positives. This bias provides the rationale for converting raw frequencies to proportions, as shown in Figure 2. These proportions take into account the population estimates, referred to as prior probabilities, of positive and negative events, or diagnoses. Swets recommended adopting a lenient criterion for events that have high prior probabilities. For example, if positive diagnosis has a high probability in the population, the instrument should employ a lenient criterion

for positive diagnosis. However, the decision criterion will also depend on the possible costs and benefits associated with correct and incorrect outcomes (Swets, 1988). In situations where the cost associated with false negatives is high, the decision criterion will be lenient, while situations with high costs for false positives would imply a strict criterion. For example, the cost of predicting a malignant cancer that does not occur (false positive) is small relative to the cost of failing to detect the cancer (false negative). Conversely, performing a dangerous operation on a patient who turns out to not have a disease (false positive) has very high costs (Swets, 1988). As Swets, Tanner, and Birdsall (1961) noted, the four probabilities are interdependent. An increase in the probability of a hit can be achieved only by accepting an increase in the probability of a false alarm, and decreases in the other probabilities. Thus, a given criterion yields a particular balance among the probabilities of the four possible outcomes. In addition, the balance desired by an observer in any instance will determine the optimal location of his criterion. This makes it necessary to find the optimum point that balances hits and misses.

Swets (1998) lists two challenges inherent in signal detection theory. Type I challenges involve balancing hits and misses. Type II challenges involve maximizing accuracy. With regard to the former, the ROC curve establishes the threshold of evidence across a variety of test cut scores that allows the instrument to maximize accurate diagnostic decisions (Swets, 1988). With regard to Type II challenges, ROC analysis plots the balance between specificity and sensitivity as a function of changing cut scores in the test instrument, producing a curve that shows the optimal position for efficient signal detection (Treisman & Faulkner, 1984; Youngstrom et al., 2001).

Prior Signal Detection Work

Signal detection comparing the CAPS to other PTSD measures. The CAPS has been used as the criterion in several signal detection studies. These studies provide excellent examples of the methodology that was employed in the current study. They also provide a benchmark against which the performance of the CPS-M may be compared. Prior signal detection studies have included the PTSD Checklist, both the original (PCL) and civilian (PCL-C) versions, the Impact of Events Scale, the Mississippi Scale for Combat-Related PTSD, the Penn Inventory for PTSD, the PTSD Symptoms Scale-Interview Version (PSS-I), the Traumatic Stress Symptoms Checklist (TSSC), and the Short PTSD Rating Interview (SPRINT). Several studies focused on examining the diagnostic efficiency of brief screening instruments for PTSD. The diagnostic efficiency of the CAPS was also assessed when comparing face-to-face to videoconferencing administrations (Porcari et al., 2009). Following this review of the literature, Table 1 provides a list of these studies and the diagnostic efficiency of the respective instruments, compared to the CAPS.

PTSD Checklist. Several studies have examined the diagnostic efficiency of both the original PTSD Checklist (PCL; Weathers et al., 1993) and the civilian version of the PCL (PCL-C; Weathers, Litz, Huska, & Keane, 1994). To date, no studies were found that used the CAPS as a gold standard against which to assess the diagnostic efficiency of the military version of the PCL (PCL-M; Weathers, Huska, & Keane, 1991).

PTSD Checklist (PCL). Blanchard, Jones-Alexander, Buckley, and Forneris (1996) administered the CAPS and PTSD Checklist (PCL) to 40 motor vehicle accident and sexual assault victims. The PCL was mailed to participants prior to the CAPS

interview, and participants were instructed to complete the PCL and bring it to the interview. Participants then completed a CAPS interview. All interviews were tape recorded, and 19 were rescored by an independent judge who was blind to the original diagnosis. The kappa for agreement between judges was .84, $p < 0.001$. The correlation between the PCL and the CAPS was .93, and overall diagnostic efficiency was 0.90. Using a cut-off score of 44 yielded a sensitivity of .94 and a specificity of .86 (Blanchard et al., 1996). Blanchard et al. recommended using the PCL as a screening instrument for posttraumatic stress disorder based on its excellent agreement with the CAPS.

Forbes, Creamer, and Biddle (2001) studied 97 male veterans who completed a questionnaire battery and the CAPS at admission to a twelve-week treatment program. The first four weeks of the program were inpatient and the remaining eight weeks consisted of one day of outpatient treatment per week. Participants were followed up at nine months post-treatment with the CAPS and PCL. The CAPS-PCL correlations were .30 at intake and .62 at follow-up. Overall diagnostic efficiency at intake was .81 with a reported sensitivity reported to be .98. Specificity was not reported for intake scores. At follow-up, a cut-off score of 50 yielded sensitivity of .91, specificity of .40, and overall diagnostic efficiency of .80 (Forbes et al., 2001).

Grubaugh, Elhai, Cusack, Wells, and Frueh (2007) examined the diagnostic efficiency of the PCL relative to the CAPS in a sample of 44 community mental health patients. They reported overall efficiency of .76, sensitivity of .69, and specificity of .78 using a cut score of 54 (Grubaugh et al., 2007). Grubaugh et al. found positive and negative predictive power values of .82 and .64, respectively.

The diagnostic efficiency of the PCL relative to the CAPS was also examined in a clinical sample of adults age 65 or older (Hudson, Beckford, Jackson, & Philpot, 2008). Using the recommended cut score of 50, Hudson et al. (2008) found sensitivity of .40, specificity of .97, and positive predictive value of .57. However, an optimal cut point of 36 resulted in values of .90, .87, and .45, respectively (Hudson et al., 2008). Hudson et al. determined that the PCL was an acceptable screening measure in older adults when an adjusted cut point was used.

Ventureyra, Yao, Cottraux, Note, and Mey-Guillard (2002) administered the CAPS and a French translation of the PCL to a group of 113 outpatients with PTSD and 31 nonclinical control participants. All participants completed the PCL, Beck Depression Inventory, Fear Questionnaire, and the Hamilton Rating Scale of Anxiety. Scores on the PCLS were compared to CAPS interview ratings. Using a cutoff score of 44 (recommended by Blanchard, Jones-Alexander, Buckley, & Forneris, 1996) they obtained a sensitivity of .97 and specificity of .87 (Ventureyra et al., 2002). Overall diagnostic efficiency of the PCL was .94.

Walker, Newman, Dobie, Ciechanowski, and Katon (2002) studied diagnostic efficiency of the PCL relative to the CAPS in a sample of 1225 women enrolled in a staff model HMO. All participants completed the PCL, and 261 were also administered the CAPS. In order to screen for childhood trauma, participants completed the Childhood Trauma Questionnaire. Using a cut-off score of 30, Walker et al. (2002) found optimal sensitivity of .82 and specificity of .76. Although this cut score is lower than those previously reported, the authors indicated that this may have been due to use of a sample

with a more limited trauma history than in previous studies (Walker et al., 2002). Thus, a lower cut score may produce more optimal diagnostic efficiency in a sub-clinical sample.

In summary, the studies assessing the utility of the PCL relative to the CAPS reported overall diagnostic efficiencies ranging from .76 to .94 (Grubaugh et al., 2007; Ventureyra et al., 2002). Sensitivity scores ranged from .40 to .98 (Hudson et al., 2008; Forbes et al., 2001), and specificity scores were between .40 and .97 (Forbes et al., 2001; Hudson et al., 2001). Recommended cut-off scores ranged from 30, which produced sensitivity of .82 and specificity of .76 (Walker et al., 2002), to 54, which resulted in sensitivity of .69 and specificity of .78 (Grubaugh et al., 2007). A cut-off score of 44, utilized by Ventureyra et al. (2002), resulted in the highest diagnostic efficiency statistics, with overall efficiency of .94, sensitivity of .97, and specificity of .87.

PTSD Checklist – Civilian version (PCL-C). In order to assess diagnostic efficiency of the PTSD Checklist – Civilian version (PCL-C), Bollinger, Cuevas, Vielhauer, Morgan, and Keane (2008) examined the diagnostic efficiency of the PTSD Checklist-Civilian version (PCL-C) in a sample of HIV seropositive individuals. Using a cut point of 52, they reported overall diagnostic efficiency of .82, but sensitivity was .71 and specificity was .84 (Bollinger et al., 2008). The recommended cut point of 50 yielded the optimal balance between sensitivity and specificity, at .86 and .79, respectively (Bollinger et al., 2008).

Dobie et al. (2002) administered the PCL-C and the CAPS to 282 female veterans. Prior to the CAPS administration, participants completed the PCL and other self-report questionnaires designed to evaluate their overall quality of life. They then completed the Alcohol Use Disorder and Associated Disabilities interview to assess for the presence of

substance use disorders. Finally, participants were administered the CAPS by a trained clinician. A ROC analysis found an AUC of .86. Using an optimal cut-point score of 38, sensitivity was .79 and specificity was .79 (Dobie et al., 2002).

Lang, Laffaye, Satz, Dresselhaus, and Stein (2003) administered the PCL-C and the CAPS to a sample of 419 women at the San Diego VA Healthcare System. Participants also completed the SF-36 (Ware & Sherbourne, 1992), a short questionnaire designed to measure health-related quality of life. Employing the recommended cut-off score of 50, they found overall diagnostic efficiency of .74, sensitivity of .39, and specificity of .74. These statistics were lower than those reported in other signal detection studies using a cut-off of 50 on the PCL-C. When Lang et al. used a cut-off of 28, they found overall efficiency of .78, sensitivity of .94, and specificity of .68. They concluded the PCL-C was an adequate screening measure of PTSD symptoms in a population of female veterans.

In summary, the studies assessing the utility of the PCL-C relative to the CAPS reported overall diagnostic efficiencies ranging from .74 to .86 (Lang et al., 2003; Dobie et al., 2002). Sensitivity scores ranged from .39 to .94 (Lang et al., 2003; Lang et al., 2003), and specificity scores were between .68 and .84 (Lang et al., 2003; Bollinger et al., 2008). Recommended cut-off scores ranged from 28, which produced sensitivity of .94 and specificity of .68 (Lang et al., 2003), to 52, which resulted in sensitivity of .71 and specificity of .84 (Bollinger et al., 2008). Bollinger et al. (2008) reported the highest balance of sensitivity and specificity at .86 and .79, respectively, using a cut score of 50. However, they did not report overall efficiency using this cut score. A cut-off score of 38,

utilized by Dobie et al. (2002), resulted in the highest diagnostic efficiency statistics, with overall efficiency of .86, sensitivity of .79, and specificity of .79.

Impact of Events Scale (IES). Coffey, Gudmundsdottir, Beck, Palyo, and Miller (2006) examined the diagnostic efficiency of the Impact of Events Scale and the PTSD Symptoms Scale-Self Report relative to the CAPS in a sample of 229 motor accident survivors. For the IES, Coffey et al. (2006) reported overall diagnostic efficiency of .80, with a sensitivity of .91 and specificity of .72 using a cut score of 27. According to Coffey et al., this cut score is somewhat lower than those in studies of other trauma populations.

Neal, Hill, Hughes, Middleton, and Busuttil (1995) used the CAPS as the criterion against which to measure several other PTSD scales in a sample of World War II Far East prisoners of war. Thirty participants completed the Impact of Events Scale, the MMPI-PTSD subscale, and the Mississippi Scale for Combat-Related Stress Disorder. They were then administered the CAPS by trained clinicians. Neal et al. (1995) found that the Mississippi Scale correlated .81 with the CAPS, and the MMPI-PTSD subscale correlated .71. A cut-off score of 81 on the Mississippi Scale produced optimal sensitivity (.78) and specificity (.57). An MMPI-PTSD subscale cutoff of 17 yielded slightly higher levels of sensitivity (.89) and specificity (.62). Results for the Impact of Events scale were less impressive. A cutoff of 35 produced sensitivity of .67 and specificity of .57 (Neal et al., 1995). Overall diagnostic efficiency scores were not reported for any of the instruments in this study.

Sondergaard, Ekblad, and Theorell (2003) compared the Health Leaflet, a screening procedure, to the Impact of Event Scale-22 (IES-22), the Harvard Trauma

Questionnaire (HTQ), and the CAPS. Participants were 86 recently resettled refugees from Iraqi ethnic groups who were interviewed by social workers. Sondergaard et al. (2003) found that a cut-off score of 77.5 on the HTQ yielded a sensitivity of .80 and specificity of .78. A cut-off of 65 on the IES-22 produced sensitivity of .72 and specificity of .71.

In summary, sensitivity scores for the original Impact of Events Scale ranged from .67 to .91 (Neal et al., 1995; Coffey et al., 2006). Specificity scores ranged from .57 to .72 (Neal et al., 1995; Coffey et al., 2006). Only the Coffey et al. (2006) study reported overall diagnostic efficiency at .80. The revised scale demonstrated moderate diagnostic efficiency in the Sondergaard et al. (2003) study, higher than that reported for the IES in the Neal et al. (1995) study and lower than that found by Coffey et al.. Overall, results for the IES do not appear to be as good as those reported for the PCL.

Mississippi Scale for Combat-Related PTSD (Mississippi Scale). Weathers et al. (1996) plotted the quality of sensitivity against the quality of specificity for the Mississippi Scale, the War-Zone PTSD Scale, and the Keane PTSD Scale of the MMPI (PK Scale). They determined PTSD diagnosis by administering either the PTSD module of the Structured Clinical Interview for DSM-III-R (SCID; Spitzer, Williams, Gibbon, & First, 1990) or the CAPS. The authors determined that the Mississippi Scale was the best PTSD predictor, as evidenced by this scale's overall diagnostic efficiency of .83, sensitivity of .83, and specificity of .83, using a cut score of 109 (Weathers et al., 1996). They also reported that the Spearman rank point-biserial correlation (r_{pb}) between the Mississippi Scale and the interview diagnosis was .69. This correlation equals the proportion of area under the ROC curve and measures the overall quality of the scale

(Weathers et al., 1996). This was higher than correlations for the WZ-PTSD scale and the PK Scale, which were .62 and .57, respectively.

Penn Inventory for PTSD. Another study that used the CAPS as a criterion was performed by Scragg, Grey, Lee, Young, and Turner (2001). The Penn Inventory for PTSD was administered to a group of 80 males and females referred to an outpatient clinic for trauma in the UK. In addition to completing the Penn Inventory and the CAPS assessment, participants also completed the General Health Questionnaire-28 (GHQ-28; Goldberg & Hillier, 1979), the Beck Depression Inventory (BDI; Beck et al., 1961), and the Beck Anxiety Inventory (BAI; Beck & Steer, 1993). Relative to the CAPS, the Penn Inventory had an overall diagnostic efficiency of .81 for male participants and .83 for females, using a cut score of 35. Sensitivity was .90 and .89 for males and females, respectively. Specificity was .55 for males and .67 for females. Scragg et al. (2002) recommended using the Penn Inventory as a screening measure to assess mental well-being after a trauma, rather than as a diagnostic tool.

PTSD Symptoms Scale. Coffey et al. (2006) examined the diagnostic efficiency of the PTSD Symptom Scale-Self Report (PSS-SR) and the Impact of Events Scale relative to the CAPS in a sample of 229 motor accident survivors. For the PSS-SR, Coffey et al. reported overall diagnostic efficiency of .74, with a sensitivity of .91 and specificity of .62 using a cut score of 14. According to Coffey et al., this cut score is consistent with those in studies of other trauma populations.

Foa and Tolin (2000) compared the PTSD Symptom Scale-Interview Version (PSS-I), the CAPS, and the SCID. Participants were 64 clinical and non-clinical adult volunteers from a community sample. They were administered the PSS-I, the CAPS, and

the Structured Clinical Interview for DSM-III-R (SCID) in counterbalanced order. Foa and Tolin (2000) used liberal, moderate, and conservative scoring rules recommended by Blanchard et al. (1995) and Weathers et al. (1999). For the CAPS, they found that using either the moderate scoring rule proposed by Weathers et al. or the conservative scoring rule proposed by Blanchard et al. yielded overall diagnostic agreement of .88, sensitivity of .71, and specificity of .94 relative to the SCID. For the PSS-I, they found overall agreement of .80, sensitivity of .86, and specificity of .78 compared with the SCID. Based on these results, they concluded that the CAPS yielded better overall agreement and specificity, while the PSS-I produced better sensitivity for PTSD diagnosis.

Traumatic Stress Symptoms Checklist (TSSC). Basoglu et al. (2001) assessed the validity of the Traumatic Stress Symptoms Checklist (TSSC) in a sample of 130 earthquake survivors in Turkey. They compared diagnoses based on the CAPS and the Major Depressive Episode module of the Semistructured Clinical Interview for DSM-IV (SCID: First, Spitzer, Gibbon, & Williams, 1996). The authors found a cut-off of 25 on total scores of 17 PTSD items to yield optimal sensitivity of .81 and specificity of .81. The authors found that the TSSC demonstrated satisfactory sensitivity and specificity in providing a diagnosis of PTSD (Basoglu et al., 2001).

Short Post-Traumatic Stress Disorder Rating Interview (SPRINT). Kim et al. (2008) assessed the diagnostic efficiency of the Korean version of the Short Post-Traumatic Stress Disorder Rating Interview (K-SPRINT). They administered a packet consisting of the K-SPRINT, CAPS, Beck Depression Inventory (BDI), and State Trait Anxiety Inventory (STAI) to a sample of 197 individuals. A cut score of 15 resulted in

overall efficiency of .92, sensitivity of .91, and specificity of .93, suggesting that the K-SPRINT is a good diagnostic measure of PTSD in this population (Kim et al., 2008).

Brief screening instruments. Kimerling, Ouimette, et al. (2006) administered Breslau's 7 item PTSD screen (Breslau et al., 1999) to 134 patients recruited from VA medical center primary care clinics to determine its diagnostic efficiency relative to the CAPS. Kimerling, Ouimette, et al. (2006) determined that a cut score of 4 resulted in optimal sensitivity of .85 and specificity of .84.

Kimerling, Trafton, and Nguyen (2006) assessed the diagnostic efficiency of a 4-item screen for PTSD, the Primary Care PTSD Screen (PC-PTSD), relative to the CAPS in a sample of 97 patients recruited from substance abuse clinics at a VA medical center. A cutoff score of 3 resulted in optimal overall efficiency of .84, sensitivity of .91, and specificity of .80 (Kimerling, Trafton, et al., 2006). Positive and negative predictive values using this cut score were .69 and .95, respectively (Kimerling, Trafton, et al., 2006).

CAPS administered face-to-face and via teleconferencing. Porcari et al. (2009) administered the CAPS to twenty male veterans seeking mental health services for PTSD. The CAPS was administered both face-to-face and via teleconferencing. They found significant correlations, ranging from .74 (subscale C: avoidance/numbing) to .92 (subscale B: reexperiencing), between the two assessment methods on the subscales and total severity scores. Confidence interval analysis revealed statistical equivalence between the two methods, and signal detection analysis resulted in overall diagnostic efficiency of .85, sensitivity of .94 and specificity of .33, suggesting a moderate agreement between the two methods (Porcari et al., 2009).

Summary of studies. Diagnostic efficiency of the PCL relative to the CAPS ranged from .76 to .94 (Grubaugh et al., 2007; Ventureyra et al., 2002). PCL-C efficiencies ranged from .74 to .86 (Lang et al., 2003; Dobie et al., 2002). Overall efficiency for the Impact of Events Scale was .80 in a study by Coffey et al. (2006), and sensitivity scores were .67 and .91 according to Coffey et al. and Neal et al. (1995), respectively. Reported specificities were .57 (Coffey et al., 2006) and .72 (Neal et al., 1995). Sondergaard et al. (2003) reported sensitivity of .72 and specificity of .71 for the Impact of Event Scale-22 in a sample of refugees.

Diagnostic efficiency of the Penn Inventory for PTSD relative to the CAPS was .81 for males and .83 for females in an outpatient trauma clinic (Scragg et al., 2001). Foa and Tolin (2000) reported overall efficiency of .80 for the PTSD Symptom Scale – Interview Version, and Coffey et al. (2006) reported overall diagnostic efficiency of .74, with a sensitivity of .91 and specificity of .62 for the PTSD Symptom Scale – Self Report. The Traumatic Stress Symptoms Checklist was compared to the CAPS and produced sensitivity of .81 and specificity of .81 in a sample of earthquake survivors (Basoglu et al., 2001). Kimerling, Trafton, et al. (2006) reported overall efficiency of .84, sensitivity of .91, and specificity of .80 on the Primary Care PTSD Screen (PC-PTSD), and Kimerling, Ouimette, et al. (2006) found optimal sensitivity of .85 and specificity of .84 on Breslau's 7 item PTSD screen. Finally, Porcari et al. (2009) administered the CAPS both face-to-face and via teleconferencing and reported overall efficiency of .85, sensitivity of .94, and specificity of .33.

Not all signal detection studies reported overall diagnostic efficiency and fewer still reported AUC data. Some reported sensitivity and specificity scores, while others

included Spearman rank point-biserial correlations or positive predictive power values. Other studies included the cut scores that yielded the optimal sensitivity and specificity. According to Youngstrom et al. (2001), sensitivity and specificity are theoretically independent of base rates, making them better than positive or negative predictive power in terms of generalizability to other samples. As such, the focus of the current study will be on diagnostic efficiency as measured by sensitivity, specificity, and the area under the curve (AUC). The optimal cut score will also be reported, as cut scores are crucial in determining an instrument's threshold for accurate diagnosis within a given population.

Table 1

Signal Detection Studies Using the CAPS as the Diagnostic Criterion

Study	Population	Test Instrument(s)	Results
Blanchard et al. (1996)	40 motor vehicle accident and sexual assault victims	PTSD Checklist	PCL-CAPS *r* = .93, diagnostic efficiency = .90. Sensitivity = .94, specificity = .86 with cut score of 44.
Forbes et al. (2001)	97 Vietnam veterans with combat-related PTSD	PTSD Checklist	At intake, PCL-CAPS *r* = .30, diagnostic efficiency = .81. At 9-month follow-up, PCL-CAPS *r* = .62, sensitivity = .94, specificity = .20 with cut score of 45.
Grubaugh et al. (2007)	44 traumatized community mental health patients	PTSD Checklist	Diagnostic efficiency = 76, sensitivity = .69, specificity = .78, positive predictive power = .82, negative predictive power = .64.
Hudson et al. (2008)	Clinical sample of adults age 65 or older	PTSD Checklist	Sensitivity = .40, specificity = .97, positive predictive power = .57 with cut score of 50; sensitivity = .90, specificity = .87, positive predictive power = .45 with cut score of 36.
Ventureyra et al. (2002)	113 outpatients suffering from PTSD, compared to 31 nonclinical controls	French translation of PTSD Checklist	Sensitivity = .97, specificity = .87, diagnostic efficiency = .94 with cut score of 44.

Study	Population	Test Instrument(s)	Results
Walker et al. (2002)	1225 women enrolled in a staff model HMO	PTSD Checklist	Sensitivity = .82, specificity = .76, cut score 30.
Bollinger et al., (2008)	HIV seropositive individuals	PTSD Checklist-Civilian Version	Diagnostic efficiency = .82, sensitivity = .71, and specificity = .84, cut score 52; sensitivity = .86, and specificity = .79, cut score 50.
Dobie et al. (2002)	282 female veterans	PTSD Checklist – Civilian version	Sensitivity = .79, specificity = .79 with cut score 38. AUC = .86.
Lang et al. (2003)	419 female veterans	PTSD Checklist – Civilian version	Diagnostic efficiency = .78, sensitivity = .94, specificity = .68, cut score 28.
Coffey et al. (2006)	229 motor vehicle accident survivors	Impact of Events Scale	IES overall diagnostic efficiency = .80, sensitivity = .91, specificity = .72, cut score 27.
Neal et al. (1995)	30 World War II Far East prisoners of war	Impact of Events Scale	IES sensitivity = .67, specificity = .57, cut score 35.
Sondergaard et al. (2003)	86 refugees from Iraqi ethnic groups	Impact of Event Scale-22	IES-22-CAPS sensitivity = .72, specificity = .71, cut score 65.
Scragg et al. (2001)	80 men and women referred to an outpatient clinic for trauma in the UK	Penn Inventory for PTSD	Diagnostic efficiency = .81 for men, .83 for women, cut score 35. Sensitivity = .90 for men, .89 for women. Specificity = .55 for men, .67 for women.

Study	Population	Test Instrument(s)	Results
Weathers et al. (1996)	202 Vietnam veterans	Mississippi Scale	Sensitivity = .83, specificity = .83, overall diagnostic efficiency = .83.
Coffey et al. (2006)	229 motor vehicle accident survivors	PTSD Symptom Scale – Self Report	PSS-SR overall diagnostic efficiency = .74, sensitivity = .91, specificity = .62, cut score 14.
Foa & Tolin (2000)	64 clinical and non-clinical adults	PTSD Symptom Scale-Interview	Diagnostic efficiency = .80, sensitivity = .86, specificity = .78, no cut score reported.
Basoglu et al. (2001)	130 earthquake survivors in Turkey	Traumatic Stress Symptoms Checklist	Sensitivity = .81, specificity = .81, cut score 25.
Kim et al. (2008)	197 patients; 87 PTSD, 47 other, 63 controls	Korean version of the Short PTSD Rating Interview	Diagnostic efficiency = .92, sensitivity = .91, and specificity = .93, cut score 15.
Kimerling, Ouimette, et al. (2006)	134 VA primary care patients	Breslau's 7 item PTSD screen	Sensitivity = .85, specificity = .84, cut score 4.
Kimerling, Trafton, et al. (2006)	97 patients recruited from VA substance abuse clinics	Primary Care PTSD Screen	Overall efficiency = .84, sensitivity = .91, specificity = .80, cut score 3.
Porcari et al. (2009)	20 male veterans	CAPS administered via teleconferencing	Overall efficiency = .85, sensitivity = .94, specificity = .33.

Signal detection comparing other PTSD instruments. Despite the fact that the CAPS is considered the "gold standard" in PTSD research, some studies have employed other instruments as the criterion against which to measure new PTSD assessment instruments. These criterion instruments have included a Chinese version of the Startle, Physiological Arousal, Anger, and Numbness scale (C-SPAN), the PTSD Checklist (PCL), the war-zone related PTSD scale (WZ-PTSD), and the SCID.

Chen, Shen, Tan, Chou, and Lu (2003) compared a Chinese version of the SPAN (C-SPAN) to a Chinese version of the Davidson Trauma Scale (DTS-C). The authors assessed 210 earthquake survivors of the 1999 Chi-Chi Earthquake. Chen et al. (2003) calculated diagnostic efficiency statistics based on varying cut points, and these statistics ranged from .33 to .98 for sensitivity and .43 to .94 for specificity. A cut score of 5 on the C-SPAN resulted in the highest overall efficiency at .80, as well as optimal sensitivity (.79) and specificity (.80).

The PTSD Checklist (PCL) was used as the criterion against which to assess the psychometric properties of the Impact of Events Scale-Revised (Creamer, Bell, & Familla, 2003). The authors administered the instruments to two samples of Vietnam veterans. One group was a treatment-seeking sample with confirmed PTSD (N = 120), and the other was a community sample with varying symptomatology (N = 154). Using total sample scores, the two scales were highly correlated (r = .84), and a cutoff on the PCL of 33 resulted in overall efficiency of .88, sensitivity of .91, and specificity of .82 (Creamer et al., 2003).

Weathers et al. (1996) developed a scale for assessing war-zone related PTSD (WZ-PTSD scale). In a psychometric study comparing the WZ-PTSD scale to the

Symptom Checklist-90-R, the instrument from which it was derived, the authors administered these two instruments, as well as several other convergent measures, to 202 male Vietnam veterans. These convergent measures included either the PTSD module of the Structured Clinical Interview for DSM-III-R (SCID) or the CAPS, the Mississippi Scale for Combat-Related PTSD, and either the MMPI or the MMPI-2. The Mississippi Scale demonstrated the highest overall efficiency (.83), followed by the WZ-PTSD. The WZ-PTSD demonstrated moderate sensitivity (.90) and specificity (.65) (Weathers et al., 1996).

Using an adaptation of the CAPS designed to assess schizophrenia (CAPS-S), Gearon, Bellack, and Tenhula (2004) reported diagnostic efficiency statistics relative to the SCID. Nineteen women with schizophrenia and drug use were administered the CAPS-S, the SCID, and the Impact of Events Scale. Gearon et al. (2004) reported overall efficiency of .78, sensitivity of .50, and specificity of 1.00. However, their small sample size significantly limited the generalizability of these findings.

Signal detection comparing psychological instruments outside of PTSD. As signal detection is a methodology ideally suited to comparing instruments and determining the diagnostic power of new instruments, it has been used in other domains of pathology. Studies employing signal detection in other domains will be briefly reviewed, as they provide further examples of "normal" performance of psychological instruments.

Signal detection has been used to evaluate the efficiency of measures assessing college maladjustment, adolescent psychiatric difficulties, substance abuse, depression, and bipolar disorder. Diagnostic efficiency statistics reported in these domains have been

similar to those found for PTSD instruments. Additionally, many studies have used signal detection techniques to evaluate the DSM-IV diagnostic criteria for attention-deficit/hyperactivity disorder (Mota & Schachar, 2000), autism (Siegel, Vukicevic, & Spitzer, 1990), and personality disorders (Farmer & Chapman, 2002; Fossati et al., 2005; Grilo, 2004; Grilo, Becker, Anez, & McGlashan, 2004). While these studies utilize signal detection techniques, their methodologies are very different from that being employed in this study.

Lauterbach, Garcia, and Gloster (2002) performed a psychometric study of the College Maladjustment Scale (Mt), a supplementary scale of the Minnesota Multiphasic Personality Inventory (MMPI-2). In addition to the Mt scale, they administered the Academic Performance Questionnaire, the Shipley Institute of Living Scale, the Traumatic Events Questionnaire (TEQ), and the PTSD Checklist-Civilian (CL-C) to 473 undergraduate students. Lauterbach et al. (2002) reported that a cut score of 29 or higher resulted in the best overall diagnostic utility. Diagnostic utility scores were .62 for individuals currently in therapy, .53 for individuals with a history of three or more therapy sessions, and .87 for individuals with a PTSD diagnosis (Lauterbach et al., 2002).

Pinto and Grilo (2004) administered selected scales from the Millon Adolescent Clinical Inventory (MACI) to 241 hospitalized adolescents. Scores on the MACI were compared to independent DSM-IV clinical discharge diagnoses. MACI scale diagnostic efficiency statistics ranged from .17 to .71 for sensitivity and from .40 to .93 for specificity, depending on the scale and the cut score used. The scale with the highest diagnostic efficiency was "substance abuse proneness," which demonstrated sensitivity

of .61 and specificity of .82 with a diagnosis of "drug use disorder," using a cut score of 75 (Pinto & Grilo, 2004).

The Neuropsychological Assessment Battery-Screening Module (NAB-SM) and the Neuropsychological Screening Battery (NSB) were administered to 84 substance abuse patients entering residential treatment (Grohman & Fals-Stewart, 2004). Using the NSB as the criterion, the authors found overall diagnostic efficiency of .88, sensitivity of .81, and specificity of .92, indicating excellent diagnostic accuracy of the test instrument.

Youngstrom et al. (2004) compared six potential bipolar disorder screening instruments to assess which produced best overall diagnostic accuracy. In addition to participating in the Semistructured Diagnostic Interview Using the Schedule of Affective Disorders and Schizophrenia for Children, scores were obtained for the 642 participants (318 aged 5-10; 324 aged 11 to 17) on the Parent Young Mania Rating Scale, the General Behavior Inventory, the Parent General Behavior Inventory, the Child Behavior Checklist, the Youth Self-Report, and the Teacher Report Form. Parent report (efficiency from .78 to .84 in both age groups) performed better than teacher report (efficiency of .57 in younger and .70 in older group) or adolescent measures (efficiencies of .67 and .71 on General Behavior Inventory and Youth Self-Report, respectively) at accurately classifying presence of bipolar disorder (Youngstrom et al., 2004).

Huprich, Sanford, and Smith (2002) compared the psychometric properties of the Depressive Personality Disorder Inventory (DPDI) to other measures of depression. In addition to the DPDI, they administered the Structured Clinical Interview for DSM-IV Axis II Disorders for Self-Report (SCID-II), the BDI-II, the Diagnostic Interview for

Depressive Personality (DIDP), the Early Life Events Questionnaire (ELEQ), the Provision of Social Relations Scale (PSR), and the Bell Object Relations and Reality Testing Inventory, Form O (BORRTI). Using the Diagnostic Interview for Depressive Personality as the criterion, Huprich et al. (2002) found sensitivity of .82, specificity of .80, and overall diagnostic efficiency of .81 with a cut score of 170 on the DPDI.

Signal detection comparing medical assessment instruments or techniques. Outside the domain of psychological assessment, signal detection theory has been mostly used to assess associative recognition, detection of auditory and visual signals, and pitch recognition. However, there is a sparse collection of literature documenting the use of signal detection to determine the diagnostic efficiency of medical assessment instruments and techniques.

McNally et al. (2009) used signal detection theory to examine methods of classifying patients as having psychogenic nonepileptic seizures versus epileptic seizures using the Wechsler Memory Scale, Third Edition, Word List Test (WMS-III WLT) as the criterion. They reported sensitivity of .59 and specificity of .62 for the logistic regression method, .52 and .74, respectively, for the routine interictal EEG, and .60 and .84, respectively, for the method using the MMPI Scales 1 and 3. Other signal detections studies have examined automated cytology (Narayanswamy & Johnson, 1998), hepatitis C, and HIV (Dragoni et al., 2005).

Diagnostic Efficiency of the CPS-M in a College Student Sample

The present study replicated an earlier phase of this project (Mainka, 2005). The previous phase assessed the diagnostic efficiency of the CPS-M in a college student sample reporting a trauma history. Participants were college students who were screened

for the presence of PTSD. Individuals who endorsed a score of 35 or higher on the Impact of Events Scale were included in the study, and participants were grouped into two severity groups based on these scores: moderate and high.

The CPS-M was compared to the CAPS to establish diagnostic efficiency. If an instrument is efficient at diagnosing PTSD, it should agree with the criterion most of the time. There were substantial interrelationships between the CAPS and the CPS-M. Pearson Product-Moment correlations revealed that all items and subscales were significantly correlated. Criterion B items on the CPS-M correlated strongly with Criterion B items on the CAPS and ranged from .42 to .82. Subscale B of the CPS-M was correlated .86 with subscale B on the CAPS. Individual items comprising Criterion C on the CPS-M were correlated strongly with CAPS Criterion C items, ranging from .42 to .89. Subscale C of the CPS-M was significantly correlated with subscale C of the CAPS (r = .89). The correlation between the Criterion D CPS-M items and the corresponding Criterion D CAPS items ranged from .55 to .88. The criterion D subscale scores from the two instruments were also significantly correlated (r = .84). Total Severity Scores on the CPS-M and CAPS were also significantly correlated at .92. High correlations indicate strong associations at the item and scale levels. These findings are not surprising given that the CPS-M was modeled after the CAPS.

The CAPS and CPS-M subscale and total scores were correlated with scores on the Purdue PTSD Scale and the Civilian Mississippi Scale (C-MISS; Kulka et al., 1990). Results showed that the CAPS and CPS-M obtained remarkably similar correlations with the convergent validity measures. CAPS subscale correlations with Purdue PTSD Scale subscales and total were .65, .74, .74 and .85 for Criteria B, C, D, and total score,

respectively. CPS-M subscale correlations with Purdue PTSD Scale subscales and total were .76, .74, .69 and .86 for Criteria B, C, D, and total score, respectively. CAPS correlations with the Civilian Mississippi Scale subscales and total were .42, .69, .68 and .72 for Criteria B, C, D and total score, respectively. CPS-M correlations with the Civilian Mississippi Scale subscales and total were .48, .63, .62 and .68 for Criterion B, C, D and total score, respectively.

Alpha for the full scale was .86, indicating good internal consistency. With the exception of items C3 (psychogenic amnesia) and D1 (sleep disturbance), removal of an item from the scale resulted in a lower alpha coefficient. Alpha coefficients for the CPS-M subscales were .86 for Criterion B, .75 for Criterion C, and .48 for Criterion D.

As an initial strategy for comparing subscale scores obtained by the CAPS and the CPS-M, effect sizes were computed. Analysis of effect sizes of mean scale differences revealed very small negative effect sizes for Criteria B and C and positive effect sizes for Criterion D and Total Severity Score on the CPS-M. None of the effect sizes exceeded a fifth of a standard deviation. Small effect sizes suggested very little mean difference in CPS-M and CAPS scores.

Using the Rogers et al. (1993) method, confidence interval testing results suggested none of the CAPS and CPS-M scales were equivalent. However, none of the scales were significantly different. This resulted in a classification of "Not Different and Not Equivalent" for all of the CPS-M scales. Given the lack of research regarding optimal equivalence range for assessment instruments, Mainka (2005) theorized that equivalence intervals of ± 20% may be more realistic in terms of the ultimate effects of scores on clinician inferences. If an equivalence range is so narrow that clinician

inferences are no different if the equivalence interval is exceeded, then a justification of the width of the interval may be needed. The small sample size in the Mainka study (*N* = 40) influenced standard error of measurement. In addition, it is not clear that these results would generalize to a clinical population. It could be that the classifications would change with true PTSD cases.

Next the diagnostic utility of the CPS-M was examined with signal detection theory. Due to the subclinical sample used in the Mainka (2005) study, only one symptom presence scoring rule resulted in enough PTSD-positive cases to perform meaningful signal detection analyses. The CAPS 1-2 symptom presence scoring rule produced eleven PTSD-positive cases and twenty-five PTSD-negative cases. Signal detection statistics revealed medium accuracy of the CPS-M in diagnosing PTSD and an AUC of .78, as demonstrated in Figure 4. A cut score of 45 produced sensitivity of .63 and specificity of .80. Although sensitivity and specificity statistics were lower than those reported by Neal et al. (1994), the two studies are not directly comparable given the subclinical sample used by Mainka and the incomplete data reporting of the Neal et al. study.

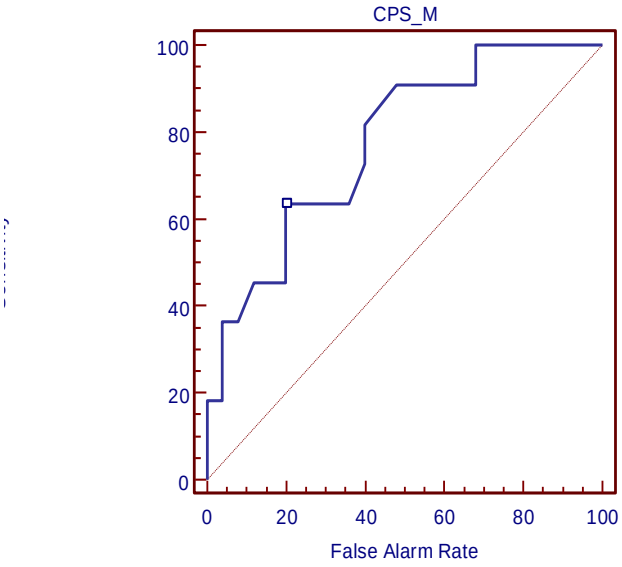

Figure 4. CPS-M ROC curve based on CAPS 1-2 symptom presence scoring rule in the Mainka (2005) study.

Although the more stringent symptom presence scoring rules were not used in the signal detection, kappa coefficients showed fair to moderate agreement across scoring rules. The Total 65 rule resulted in the highest kappa, at 1.00. However, this is not surprising given that the scoring rule resulted in only one PTSD-positive case. Based on this, Mainka (2005) determined that future research with a clinical sample should use more stringent scoring rules to produce meaningful signal detection.

Signal detection revealed that a cut score of 45 produced the best sensitivity and specificity. Cut scores tend to be very sample-specific, however, and a cut score of 45 in a nonclinical sample may not be the optimal cut score in a clinical sample. Mainka (2005) hypothesized that, in a sample with more psychopathology and PTSD-positive cases, the cut score would most likely need to be raised to produce optimum sensitivity, specificity, and overall efficiency.

In the final phase of analysis, the ROC curve of the CPS-M was compared to ROC curves produced by the Civilian version of the Mississippi PTSD scale and the Purdue PTSD Scale. It was hypothesized that the ROC curve produced by the CPS-M would be slightly better or would not be significantly different from those produced by these other instruments. However, the ROC curves for all the instruments did not significantly differ from one another. Figure 5 compares all of the ROC curves. This suggested that the CPS-M was as efficient as other PTSD screening measures in correctly identifying PTSD and non-PTSD cases. Although there was no difference in detection efficiency, these results were constrained to just one scoring rule.

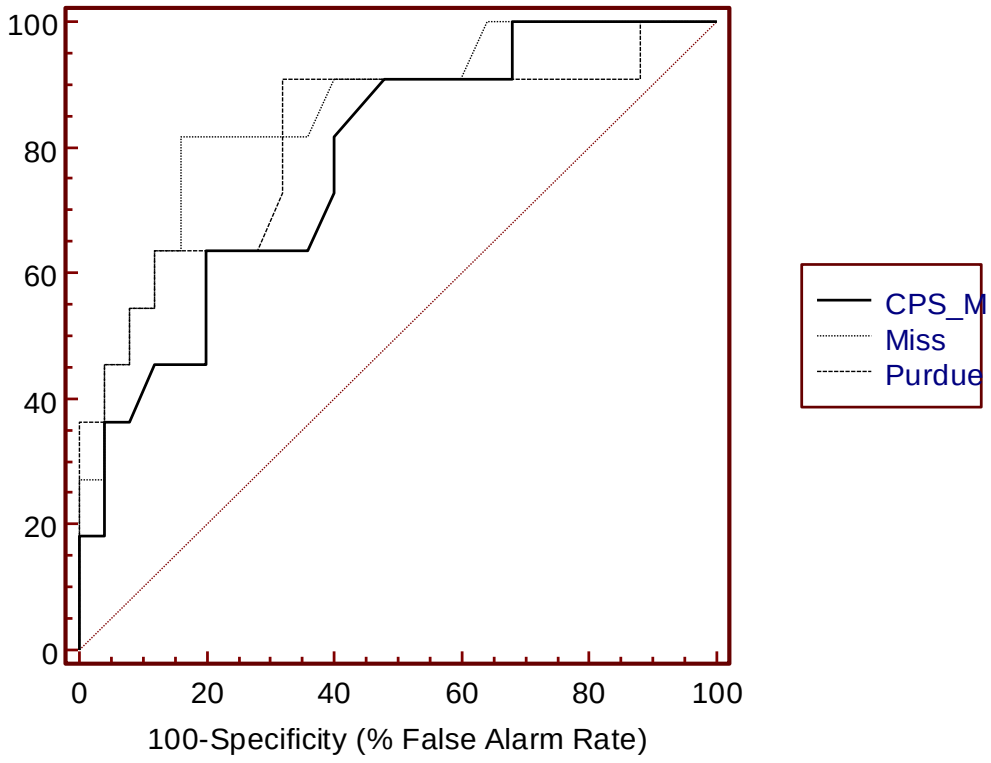

Figure 5. ROC curves for the CPS-M, C-Miss, and Purdue PTSD Scale.

Overall, the findings in the Mainka (2005) study were encouraging regarding the CPS-M. Correlations between the CAPS and CPS-M were high at the item, subscale, and

full-scale levels. Confidence interval analysis revealed that the CPS-M scales were not significantly different from their CAPS counterparts. However, the conservative ± 10% criterion for equivalence virtually guaranteed that the scales could not be judged equivalent to their CAPS counterparts. ROC curve analysis for the CPS-M revealed medium diagnostic accuracy. The CPS-M's ROC curve was not significantly different from those of the C-MISS and Purdue PTSD Scale, indicating similar diagnostic efficiency across measures. Signal detection was limited by the use of only the most liberal scoring rule to produce PTSD-positive cases. The present study used a clinical sample in which all scoring rules produced both PTSD positive and negative cases.

Overview of the Current Study

The current study aimed to accomplish several goals. One goal was to establish correlational relationships between the CAPS and the CPS – M at the item, subscale, and total score levels. Another goal was to determine statistical difference and equivalence of the two measures. The third goal was to examine the diagnostic utility of the CPS – M in a clinical population and identify the optimum cut point for the CPS – M relative to the CAPS. The overall efficiency and cut point were determined by the area under the ROC curve (AUC) for the CPS – M. The final aim was to compare the diagnostic utility of the CPS – M to several convergent measures when using the CAPS as the criterion. The convergent validity measures included the PTSD Checklist, the Purdue PTSD Scale-Revised, the Beck Depression Inventory, and the Trauma Related Dissociation Scale. Because depression and dissociation are very closely related to PTSD, these were included to compare the CPS-M to measures assessing similar, but not identical, constructs. The PCL and Purdue Scale were included to evaluate ROC curves of several PTSD instruments, when compared to the CAPS. Because the CPS-M is modeled after the CAPS, it was expected that it would be more highly correlated with the CAPS than with the PCL or Purdue PTSD Scale-Revised, and it would be more strongly associated with the measures of PTSD than with those assessing depression or dissociation.

Method

Recruitment of Participants

Participants were drawn from three clinics at a Veterans Affairs Hospital: the mental health clinic, the PTSD clinic, and primary care. These three clinics were selected for the purpose of obtaining a heterogeneous sample, and participants from the clinics were not compared. Participants received a $10 gift card to a local retail store for completion of the study. Potential participants completed a screening packet that assessed for the presence of a traumatic event and the severity of post-traumatic symptomatology.

In order for signal detection to provide a clear picture of diagnostic efficiency, it is necessary to utilize participants with and without PTSD. To obtain an adequate sample size, a sample of 90 participants was needed for the study. Of these participants, 45 were individuals who endorsed a trauma but little or no PTSD symptomatology, as demonstrated by a score of 34 or lower on the PTSD Checklist. Forty-five were individuals who endorsed a trauma and reported moderate to high symptom severity, as evidenced by a score of 44 or higher on the PTSD Checklist. A cut score of 44 yielded overall efficiency of .90, sensitivity of .94, and specificity of .86 in a sample of motor vehicle accident and sexual assault victims (Blanchard et al., 1996).

Study Procedure

Potential participants were administered a screening packet (described below) prior to the test session. Eligible participants then completed one subsequent test session. During the test session, the CAPS interview and other scales were administered in a quiet room at the administration site. A trained interviewer performed the CAPS interview.

The CPS – M was administered via a laptop computer with headphones and a mouse. Average CPS – M administration time to completion was 30 minutes. The administration of the CAPS took approximately an hour and a half. Completion of the remaining self-report questionnaires took approximately 20 minutes. Participants spent approximately 60-150 minutes completing the study. After their session, participants were debriefed with a written explanation of the study and provided a list of counseling resources in case of emotional difficulties due to participation in the study (described below).

Training Protocol

In addition to the principal investigator, three research assistants administered study sessions. Research assistants attended several training sessions on administering CAPS interviews, where CAPS training modules from the National Center for PTSD were followed. They then sat in on several interviews performed by the principal investigator and took the lead while being observed. Additionally, a comprehensive training manual was created for study procedures, including instructions for recruitment, obtaining informed consent, and administration of the study session.

Screening Instruments

The screening packet consisted of a cover page (see Appendix A) that asked for contact information, a Risk Assessment Form (Appendix B) that assessed suicidality and homicidality, the Life Events Checklist (see Appendix C), and the PTSD Checklist (see Appendix D). A study Exclusion Form was also completed for each potential participant (see Appendix E).

Life Events Checklist (LEC). The LEC was included in the screening battery to assess participants' history of exposure to traumatic events. The LEC is a 17-item

questionnaire developed by the National Center for PTSD. Gray, Litz, Hsu, and Lombardo (2004) compared the LEC to the Traumatic Life Events Questionnaire (TLEQ) in a sample of college undergraduates. They reported a mean kappa of .61 across the seven items assessing direct trauma exposure, as well as adequate agreement with the TLEQ, $r = -.55$, $p < .001$ (lower scores on the TLEQ indicate more exposure). The mean kappa was .47 for all items, and test-retest reliability was adequate, $r = .82$, $p < .001$.

PTSD Checklist (PCL). The PCL is a 17-item self-report checklist that measures frequency and severity of symptoms. Respondents score symptom intensity over the two previous weeks. Blanchard et al. (1996) reported overall diagnostic efficiency relative to the CAPS of .90, sensitivity of .94, and specificity of .94 using a cut score of 44 in a sample of motor vehicle and sexual assault victims. The kappa for agreement between judges was .84, $p < 0.001$. Blanchard et al. (1996) recommended using the PCL as a screening instrument for posttraumatic stress disorder based on its excellent agreement with the CAPS. Using a cut-point score of 38, Dobie et al. (2002) reported sensitivity of .79 and specificity of .79 in a sample of female veterans.

Test Session

After screening, individuals who were eligible for participation had the option of either completing the test session immediately or were scheduled to return within two weeks. Participants were excluded if they had not experienced a trauma, were under the age of 18, had current or life-time history of a formal thought disorder, or voluntarily declined. Seventeen veterans were excluded based on these criteria.

Before beginning the test session, participants completed an Informed Consent Form (See Appendix F). The test session consisted of two assessment components: 1)

CPS-M and paper-and-pencil questionnaires and the 2) CAPS. Components 1 and 2 were counterbalanced in their administration, with the CPS-M and paper-and-pencil measures also being counterbalanced within component 1. Half of the participants completed the CPS-M and paper-and-pencil questionnaires first and CAPS second, while the other half completed the CAPS first and CPS-M and paper-and-pencil measures second.

The paper-and-pencil questionnaires were used as convergent measures, against which to evaluate the performance of the CPS-M in diagnosing PTSD. The previous study, which assessed diagnostic efficiency in a college student sample, required participants to complete a paper-and-pencil packet with two PTSD questionnaires. In response to complaints that participants were answering the same questions repeatedly, the present study only included one PTSD questionnaire as part of the convergent battery (in addition to the PCL in the screening packet, which was also used as a convergent measure). Due to the fact that the present study included military participants, the Civilian version of the Mississippi Scale for PTSD, which was used in the previous study, was not used. The Purdue PTSD Scale-Revised provided a convergent measure for PTSD. The other two questionnaires assessed similar constructs: depression and dissociation.

Test Session Instruments

Paper-and-pencil measures

Demographic Questionnaire (See Appendix G). This form queried participants about age, sex, employment status, ethnicity, psychiatric history, and other demographic variables. It was used for determining and reporting descriptives for the present study.

Purdue Posttraumatic Stress Disorder Scale – Revised (See Appendix H). The Purdue Scale (Lauterbach & Vrana, 1996) was included in the study to compare its diagnostic efficiency statistics to those obtained by the CPS - M. The Purdue is a 17-item scale measuring frequency of PTSD symptoms to produce four scales: Reexperiencing, Avoidance, Arousal, and Total. Lauterbach and Vrana (1996) found the Purdue Scale to be highly correlated with other measures of PTSD in a sample of 562 undergraduates. Lauterbach and Vrana found adequate test-retest reliability over two weeks (r = .72) and excellent internal consistency (α = .91). They concluded that, due to its ability to discriminate between people who were and were not traumatized, the PPTSD – R is a promising measure of PTSD symptoms in a college population.

Trauma Related Dissociation Scale (TRDS) (See Appendix I). The TRDS (Carlson & Waelde, 1999) is a 24-item questionnaire that assesses five dimensions of trauma-related dissociation (depersonalization, derealization, gaps in awareness, amnesia, and gaps in awareness plus reexperiencing). In a sample of 30 outpatients and 62 veterans, Carlson and Waelde (1999) reported internal consistency alphas ranging from .60 for derealization to .87 for amnesia, with an alpha of .93 for the total score. The TRDS is significantly correlated with the Dissociative Experiences Scale (DES). The TRDS-DES correlations ranged from .35 to .55 for subscales, and the total scale correlation between the DES and TRDS was .56. They also reported a correlation of .51 between the TRDS and the CAPS (Carlson & Waelde, 1999).

Beck Depression Inventory (BDI-II) (See Appendix J). The BDI-II (Beck, Steer, & Brown, 1996) is a 21-item questionnaire used to assess the severity of depression in adolescents and adults. It asks questions about the physical, cognitive, and behavioral

symptoms of depression. This self-report measure is effective in classifying participants as depressed or nondepressed, and it was included in the study to compare correlations between the measures assessing PTSD and the measures assessing the similar construct of depression. The BDI-II has been found to have good internal consistency with values for coefficient alpha ranging from .90 to .91 (Beck, Steer, Ball, & Ranieri, 1996; Dozois, Dobson, & Ahnberg, 1998; Osman et al., 1997).

Computerized PTSD Scale – Multimedia Version (CPS – M)

The CPS – M is a computerized version of the CAPS, in which participants respond to computerized prompts and report frequency and intensity of PTSD symptoms. Mayo et al. (2000) found that the CPS – M correlated .87 with the Civilian Mississippi Scale, .79 with the Beck Depression Inventory, .79 with the Beck Anxiety Inventory, and .13 with the Antisocial Behavior Inventory. Test-retest reliability was .92 for the full scale, and the alpha coefficient was .91.

Clinician-Administered PTSD Scale (See Appendix K)

The CAPS is a structured interview that provides an assessment of the frequency and intensity of PTSD symptoms. This was included in the study as the criterion against which to assess diagnostic agreement of the CPS-M.

Results

Descriptive Statistics

Participants. One hundred and eight veterans were screened with the LEC and PCL. Seventeen veterans were screened out of the study due to a variety of reasons. Eight were screened out due to the presence of a thought disorder, three reported thoughts of harming self or others and were directly referred for follow-up care, one did not meet Criterion A, one completed an earlier version of the study, one found the study too distressing, and three obtained exclusionary PCL scores between 35 and 43 (these scores fall outside of either the low/moderate or high symptom groups). One participant completed part of the study but removed himself due to emotional distress.

The protocol for participants who reported any kind of emotional distress or desired follow-up care was to escort them to the mental health clinic at the VA hospital, where they would be referred for a triage visit or a meeting with their mental health or primary care provider. No veterans requested follow-up care due to emotional distress from the study. The one participant who discontinued the study declined to seek follow-up care and reported feeling fine by the time he left the room where he was participating in the study.

Demographic variables. Ninety participants completed all phases of the study (77 men, 13 women). The mean age of participants was 58.8 years (*SD* = 10.6) with the majority of participants over age 50 (81.1%). Seventy-six participants (83.5%) were Caucasian, 9 (9.9%) were African-American, 4 (4.4%) were American Indian/Alaskan, and 1 (1.1%) was Latino.

Forty-five participants had scores greater than 44 on the PCL and were classified as having severe symptoms. Forty-five were classified as having low/moderate stress symptoms (i.e., scores ranging from 17 to 34 on the PCL). These cutoff scores on the PCL were used in order to increase the chances of obtaining both a signal group (PTSD-positive) and a non-signal group (PTSD-negative) that were of equal size. Participants who endorsed low/moderate stress symptoms were included because they had experienced a trauma but would most likely not qualify for a PTSD diagnosis. Participants from the high severity group were believed to be more likely to receive a PTSD diagnosis.

Of the participants in the low/moderate group, 41 were men (91.1%) and four (8.9%) were women. There were 36 (80.0%) men and nine (20.0%) women in the high severity group. Thus 30.8% of women who participated were in the low/moderate group and 69.2% were in the high group. Men were far more balanced between severity groups, with 53.2% in the low/moderate group and 46.7% in the high severity group.

Additional demographic information was collected on all participants who complted the study. Some information was reported by participants on a demographic form and some was obtained through the Computerized Patient Record System (CPRS), the VA database containing patient information. Participants endorsed a range of religious preferences. The most commonly cited religion or denomination was Roman Catholic (n = 22; 24.4%), followed by Protestant (n = 19; 21.1%). Seventeen did not endorse a religious preference (18.8%). The next most common preferences were Lutheran (n = 11; 12.2%) and Baptist (n = 10; 11.1%).

History of treatment. Thirteen participants (14.4%) reported that they had received inpatient psychiatric care in their past, and 59 (65.5%) had received outpatient care. Of participants endorsing that they had received inpatient care, three were women and 10 were men. Twelve women (92.3%) endorsed receiving outpatient care, compared to 47 (52.2%) men who reported receiving outpatient care. Far fewer participants in the low/moderate severity group reported receiving inpatient care (n = 3) than in the high severity group (n = 10). Similar findings were obtained when comparing participants who had received outpatient care (low/moderate: n = 15; high: n = 44).

Participants reported how many psychiatric or mental health therapy sessions they had attended in the past year. The modal number of sessions attended was zero, which was endorsed by 41.8% of participants. Half of all participants had attended two or more sessions in the past year, and over 25% had attended 11 or more sessions. Number of sessions attended ranged from zero to 273, and there were four participants who had attended 65 or more sessions in the past year.

Participants also reported which psychotropic medications they had prescriptions for at the time of the study. Thirty-seven participants did not report any prescriptions, 54 endorsed only one medication, 42 endorsed two or more medications, 21 had prescriptions for three or more, nine endorsed four or more medications, three endorsed five or more, and one participant had prescriptions for six psychotropic medications. The most commonly reported medication was citalopram, which was endorsed by 19 participants (21.1%). The next most frequently reported medications were bupropion (n = 16; 17.7%) and trazodone (n = 10; 11.1%).

Service connection status. In the VA population, service connection is commonly examined, especially as it relates to the diagnosis of PTSD. Aribisi, Murdoch, Mcnulty, and Fortier (2004) note that veterans achieve service-connected status by having a documented, compensable condition that was directly caused by or aggravated by military service. Service connection can result in priority in receiving medical care through the VA healthcare system, disability payments, and sometimes college payments for dependents (Aribisi et al., 2004). Table 2 shows the distribution of service connection percentages. In the present study, just over half of participants (n = 48) received at least 10% total service connection, and 46.2% indicated that they did not receive service connected status. Only 13 participants (15.4%) received service connection for PTSD. Slightly more participants received mental health service connection (n = 19).

Table 2

Service Connection Percentages

	Service Connection Percentage										
Service Connection Type	0	10	20	30	40	50	60	70	80	90	100
Total[a]	42	9	4	3	5	3	4	7	5	2	6
PTSD	77	0	0	5	0	3	0	3	0	0	2
Mental Health	71	0	0	7	0	5	0	4	0	0	3

[a]Cell values reflect numbers of participants

Differences based on sex. Differences between men and women were examined. Men and women did not differ on frequency of therapeutic contact (i.e., number of therapy sessions/number of medication consultation sessions attended in the past year)

(women: *M* = 18.92, *SD* = 10.01; men: *M* = 12.13, *SD* = 35.52). Men and women did not differ on total service connection percentage, PTSD service connection, or mental health service connection.

Significant differences were found between men and women on all of the subscale means and total severity scores on the CAPS and the CPS –M, with women obtaining significantly higher scores. These subscale and total severity score means are reported in Table 3.

Table 3

Comparisons of Men and Women on Subscale and Total Severity Scores

	Men		Women		
Scale	*M*	*SD*	*M*	*SD*	*t*
CPS-M					
B	9.12	10.41	21.31	14.81	2.85*
C	15.27	15.46	28.85	18.87	2.46*
D	15.32	11.77	27.69	10.27	3.93***
TSS	39.71	35.82	77.85	42.58	3.05**
CAPS					
B	7.85	8.86	16.69	11.55	2.63*
C	15.04	13.85	28.08	13.70	3.16**
D	14.89	10.27	22.46	9.20	2.69*
TSS	37.78	30.31	67.23	30.20	3.24**

Note. *** = $p < .001$, ** = $p < .01$, * = $p < .05$

Manipulation Check

To determine if the prescreen had the desired effect of creating groups that differed on CAPS and CPS-M total scores, an initial set of between-groups contrasts was conducted. Levene's test for homogeneity of variance revealed significant between group differences. Consequently, independent samples *t*-tests using non-pooled variances were conducted. The degrees of freedom used reflect this adjustment. There were significant between-group differences for Total Severity on the CAPS, $t(74) = -15.6$, $p < .001$, and on the CPS-M, $t(65) = -14.8$, $p < .001$. Participants in the severe symptomatology group reported higher total severity scores than the low/moderate severity group on the CPS-M (high: $M = 78.0$, $SD = 26.5$; low/moderate: $M = 12.4$, $SD = 13.2$) and the CAPS (high: $M = 69.2$, $SD = 19.3$; low; $M = 14.6$, $SD = 12.7$).

To determine if severity of symptoms was related to variables such as psychiatric or mental health sessions attended and service connection, between-groups contrasts were conducted. Levene's test for homogeneity of variance revealed significant between group differences. Consequently, independent samples *t*-tests using non-pooled variances were conducted. The degrees of freedom used reflect this adjustment. Between-groups contrasts revealed no significant differences for number of therapy or medications sessions attended in the past year. Participants in the low/moderate severity group reported fewer sessions in the past year ($M = 5.78$, $SD = 20.37$) than those in the severe symptomatology group ($M = 20.44$, $SD = 44.96$), but this comparison did not reach conventional levels of significance, $t(61) = -1.99$, $p > .05$.

Significant between-groups differences were found for PTSD service connection percentage, $t(44) = -3.73$, $p < .001$, as well as for mental health service connection, $t(75)$

= -2.78, $p < .01$. Participants in the low/moderate group received significantly lower percentages of PTSD service connection ($M = 0.00$, $SD = 0.00$) than those in the severe symptomatology group ($M = 15.78$, $SD = 28.40$). Similarly, significantly lower percentages were found for mental health service connection when comparing the low/moderate group ($M = 4.44$, $SD = 18.41$) to the severe symptomatology group ($M = 18.67$, $SD = 28.97$). Total service connection percentage did not differ significantly between groups (low/moderate: $M = 22.22$, $SD = 32.47$; high: $M = 33.78$, $SD = 36.20$).

Item-level descriptive statistics. Table 4 reports means and standard deviations for each CAPS and CPS-M item. To assess level of skew and kurtosis, the parameter estimates were divided by the corresponding standard errors. These values appear in Table 4. Computed values exceeding 1.96 are significant. On the CAPS, all items on subscale B exceeded the threshold for skewness. Subscale C items C2, C3, C4, and C7 exceeded 1.96. No subscale D items on the CAPS exceeded the threshold for skewness. All items were positively skewed, indicating a higher distribution of scores at the lower end of the distribution. Items B1 and B2 on the CAPS exceeded 1.96 for kurtosis. Subscale C items C1, C3, C4, C5, and C6 exceeded the threshold, and all Subscale D items values were beyond criterion for kurtosis. All items exceeding threshold for kurtosis on the CAPS were platykurtotic, except for item C3, which was leptokurtotic. Measures of skewness and kurtosis indicate that several items on the CAPS are not normally distributed.

All CPS-M subscale B items were skewed beyond criterion. Items C1, C2, C3, and C7 exceeded the threshold, and no items from subscale D were skewed beyond criterion. Like CAPS items, all CPS-M items were positively skewed, indicating a

clustering of scores at the lower end of the distribution. No CPS-M subscale B items were kurtotic beyond the threshold. Items C3, C4, C5, and C6 all displayed values above 1.96, and all subscale D items were beyond the threshold for kurtosis. Similar to the findings on the CAPS, all CPS-M items exceeding criterion for kurtosis were platykurtotic except for item C3, which was leptokurtotic. Because several CPS-M items were skewed or kurtotic beyond criterion, normality cannot be assumed for these scores.

Though many items on the CAPS and CPS-M were skewed beyond criterion, the data were not transformed. By not assuming normality, we employed a conservative standard, and correlations between items remained high despite skewness. Transformation would likely yield even higher levels of significance, and because significance was already achieved, we presented raw data to facilitate ease of interpretation.

Table 4

Instrument Comparisons on Item Scores

	CAPS				CPS-M				
Item	*M*	*SD*	*skewness*	*kurtosis*	*M*	*SD*	*skewness*	*kurtosis*	r_{xy}
B1	2.28	2.52	2.33	-1.98	2.67	2.60	2.07	-1.69	.81***
B2	2.22	2.53	2.37	-2.04	2.23	2.40	3.39	-0.56	.87***
B3	1.10	2.07	5.99	1.54	1.67	2.67	5.07	0.25	.73***
B4	1.99	2.34	2.52	-1.94	2.17	2.45	3.29	-0.79	.79***
B5	1.59	2.22	4.09	-0.31	2.14	2.59	3.31	-1.18	.77***
C1	2.54	2.68	1.65	-2.68	2.44	2.72	2.68	-1.88	.64***
C2	1.80	2.66	4.29	-0.58	2.33	2.86	3.09	-1.76	.69***
C3	1.11	2.10	6.56	3.05	1.31	2.19	6.29	2.91	.59***
C4	2.84	3.01	1.97	-2.55	2.69	2.91	1.65	-2.86	.69***
C5	3.43	3.08	0.33	-3.22	2.9	3.12	1.86	-2.69	.83***
C6	3.22	3.04	0.71	-3.15	3.33	3.13	1.02	-3.06	.79***
C7	2.05	2.77	3.66	-1.26	2.17	2.75	3.09	-1.75	.75***
D1	4.24	3.08	-1.28	-2.93	4.36	3.22	-1.15	-3.02	.93***
D2	2.60	2.52	1.00	-2.83	3.01	2.71	0.58	-2.82	.70***
D3	3.23	2.82	0.18	-3.17	3.22	2.99	0.49	-3.23	.84***
D4	3.75	2.92	-0.99	-3.04	3.68	3.09	-.13	-3.15	.78***
D5	2.21	2.33	1.94	-2.21	2.84	2.59	1.53	-2.12	.72***

Note. *** = $p < .001$

Scale-level descriptive statistics. Table 5 reports means and standard deviations for each CAPS and CPS-M subscale. On the CAPS, subscale B exceeded the customary 1.96 threshold for skewness, while subscales C and D and Total Severity Scores were within range. Subscales C and D were kurtotic beyond criterion. Total Severity Scores on the CAPS were also beyond threshold for kurtosis.

CPS-M subscales B and C were skewed beyond criterion, as were Total Severity Scores. Subscales C and D on the CPS-M were kurtotic beyond the 1.96 threshold, while subscale B and Total Severity Scores were within range for kurtosis. Because several subscales fell outside of range for skewness or kurtosis, normality was not assumed for subscale means on the CAPS or CPS-M. As stated previously regarding item-level skewness, data were not transformed. Because all subscales and Total Severity Scores were significantly correlated, we presented raw data to facilitate ease of interpretation.

Table 5

CAPS-CPS-M Scale Score Comparisons

	CAPS				CPS-M				
Scale	*M*	*SD*	*skewness*	*kurtosis*	*M*	*SD*	*skewness*	*kurtosis*	r_{xy}
B	9.17	9.76	3.44	0.06	10.88	11.86	3.54	-0.57	.93***
C	16.99	14.52	1.38	-2.52	17.23	16.58	2.17	-1.98	.88***
D	16.02	10.43	0.23	-2.28	17.11	12.31	0.14	-2.34	.92***
TSS	42.18	31.92	1.00	-2.40	45.22	39.01	1.98	-1.84	.95***

Note. *** = $p < .001$

Table 6

CAPS and CPS-M Subscale and Total Severity Score Correlations

Scale		CAPS				CPS-M			
		B	C	D	TSS	B	C	D	TSS
CAPS	B	1.00							
	C	.78	1.00						
	D	.72	.78	1.00					
	TSS	$.79_a$	$.84_a$	$.79_a$	1.00				
CPS-M	B	.93	.80	.75	.89	1.00			
	C	.86	.88	.77	.92	.90	1.00		
	D	.78	.84	.92	.92	.84	.87	1.00	
	TSS	.89	.88	.85	.95	$.91_a$	$.92_a$	$.88_a$	1.00

Note.[a]Values reflect corrected item-total correlations. All values significant at $p < .001$ level

Reliability: Alpha Coefficients

Clinician-Administered PTSD Scale

Subscale and total alphas. Alpha coefficients for the CAPS subscales were all above .80. Alphas were .89 for Criterion B, .87 for Criterion C, and .82 for Criterion D. Removal of any item from subscales B and D resulted in a lower alpha coefficient for that subscale. A similar pattern of findings emerged for criterion C. However, removal of C3 (inability to recall) resulted in a slight increase in alpha from .87 to .88. Alpha for the full scale was .94[1]. Removal of any item from the scale resulted in a lower alpha coefficient.

[1] Alpha was computed at the subscale and item level. When calculated at the subscale level, Chronbach's alpha was .89, slightly lower than when computed at the item level. Unless otherwise indicated, throughout the paper, when Chronbach's alpha is reported for the entire scale, this refers to the item-level calculation.

CPS-M

Subscale and total alphas. Alpha coefficients for the CPS-M subscales were all above .90. Alphas were .96 for Criterion B, .93 for Criterion C, and .90 for Criterion D. Removal of any item from subscales B, C, and D resulted in lower alpha coefficients. Alpha for the full scale was .97. Removal of any item from the scale resulted in a lower alpha coefficient.

Validity Coefficients

To examine the relationship between the CAPS and CPS-M at the item, subscale, and total score levels, a series of Pearson Product-Moment correlations were conducted.

Item-level correlations. Table 4 reports correlations between CAPS items and the corresponding CPS-M items. All items were significantly correlated and ranged from .64 (Item C1: avoidance of thoughts) to .93 (Item D1: difficulty sleeping).

Scale correlations. Correlations between the CAPS subscales and the corresponding CPS-M subscales are included in Table 5, and the intercorrelations among all CAPS and CPS-M subscales are reported in Table 6. Subscale scores were significantly correlated and ranged from .88 (Criterion C) to .93 (Criterion B). Total Severity Scores on the CAPS and CPS-M were significantly correlated at .95. As expected, the subscales of the CPS-M were more strongly correlated with the corresponding CAPS subscale than with other CAPS subscales. These findings are consistent with the hypothesis that CAPS and CPS-M subscales should be more similar to each other than to other subscales.

Convergent validity correlations. Table 7 details the correlations among the CAPS, the CPS-M, and all convergent validity measures. Specific findings will be highlighted in the sections that follow.

Purdue PTSD Scale-Revised. The Purdue PTSD Scale-Revised total score mean was 44.98 (*SD* = 22.27). All CAPS subscale scores and the total score were significantly correlated with the Purdue PTSD Total Severity Score: *r* = .82, .87, .83, and .92 for Criterion B (reexperiencing), C (avoidance/numbing), D (hyperarousal), and total score, respectively.

All CPS-M subscale scores and the total score were correlated significantly with the Purdue PTSD Scale Total Score: *r* =.87, .91, .90, and .93 for Criterion B (reexperiencing), C (avoidance/numbing), D (hyperarousal), and total score, respectively.

PTSD Checklist (PCL). The PCL total score mean was 43.54 (*SD* = 22.68). All CAPS subscale scores and total score were significantly correlated with the PCL Total Severity Score: *r* = .84, .85, .78, and .90 for Criterion B (reexperiencing), C (avoidance/numbing), D (hyperarousal), and total score, respectively.

All CPS-M subscale scores and the total score correlated significantly with the PCL Total Severity Score; *r* = .87, .87, .84, and .90 for Criterion B (reexperiencing), C (avoidance/numbing), D (hyperarousal), and total score, respectively.

Trauma Related Dissociation Scale (TRDS). The TRDS total score mean was 16.99 (*SD* = 20.24). All CAPS subscale scores and the total score were significantly correlated with the TRDS Total Severity Score: *r* = .66, .68, .64, and .72 for Criterion B (reexperiencing), C (avoidance/numbing), D (hyperarousal), and total score, respectively.

All CPS-M subscale scores and the total score correlated significantly with the TRDS Total Severity Score: r = .77, .78, .75, and .80 for Criterion B (reexperiencing), C (avoidance/numbing), D (hyperarousal), and total score, respectively.

Beck Depression Inventory (BDI). The BDI total score mean was 18.51 (*SD* = 14.30). All CAPS subscale scores and the total score were significantly correlated with the BDI Total Severity Score: r = .72, .82, .77, and .84 for Criterion B (reexperiencing), C (avoidance/numbing), D (hyperarousal), and total score, respectively.

All CPS-M subscale scores and the total score correlated significantly with the BDI Total Score: r = .76, .82, .83, and .72 for Criterion B (reexperiencing), C (avoidance/numbing), D (hyperarousal), and total score, respectively.

Convergent validity findings. Overall, the CAPS was most strongly related to the other measures of PTSD, less strongly related to the measure of depression, and the least related to the measure of dissociation. The average correlation between the CAPS subscales and total score and Purdue PTSD Scale-Revised was .86, followed by an average correlation with the PTSD Checklist of .84. The average of the CAPS subscales and total score correlations was lower with the Beck Depression Inventory at .79 and the lowest with the Trauma Related Dissociation Scale at .67.

The CPS-M showed the same pattern of associations, displaying the highest correlations with measures of PTSD and lower with those of depression and dissociation. The average correlation between the CPS-M subscales and total score and Purdue PTSD Scale-Revised was .90, followed by an average correlation with the PTSD Checklist of .87. The average of the CPS-M subscales and total score correlations was lower with

the Beck Depression Inventory and the Trauma Related Dissociation Scale at .78 with each.

Table 7

CAPS and CPS-M Correlations with Other Measures

Measure	Purdue PTSD Scale	Trauma Related Dissociation Scale	Beck Depression Inventory-II	PTSD Checklist
CAPS Crit B	.82***	.66***	.72***	.84***
CAPS Crit C	.87***	.68***	.82***	.85***
CAPS Crit D	.83***	.64***	.77***	.78***
CAPS Total Severity Score	.92***	.72***	.84***	.90***
CPS-M Crit B	.87***	.77***	.76***	.87***
CPS-M Crit C	.91***	.78***	.82***	.87***
CPS-M Crit D	.90***	.75***	.79***	.84***
CPS-M Total Severity Score	.93***	.80***	.83***	.90***

Note. *** $p < .001$

Equivalence Analyses

Scale-level effect sizes. Effect size estimates were made by calculating Cohen's *d* for each subscale and the Total Severity Score ([CPS-M tss – CAPS tss]/pooled SD). Positive effect sizes imply greater CPS-M mean scale scores. According to Cohen (1992), when comparing two independent means, effect sizes of .20 or below are small, .21-.50 are medium, and .51-.80 are large. Cohen's *d* was .16 for Criterion B, .02 for Criterion C, .10 for Criterion D, and .09 for the Total Severity Score. These small effect sizes

indicate that CPS-M mean scale scores were only slightly higher than CAPS mean scale scores, indicating almost no effect based on the instrument used, or no significant difference between CAPS and CPS-M mean scores.

Confidence interval analysis. Table 8 provides the data used for the confidence interval analysis with the CPS-M and CAPS. For the Total Severity Score, the 90% confidence interval for the CPS-M was -11.87 to 5.79. This corresponds with a lower confidence limit of -28.14% and an upper confidence limit of 13.73%. These limits are expressed as percentages of the CAPS mean, which standardizes the confidence intervals. Because the CPS-M confidence interval is not contained within the ±10% equivalence interval, the Total Severity Scores on the two instruments are not equivalent. However, because the 95% confidence interval includes zero (-13.56-7.48), we can conclude that Total Severity Scores are not statistically different. The Total Severity Score is classified as "Not Different and Not Equivalent."

For Criterion B, the 90% confidence interval for the CPS-M was -4.40 to 0.98 (-47.98% - 10.69%). Because the CPS-M confidence interval is not contained within the CAPS equivalence interval of ±10%, Criterion B on the two instruments is not equivalent. Because the 95% confidence interval includes zero (-4.92-1.49), we can conclude that Criterion B scores are not statistically different. Criterion B is classified as "Not Different and Not Equivalent."

For Criterion C, the 90% confidence interval for the CPS-M was -4.09 to 3.62 (-24.13% - 21.30%). Because the CPS-M confidence interval is not contained within the 10% equivalence interval, Criterion C on the two instruments is not equivalent. Because the 95% confidence interval includes zero (-4.84-4.36), we can conclude that Criterion C

scores are not statistically different. Criterion C is classified as "Not Different and Not Equivalent." For Criterion D, the 90% confidence interval for the CPS-M was -3.92 to 1.74 (-24.44% - 10.84%). Because the CPS-M confidence interval is not contained within the CAPS equivalence interval of ±10%, Criterion D on the two instruments is not equivalent. Because the 95% confidence interval includes zero (-4.46-2.28), we can conclude that Criterion D scores are not statistically different. Criterion D is classified as "Not Different and Not Equivalent."

Because there has been little research on using confidence interval analysis with psychological assessment instruments, we theorized that the ± 10% equivalence interval recommended by Rogers et al. (1993) may be too stringent. When we expanded this to a ± 20% equivalence interval, we found similar results. All subscales were classified as "Not Different and Not Equivalent." Therefore, we are basing our conclusions on the classifications found using the ± 10% equivalence interval. Figure 6 visually displays results from Table 8 for easier interpretation.

Table 8

Confidence Interval Analysis Statistics

		Subscales		
	Total Severity Score	Criterion B	Criterion C	Criterion D
CAPS				
CAPS Mean	42.18	9.17	16.99	16.02
10 % Equivalency Interval	±4.22	±0.92	±1.70	±1.60
20% Equivalency Interval	±8.44	±1.84	±3.40	±3.20
CPS-M				
90% CI	-11.87-5.79	-4.40-0.98	-4.09-3.62	-3.92-1.74
95% CI	-13.56-7.48	-4.92-1.49	-4.84-4.36	-4.46-2.28
Upper Confidence Limit (UCL) as % of CAPS Mean	13.73%	10.69%	21.30%	10.84%
Lower Confidence Limit (LCL) as % of CAPS Mean	-28.14%	-47.98%	-24.13%	-24.44%
CPS-M 90% CI Within CAPS 10% Equiv. Range?	No	No	No	No
CPS-M 90% CI Within CAPS 20% Equiv. Range?	No	No	No	No
CAPS Mean Within CPS-M 95% CI?	Yes	Yes	Yes	Yes
Classification per Rogers, Howard, & Vessey (1993)	Not Different, Not Equivalent	Not Different, Not Equivalent	Not Different, Not Equivalent	Not Different, Not Equivalent

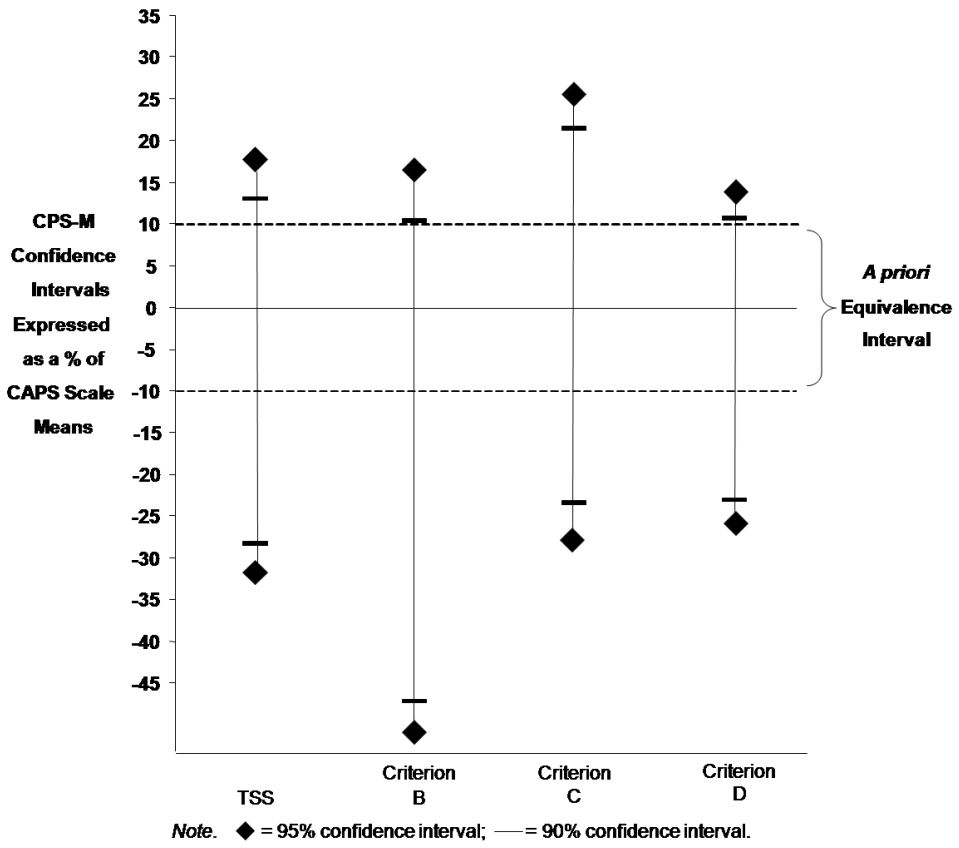

Figure 6. Confidence interval analysis.

A final complication relates to the formula used to compute the confidence interval. In the standard equation, which is used in the pharmaceutical industry, the assumption is that a between-groups standard error should be computed. However, in the current investigation, participants were administered both the CPS-M and the CAPS. Consequently, use of a within-groups standard error in the computation of the confidence may result in these measures being judged equivalent. To test this notion, confidence intervals for Total Severity Scores using the within-groups standard error were calculated. Table 9 shows these results. This resulted in Total Severity Scores being classified as "Statistically Different and Not Equivalent." When the range was expanded to the ± 20%

equivalence interval, the classification changed to "Statistically Different and Equivalent."

Table 9

Confidence Interval Analysis Statistics Using Adjusted Standard Error

	Total Severity Score
CAPS	
CAPS Mean	42.18
10 % Equivalency Interval	±4.22
20% Equivalency Interval	±8.44
CPS-M	
90% CI	-5.31- -0.78
95% CI	-5.74- -0.34
Upper Confidence Limit (UCL) as % of CAPS Mean	-1.84%
Lower Confidence Limit (LCL) as % of CAPS Mean	-12.58%
CPS-M 90% CI Within CAPS 10% Equiv. Range?	No
CPS-M 90% CI Within CAPS 20% Equiv. Range?	Yes
CAPS Mean Within CPS-M 95% CI?	No
Classification	Statistically Different, Not Equivalent

The mean difference between CAPS and CPS-M Total Severity Scores was -3.18 (*SD* = 12.85). Almost half of the CPS-M Total Severity Scores fell within ± 5 points of CAPS scores (49.2%), and 70.9% fell within ± 10 points of CAPS Total Severity Scores. The distribution of difference scores was not skewed or kurtotic beyond the 1.96 criterion (skewness: -1.065, *SE* = .258; kurtosis: 1.605, *SE* = .511), suggesting a normal

distribution. Modal score differences of 0 and -2 were found between the two instruments. See Figure 7 for distribution of differences between the paired CAPS and CPS-M Total Severity Scores.

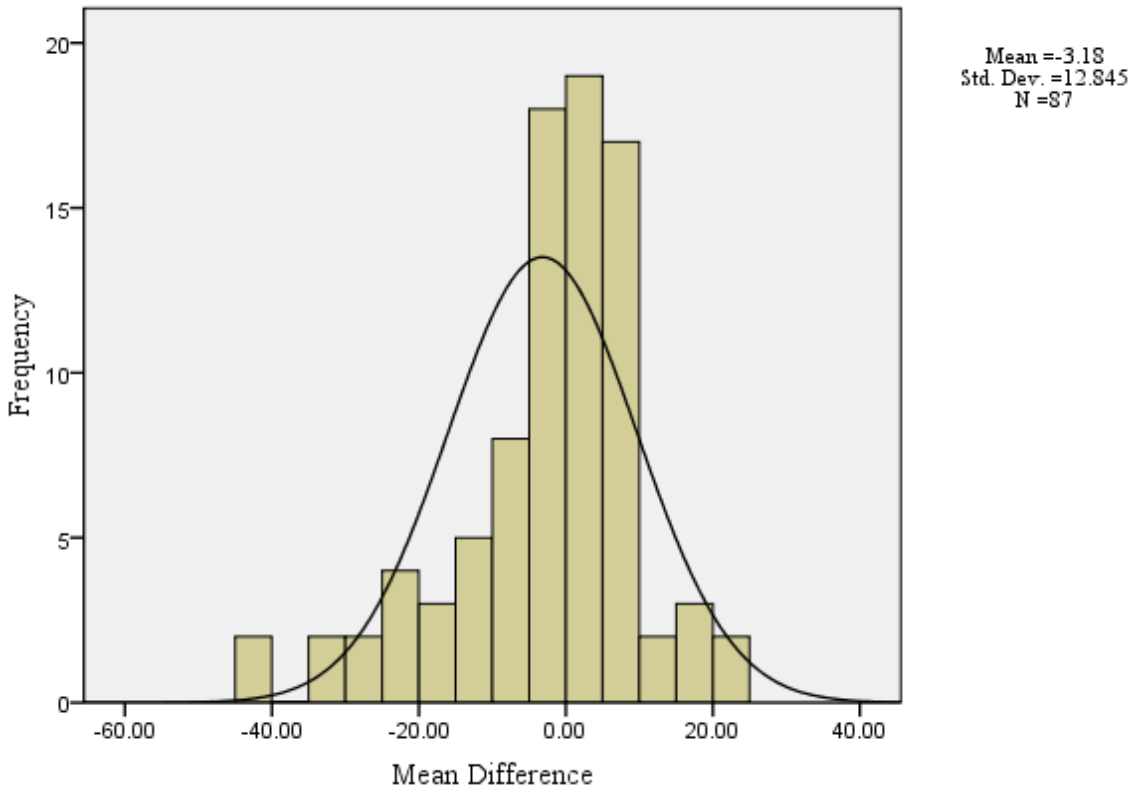

Figure 7. Distribution of CAPS-CPS-M Total Severity Score differences.

Signal Detection Statistics

Kappa coefficients. The CPS-M and CAPS were scored using the 1-2 symptom presence rule (i.e., score of 1 or higher on frequency and score of 2 or higher on intensity for an individual item), the 2-2 symptom presence rule (i.e., scores of 2 or higher on both frequency and intensity for individual items), the Sum 4 symptom presence rule (i.e., sum of 4 of higher when adding frequency and intensity scores on an individual item), and the Total 65 rule (i.e., total severity scores of 65 or higher on the scale).

The 1-2 symptom presence scoring rule resulted in 35 positive cases and 55 negative cases on the CAPS and 39 positive and 51 negative cases on the CPS-M. The 2-2 symptom presence scoring rule resulted in 29 positive and 61 negative cases on the CAPS and 35 positive and 55 negative cases on the CPS-M. The Sum 4 symptom presence scoring rule resulted in 31 positive and 59 negative cases on the CAPS and 36 positive and 54 negative cases on the CPS-M. The Total 65 scoring rule resulted in 29 positive and 58 negative cases on the CAPS and 29 positive and 61 negative cases on the CPS-M. Because three participants did not meet criterion A after beginning the CAPS interview, they were considered negative for PTSD diagnosis, but Total Severity scores were not obtained. Thus, the *n* for CAPS Total 65 scoring rule calculations was 87 instead of 90 as with the other scoring rules, which were based on presence or absence of individual symptoms.

When using the same symptom presence rule on both instruments, the Total 65 rule produced the lowest kappa (.71), and the Sum 4 rule produced the highest kappa (.79). According to Altman (1991), kappa scores that fall within the range of .61 to .80 show good agreement. There was good agreement across all scoring rules, indicating that the CPS-M agreed with the CAPS well regardless of scoring rule used. See Table 10 for kappa coefficients. The 1-2 symptom presence rule was used for signal detection analyses because positive and negative cases were most closely balanced when using this rule to determine PTSD diagnosis on the CAPS. Diagnostic agreement between the CAPS and CPS-M based on this rule was also good, with a kappa coefficient of .77.

Table 10

CAPS-CPS-M Diagnostic Agreement: Kappa Coefficients

	CPS-M symptom presence scoring rule			
CAPS symptom presence scoring rule	1-2 Rule	2-2 Rule	Sum 4 Rule	Total 65 Rule
1-2 Rule	.77	.77	.79	.66
2-2 Rule	.72	.76	.79	.75
Sum 4 Rule	.72	.76	.79	.70
Total 65 Rule	.74	.68	.71	.71

CPS-M ROC curve analysis. Receiver Operator Characteristics (ROC) curves were conducted to find the optimal cut point on the CPS-M that would maximize diagnostic sensitivity and specificity on the CAPS. Signal detection analyses were conducted using the MedCalc v11.1.1 statistical program. The CAPS 1-2 symptom presence scoring rule, in which participants must receive at least a rating of 1 or higher on symptom frequency and a 2 or higher on intensity, resulted in 35 participants who were positive for PTSD diagnosis and 55 who were negative. The area under the ROC curve was .95 ($p < .001$), which, according to Swets (1988), is considered high diagnostic accuracy. A cut score of 40 produced 100.0% sensitivity and 83.6% specificity. Figure 8 shows the receiver operating characteristic (ROC) curve with sensitivity plotted against false alarm rate. The optimal cut score is located on the graph at the data point closest to

the upper left corner. This cut point was selected since it represents the point in the ROC distribution that best balances sensitivity in detection against false alarms (or, 1-specificity).

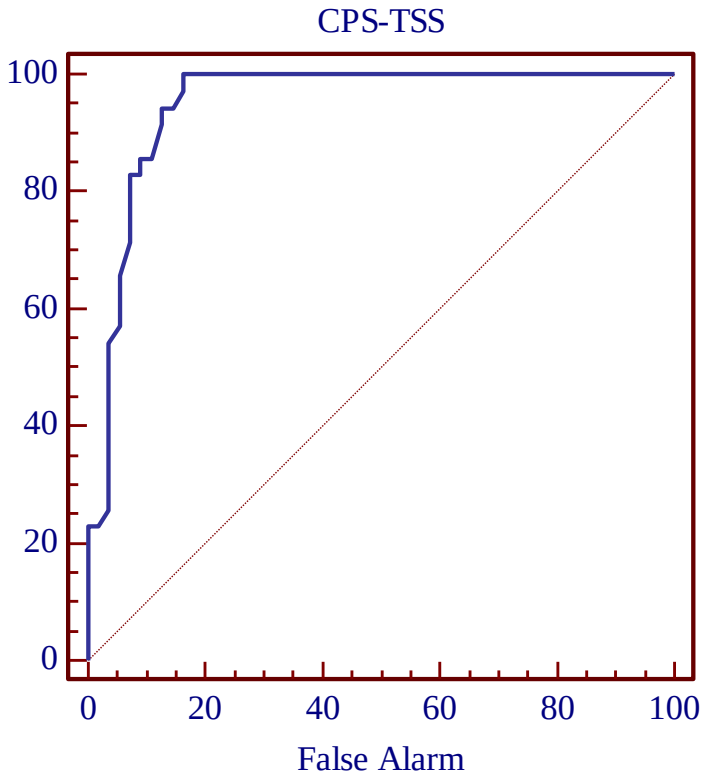

Figure 8. CPS-M ROC curve based on CAPS 1-2 symptom presence scoring rule.

Figure 9 shows the ROC curve for the CPS-M based on the CAPS 2-2 symptom presence scoring rule, in which participants must receive a rating of 2 or higher on both frequency and intensity. The CAPS 2-2 rule resulted in 29 participants who were positive for PTSD diagnosis and 61 who were negative. The area under the ROC curve was .96 ($p < .001$), which, according to Swets (1988), is considered high accuracy. A cut score of 59 produced 93.1% sensitivity and 90.2% specificity.

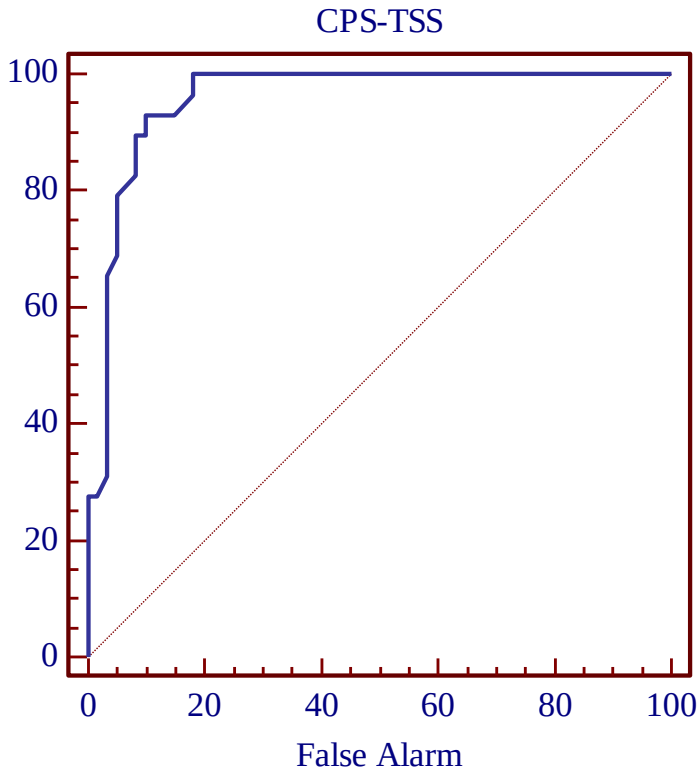

Figure 9. CPS-M ROC curve based on CAPS 2-2 symptom presence scoring rule.

The CAPS Sum 4 rule, in which a symptom is considered present when frequency plus intensity scores equal 4 or higher, resulted in 31 participants who were positive for PTSD diagnosis and 59 who were negative. The area under the ROC curve was .94 ($p < .001$), which is considered high accuracy. A cut score of 50 produced 96.8% sensitivity and 83.1% specificity. Figure 10 shows the ROC curve for the CPS-M based on the CAPS Sum 4 symptom presence scoring rule.

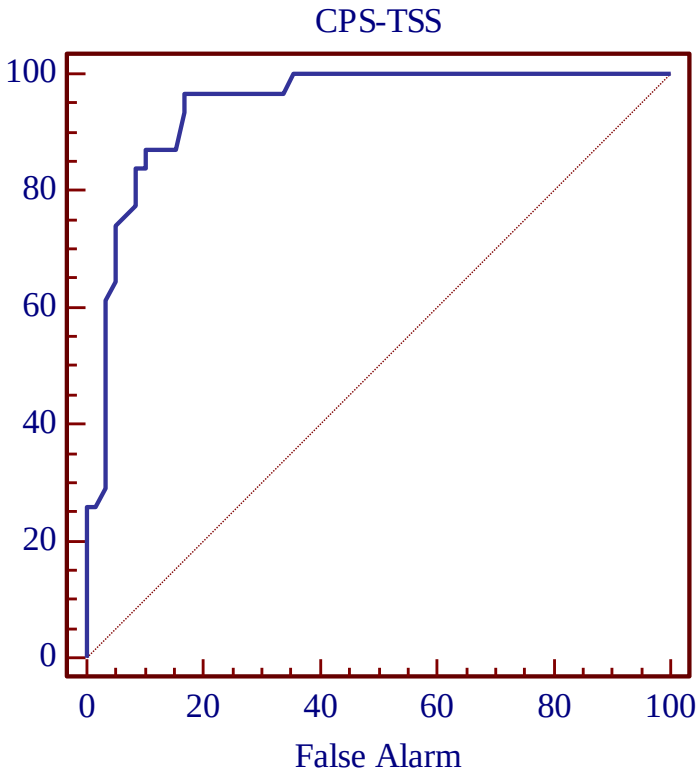

Figure 10. CPS-M ROC curve based on CAPS Sum 4 symptom presence scoring rule.

Figure 11 shows the ROC curve for the CPS-M based on the CAPS Total 65 scoring rule, in which participants are positive for PTSD diagnosis when they receive a total severity score of 65 or higher. The CAPS Total 65 rule resulted in 29 participants who were positive for PTSD diagnosis and 58 who were negative. The area under the ROC curve was .96 ($p < .001$), which is considered high diagnostic accuracy. A cut score of 50 produced 100.0% sensitivity and 82.8% specificity.

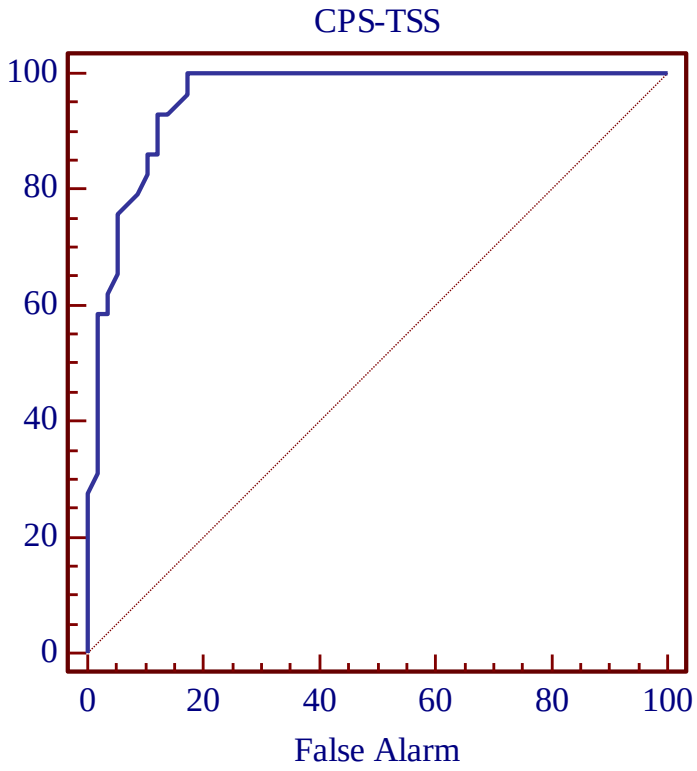

Figure 11. CPS-M ROC curve based on CAPS Total 65 scoring rule.

Figure 12 provides a distribution of CPS-M scores based on the CAPS 1-2 symptom presence scoring rule. The horizontal line on the graph represents the CPS-M Total Severity Score cut point of 40 that produced the best sensitivity and specificity. Points above the line represent participants classified as PTSD-positive on the CPS-M, while points below the line represent PTSD-negative cases. Points plotted in the column labeled "1" indicate PTSD-positive classifications on the CAPS, whereas the "0" column indicates PTSD-negative cases. This graph reproduces the 2 x 2 signal detection contingency table where points in the upper right quadrant and lower left quadrant are correct classifications and points in the upper left and lower right quadrants are misclassifications. Comparing the CPS-M to the CAPS using the 1-2 scoring rule with a

cut score of 40 resulted in 100% sensitivity and 83.6% specificity. There were zero false negatives and nine false positives.

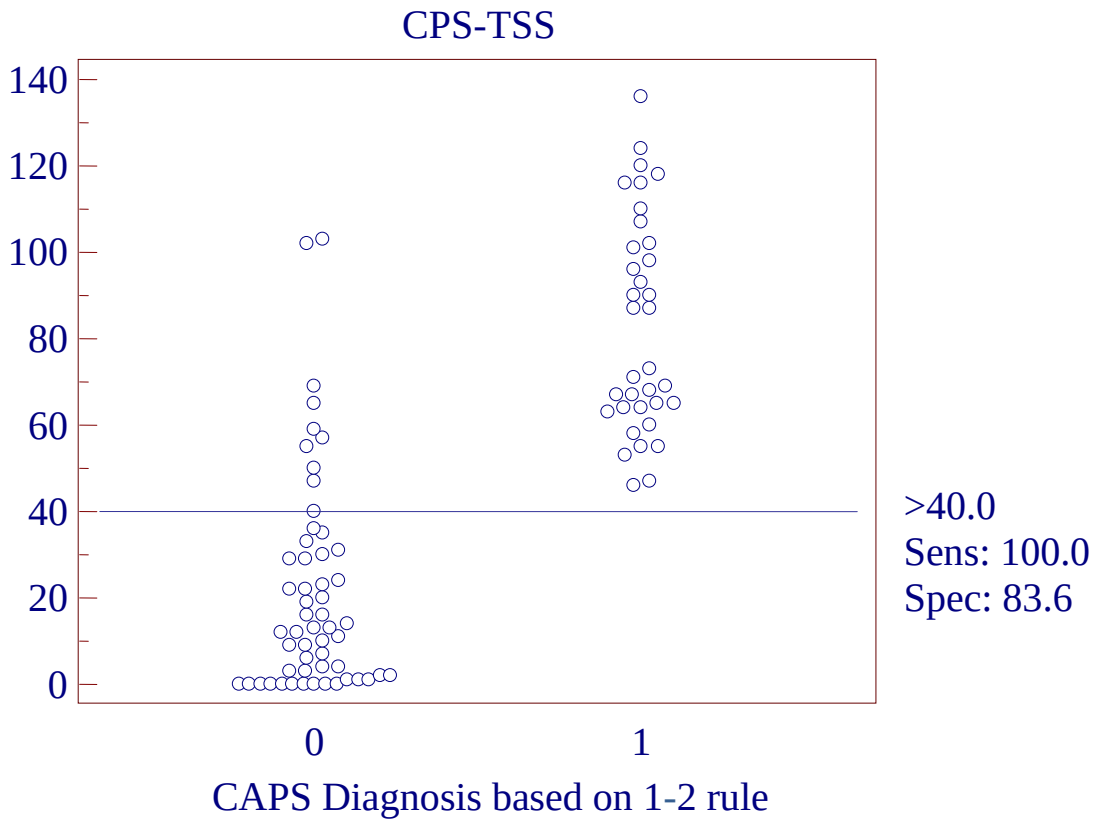

Figure 12. Distribution of CPS-M Scores based on CAPS 1-2 scoring rule.

Figure 13 provides a distribution of CPS-M scores based on the CAPS 2-2 symptom presence scoring rule. A cut score of 59 resulted in 93.1% sensitivity and 90.2% specificity. Two false negatives resulted, and six false positives were found.

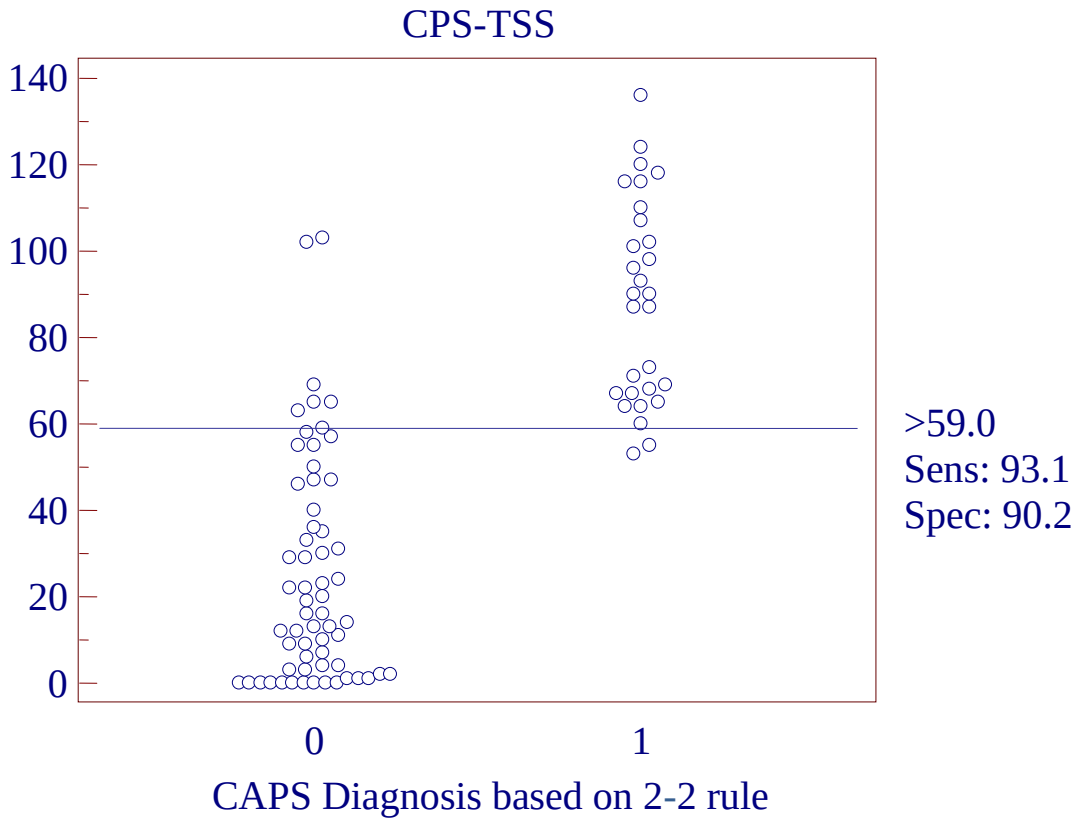

Figure 13. Distribution of CPS-M Scores based on CAPS 2-2 scoring rule.

Figure 14 provides a distribution of CPS-M scores based on the CAPS sum 4 scoring rule. A cut score of 50 resulted in 96.8% sensitivity and 83.1% specificity. One false negative resulted, and ten false positives were found.

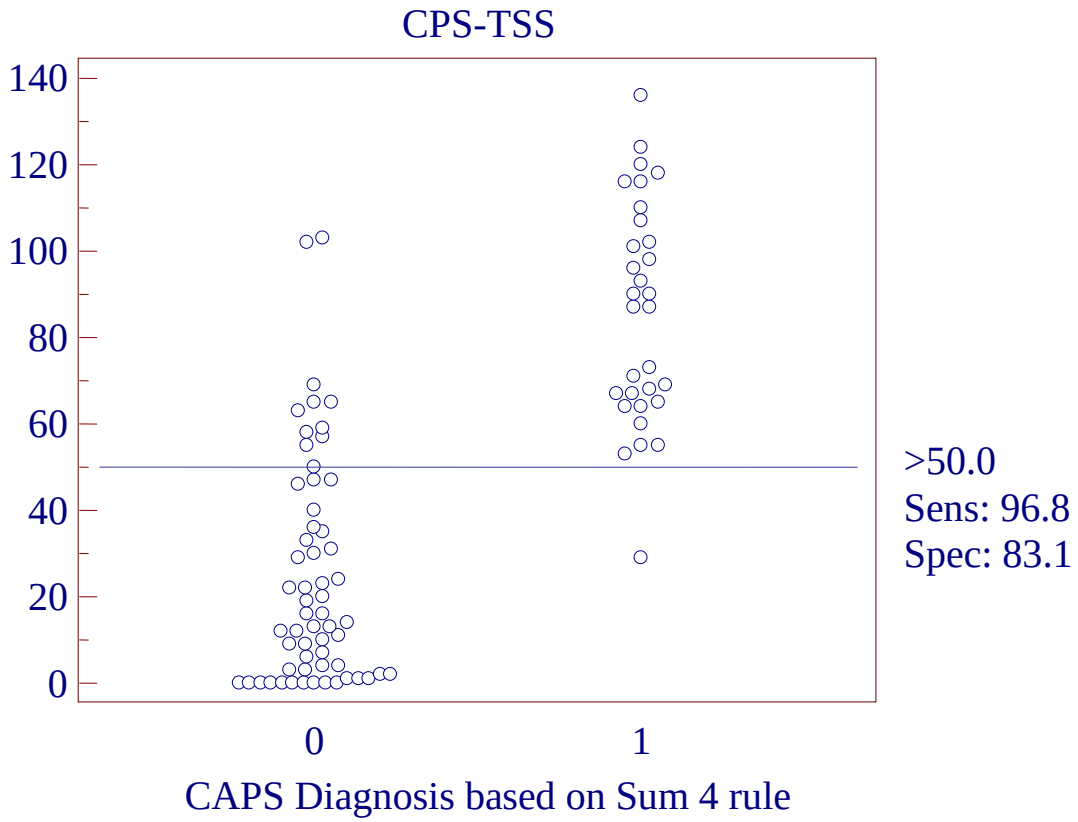

Figure 14. Distribution of CPS-M Scores based on CAPS Sum 4 scoring rule.

Figure 15 provides a distribution of CPS-M scores based on the CAPS total 65 scoring rule. A cut score of 50 resulted in 100% sensitivity and 82.8% specificity. There were zero false negatives and ten false positives found.

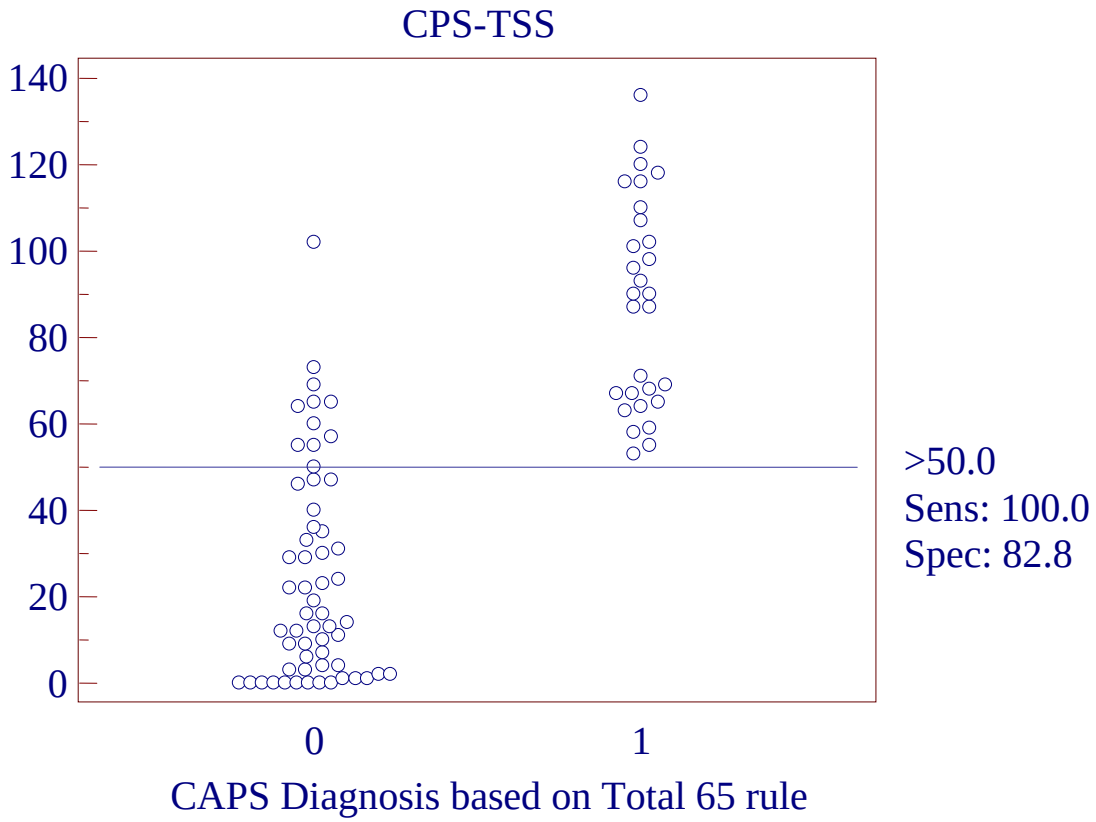

Figure 15. Distribution of CPS-M Scores based on CAPS Total 65 scoring rule.

When analyzing signal detection statistics, it is important to consider the costs of false positives or negatives. In some situations, the cost of a false negative (i.e., missing a diagnosis of cancer) would be far higher than that of a false positive, or false alarm (i.e., incorrectly diagnosing cancer that is not present). This is an important factor in determining a cut point score. Varying the cut point changes the sensitivity and specificity, so the research can determine how best to balance false positives and negatives. In the diagnosis of PTSD, it would seem that finding a cut point that minimizes false negatives (i.e., not diagnosis PTSD when it is present) would be responsible. As such, the cut score of 40 when using the CAPS 1-2 scoring rule seems to

have clinical relevance, as it produces high sensitivity and specificity, with zero false negatives.

CPS-M signal detection compared to other instruments. The final phase of the analysis involved comparing the ROC curves of the CPS-M, PTSD Checklist, Purdue PTSD Scale, Beck Depression Inventory, and the Trauma-Related Dissociation Scale. We hypothesized that the ROC curve for the CPS-M would obtain an AUC estimate that was either not significantly different than those of the convergent validity measures or accounted for slightly more area than the AUCs for the convergent validity measures.

Total scores for all five instruments were compared, resulting in AUCs of .94 (*SE* = .02, CI = 0.86 to 0.98) for the PCL, .92 (*SE* = .03, CI = 0.85 to 0.97) for the Purdue, .94 (*SE* = .02, CI = 0.87 to 0.98) for the BDI, .91 (*SE* = .03, CI = 0.83 to 0.96) for the TRDS, and scores ranging from .94 to .96 for the CPS-M, based on the different scoring rules. Figures 16, 17, 18, and 19 show the ROC curves for each of the five measures based on the four scoring rules (i.e., 1-2, 2-2, Sum 4, and Total 65). The AUC for the CPS-M was highest when using the 2-2 and Total 65 symptom presence scoring rules (.96) and lowest when using the Sum 4 rule (.94).

There were no significant differences between the CPS-M and the PTSD Checklist based on any scoring rules. Significant differences were found between AUCs when comparing the CPS-M to the TRDS based on all CAPS scoring rules except the 1-2 rule. Between the TRDS and CPS-M, differences between areas under the curves were 0.06 (*SE* = .03, 95% CI = 0.00 to 0.11, $p < .05$) using the CAPS 2-2 rule, 0.06 (*SE* = .02, 95% CI = 0.00 to 0.10, $p < .01$) using the CAPS Sum 4 rule, and 0.09 (*SE* = .03, 95% CI = 0.03 to 0.15, $p < .05$) using the CAPS Total 65 rule. These significant differences

suggest that the CPS-M is more efficient at diagnosing PTSD than the TRDS, which was hypothesized. The AUC for the CPS-M was also significantly higher than those of the Purdue and BDI when using the CAPS Sum 4 rule. The difference between the areas under the curve using the CAPS Sum 4 rule were .04 (*SE* = .02, 95% CI = 0.01 to 0.08, $p < .05$) between the Purdue and CPS-M and .06 (*SE* = .03, 95% CI = 0.01 to 0.11, $p < .05$) between the BDI and CPS-M. These differences suggest that the CPS-M was slightly more efficient at diagnosing PTSD based on the Sum 4 rule. When using any other rule, AUCs were not significantly different between the CPS-M and other measures, suggesting they were similarly efficient in their ability to detect PTSD.

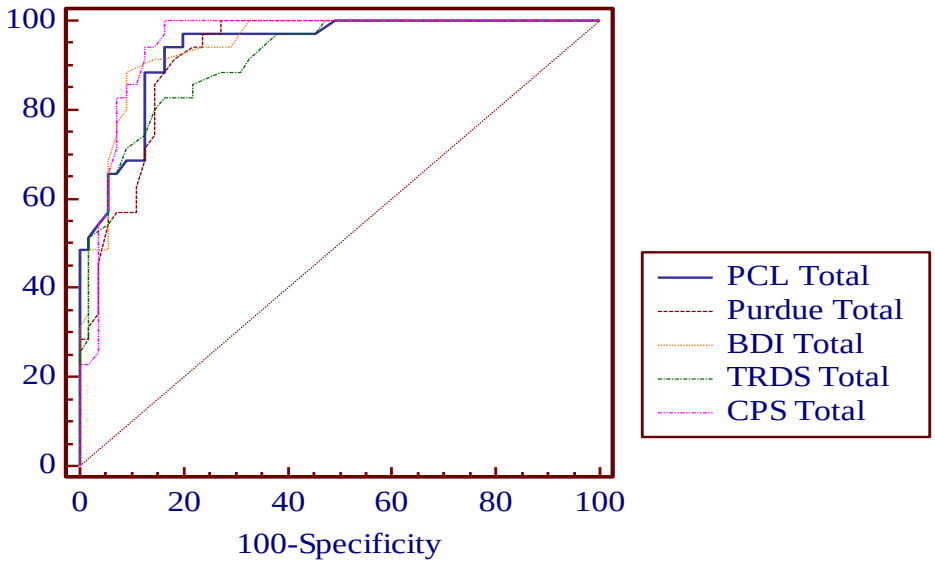

Figure 16. ROC curves for the PCL, Purdue, BDI, TRDS, and CPS-M (based on CAPS 1-2 scoring rule).

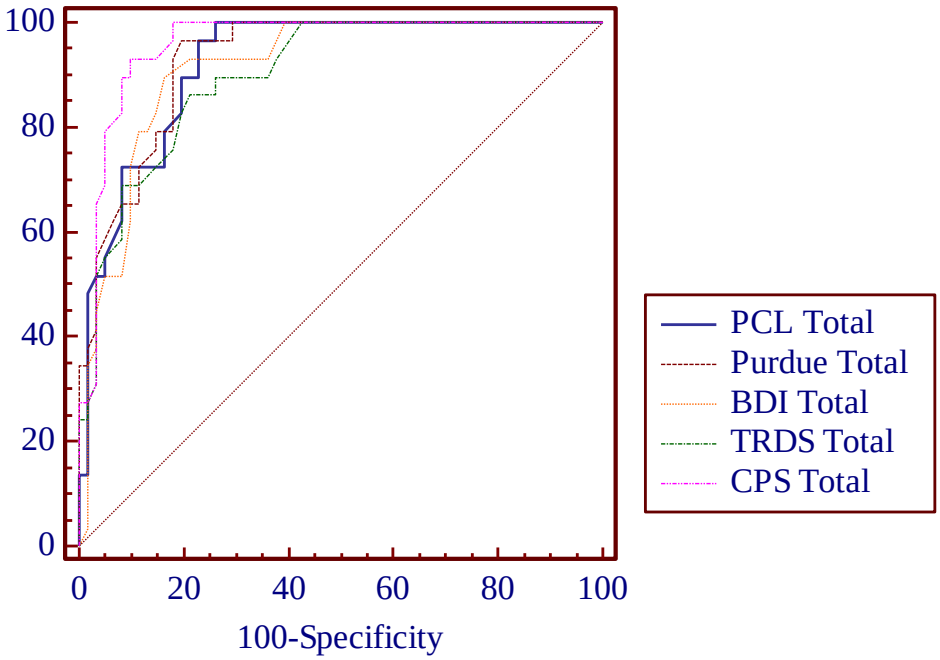

Figure 17. ROC curves for the PCL, Purdue, BDI, TRDS, and CPS-M (based on CAPS 2-2 scoring rule).

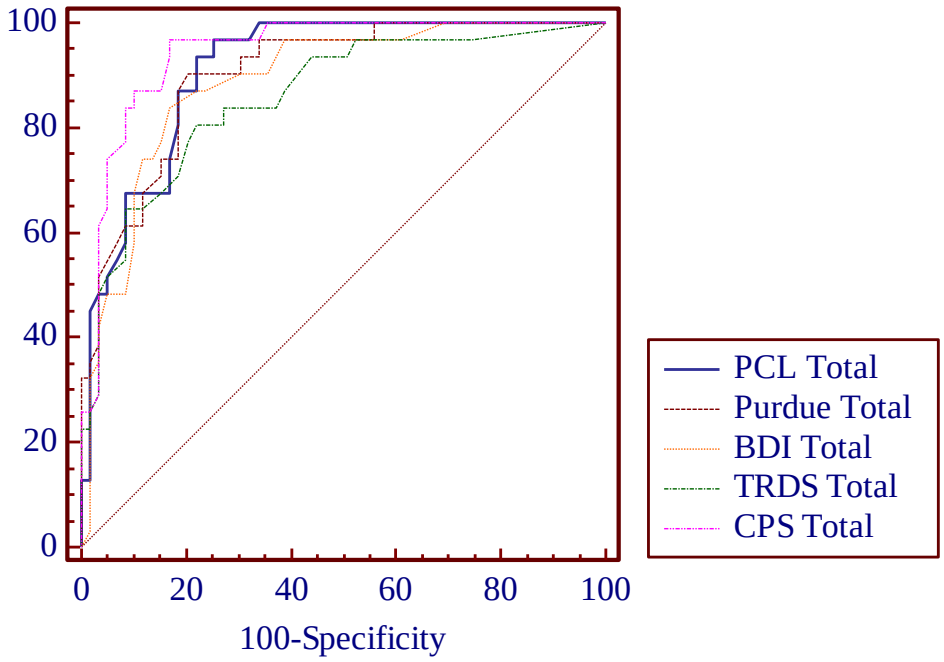

Figure 18. ROC curves for the PCL, Purdue, BDI, TRDS, and CPS-M (based on CAPS Sum 4 scoring rule).

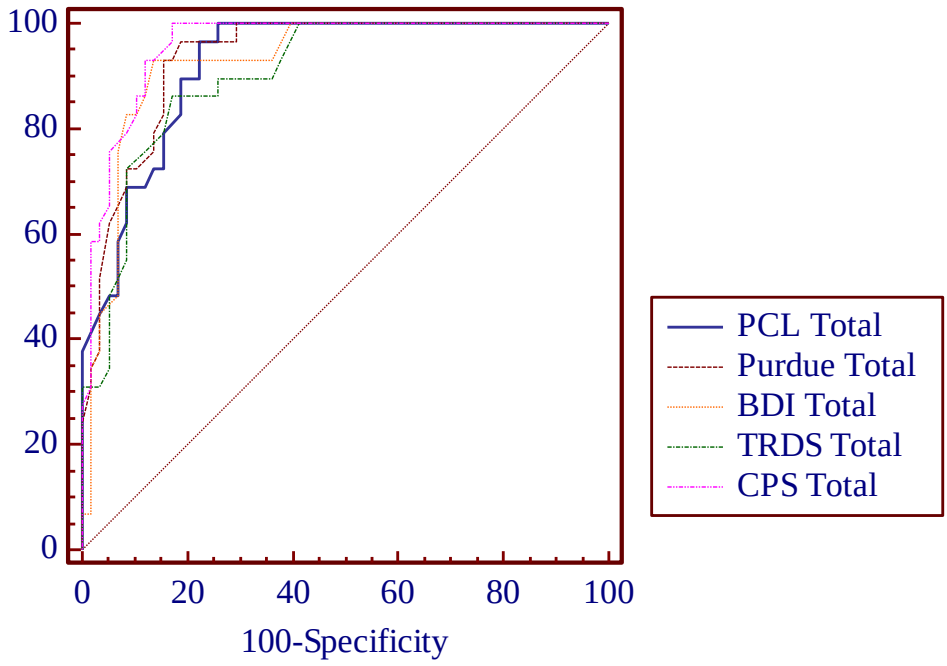

Figure 19. ROC curves for the PCL, Purdue, BDI, TRDS, and CPS-M (based on CAPS Total 65 scoring rule).

Discussion

This study compared the CPS-M to the CAPS in order to test the convergent and discriminant validity of the former and to examine its diagnostic efficiency with regard to PTSD diagnosis. All items and subscales were significantly correlated, which is not surprising given that the CPS-M was modeled after the CAPS. Total severity scores on the CAPS and CPS-M were most strongly correlated, followed by subscale B (reexperiencing), D (hyperarousal), and C (avoidance/numbing). As predicted, subscales on the CPS-M were better correlated with their corresponding subscale on the CAPS than with others. As hypothesized, the CPS-M was most strongly correlated with the CAPS, moderately correlated with other measures of PTSD, and lowest with measures outside of PTSD (i.e., measures of depression and dissociation). This pattern of correlations supports the construct validity of the CPS-M.

Total severity scores on the CPS-M were slightly higher than those on the CAPS, though this difference was not significant. Difference scores were normally distributed, though several outliers were present. It is hypothesized that the few outliers in difference scores can be accounted for by veteran confusion about trauma reporting. A few veterans began responding on either the CAPS or CPS-M based on a different trauma than that they had discussed in other portions of study and were reminded to respond based on the same trauma. However, it is possible that several veterans continued responding based on a different trauma, which could result in difference between CAPS and CPS-M severity.

The CPS-M demonstrated good internal consistency, with alpha coefficients for the subscales all above .90 and a full scale alpha of .97. These values compared

favorably with the CAPS, with subscale alphas above .80 and a full scale alpha of .94. Criterion D on the CPS-M yielded the lowest value for coefficient alpha. This was also true of the CAPS. This somewhat lower alpha value for Criterion D across both was most likely due to a variety of factors. First, the sample included veterans, many of whom tend to have irregular sleep patterns frequently caused by variables not related to trauma. Many respondents indicated during the CAPS interview that they had encountered sleep problems but were not sure if these were related to their trauma so much as other medical issues. Second, many respondents reported difficulties hearing. Several stated during the interview that they were often startled when someone entered a room because they "could not hear them coming." These respondents expressed that the startle response had more to do with their hearing loss than with trauma. Conceptually, these findings suggest that the CAPS and the CPS-M may both yield false positives for this symptom category. These data also suggest that a possible failing of the CAPS was recreated in the CPS-M. Unfortunately, the absence of a "true" gold standard prevents a careful examination of the diagnostic utility of the CAPS.

Confidence interval testing revealed that, although the CPS-M scales were not statistically different from the corresponding CAPS scales, none of the CAPS and CPS-M scales were equivalent. These findings provide partial support for the hypothesized equivalence. Confidence interval analysis is a method that has still not been used much in the field of psychological diagnosis. As it has been used in medicine, stringent equivalence intervals have been employed. It is possible that the equivalence intervals used in the current study were too stringent. This notion was tested by using a broader equivalence interval of ±20%. However, even when the equivalence interval was

expanded, the findings indicate that the CPS-M and the CAPS are not equivalent. It is possible that this may still be too stringent for determining equivalence of two psychological assessment instruments.

Another possibility is that the tests were not equivalent due to factors inherent in our sample. Many of the participants were older adults (M = 58.8; SD = 10.6) who expressed great discomfort with computer use. Several reported confusion and anxiety when completing the CPS-M. These feelings could have led to carelessness in choosing responses as well as attempts to quickly finish, resulting in responses that were more reflective of their mood at the time than of their symptoms related to the trauma. This could have also contributed to the slightly higher mean for CPS-M total severity scores, although this difference was not significant.

This reported anxiety is inconsistent with findings from several studies in which users preferred computerized assessment over face-to-face or paper-and-pencil measures. However, the samples in these studies were very different from the current study in both age and familiarity with computers, including college students (Vispoel, 2000; Vispoel, Boo, & Bleiler, 2001), retail managers with a mean age of 42 years (Richman-Hirsh, Olson-Buchanan, and Drasgow, 2000), and psychiatric inpatients with a mean age of 43.5 years (Weber et al., 2003). Weber et al. found that participants who were older or less educated reported more difficulty with the computerized assessment than did younger or more educated participants.

In two previous phases of the current study, the Computer Anxiety Rating Scale (CARS: Miller & Rainer, 1995) was administered to participants to determine their level of computer-related anxiety. On the CARS, participants rate the level of computer-

related anxiety on a seven-point Likert-type scale anchored by 1 (*less anxious*) and 7 (*most anxious*). CARS scores can a range from 7 to 35. Mason (2005) reported a CARS mean total score of 13.74 (*SD* = 5.02), suggesting negligible levels of computer-related anxiety in a sample of college students endorsing a trauma history. In a clinical sample of 161 participants from two sites, a VA outpatient clinic (*n* = 56) and a large urban outpatient clinic (*n* = 105), Mason (2007) reported a CARS total score mean of 15.04 (*SD* = 6.35), suggesting that aggregate levels of computer-related anxiety were relatively low. The mean age of participants (*M* = 50.12) in the Mason (2007) study was lower than that in the current study (*M* = 58.8), which may contribute to the disparity in reported computer anxiety.

A final consideration was that the tests may not have been equivalent due to the equivalence methodology used. Because there is no standard for calculating confidence intervals in psychological assessment, the methodology detailed by Rogers et al. (1993), which assumes a between-groups design, was followed closely. Because equivalence testing has been mainly used in pharmaceutical research, there is no guideline for within-groups designs.

Next the diagnostic utility of the CPS-M was examined. There was good agreement between the CPS-M and the CAPS across all scoring rules. The 1-2 symptom presence rule was used for signal detection analyses because positive and negative cases were most closely balanced when using this rule to determine PTSD diagnosis on the CAPS. As hypothesized, signal detection statistics revealed high accuracy of the CPS-M in diagnosing PTSD and an AUC of .95. A cut score of 40 produced 100.0% sensitivity and 83.6% specificity. This is considered high diagnostic accuracy, which improves

upon the medium accuracy reported in the Mainka (2005) study, in which a cut score of 45 using the CAPS 1-2 symptom presence scoring rule yielded overall diagnostic efficiency of .78, sensitivity of .63 and specificity of .80.

These diagnostic efficiency statistics compared favorably to previous signal detection studies comparing PTSD instruments to the CAPS. The CPS-M's overall diagnostic efficiency of .95 was higher than any studies detailed in our review of the literature, with the next highest reported overall efficiency being .94, which resulted in .97 sensitivity of specificity of .87 on the PTSD Checklist (Ventureyra et al., 2002). No other studies reported 1.00 sensitivity, and the next highest reported sensitivity was .98 on the PTSD Checklist (Forbes et al., 2001). The CPS-M specificity score of .84 was higher than that reported in most reviewed studies, though higher scores were reported for the PTSD Checklist (.97; Hudson et al., 2001) and the Korean version of the Short PTSD Rating Interview (.93; Kim et al., 2008). The same specificity was reported for the PTSD Checklist - Civilian version (Bollinger et al., 2008) and Breslau's 7-item PTSD screen (Kimerling, Ouimette et al., 2006).

These findings are encouraging with regard to the ability of the CPS-M to accurately diagnose PTSD. Though we focused on the diagnostic efficiency based on the CAPS 1-2 symptom presence scoring rule, the diagnostic efficiency statistics produced by the other scoring rules also favorably compared to previous single detection studies, with overall diagnostic efficiencies on the CPS-M ranging from .94-.96, sensitivity scores ranging from .93-1.00, and specificity scores ranging from .83-.90. Because the CPS-M was directly modeled after the CAPS, we would expect that its diagnostic efficiency relative to the criterion would be higher than that found for other PTSD instruments.

Though some of these instruments are similar to the CAPS, the CPS-M asks the same questions and provides the same prompts as a computerized adaptation of the parent instrument.

A limitation inherent in signal detection work is that cut scores are very sample-specific. Cut scores vary widely across studies on the same instrument due to factors such as sample age, trauma history, and history of treatment. This makes it difficult to examine the clinical utility of the cut score provided. It is therefore necessary to be cautious when generalizing a specific cut score outside of the studied population. Clinicians are encouraged to examine research in which the sample studied closely resembles their patient population. Additionally, this calls for more research in large samples with heterogeneous variables (i.e.-wider age range, several study sites).

In the final phase of analysis, the ROC curve of the CPS-M was compared to ROC curves produced by the Purdue PTSD Scale, PTSD Checklist, BDI, and TRDS. It was hypothesized that the ROC curve produced by the CPS-M would be slightly better or would not be significantly different from those produced by these other instruments. This hypothesis was supported by the finding that there were not significant differences between the CPS-M and the PCL using any scoring rule, the CPS-M was as efficient as the Purdue and BDI on all but one scoring rule, and the CPS-M was more efficient than the TRDS using all but one rule. As expected, the CPS-M was better at diagnosing PTSD than the TRDS, which is a measure of dissociation. It was as efficient as the other two measures of PTSD, the Purdue and the PCL.

The CPS-M was not more efficient at diagnosing PTSD than the BDI, a measure of depression. One possible reason for the lack of significant difference in efficiency

between these two instruments is the high comorbidity between depression and PTSD, especially in veterans. When reviewing the literature, Chan, Cheadle, Reiber, Unutzer, and Chaney (2009) reported that, among veterans with a PTSD diagnosis, rates of comorbid depression ranged from 29% to 68%. Among veterans with major depression diagnoses, rates of comorbid PTSD ranged from 36% to 51% (Chan et al., 2009). As many of the study participants were patients recruited from the mental health clinic, it is likely that many had diagnoses of major depression and/or PTSD.

A surprising finding was that only 13 participants had service-connected status for PTSD, when the number of participants classified as PTSD-positive based on CAPS diagnosis ranged from 29 to 35 across scoring rules. This may have occurred for several reasons. First, at the time of the study, the wait time for a PTSD evaluation to determine service connection ranged from two to four months. It is possible that many veterans were waiting to be evaluated, though we do not have data on how many actually applied for service connection. However, given the age of the participants (M = 58.8, SD = 10.6), this explanation is somewhat less likely.

Second, many of the participants were recruited from primary care and had not had any connection with the mental health or PTSD clinics. While many veterans enjoy close interpersonal connections, several studies have documented increased feelings of alienation and isolation among combat veterans with PTSD (Egendorf, Laufer, & Sloane, 1981; Rippy, 2008). As many veterans receive information from each other (e.g., how to obtain service connected status, treatment options, outreach resources), it is possible that limited interpersonal connection has precluded some veterans from obtaining information about the treatment or service connection options available to them.

Third, several participants noted that they did not want any kind of assistance from the government for their PTSD, indicating the desire to "handle it." There is a long history of veteran disenfranchisement from the VA Medical Center. The Veterans Healthcare Amendment Act of 1979 (PL-96-22), also known as Operation Outreach, was developed to provide comprehensive psychosocial services to Vietnam-era veterans. This resulted in the development of veteran outreach centers across the nation, often housed in storefronts and other nongovernmental buildings and staffed by Vietnam veterans (Blank, 1982). Giles (1981) estimated that these centers delivered services to about 20 to 50 percent of Vietnam veterans suffering from PTSD symptoms.

In 1983, Congress passed Public Law 98-160, mandating a study of readjustment and mental health status of Vietnam veterans, which became known as the National Vietnam Veterans Readjustment Study (NVVRS; Kulka et al., 1990). One of the goals of this study was that veterans would become connected with other veterans and would more readily seek services. While many veterans have taken advantage of these connections, responses from participants in the present study indicate that some individuals still prefer to keep private their symptoms and need for assistance.

Finally, the possibility exists that participants, though assured that the study had nothing to do with their diagnoses, benefits, or service connection, still hoped that the results of the study would be reported to the VA. It is possible that some of these individuals had been evaluated and determined to not meet criteria for PTSD service connection. This hypothesis is reinforced by several participants who asked at the completion of the study how they had performed or when the results would be shared with their provider.

Conclusions

The current study had a number of limitations. First, the demographic compilation of our sample somewhat limits the generalizability of our findings. Participants were mostly men (77 men, 13 women) who were over the age of 50 (81.1%), which is representative of the VA population. The cut score and resulting diagnostic efficiency are strongly supported in older adult males with a wide range of traumas. Future research should include a clinical sample from several sites with a younger age range, as well as more equal representation of men and women. Additionally, the majority of the participants were Caucasian (83.5%), which was representative of the Ann Arbor VA population, but this limits the generalizability of findings to other racial groups.

A second limitation of the study is the high rate of comorbidity in veterans, especially those suffering from depression or PTSD. While the current study did not assess comorbidity, rates of comorbid depression range from 29% to 68% for veterans with a primary PTSD diagnosis and from 36% to 51% for veterans with a primary major depressive diagnosis (Chan et al., 2009). Because participants were not asked to report diagnoses, the current study was unable to examine the diagnostic efficiency of the CPS-M based on presence or absence of depression. In examining BDI-II scores, we found that over half of veterans reported at least mild depression (scores of 14 or higher), though these scores were not converted into diagnoses. While the comorbidity likely inherent in the sample is representative of much of the PTSD population, future validation of the CPS-M should also focus on individuals without comorbid diagnoses or should assess for comorbidity at the time of the study.

Overall, findings were encouraging regarding the CPS-M. It displayed good internal consistency at both the full scale and subscale levels. It was significantly correlated with the CAPS and was also more strongly associated with other measures of PTSD than with the measures of depression and dissociation. This lends support to its use for the diagnosis of PTSD.

Confidence interval analysis revealed that, though not statistically different, the CPS-M and CAPS were not equivalent, which did not support the original prediction. It was theorized that reasons for this include the potential that the equivalence intervals used were too stringent for psychological diagnosis or that respondents' discomfort and anxiety with using a computer could have led to slightly elevated responses on the CPS-M that were more reflective of their mood than of their symptomatology.

Signal detection statistics were excellent for the CPS-M and compared favorably to other studies using the CAPS as the criterion. The overall diagnostic efficiency and sensitivity obtained were higher than those reported in all other reviewed studies, and the CPS-M specificity score was higher than that reported in all but two other studies. Regardless of the scoring rule used, the diagnostic efficiency statistics revealed high diagnostic accuracy of the CPS-M relative to the CAPS.

ROC curve analysis revealed that the CPS-M was more efficient at PTSD diagnosis than the measure of trauma, and as efficient at the measures of PTSD and depression. While this matched the initial hypothesis, it was suggested that future research that separates participants by comorbid diagnoses would yield a ROC curve that is better than that for the measure of depression.

The CPS-M is a computerized adaptation of the CAPS that addresses several of the limitations of the CAPS. According to Weathers et al. (2001), limitations of the CAPS include its lengthy administration time, cumbersome training, overlap of the frequency and intensity prompts, and the hesitation of individuals to disclose sensitive information immediately upon meeting the administrator or clinician. With regard to administration time, average CPS-M completion time in the present study was 30 minutes, compared to average administration time of an hour and a half for the CAPS, making it far more efficient. This provides support for the CPS-M as an excellent assessment instrument for use in the field or in cases in which individuals are not able to sit through a lengthy interview.

The second limitation of the CAPS is also addressed by the CPS-M. The only training required for the CPS-M is a practice administration so the examiner is familiar with the way the program runs. Interviewers are not required to assist with or even be present during the administration. The complaint of overlap of the frequency and intensity prompts is not entirely addressed in the CPS-M, though the program is structured differently, so that this may seem less redundant. Participants first endorse presence or absence of each symptom. The program then cycles through only the symptoms they endorsed and assesses frequency, and a final round assesses intensity.

The final criticism of the CAPS is that it often requires individuals to disclose sensitive information to an examiner immediately upon meeting that person. This is the case with all interview-based assessment, and the CPS-M sidesteps this criticism by allowing individuals to report their trauma and endorse symptoms privately. Turner, Ku, Rogers, Lindberg, Pleck, and Sonenstein (1998) suggested that computerized interviews

encourage the disclosure of more sensitive information than face-to-face interviews. According to Mason (2007), this is particularly advantageous in the assessment of traumas that are difficult or embarrassing to discuss, such as unwanted sexual experiences or death caused to another person (i.e., killing another soldier in combat). A disadvantage of this approach is that individuals are not provided with any type of empathy or understanding that an examiner may show during the interview. While some respondents may feel more comfortable reporting traumatic events to a computer rather than a person, there are certainly some who would prefer human contact and interaction.

Because the CPS-M is modeled after the CAPS and DSM-IV-TR diagnostic criteria, changes in diagnostic criteria will necessitate changes in the CPS-M, which presents a limitation of this instrument. However, the CPS-M addresses many of the limitations of the CAPS and provides an excellent alternative for PTSD assessment remotely and without the presence of clinicians. Based on its high associations with the screening measures in this study, the Purdue and PCL, arguments can be made for its use as a remote screening instrument.

It also shows promise as a diagnostically accurate instrument that can be used in VA medical centers, which typically have long wait times for PTSD evaluations. The reduction in administration time and removal of the need for a clinician allow it to be used more quickly, more often, and with more patients. These advantages also lend support to its use in settings such as domestic violence shelters, college counseling centers, and emergency rooms. Comparison of the CPS-M to the gold standard CAPS and other PTSD measures lead to the conclusion that the CPS-M is an excellent

diagnostic tool that not only sidesteps many of the limitations of the CAPS but fits well into the emerging world of computerized assessment.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC.

Aribisi, P. A., Murdoch, M., McNulty, J., & Fortier, L. (2004). MMPI-2 validity and award of service connection for PTSD during the VA compensation and pension evaluation. *Psychological Services, 1,* 56-67.

Asmundson, G. J. G., Wright, K. D., McCreary, D. R., & Pedlar, D. (2003). Post-traumatic stress disorder symptoms in United Nations peacekeepers: An examination of factor structure in peacekeepers with and without chronic pain. *Cognitive Behaviour Therapy, 32,* 26–37.

Baschnagel, J. S., O'Connor, R. M., Colder, C. R., & Hawk, L. W. (2005). Factor structure of posttraumatic stress among Western New York undergraduates following the September 11th terrorist attack on the World Trade Center. *Journal of Traumatic Stress, 18,* 677–684.

Basoglu, M., Salcioglu, E., Livanou, M., Ozeren, M., Aker, T., Kihc, C., & Mestcioglu, O. (2001). A study of the validity of a screening instrument for traumatic stress in earthquake survivors in Turkey. *Journal of Traumatic Stress, 14,* 491-509.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of Beck Depression Inventories-IA and -II in psychiatric outpatients. *Journal of Personality Assessment, 67*(3), 588-597.

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Klauminzer, G., Charney, D. S., & Keane, T. M. (1990). A clinician rating scale for assessing current and lifetime PTSD: The CAPS-1. *The Behavior Therapist, 13,* 187-188.

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., & Keane, T. M. (1995). The development of a Clinician-Administered PTSD Scale. *Journal of Traumatic Stress, 8*, 75-90.

Blanchard, E. B., Hickling, E. J., Taylor, A. E., Forneris, C. A., Loos, W., & Jaccard, J. (1995). Effects of varying scoring rules of the Clinician-Administered PTSD Scale (CAPS) for the diagnosis of posttraumatic stress disorder in motor vehicle accident victims. *Behaviour Research & Therapy, 33*, 471-475.

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

Blank, A. (1982). Apocalypse terminable and interminable: Operation Outreach for Vietnam veterans. *Hospital & Community Psychiatry, 33(11),* 913-918.

Blume, J. (2009). Bounding sample size projections for the area under a ROC curve. *Journal of Statistical Planning & Inference, 139,* 711-721.

Bollinger, A. R., Cuevas, C. A., Vielhauer, M. J., Morgan, E. E., & Keane, T. M. (2008). The operating characteristics of the PTSD checklist in detecting PTSD in HIV+ substance abusers. *Journal of Psychological Trauma, 7*(4), 213-234.

Buckley, T. C., Blanchard, E. B., & Hickling, E. J. (1998). A confirmatory factor analysis of posttraumatic stress symptoms. *Behavior Research and Therapy, 36,* 1091–1099.

Carlson, E. B. & Waelde, L. C. (1999). *Preliminary psychometrics of the Trauma-Related Dissociation Scale (TRDS).* A poster presented at the annual meeting of the International Society for Traumatic Stress Studies.

Chan, D., Cheadle, A. D., Reiber, G., Unutzer, J., & Chaney, E. F. Health care utilization and its costs for depressed veterans with and without comorbid PTSD symptoms. *Psychiatric Services, 60(12),* 1612-1617.

Charney, M. E., & Keane, T. M. (2007). Psychometric analyses of the clinician-administered PTSD scale (CAPS)--Bosnian translation. *Cultural Diversity and Ethnic Minority Psychology, 13*(2), 161-168.

Chen, C., Shen, W. W., Tan, H. K., Chou, J., & Lu, M. (2003). The validation study and application of stratum-specific likelihood rations in the Chinese version of SPAN. *Comprehensive Psychiatry, 44*, 78-81.

Coffey, S. R., Gudmundsdottir, B., Beck, J. G., Palyo, S. A., & Miller, L. (2006). Screening for PTSD in motor vehicle accident survivors using the PSS-SR and IES. *Journal of Traumatic Stress, 19*(1), 119-128.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Creamer, M., Bell, R., & Failla, S. (2003). Psychometric properties of the Impact of Event Scale – Revised. *Behaviour Research and Therapy, 41*, 1489-1496.

DeCarlo, L. T. (1998). Signal detection theory and generalized linear models. *Psychological Methods, 3*, 186-205.

Dobie, D. D., Kivlahan, D. R., Maynard, C., Bush, K. R., McFall, M., Epler, A. J., & Bradley, K. A. (2002). Screening for post-traumatic stress disorder in female Veteran's Affairs patients: Validation of the PTSD Checklist. *General Hospital Psychiatry, 24*, 367-374.

Dozois, D. J. A., Dobson, K. S., & Ahnberg, J. L. (1998). A psychometric evaluation of the Beck Depression Inventory-II. *Psychological Assessment, 10*(2), 83-89.

Dragoni, F., Gentile, G., Cartoni, C., Chiarotti, F., Puopolo, M., Mazzucconi, M. G., De Matteis, L. M., Pirillo, M., & Martino, P. (2005). The role of real-time ultrasonography in predicting esophageal varices in hemophiliacs co-infected with hepatitis C and human immunodeficiency virus. *The Hematology Journal*, *90*, 207-213.

Egendorf, A., Laufer, R., & Sloan, L. (1981). *Legacies of Vietnam: Comparative adjustment of Vietnam veterans and their peers: Overview of the report*. Washington, DC: Government Printing Office.

Farmer, R. F. & Chapman, A. L. (2002). Evaluation of the DSM-IV personality disorder criteria as assessed by the structured clinical interview for DSM-IV personality disorders. *Comprehensive Psychiatry, 43 (4)*, 285-300.

Foa, E. B., & Tolin, D. F. (2000). Comparison of the PTSD Symptom-Scale - Interview Version and the Clinician-Administered PTSD Scale. *Journal of Traumatic Stress, 13*, 181-191.

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

Forbes, D., Creamer, M., & Biddle, D. (2001). The validity of the PTSD checklist as a measure of symptomatic change in combat-related PTSD. *Behaviour Research & Therapy, 39*, 977-986.

Fossati, A., Beauchaine, T. P., Grazioli, F., Carretta, I., Cortinovis, F., & Maffei, C. (2005). A latent structure analysis of Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Narcissistic Personality Disorder criteria. *Comprehensive Psychiatry, 46*, 361-367.

Gearon, J. S., Bellack, A. S., & Tenhula, W. N. (2004). Preliminary reliability and validity of the Clinician-Administered PTSD Scale for Schizophrenia. Journal of *Consulting and Clinical Psychology, 72*, 121-125.

Giles, S. (1981). Appalachia's "forgotten warriors": outreach services for Vietnam veterans. *New Dimensions in Mental Health, 9*, 23-31.

Gray, M. J., Litz, B. T., Hsu, J. L., & Lombardo, T. W. (2004). Psychometric properties of the Life Events Checklist. *Assessment, 11*, 330-341.

Greig, G. L. (1990). On the shape of energy-detection ROC curves. *Perception & Psychophysics, 48*, 77-81.

Grilo, C. M., Becker, D. F., Anez, L. M., & McGlashan, T. H. (2004). Diagnostic efficiency of DSM-IV criteria for Borderline Personality Disorder: An evaluation in Hispanic men and women with substance use disorders. *Journal of Consulting and Clinical Psychology, 72*, 126-131.

Grilo, C. M. (2004). Factorial structure and diagnostic efficiency of DSM-IV criteria for Avoidance Personality Disorder in patients with Binge Eating Disorder. *Behaviour Research and Therapy, 42*, 1149-1162.

Groham, K., & Fals-Stewart, W. (2004). The detection of cognitive impairment among substance-abusing patients: The accuracy of the Neuropsychological Assessment Battery – Screening Module. *Experimental and Clinical Psychopharmacology, 12*, 200-207.

Grubaugh, A. L., Elhai, J. D., Cusack, K. J., Wells, C., & Frueh, B. C. (2007). Screening for PTSD in public-sector mental health settings: The diagnostic utility of the PTSD checklist. *Depression and Anxiety, 24*(2), 124-129.

Hinton, D. E., Chhean, D., Pich, V., Pollack, M. H., Orr, S. P., & Pitman, R. K. (2006). Assessment of posttraumatic stress disorder in Cambodian refugees using the clinician-administered PTSD scale: Psychometric properties and symptom severity. *Journal of Traumatic Stress, 19*(3), 405-409.

Hudson, S. A., Beckford, L. A., Jackson, S. D., & Philpot, M. P. (2008). Validation of a screening instrument for post-traumatic stress disorder in a clinical sample of older adults. *Aging & Mental Health, 12*(5), 670-673.

Huprich, S. K., Sanford, K., & Smith, M. (2002). Psychometric evaluation of the Depressive Personality Disorder Inventory. *Journal of Personality Disorders, 16,* 255-269.

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.

Kim, T., Chung, M. Y., Kim, W., Koo, Y. J., Ryu, S. G., Kim, E., Woo, J., Kim, T., Yang, J., Choi, K., Pae, C., Seo, H., Lim, H., & Chae, J. & Disaster Psychiatry Committee in Korean Academy of Anxiety Disorder, Korea. (2008). Psychometric properties of the Korean version of the short post-traumatic stress disorder rating interview (K-SPRINT). *Psychiatry and Clinical Neurosciences, 62(1),* 34-39.

Kimerling, R., Trafton, J. A., & Nguyen, B. (2006). Validation of a brief screen for post-traumatic stress disorder with substance use disorder patients. *Addictive Behaviors, 31*(11), 2074-2079.

Kimerling, R., Ouimette, P., Prins, A., Nisco, P., Lawler, C., Cronkite, R., & Moos, R. H. (2006). Utility of a short screening scale for DSM-IV PTSD in primary care. *Journal of General Internal Medicine, 21*(1), 65-67.

King, D. W., Leskin, G. A., King, L. A., & Weathers, F. W. (1998). Confirmatory factor analysis of the Clinician-Administered PTSD Scale: Evidence for the dimensionality of posttraumatic stress disorder. *Psychological Assessment, 10,* 90–96.

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

Lang, A. J., Laffaye, C., Satz, L. E., Dresselhaus, T. R., & Stein, M. B. (2003). Sensitivity and specificity of the PTSD Checklist in detecting PTSD in female veterans in primary care. *Journal of Traumatic Stress, 16*, 257-264.

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., Garcia, M., & Gloster, A. (2002). Psychometric properties and predictive validity of the Mt Scale of the MMPI-2. *Assessment, 9,* 390-400.

Mainka, J. B., & Richard, D. C. S. (May, 2005). *Diagnostic efficiency of the CPS-M with a college student sample reporting a trauma history*. A poster presented at the annual meeting of the Association for Behavioral and Cognitive Therapies.

Mason, S. T. (2005). *Psychometric properties of the Computerized PTSD Scale: Multimedia Version using a student sample.* Eastern Michigan University, Ypsilanti.

Mason, S. T. (2007). *Psychometric properties and factor structure of the Computerized PTSD Scale – Multimedia version among adult samples reporting a trauma*. Eastern Michigan University, Ypsilanti.

Mayo, S., Richard, D. C. S., Raglin, P., Graven, J. J., Henslee, A. M., & Graham, G. (May, 2000). *A multimedia adaptation of the Computerized PTSD Scale: Initial results.* A poster presented at the annual meeting of the Midwest Psychological Association.

McNally, K. A., Schefft, B. K., Szaflarski, J. P., Howe, S. R., Yeh, H., & Privitera, M. D. (2009). Application of signal detection theory to verbal memory testing to distinguish patients with psychogenic nonepileptic seizures from patients with epileptic seizures. *Epilepsy & Behavior, 14*(4), 597-603.

Miller, M. D., & Rainer, R. K. (1995). Assessing and improving the dimensionality of the Computer Anxiety Rating Scale. *Educational & Psychological Measurement, 55*(4), 652-657.

Mota, V. L., & Schachar, R. J. (2000). Reformulating attention-deficit/hyperactivity disorder according to signal detection theory. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*, 1144-1151.

Narayanswamy, R., & Johnson, K. M. (1998). Optoelectronic region of interest detection: An application in automated cytology. *Applied Optics, 37*, 6011-6025.

Neal, L. A., Busuttil, W., Herepath, R., & Strike, P. W. (1994). Development and validation of the computerized Clinician Administered Post Traumatic Stress Disorder Scale–1–Revised. *Psychological Medicine, 24*, 701-706.

Neal, L. A., Hill, N., Hughes, J., Middleton, A, & Busuttil, W. (1995). Convergent validity of measure of PTSD in an elderly population of former prisoners of war. *International Journal of Geriatric Psychiatry, 10*, 617-622.

Osman, A., Downs, W. R., Barrios, F. X., Kopper, B. A., Gutierrez, P. M., & Chiros, C. E. (1997). Factor structure and psychometric characteristics of the Beck Depression Inventory-II. *Journal of Psychopathology & Behavioral Assessment, 19*(4), 359-376.

Palmieri, P. A., Weathers, F. W., Difede, J., & King, D. W. (2007). Confirmatory factor analysis of the PTSD checklist and the clinician-administered PTSD scale in disaster workers exposed to the world trade center ground zero. *Journal of Abnormal Psychology, 116*(2), 329-341.

Pinto, M., & Grilo, C. M. (2004). Reliability, diagnostic efficiency, and validity of the Millon Adolescent Clinical Inventory: Examination of selected scales in psychiatrically hospitalized adolescents. *Behaviour Research and Therapy, 42*, 1505-1519.

Porcari, C. E., Amdur, R. L., Koch, E. I., Richard, D. C. S., Favorite, T., Martis, B., & Liberzon, I. (2009). Assessment of post-traumatic stress disorder in veterans by videoconferencing and by face-to-face methods. *Journal of Telemedicine and Telecare, 15*(2), 89-94.

Renner, W., Salem, I., & Ottomeyer, K. (2006). Cross-cultural validation of measures of traumatic symptoms in groups of asylum seekers from Chechnya, Afghanistan, and West Africa. *Social Behavior and Personality, 34*(9), 1101-1114.

Richard, D. C. S., Mayo, S., Bohn, J., Haynes, S., & Kolman, L., (1997). *Development of the Computerized PTSD Scale: Multimedia Version.* Poster presented at the annual conference of the Association for the Advancement of Behavior Therapy. Miami, FL.

Richard, D. C. S. (1999). Development and psychometric evaluation of the computerized PTSD scale. *Dissertation Abstracts International: Section B: The Sciences & Engineering, 60*(5-B) (Dec 1999).

Richman-Hirsch, W. L., Olson-Buchanan, J. B., & Drasgow, F. (2000). Examining the impact of administration medium on examinee perceptions and attitudes. *Journal of Applied Psychology, 85*(6), 880-887.

Rippy, A. (2008). The relationship between combat-related posttraumatic stress disorder and subclinical paranoia among Vietnam veterans. *Dissertation Abstracts International: Section B. Sciences and Engineering, 69(1-B),* 695.

Rogers, J. L., Howard, K. I., & Vessey, J. T. (1993). Using significance tests to evaluate equivalence between two experimental group. *Psychological Bulletin, 113*, 553-565.

Scragg, P., Grey, N., Lee, D., Young, K., & Turner, S. (2001). A brief report on the Penn Inventory for Posttraumatic Stress Disorder. *Journal of Traumatic Stress, 14*, 605-611.

Seigel, B., Vukicevic, J., & Spitzer, R. L. (1990). Using signal detection methodology to revise DSM-III-R: Re-analysis of the DSM-III-R national field trials for autistic disorder. *Journal of Psychiatric Research, 24*, 294-311.

Siegel, B., Vukicevic, J., Elliott, G. R., & Kraemer, H. C. (1989). The use of signal detection theory to assess DSM-III—R criteria for autistic disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 28*, 542-548.

Simms, L. J., Watson, D., & Doebbeling, B. N. (2002). Confirmatory factor analyses of posttraumatic stress symptoms in deployed and nondeployed veterans of the Gulf War. *Journal of Abnormal Psychology, 111*, 637–647.

Sondergaard, H. P., Ekblad, S., & Theorell, T. (2003). A longitudinal study of hormonal reactions accompanying life events in recently resettled refugees. *Psychotherapy & Psychosomatics, 72*, 49-58.

Swets, J. A., Tanner, W. P., & Birdsall, G. (1961). Decision processes in perception. *Psychological Review, 68(5)*, 301-340.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science, 240*, 1285-1293.

Swets, J. A. (1998). Enhancing diagnostic decisions. In R.R. Hoffman & M.F. Sherrick (Eds), *Viewing psychology as a whole: The integrative science of William N. Dember* (pp. 559-577). Washington, DC: American Psychological Association.

Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review, 61*, 401-409.

Taylor, S., Kuch, K., Koch, W. J., Crockett, D. J., & Passey, G. (1998). The structure of posttraumatic stress symptoms. *Journal of Abnormal Psychology, 107*, 154–160.

Treisman, M. & Faulkner, A. (1984). The effect of signal probability on the slope of the receiver operating characteristic given by the rating procedure. *British Journal of Mathematical & Statistical Psychology, 37*, 199-215.

Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., & Pleck, J. H. (1998). Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science, 280*(5365), 867-873.

Ventureyra, V. A. G., Yao, S., Cottraux, J., Note, I., & Mey-Guillard, C. D. (2002). The validation of the Posttraumatic Stress Disorder Checklist Scale in Posttraumatic Stress Disorder and nonclinical subjects. *Psychotherapy Psychosomatics, 71*, 47-53.

Vispoel, W. P. (2000). Computerized versus paper-and-pencil assessment of self-concept: Score comparability and respondent preferences. *Measurement & Evaluation in Counseling & Development, 33*(3), 130-143.

Vispoel, W. P., Boo, J., & Bleiler, T. (2001). Computerized and paper-and-pencil versions of the Rosenberg Self-Esteem Scale: A comparison of psychometric features and respondent preferences. *Educational & Psychological Measurement, 61*(3), 461-474.

Walker, E. A., Newman, E., Dobie, D. J., Ciechanowski, P, & Katon, W. (2002). Validation of the PTSD Checklist in an HMO sample of women. *General Hospital Psychiatry, 24*, 375-380.

Weathers, F. W., Huska, J., & Keane, T. M. (1991). The PTSD Checklist – Military Version (PCL-M). Boston, MA: National Center for PTSD.

Weathers, F. W., Keane, T. M., & Davidson, J. R. T. (2001). Clinician-Administered PTSD Scale: A review of the first ten years of research. *Depression & Anxiety, 13*, 132-156.

Weathers, F. W., Keane, T. M., King, L. D., & King, D. W. (2001). Psychometric theory in development of PTSD assessment tools. In J.P. Wilson & T.M. Keane (Eds), *Assessing psychological trauma and PTSD* (pp. 98-135). New York: Guilford Press.

Weathers, F. W., Litz, B. T., Keane, T. M., Herman, D. S., Steinberg, H. R., Huska, J. A., & Kraemer, H. C. (1996). The utility of the SCl-90-R for the diagnosis of war-zone related Posttraumatic Stress Disorder. *Journal of Traumatic Stress, 9*, 111-128.

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

Weber, B., Schneider, B., Fritze, J., Gille, B., Hornung, S., Kuehner, T., & Maurer, K. (2003). Acceptance of computerized compared to paper-and-pencil assessment in psychiatric inpatients. *Computers in Human Behavior, 19*(1), 81-93.

Youngstrom, E. A., Findling, R. L., Calabrese, J. R., Gracious, B. L., Demeter, C., Bedoya, D. D., & Price, M. (2004). Comparing the diagnostic accuracy of six potential screening instruments for Bipolar Disorder in youths aged 5 to 17 years. *Journal of the American Academy of Child & Adolescent Psychiatry, 43*, 847-858.

Youngstrom, E. A., Findling, R. L., Danielson, C. K., & Calabrese, J. R. (2001). Discriminative validity of parent report of hypomanic and depressive symptoms on the General Behavior Inventory. *Psychological Assessment, 13*, 267-276.

APPENDICES

APPENDIX A: Screening Packet Cover Page

The purpose of this research is to examine the properties of a computerized interview for posttraumatic stress disorder (PTSD). Posttraumatic stress symptoms sometimes occur after individuals have been exposed to an especially stressful event. The purpose of this screening packet is to identify potential research participants for our study. Please complete the information below and the attached questionnaires. The demographic information will not influence your participation in the study and will be detached from your responses.

In you qualify for participation in the study, one of our lab research assistants will call you and set up an appointment time for you to come into the clinic or hospital. At that point, you will complete the computerized interview and a face-to-face interview conducted by one of the research staff. The expected total participation time for all aspects of the research is between 60-90 minutes.

Any participant who completes the entire study will receive a $10 gift card to a local retail store. Your participation or responses will not in any way impact your diagnosis or compensation for psychological or psychiatric treatment at this time or in the future.

Thank you for your assistance!

Your First Name	
Your Last Name	
Your Home Phone Number (for appointment scheduling purposes)	
Your Cell Phone Number (for appointment scheduling purposes)	
Best way to reach you. May we leave a message at this number?	
Best times to call you to set up an appointment?	
Best appointment times for you? (we will work around your schedule)	

To keep your information confidential, this page will be detached from the rest of your screening data after scores on the screening instruments or participation is complete.

Appendix B: Risk of Harm Assessment

Risk of Harm Assessment

1) "**In the past week have you had thoughts about harming yourself?"** Yes No
If No Skip to #2

"**Can you tell me today that you will not harm yourself**?" Yes No

If no, then the veteran is not eligible for the study and should be connected with a triage clinician (Dr. Rauch or scheduled provider) immediately for follow-up of suicidal risk.

If Yes, let the veteran know that should he/she feel that they are a risk to harm themselves, they should either come immediately to urgent care at the VA, dial 911, or go to the nearest emergency room.

2) "**In the past week have you had thoughts about harming others?"** Yes No
If No Skip to #2

"**Can you tell me today that you will not harm others**?" Yes No

If no, then the veteran is not eligible for the study and should be connected with a triage clinician (Dr. Rauch or scheduled provider) immediately for follow-up of risk to harm others.

If Yes, let the veteran know that should he/she feel that they are a risk to harm others, they should either come immediately to urgent care at eth VA, dial 911, or go to the nearest emergency room.

APPENDIX C: Life Events Checklist

Listed below are a number of difficult or stressful things that sometimes happen to people. For each event check one or more of the boxes to the right to indicate that: (a) it happened to you personally, (b) you witnessed it happen to someone else, (c) you learned about it happening to someone close to you, (d) you're not sure if it fits, or (e) it doesn't apply to you. Be sure to consider your entire life (growing up as well as adulthood) as you go through the list of events.

Event	Happened to me	Witnessed it	Learned about it	Not Sure	Doesn't apply
1. Natural disaster (for example, flood, hurricane, tornado, earthquake)					
2. Fire or explosion					
3. Transportation accident (for example, car accident, boat accident, train wreck, plane crash)					
4. Serious accident at work, home, or during recreational activity					
5. Exposure to toxic substance (for example, dangerous chemicals, radiation)					
6. Physical assault (for example, being attacked, hit, slapped, kicked, beaten up)					
7. Assault with at weapon (for example, being shot, stabbed, threatened with a knife, gun, bomb)					
8. Sexual assault (rape, attempted rape, made to perform any type of sexual act through force or threat of harm)					
9. Other unwanted or uncomfortable sexual experience					
10. Combat or exposure to a war-zone (in the military or as a civilian)					
11. Captivity (for example, being kidnapped, abducted, held hostage, prisoner of war)					
12. Life-threatening illness or injury					
13. Severe human suffering					
14. Sudden, violent death (for example, homicide, suicide)					
15. Sudden, unexpected death of someone close to you					
16. Serious injury, harm, or death you caused to someone else					
17. Any other very stressful event or experience					

APPENDIX D: PTSD Checklist

Instructions: Below is a list of problems and complaints that people sometimes have in response to stressful life experiences. Please read each one carefully, then circle one of the numbers to the right to indicate how much you have been bothered by that problem in the past month.

The event you experienced was _____ on _____.
(event) (date)

	Not at all	A little bit	Moderately	Quite a bit	Extremely
1. Repeated, disturbing memories, thoughts, or images of the stressful experience?	1	2	3	4	5
2. Repeated, disturbing dreams of the stressful experience?	1	2	3	4	5
3. Suddenly acting or feeling as if the stressful experience were happening again (as if you were reliving it)?	1	2	3	4	5
4. Feeling very upset when something reminded you of the stressful experience?	1	2	3	4	5
5. Having physical reactions (e.g., heart pounding, trouble breathing, sweating) when something reminded you of the stressful experience?	1	2	3	4	5
6. Avoiding thinking about or talking about the stressful experience or having feelings related to it?	1	2	3	4	5
7. Avoiding activities or situations because they reminded you of the stressful experience?	1	2	3	4	5
8. Trouble remembering important parts of the stressful experience?	1	2	3	4	5
9. Loss of interest in activities that you used to enjoy?	1	2	3	4	5
10. Feeling distant or cut off from other people?	1	2	3	4	5
11. Feeling emotionally numb or unable to have loving feelings for those close to you?	1	2	3	4	5
12. Feeling as if your future somehow will be cut short?	1	2	3	4	5
13. Trouble falling or staying asleep?	1	2	3	4	5
14. Feeling irritable or having angry outbursts?	1	2	3	4	5
15. Having difficulty concentrating?	1	2	3	4	5
16. Being “superalert” or watchful or on guard?	1	2	3	4	5
17. Feeling jumpy or easily startled?	1	2	3	4	5

Appendix E: Study Exclusion Form

Inclusion/Exclusion Summary Form

Research ID _____

Needs "Yes" to one of the following:

YES **NO** Referred for PCT evaluation

YES **NO** Experienced significant trauma on LEC

Needs YES to all of the following:

YES **NO** Fits in a PTSD Severity group that is not closed to recruitment (n = 70)

YES **NO** Veteran able to hear the computer administration.

YES **NO** Veteran speaks English

YES **NO** Veteran not currently reporting extreme distress or significant suicidal or homicidal intent.

YES **NO** Veteran does not report a history or presence of thought disorder

Appendix F: Informed Consent Form

Department of Veterans Affairs VA Research Consent Form

Subject Name: Date:
Title of Study: Psychometric Properties of the Computerized PTSD Scale – Multimedia Version(CPS-M) Among Veterans
Principal Investigator: Sheila Rauch, PhD VAMC: VA Ann Arbor Healthcare System

PURPOSE OF RESEARCH STUDY:
The purpose of the study is to develop a computerized Posttraumatic Stress Disorder assessment instrument. In order to conduct this investigation, we need to determine the relationship between responses given to a computerized questionnaire, a face-to-face interview, and other written questions. Your involvement will be for one session that lasts about 120 to 180 minutes.

DESCRIPTION:
You have been found eligible to participate in the study based on the screening you have completed. Up to 210 male/female veterans who are eligible will participate in the study. Veterans will be assigned to groups based on the severity of their symptoms. Seventy veterans in each of 3 symptom groups (e.g., mild /no symptoms, moderate symptoms, and severe symptoms) will be enrolled. Veterans will be eligible on a first come basis until the groups are filled (70 patients for each group).

During your participation in the study, you will sit in front of a computer for a computerized assessment, fill out some paper-and-pencil forms, and complete a face-to-face interview. The order may vary; meaning, some people will complete the computer segment first and others will complete the paper forms and interview first. For the computer segment, you will answer questions using a computer mouse. This software has sound files, so most questions will be read to you by the computer. This usually takes about 30 minutes and the computer will let you know when it is finished. The other segment involves completing paper-and-pencil forms and an interview. This usually takes about 60-90 minutes. If any of the language in these forms is confusing, please ask the research assistant for help. In each of these sections, you will be asked about questions regarding past traumatic events and your reactions to them. Some of the paper-and-pencil forms ask other questions about depression and anxiety.

RISKS:
Some people find it unpleasant to fill out the surveys or report upsetting memories. However, this is a standard part of the assessment of traumatic events and PTSD. Some questions may remind you of painful memories and cause some emotional discomfort. There may be other risks that are unforeseeable at this time.

If you become distressed at any time during the interview or other assessments, you may pause or discontinue participation in the study. Additionally, the study personnel conducting the session may work with you to reduce negative reactions. If needed, he/she will contact the principle investigator or other PCT clinicians in order to assist with your care. Referral to psychiatry triage may be made as determined necessary.

The magnitude of harm if there is loss of confidentiality potentially includes social damage to relationships with friends and peers, and secondly, damage to business relationships that may decrease economic gains. In order to protect against breach of confidentiality, all policies regarding training of research study staff and research data management will be followed. All research data will be housed and secured at the VA to ensure confidentiality and later destroyed by Dr. Rauch. Funding for this study is provided through Eastern Michigan University. Your name and social security number are required to be maintained and may be disclosed to research staff at Eastern Michigan University for the purpose of reporting payment.

BENEFITS:
You are not likely to directly benefit by participating in this study. Your participation will assist in the development of a new assessment tool for the improvement of treatment for other people who have suffered from traumatic events.

ALTERNATE COURSES OF ACTION:
You do not have to participate in this study. You may drop out at any time without penalty or loss of benefits entitled to you. If you consent to participate in this research study, you may stop and leave at any time with no penalty to you. Your participation is strictly voluntary. Your responses will not affect your eligibility for clinical care at the VA Ann Arbor Healthcare System and cannot be used for service connection. The results will not be entered into your medical record except in the instance of reported danger to yourself or others (see below).

If participating in this study does bother you, you can stop and leave at any time without any impact on your care at the VAAAHCS. You may also choose to take a

break or discuss your feelings with study staff. If you are distressed, study staff may ask that you meet briefly with a VA clinician face-to-face.

STATEMENT OF RESEARCH RESULTS:
Your identifying information (e.g., name) will be removed from the file in order to protect your privacy. Your data will be assigned a research ID number. The research data will be stored in a locked office and in a password protected computer at the VAAAHCS. Data will be encrypted to provide additional protection. This information will be destroyed after the all the data has been collected. To prevent any potential negative consequences to you, any information gathered during the study will not be included in your medical records unless you report risk of harm to self or others (see below).

If the research in this study is published in journals or presented at conferences, it will not be connected with your identifying information. As a participant, you are entitled to a summary of the results, and if desired, this may be obtained from Dr. Sheila Rauch at the VA PTSD Clinic (734-845-3545) or Dr. Dean Lauterbach at Eastern Michigan University (734-487-0785).

We will let you know of any important discoveries made during this study which may affect you, your condition, or your willingness to participate in this study. The study includes surveys which may elicit information concerning suicidal and homicidal intent, depression, or other major clinical findings. The research investigators will notify your primary mental health provider and/or your treating psychologist if you express these concerns. This contact will also be documented in your medical record.

SPECIAL CIRCUMSTANCES:
There will be no costs to you for any of the assessments done as part of this research study. You may withdraw from the study at any time. There are no consequences for discontinuing.

COMPENSATION:
After completion of the study session, you will receive a $10 gift card to a local department store. However, completion of the individual study session is required to receive the ten-dollar gift card.

REQUEST FOR PATIENT AUTHORIZATION FOR ACCESS TO PROTECTED HEALTH INFORMATION

1. By signing this document, you authorize the Veterans Health Administration (VHA) to provide Sheila Rauch, Ph.D. and the research team permission to access your Protected Health Information (medical chart data) for research purposes. This information may include the following: Hospital records and reports; admission history, and physical; X-ray films and reports; operative reports; laboratory reports; treatment and test results; dental notes; immunizations; allergy reports; prescriptions; consultations; clinic notes; and any other medical records needed by the research team. **The investigators may view restricted information about you including: HIV infection, Sickle Cell Anemia, drug and/ or alcohol abuse treatment.**

2. The research investigators will collect your Protected Health Information for the following specific reasons: to determine your appointment history in either the PCT or MHC clinics and to collect additional information about your mental health.

3. Your Protected Health Information, the research data and any identifying linkage will be stored in a secure location. Your data will be assigned a research number and will be encrypted to provide additional security. All data will be housed and stored within a locked office and in a password protected computer on the VA network at the VAAAHS.

4. You may refuse to sign this authorization and refuse to allow the disclosure of your Protected Health Information. Your refusal will not affect your ability to receive medical care or benefits at the VA Ann Arbor Healthcare System, but you will not be eligible to participate in the study.

5. This authorization expires at the end of the study.

6. This authorization may be revoked at any time by sending a written request to: Sheila Rauch, Ph.D., PTSD Clinical Team VA Ann Arbor Healthcare System 2215 Fuller Road (116c) Ann Arbor, MI 48105 If you revoke this authorization, Dr. Rauch and the research team can continue to use information about you that has been collected. No information will be collected after you revoke the authorization.

7. The Ann Arbor VAMC complies with the requirements of the Health Insurance Portability and Accountability Act of 1996 and its privacy regulations and all other

applicable laws that protect your privacy. We will protect your information according to these laws. Despite these protections, there is a possibility that your information could be used or disclosed in a way that it will no longer be protected.

8. Payments to research subjects are funded by a grant from Eastern Michigan University. Eastern Michigan University research team staff will access your name and social security number in order to process your payment.

RESEARCH SUBJECT'S RIGHTS:

_____ has explained this research study and answered all questions. The risks or discomforts and possible benefits of the study have been described. Other choices of available treatment have been explained. Some veterans are required to pay co-payments for medical care and services provided by the VA. These co-payment requirements will continue to apply for VA care and services that are not part of this research study.

Participation in this study is entirely voluntary. You may refuse to participate. Refusal to participate will involve no penalty or loss of rights to which individuals are entitled. Participants may withdraw from this study at any time without penalty or loss of VA or other benefits. In the event that you sustain an injury or illness as a result of your participation in this VA approved research study, all necessary medical treatment (except in limited circumstances), will be provided in a VA medical facility. You will be treated for the injury at no cost to you. However, no additional compensation has been set aside. You have not waived any legal rights or released the hospital or its agents from liability for negligence by signing this form.

In case there are medical problems or if you have questions, concerns or complaints about the research study, you can contact member(s) of the research team: Sheila Rauch, Ph.D. can be called at 734-845-3545 during the day and can be contacted after hours by paging (734) 651-9379. You may contact the VA IRB coordinator (at 734-845-3440) when staff members of the research study are not available to discuss questions or concerns with someone other than research study staff. Research subjects may learn more about research at the VA Ann Arbor Healthcare System at this website: www1.va.gov/aavaresearch

I am informed about my rights as a research subject, and I voluntarily consent to participate in this study. I will receive a signed copy of this consent form.

x_____ X_____
Signature of Subject Date

x_____
Signature of Witness (A witness must observe the subject's signature)

x_____ X_____
Witness (Print Name) Date

x_____
Signature of person obtaining consent (Study personnel must be approved by VA IRB.)

x_____ X_____
Study Personnel (Print Name) Date

Appendix G: Demographic Questionnaire

Participant Information

For Staff Use Only

ID: _____

LOCATION: _____

INTERVIEWER: _____

STATUS: IP OP SCR STU

FIRST ADMIN: _____ _____

SECOND AD: _____ _____

This information is completely confidential. The coding system that is used makes it impossible for the project research team to associate you with the information you will be providing. The informed consent that you completed will be removed from your folder when your participation is complete and kept in a separate location.

1. What are the last four digits of your social security number? ☐☐☐☐

2. What is your date of birth? ___/___/19___

3. How old are you? _____ years old

4. Sex? MALE FEMALE

5. Primary Ethnic Background (circle the appropriate code):

01 White, not Hispanic
02 Black, not Hispanic
03 Hispanic, White
04 Hispanic, Black
05 American Indian / Alaskan
06 Asian
07 Pacific Islander / Hawaiian
08 Other

6. Are you working at all now? YES NO

If yes, how many hours per week? _____

7. Circle the highest educational level that you have ***completed*** in school:

01 Grade School
02 Junior High School
03 Some High School
04 High School
05 Some College
06 4 Year College (e.g., B.A., B.S.)
07 Some Graduate work (e.g., master's degree)
08 Doctorate/Professional degree (e.g., M.D., Ph.D., J.D.)

Psychiatric History

8. Have you ever received professional treatment as an outpatient or inpatient for an emotional or substance use problem?

_____ NO _____ Outpatient

_____ Inpatient _____ Inpatient and outpatient

9. Approximate number of counseling/ therapy sessions in the last year? ☐

9a If you have been in counseling, please estimate the number of counseling sessions that have focused on traumatic event(s) that you have experienced . ☐

10. In the last thirty days, have you been taking a prescribed medication for a psychological or emotional problem? YES NO

If yes, which drugs?

_____

APPENDIX H: Purdue PTSD Scale - Revised

Purdue Scale

In **the last month**, how often . . .	not at all		sometimes		often
1. Were you bothered by memories or thoughts of the event when you didn't want to think about it?	○	○	○	○	○
2. Have you had upsetting dreams about the event?	○	○	○	○	○
3. Have you suddenly felt as if you were experiencing the event again?	○	○	○	○	○
4. Did you feel very upset when something happened to remind you of the event?	○	○	○	○	○
5. Did you avoid activities or situations that might remind you of the event?	○	○	○	○	○
6. Did you avoid thoughts or feelings about the event?	○	○	○	○	○
7. Did you have difficulty remembering important aspects of the event?	○	○	○	○	○
8. Did you react physically (heart racing, breaking out in a sweat) to things that reminded you of the event?	○	○	○	○	○

Since the event . . .	not at all		sometimes		often
9. Have you lost interest in one or more of your usual activities (e.g., work, hobbies, entertainment)?	○	○	○	○	○
10. Have you felt unusually distant or cut off from people?	○	○	○	○	○
11. Have you felt emotionally "numb" or unable to respond to things emotionally the way you used to?	○	○	○	○	○
12. Have you been less optimistic about your future?	○	○	○	○	○
13. Have you had more trouble sleeping?	○	○	○	○	○
14. Have you been more irritable or angry?	○	○	○	○	○
15. Have you had more trouble concentrating?	○	○	○	○	○
16. Have you found yourself watchful or on guard, even when there was no reason to be?	○	○	○	○	○
17. Are you more jumpy or easily startled by noises?	○	○	○	○	○

APPENDIX I: Trauma Related Dissociation Scale

TRDS

(Carlson & Waelde, 1999)

For each statement below, circle one of the choices to show how many times each thing has happened to you in the past week.

	NOT AT ALL	ONCE OR TWICE	3-6 TIMES	7-10 TIMES	MORE THAN 10 TIMES
			(IN THE PAST WEEK)		
1. My body felt strange or unreal.	0	1-2	3-6	7-10	10+
2. Things around me seemed strange or unreal.	0	1-2	3-6	7-10	10+
3. I got reminded of something upsetting and then spaced out for a while.	0	1-2	3-6	7-10	10+
4. I had moments when I lost control and acted like I was back in an upsetting time in my past.	0	1-2	3-6	7-10	10+
5. I noticed that I couldn't remember the details of something upsetting that happened to me.	0	1-2	3-6	7-10	10+
6. Familiar places seemed strange or unreal.	0	1-2	3-6	7-10	10+
7. I felt like I was outside myself, watching myself do things.	0	1-2	3-6	7-10	10+
8. I heard something that I know really wasn't there.	0	1-2	3-6	7-10	10+
9. I got upset about something and can't remember what happened next.	0	1-2	3-6	7-10	10+
10. I felt like I was in a movie - like nothing that was happening was real.	0	1-2	3-6	7-10	10+
11. I didn't feel pain when I was hurt and should have felt something.	0	1-2	3-6	7-10	10+
12. A memory came back to me that was so strong that I lost track of what was going on around me.	0	1-2	3-6	7-10	10+
13. I found myself staring into space and thinking of nothing.	0	1-2	3-6	7-10	10+
14. I couldn't remember things that had happened during the day even when I tried to.	0	1-2	3-6	7-10	10+
15. I felt like I wasn't myself.	0	1-2	3-6	7-10	10+
16. I felt like I was in a daze and couldn't make sense of what was going on around me.	0	1-2	3-6	7-10	10+
17. I saw something that seemed real, but was not.	0	1-2	3-6	7-10	10+
18. I suddenly realized that I hadn't been paying attention to what was going on around me.	0	1-2	3-6	7-10	10+
19. I felt cut off from what was going on around me.	0	1-2	3-6	7-10	10+
20. Parts of my body seemed distorted - like they were bigger or smaller than usual.	0	1-2	3-6	7-10	10+
21. I reacted to people or situations as if I were back in an upsetting time in my past.	0	1-2	3-6	7-10	10+
22. I got so focused on something going on in my mind that I lost track of what was happening around me.	0	1-2	3-6	7-10	10+
23. I noticed there were gaps in my memory for things that happened to me that I should be able to remember.	0	1-2	3-6	7-10	10+
24. I smelled something that I know really wasn't there.	0	1-2	3-6	7-10	10+

APPENDIX J: Beck Depression Inventory-II

BDI-II

This questionnaire consists of 21 groups of statements. After reading each group of statements carefully, circle the number(0,1,2 or 3) next to the one statement in each group which **best** describes the way you have been feeling the **past week, including today.** If several statements within a group seem to apply equally well, circle each one. **Be sure to read all the statements in each group before making your choice.**

1. 0 I do not feel sad.
   1 I feel sad.
   2 I am sad all the time and I can't snap out of it.
   3 I am so sad or unhappy that I can't stand it.

2. 0 I am not particularly discouraged about the future.
   1 I feel discouraged about the future.
   2 I have nothing to look forward to.
   3 I feel I am a complete failure as a person.

3. 0 I do not feel like a failure.
   1 I feel I have failed more the average person.
   2 I As I look back on my life, all I can see is a lot of failures.
   3 I feel I am a complete failure as a person.

4. 0 I get as much satisfaction out of things as I used to.
   1 I don't enjoy things the way I used to.
   2 I don't get real satisfaction out of anything anymore
   3 I am dissatisfied or bored with everything

5. 0 I don't feel particularly guilty.
   1 I feel guilty a good part of the time.
   2 I feel quite guilty most of the time.
   3 I feel guilty all of the time.

6. 0 I don't feel I am being punished.
   1 I feel I may be punished.
   2 I expect to be punished.
   3 I feel I am being punished.

7. 0 I don't feel disappointed in myself.
   1 I am disappointed in myself.
   2 I am disgusted with myself.
   3 I hate myself.

8. 0 I don't feel I am any worse than anybody else.
1 I am critical of myself for my weaknesses or mistakes.
2 I blame myself all the time for my faults.
3 I blame myself for everything bad that happens.

9. 0 I don't have any thoughts of killing myself.
1 I have thoughts of killing myself, but I would not carry them out.
2 I would like to kill myself.
3 I would like to kill myself if I had the chance.

10. 0 I don't cry any more than usual.
1 I cry more now than I used to.
2 I cry all the time now.
3 I used to be able to cry, but now I can't cry even though I want to.

11. 0 I am no more irritated now than I ever am.
1 I get annoyed or irritated more easily than I used to.
2 I feel irritated all the time now.
3 I don't get irritated at all by the things that used to irritate me.

12. 0 I have not lost interest in other people.
1 I am less interested in other people than I used to be.
2 I have lost most of my interest in other people.
3 I have lost all of my interest in other people.

13. 0 I Make decisions about as well as I ever could.
1 I put off making decisions more than I used to.
2 I have greater difficulty in making decisions than before.
3 I can't make decisions at all anymore.

14. 0 I Don't feel I look any worse than I used to.
1 I am worried that I am looking old or unattractive.
2 I feel that there are permanent changes in my appearance
that make me look unattractive.
3 I believe that I look ugly.

15. 0 I can work about as well as before.
1 It takes an extra effort to get started at doing something.
2 I have to push myself very hard to do anything.
3 I can't do any work at all.

16. 0 I can sleep as well as usual.
1 I don't sleep as well as I used to.
2 I wake up 1-2 hours earlier than usual and find it hard to get back to sleep.
3 I wake up several hours earlier than I used to and cannot
get back to sleep.

17. 0 I don't get tired any more than usual.
1 I get tired more easily than I used to.
2 I get tired from doing almost anything.
3 I am too tired to do anything.

18. 0 My appetite is no worse than usual.
1 My appetite is not as good as it used to be.
2 My appetite is much worse now.
3 I have no appetite at all anymore.

19. 0 I haven't lost much weight, if any, lately.
1 I have lost more than 5 pounds.
2 I have lost more than 10 pounds.
3 I have lost more than 15 pounds.

I am purposely trying to lose weight by eating less.
YES_____ NO_____

20. 0 I am no more worried about my health than usual.
1 I am worried about physical problems such as aches and pains; or upset stomach; or constipation.
2 I am very worried about physical problems and it's hard to think of much else.
3 I am so worried about my physical problems that I cannot think about anything else.

21 0 I have not noticed any recent change in my interest in sex.
1 I am less interested in sex than I used to be.
2 I am much less interested in sex now.
3 I have lost interest in sex completely.

APPENDIX K: Clinician-Administered PTSD Scale

National Center for PTSD

CLINICIAN-ADMINISTERED PTSD SCALE FOR DSM-IV

(CAPS-DX)

Research # _____ Date: _____

Interviewer: _____ Study: CBT for Acute Stress Disorder_

Session: Posttest 3-Month Follow-up 6-Month Follow-up

Dudley D. Blake, Frank W. Weathers, Linda M. Nagy,
Danny G. Kaloupek, Dennis S. Charney, & Terence M. Keane

National Center for Posttraumatic Stress Disorder

Behavioral Science Division – Boston VA Medical Center
Neurosciences Division – West Haven VA Medical Center

Revised January 1997

Criterion A. The person has been exposed to a traumatic event 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
(2) the person's response involved intense fear, helplessness, or horror. Note: In children, this may be expressed instead by disorganized or agitated behavior

I'm going to be asking you about some difficult or stressful things that sometimes happen to people. Some examples of this are being in some type of serious accident; being in a fire, a hurricane, or an earthquake; being mugged or beaten up or attacked with a weapon; or being forced to have sex when you didn't want to. I'll start by asking you to look over a list of experiences like this and check any that apply to you. Then, if any of them do apply to you, I'll ask you to briefly describe what happened and how you felt at the time.

Some of these experiences may be hard to remember or may bring back uncomfortable memories or feelings. People often find that talking about them can be helpful, but it's up to you to decide how much you want to tell me. As we go along, if you find yourself becoming upset, let me know and we can slow down and talk about it. Also, if you have any questions or you don't understand something, please let me know. Do you have any questions before we start?

ADMINISTER CHECKLIST, THEN REVIEW AND INQUIRE UP TO THREE EVENTS. IF MORE THAN THREE EVENTS ENDORSED, DETERMINE WHICH THREE EVENTS TO INQUIRE (E.G., FIRST, WORST, AND MOST RECENT EVENTS; THREE WORST EVENTS; TRAUMA OF INTEREST PLUS TWO OTHER WORST EVENTS, ETC.)

NO EVENTS ENDORSED ON CHECKLIST: (Has there ever been a time when your life was in danger or you were seriously injured or harmed?)

IF NO: (What about a time when you were threatened with death or serious injury, even if you weren't actually injured or harmed?)

IF NO: (What about witnessing something like this happen to someone else or find out that it happened to someone close to you?)

IF NO: (What would you say are some of the most stressful experiences you have had over your life?)

EVENT #1

What happened? (How old were you? Who else was involved? How many times did this happen? Life threat? Serious injury? **How did you respond**	Describe (e.g., event type, victim, perpetrator, age, frequency): A.(1) Life threat? NO YES [self____other___]

emotionally? (Were you very anxious or frightened? Horrified? Helpless? How so? Were you stunned or in shock so that you didn't feel anything at all? What was that like? What did other people notice about your emotional response? What about after the event--how did you respond emotionally?)	Serious injury? NO YES [self____other___] Threat to physical integrity? NO YES [self____other___] A.(2) Intense fear/help/horror? [during ____ after ____] Criterion A met? NO PROBABLE YES

EVENT #2

What happened? (How old were you? Who else was involved? How many times did this happen? Life threat? Serious injury? **How did you respond emotionally?** (Were you very anxious or frightened? Horrified? Helpless? How so? Were you stunned or in shock so that you didn't feel anything at all? What was that like? What did other people notice about your emotional response? What about after the event--how did you respond emotionally?)	Describe (e.g., event type, victim, perpetrator, age, frequency): A.(1) Life threat? NO YES [self____other___] Serious injury? NO YES [self____other___] Threat to physical integrity? NO YES [self____other___] A.(2) Intense fear/help/horror? [during ____ after ____] Criterion A met? NO PROBABLE YES

EVENT #3

What happened? (How old were you? Who else was involved? How many times did this happen? Life threat? Serious injury? **How did you respond emotionally?** (Were you very anxious or frightened? Horrified? Helpless? How so? Were you stunned or in shock so that you didn't feel anything at all? What was that like? What did other people notice about your emotional response? What about after the	Describe (e.g., event type, victim, perpetrator, age, frequency): A.(1) Life threat? NO YES [self____other___] Serious injury? NO YES [self____other___] Threat to physical integrity? NO YES [self____other___] A.(2) Intense fear/help/horror? [during ____

event--how did you respond emotionally?)	after ____] Criterion A met? NO PROBABLE YES

For the rest of the interview, I want you to keep (EVENTS) in mind as I ask you some questions about how they may have affected you.

I'm going to ask you about twenty-two questions altogether. Most of them have two parts. First, I'll ask if you've ever had a particular problem, and if so, about how often in the past month. Then I'll ask you how much distress or discomfort that problem may have caused you.

CRITERION b. The traumatic event is persistently reexperienced in one (or more) of the following ways:

1. **(B-1)** recurrent and intrusive distressing recollection of the event, including images, thoughts, or perceptions. **Note**: In young children, repetitive play may occur in which themes or aspects of the trauma are expressed.

Frequency	**Intensity**	**Past month**
In the past month have you had unwanted memories of (EVENT)? What were they like? (What did you remember?) [IF NOT CLEAR:] (Did they ever occur while you were awake, or only in dreams?) [EXCLUDE IF MEMORIES OCCURRED ONLY DURING DREAMS] **How often?** 0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day **Description/Examples**	**How much distress or discomfort did these memories cause you? Were you able to put them out of your mind and think about something else?** (How hard did you have to try?) **How much did they interfere with your life?** 0. None 1. Mild, minimal distress or disruption of activities 2. Moderate, distress clearly present but still manageable, some disruption of activities 3. Severe, considerable distress, difficulty dismissing memories, marked disruption of activities 4. Extreme, incapacitating distress, cannot dismiss memories, unable to continue activities **QV (specify)**_____	**F** ____ **I** ____ **Sx: Y N**

2. (B-2) recurrent distressing dreams of the event. **Note:** In children, there may be frightening dreams without recognizable content.

Frequency	**Intensity**	**Past month**
In the past month have you had unpleasant dreams about (EVENT)? Describe a typical dream. (What happens in them?) **How often?** 0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day **Description/Examples**	**How much distress or discomfort did these dreams cause you? Did they ever wake you up? [IF YES:]** What happened when you woke up? How long did it take you to get back to sleep?) **[LISTEN FOR REPORT OF ANXIOUS AROUSAL, YELLING, ACTING OUT THE NIGHTMARE]** (Did your dreams ever affect anyone else? How so?) 0. None 1. Mild, minimal distress, may not have awoken 2. Moderate, awoke in distress but readily returned to sleep 3. Severe, considerable distress, difficulty returning to sleep 4. Extreme, incapacitating distress, did not return to sleep **QV (specify)**_____	**F** ____ **I** ____ **Sx: Y N**

3. (B-3) acting or feeling as if the traumatic event were recurring (includes a sense of reliving the experience, illusions, hallucinations, and dissociative flashback episodes, including those that occur on awakening or when intoxicated).

Note: In young children, trauma-specific reenactment may occur.

Frequency	**Intensity**	**Past month**
In the past month have you suddenly acted or felt as if (EVENT) were happening again (Have you ever had flashbacks about [EVENT]?) [IF NOT CLEAR:] (Did this ever occur while you were awake, or only in dreams?) [EXCLUDE IF OCCURRED ONLY DURING DREAMS] **Tell me more about that. How often?**	**How much did it seem as if (EVENT) were happening again?** (Were you confused about where you actually were or what you were doing at the time?) **How long did it last? What did you do while this was happening?** (Did other people notice your behavior? What did they say?) 0. None 1. Mild, somewhat more realistic than	**F** ____ **I** ____ **Sx: Y N**

0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day **Description/Examples**	just thinking about event 2. Moderate, definite but transient dissociative quality, still very aware of surroundings, daydreaming quality 3. Severe, strongly dissociative (reports images, sounds, or smells) but retained some awareness of surroundings 4. Extreme, complete dissociation (flashback), no awareness of surroundings, may be unresponsive, possible amnesia for the episode (blackout) **QV (specify)**_____	

4. **(B-4)** intense psychological distress at exposure to internal or external cues that symbolize or resemble an aspect of the traumatic event

Frequency	**Intensity**	**Past month**
In the past month have you gotten emotionally upset when something reminded you of (EVENT)? (Has anything ever triggered bad feelings related to [EVENT]?) **What kinds of reminders made you upset? How often?** 0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day **Description/Examples**	**How much distress or discomfort did (REMINDERS) cause you? How long did it last? How much did it interfere with your life?** 0. None 1. Mild, minimal distress or disruption of activities 2. Moderate, distress clearly present but still manageable, some disruption of activities 3. Severe, considerable distress, difficulty dismissing memories, marked disruption of activities 4. Extreme, incapacitating distress, unable to continue activities **QV (specify)**_____	**F** ____ **I** ____ **Sx: Y N**

5. (B-5) physiological reactivity on exposure to internal or external cues that symbolize or resemble an aspect of the traumatic event

Frequency	**Intensity**	**Past month**
In the past month have you had any physical reactions when something reminded you of (EVENT)? (Did your body ever react in some way when something reminded you of {EVENT}? **Can you give me some examples?** (Did your heart race or your breathing change? What about sweating or feeling really tense or shaky?) **What kinds of reminders triggered these reactions? How often?** 0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day **Description/Examples**	**How strong were (PHYSICAL REACTIONS)? How long did they last?** (Did they last even after you were out of the situation?) 0. No physical reactivity 1. Mild, minimal reactivity 2. Moderate, physical reactivity clearly present, may be sustained if exposure continues 3. Severe, marked physical reactivity, sustained throughout exposure 4. Extreme, dramatic physical reactivity, sustained arousal even after exposure has ended **QV (specify)**_____	**F** ____ **I** ____ **Sx: Y N**

Criterion C. Persistent avoidance of stimuli associated with the trauma and numbing of general responsiveness (not present before the trauma), as indicated by three (or more) of the following:

6. (C-1) efforts to avoid thoughts, feelings or conversations associated with trauma

Frequency	**Intensity**	**Past month**
In the past month have you tried to avoid thoughts or feelings about (EVENT)? (What kinds of thoughts or feelings did you try to avoid?) **What about trying to avoid talking with other people about it?** (Why is that?) **How often?** 0. Never	How much effort did you make to avoid (THOUGHTS/FEELINGS/CONVERSATIONS)? (What kinds of things did you do? What about drinking or using medication or street drugs?) [CONSIDER ALL ATTEMPTS AT AVOIDANCE, INCLUDING DISTRACTION, SUPPRESSION, AND USE OF ALCOHOL/DRUGS] **How much did that**	**F** ____ **I** ____ **Sx: Y N**

1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day **Description/Examples**	**interfere with your life?** 0. None 1. Mild, minimal effort, little or no disruption of activities 2. Moderate, some effort, avoidance definitely present, some disruption of activities 3. Severe, considerable effort, marked avoidance, marked disruption of activities, or involvement in certain activities as avoidance strategy 4. Extreme, drastic attempts at avoidance, unable to continue activities, or excessive involvement in certain activities as avoidance strategy. **QV** **(specify)**_____ ___	

7. (C-2) efforts to avoid activities, places, or people that arouse recollections of the trauma

Frequency	**Intensity**	**Past month**
In the past month have you tried to avoid certain activities, places, or people that reminded you of (EVENT)? (What kinds of things did you avoid? Why is that?) **How often?** 0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day **Description/Examples**	**How much effort did you make to avoid (ACTIVITIES/PLACES/PEOPLE)?** (What did you do instead?) **How much did that interfere with your life?** 0. None 1. Mild, minimal effort, little or no disruption of activities 2. Moderate, some effort, avoidance definitely present, some disruption of activities 3. Severe, considerable effort, marked avoidance, marked disruption of activities, or involvement in certain activities as avoidance strategy 4. Extreme, drastic attempts at avoidance, unable to continue activities, or excessive involvement in certain activities as avoidance strategy *QV*	**F** ____ **I** ____ **Sx: Y N**

	*(specify)*_____ _	

8. (C-3) inability to recall an important aspect of the trauma

Frequency	**Intensity**	**Past month**
In the past month have you had difficulty remembering some important parts of (EVENT)? Tell me more about that. (Do you feel you should be able to remember these things? Why do you think you can't?) **How much of the important parts of (EVENT) have you had difficulty remembering?** (What parts do you still remember?) 0. None, clear memory 1. Few aspects not remembered (less than 10%) 2. Some aspects not remembered (approx. 20-30%) 3. Many aspects not remembered (approx. 50-60%) 4. Most or all aspects not remembered (more than 80%) **Description/Examples**	**How much effort difficulty did you have recalling important parts of (EVENT)** (Were you able to recall more if you tried?) 0. None 1. Mild, minimal difficulty 2. Moderate, some difficulty, could recall with effort 3. Severe, considerable difficulty, even with effort 4. Extreme, completely unable to recall important aspects of event *QV* *(specify)*_____ _	**F** ____ **I** ____ **Sx: Y N**

9. (C-4) markedly diminished interest or participation in significant activities

Frequency	**Intensity**	**Past month**
In the past month have you been less interested in activities that you used to enjoy? **(What kinds of things have you lost interest in? Are there some things you don't do at all anymore? Why is that?) [EXCLUDE IF NO OPPORTUNITY, IF PHYSICALLY UNABLE, OR IF DEVELOPMENTALLY APPROPRIATE CHANGE IN PREFERRED ACTIVITIES]** How many activities have you been less interested in? **(What kinds of things do you still enjoy doing?)** When did you first start to feel that way? **(After the [EVENT]?)** 0. None 1. Few activities (less than 10%) 2. Some activities (approx. 20-30%) 3. Many activities (approx. 50-60%) 4. Most or all activities (more than 80%) **Description/Examples**	**How strong was your loss of interest?** (Would you enjoy [ACTIVITIES] once you got started?) 0. No loss of interest 1. Mild, slight loss of interest, probably would enjoy after starting activities 2. Moderate, definite loss of interest, but still has some enjoyment of activities 3. Severe, marked loss of interest in activities 4. Extreme, complete loss of interest, no longer participates in any activities *QV* *(specify)*_____ __ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	**F** ____ **I** ____ **Sx: Y N**

10. (C-5) feeling of detachment or estrangement from others

Frequency	**Intensity**	**Past month**
In the past month have you felt distant or cut off from other people? What was that like? How much of the time? When did you first start to feel that way? **(After the [EVENT]?)** 0. None 1. Very little of the time (less than 10%)	**How strong were your feelings of being distant or cut off from others?** (Whom do you feel closest to? How many people do you feel comfortable talking with about personal things?) 0. No feelings of detachment or estrangement 1. Mild, may feel "out of synch" with others 2. Moderate, feelings of detachment	**F** ____ **I** ____ **Sx: Y N**

2. Some of the time (approx. 20-30%) 3. Much of the time (approx. 50-60%) 4. Most or all of the time (more than 80%) **<u>Description/Examples</u>**	clearly present, but still feels some interpersonal connection 3. Severe, marked feelings of detachment or estrangement from most people, may feel close to only one or two people 4. Extreme, feels completely detached or estranged from others, not close with anyone *QV (specify)*_____ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	

11. (C-6) restricted range of affect (e.g., unable to have loving feelings)

Frequency	**Intensity**	**Past month**
In the past month have there been times when you felt emotionally numb or had trouble experiencing feelings like love or happiness? What was that like? **(What feelings did you have trouble experiencing?)** How much of the time? When did you first start having trouble experiencing (EMOTIONS)? **(After the [EVENT]?)** 0. None of the time 1. Very little of the time (less than 10%) 2. Some of the time (approx. 20-30%) 3. Much of the time (approx. 50-60%) 4. Most or all of the time (more than 80%) **Description/Examples**	**How much trouble did you have experiencing (EMOTIONS)? (**What kinds of feelings are you still able to experience?) [INCLUDE OBSERVATIONS OF RANGE OF AFFECT DURING INTERVEW**]** 0. No reduction of emotional experience 1. Mild, slight reduction of emotional experience 2. Moderate, definite reduction of emotional experience, but still able to experience most emotions 3. Severe, marked reduction of experience of at least two primary emotions (e.g., love, happiness) 4. Extreme, completely lacking emotional experience *QV* *(specify)*_____ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	**F** ____ **I** ____ **Sx: Y N**

12. (C-7) sense of a foreshortened future (e.g. does not expect to have a career, marriage, children, or a normal life

Frequency	**Intensity**	**Past month**
In the past month have there been times when you felt there is no need to plan for the future, that somehow your future will be cut short? Why is that? **[RULE OUT REALISTIC RISKS SUCH AS LIFE-THREATENING MEDICAL CONDITONS**] How much of the time? When did you first start to gel that way? **(After the [EVENT]?)**	**How strong was this feeling that your future will be cut short? (**How long do you think you will live? How convinced are you that you will die prematurely?) 0. No sense of a foreshortened future 1. Mild, slight sense of a foreshortened future 2. Moderate sense of a foreshortened future but no specific prediction	**F** ____ **I** ____ **Sx: Y N**

0. None of the time 1. Very little of the time (less than 10%) 2. Some of the time (approx. 20-30%) 3. Much of the time (approx. 50-60%) 4. Most or all of the time (more than 80%) **Description/Examples**	about longevity 3. Severe, marked sense of a foreshortened future, may make specific prediction about longevity 4. Extreme, overwhelming sense of a foreshortened future, completely convinced of premature death *QV* *(specify)*_____ _ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	
Criterion D. Persistent symptoms of increased arousal (not present before the trauma), as indicated by two (or more) of the following:		

13. (D-1) difficulty falling or staying asleep

Frequency	**Intensity**	**Past month**
In the past month have you had any problems falling or staying asleep? How often? When did you first start having problems sleeping? **(After the [EVENT]?)** 0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day Sleep onset problems? Y N Mid-sleep awakening? Y N Early a.m. awakening? Y N	**How much of a problem did you have with your sleep?** (How long did it take for you to fall asleep? How often did you wake up in the night? Did you often wake up earlier than you wanted to? How many total hours did you sleep each night**?** 0. No sleep problems 1. Mild, slightly longer latency, or minimal difficulty staying asleep (up to 30 minutes loss of sleep) 2. Moderate, definite sleep disturbance, clearly longer latency, or clear difficulty staying asleep (30-90 minutes loss of sleep) 3. Severe, much longer latency, or marked difficulty staying asleep (90 min to 3 hrs loss of sleep) 4. Extreme, very long latency, or profound difficulty staying asleep (> 3 hrs loss of sleep)	**F** ____ **I** ____ **Sx: Y N**

Total # hrs sleep/night _____ Desired # hrs sleep/night _____	*QV* *(specify)*_____ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	

14. (D-2) irritability or outbursts of anger

Frequency	**Intensity**	**Past month**
In the past month have there been times when you felt especially irritable or showed strong feelings of anger? Can you give me some examples? How often? When did you first start feeling that way? **(After the [EVENT]?)** 0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day *Description/Examples*	**How strong was your anger?** (How did you show it?) [IF REPORTS SUPPRESSION:] (How hard was it for you to keep from showing your anger?) **How long did it take you to calm down? Did your anger cause you any problems?** 0. No irritability or anger 1. Mild, minimal irritability, may raise voice when angry 2. Moderate, definite irritability or attempts to suppress anger, but can recover quickly 3. Severe, marked irritability or marked attempts to suppress anger, may become verbally or physically aggressive when angry 4. Extreme, pervasive anger or drastic attempts to suppress anger, may have episodes of physical violence *QV (specify)*_____ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	**F** ____ **I** ____ **Sx: Y N**

15. (D-3) difficulty concentrating

Frequency	**Intensity**	**Past month**
In the past month have you found it difficult to concentrate on what you were doing or on things going on around you? What was that like?	**How difficult was it for you to concentrate?** [INCLUDE OBSERVATIONS OF CONCENTRATION AND ATTENTION IN INTERVIEW]? **How**	**F** ____ **I** ____

How much of the time? When did you first start having trouble concentrating? **(After the [EVENT]?)** 0. None of the time 1. Very little of the time (less than 10%) 2. Some of the time (approx. 20-30%) 3. Much of the time (approx. 50-60%) 4. Most or all of the time (more than 80%) *Description/Examples*	**much did that interfere with your life?** 0. No difficulty with concentration 1. Mild, only slight effort needed to concentrate, little or no disruption of activities 2. Moderate, definite loss of concentration but could concentrate with effort, some disruption of activities 3. Severe, marked loss of concentration even with effort, marked disruption of activities 4. Extreme, complete inability to concentrate, unable to engage in activities *QV* *(specify)*_____ __ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	**Sx: Y N**

16. (D-4) hypervigilance

Frequency	**Intensity**	**Past month**
In the past month have you been especially alert or watchful, even when there was no real need to be? (**Have you felt as if you were constantly on guard?)** What is that? How much of the time? When did you first start acting that way? **(After the [EVENT]?)** 0. None of the time 1. Very little of the time (less than 10%) 2. Some of the time (approx. 20-30%) 3. Much of the time (approx.	**How hard did you try to be watchful of things going on around you** [INCLUDE OBSERVATIONS OF HYPERVIGILANCE IN INTERVIEW]? **Did your (HYPERVIGILANCE) cause you any problems?** 0. No hypervigilance 1. Mild, minimal hypervigilance, slight heightening of awareness 2. Moderate, hypervigilance clearly present, watchful in public (e.g., chooses safe place to sit in a restaurant or movie theater) 3. Severe, marked hypervigilance, very	**F** ____ **I** ____ **Sx: Y N**

50-60%) 4. Most or all of the time (more than 80%) *Description/Examples*	alert, scans environment for danger, exaggerated concern for safety of self/family/home 4. Extreme, excessive hypervigilance, efforts to ensure safety consume significant time and energy and may involve extensive safety/checking behaviors, marked watchfulness during interview *QV* *(specify)*_____ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	

17. (D-5) exaggerated startle response

Frequency	**Intensity**	**Past month**
In the past month have you had any strong startle reactions? When did that happen? (**What kinds of things made you startle?)** How often? When did you first have these reactions? **(After the [EVENT]?)** 0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day *Description/Examples*	**How strong were these startle reactions?** (How strong were they compared to how most people would respond? **How long did they last?** 0. No startle reaction 1. Mild, minimal startle reaction 2. Moderate, definite startle reaction, feels "jumpy" 3. Severe, marked startle reaction, sustained arousal following initial reaction 4. Extreme, excessive startle reaction, overt coping behavior j(e.g., combat veteran who "hits the dirt") *QV* *(specify)*_____ ___ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	**F** ____ **I** ____ **Sx: Y N**

CRITERION E. Duration of the disturbance (symptoms in Criteria B, C, and D) is more than 1 month.

18. onset of symptoms

[IF NOT ALREADY CLEAR:] When did you first start having (PTSD SYMPTOMS) you've told me about? (How long after the trauma did they start? More than six months?)	_____ total # months delay in onset With delayed onset (≥ 6 months)? NO YES

19. duration of symptoms

How long have these (PTSD SYMPTOMS) lasted altogether?	Duration more than 1 month? Total # months duration	NO YES

	Acute (< 3 months) or chronic (≥ 3 months)? **Duration > 6 months?**	_____ acute chronic NO YES

CRITERION F. The disturbance causes clinically significant distress or impairment in social, occupational, or other important areas of functioning.

20. subjective distress

*Overall, in the past month how much have you been bothered by these (PTSD SYMPTOMS) you've told me about****? [CONSIDER DISTRESS REPORTED ON EARLIER ITEMS]***	0. None 1. Mild, minimal distress 2. Moderate distress clearly present but still manageable 3. Severe, considerable distress 4. Extreme, incapacitating distress	**Past month** _____

21. impairment in social functioning

*In the past month have these (PTSD SYMPTOMS) affected your relationships with other people? **[CONSIDER IMPAIRMENT IN SOCIAL FUNCTIONING REPORTED ON EARLIER ITEMS]***	0. No adverse impact 1. Mild impact, minimal impairment in social functioning 2. Moderate impact, definite impairment, but many aspects of social functioning still intact 3. Severe impact, marked impairment, few aspects of social functioning still intact 4. Extreme, impact, little or no social functioning	**Past month** _____

22. impairment in occupational or other important area of functioning

[IF NOT ALREADY CLEAR] *Are you working now?* IF YES: **Have these (PTSD SYMPTOMS) affected your work or your ability to work in the past month? How so?** [CONSIDER REPORTED WORK HISTORY, INCLUDING NUMBER AND DURATION OF JBOS, AS WELL AS THE QUALITY OF WORK RELATIONSHIPS. IF PREMORBID FUNCTIONING IS UNCLEAR, INQUIRE ABOUT WORK EXPERIENCES BEFORE THE TRAUMA. FOR CHILD/ADOLESCENT TRAUMAS, ASSESS PRE-TRAUMA SCHOOL PERFORMANCE AND POSSIBLE PRESENCE OF BEHAVIOR PROBLEMS] IF NO: In **the past month, have these (PTSD SYMPTOMS) affected any other important part of your life?** [AS APPROPRIATE, SUGGEST EXAMPLES SUCH AS PARENTING, HOUSEWORK, SCHOOLWORK, VOLUNTEER WORK, ETC.] **How so?**	0. No adverse impact 1. Mild impact, minimal impairment in occupational/other important functioning 2. Moderate impact, definite impairment, but many aspects of occupational/other important functioning still intact 3. Severe impact, marked impairment, few aspects of occupational/other important functioning still intact 4. Extreme impact, little or no occupational/other important functioning

GLOBAL RATINGS

23. global validity

ESTIMATE THE OVERALL VALIDITY OF RESPONSES. CONSIDER SUCH ISSUES AS COMPLAINCE WITH THE INTERVIEW, PATIENT STATUS (E.G., PROBLEMS WITH CONCENTRATION, COMPREHENSION OF ITEMS, DISSOCIATION), AND EVIDENCE OF EFFORTS TO EXAGGERATE OR MINIMIZE SYMPTOMS	0. Excellent, no reason to suspect invalid responses 1. Good, factors present that may adversely affect validity 2. Fair, factors present that definitely reduce validity 3. Poor, substantially reduced validity 4. Invalid responses, severely impaired mental status or possible deliberate "faking bad" or "faking good"

24. global severity

ESTIMATE THE OVERALL SEVERITY OF PTSD SYMPTOMS. CONSIDER DEGREE OF SUBJECTIVE DISTRESS, DEGREE OF FUNCITONAL IMPARIMENT, OBSERVATIONS OF BEHAVIORS IN INTERVIEW, AND JUDGMENT REGARDING REPORTING STYLE.	0. No clinically significant symptoms, no distress and no functional impairment 1. Mild, minimal distress or functional impairment 2. Moderate, definite distress or functional impairment but functions satisfactorily with effort 3. Severe, considerable distress or functional impairment, limited functioning even with effort 4. Extreme, marked distress or marked impairment in two or more major areas of functioning

25. global improvement

RATE TOTAL OVERALL IMPROVEMENT PRESENT SINCE THE INITIAL RATING. IF NO EARLIER RATING, ASK HOW THE SYMPTOMS ENDORESED HAVE CHANGED OVER THE PAST 6 MONTHS. RATE THE DEGREE OF CHANGE, WHETHER OR NOT, IN YOUR JUDGEMENT, IT IS DUE TO TREATMENT	0. Asymptomatic 1. Considerable improvement 2. Moderate improvement 3. Slight improvement 4. No improvement 5. Insufficient information

Associated Features

26. guilt over acts of commission or omission

Frequency	**Intensity**	**Past month**
In the past month have you felt guilty about anything you did or didn't do during (EVENT)? Tell me more about that. (What do you feel guilty about?) **How much of the time?** 0. None of the time 1. Very little of the time (less than 10%) 2. Some of the time (approx. 20-30%) 3. Much of the time (approx. 50-60%)	**How strong were these feelings of guilt? How much distress or discomfort did they cause?** 0. No feelings of guilt 1. Mild, slight feelings of guilt 2. Moderate, guilt feelings definitely present, some distress but still manageable 3. Severe, marked feelings of guilt, considerable distress 4. Extreme, pervasive feelings of guilt, self-condemnation regarding behavior, incapacitating distress	**F** ____ **I** ____

4. Most or all of the time (more than 80%) *Description/Examples*	**QV (specify)**_____ _____	

27. survivor guilt [APPLICABLE ONLY IF MULTIPLE VICTIMS]

Frequency	**Intensity**	**Past month**
In the past month have you felt guilty about SURVIVING (EVENT) WHEN OTHERS DID NOT? Tell me more about that. (What do you feel guilty about?) **How much of the time?** 0. None of the time 1. Very little of the time (less than 10%) 2. Some of the time (approx. 20-30%) 3. Much of the time (approx. 50-60%) 4. Most or all of the time (more than 80%) *Description/Examples*	**How strong were these feelings of guilt? How much distress or discomfort did they cause?** 0. No feelings of guilt 1. Mild, slight feelings of guilt 2. Moderate, guilt feelings definitely present, some distress but still manageable 3. Severe, marked feelings of guilt, considerable distress 4. Extreme, pervasive feelings of guilt, self-condemnation regarding behavior, incapacitating distress **QV (specify)**_____ _____	**F** ____ **I** ____

28. a reduction in awareness of his or her surrounding (e.g., "being in a daze)

Frequency	**Intensity**	**Past month**
In the past month have there been times when you felt out of touch with things going on around you, like you were in a daze? What was that like? **[DISTINGUISH FROM FLASHBACK EPISODERS]** How often? **[IF NOT CLEAR:] (Was it due to an illness or the effects of drugs or alcohol?)** When did you first start feeling that way? **(After the [EVENT]?)** 0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day *Description/Examples*	**How strong was this feeling of being out of touch or in a daze?** (Were you confused about where you actually were or what you were doing at the time?)? **How long did it last? What did you do while this was happening?** (Did other people notice your behavior? What did they say?) 0. No reduction in awareness 1. Mild, slight reduction in awareness 2. Moderate, definite but transient reduction in awareness, may report feeling "spacey" 3. Severe, marked reduction in awareness, may persist for several hours 4. Extreme, complete reduction in awareness of surroundings, may be unresponsive, possible amnesia for the episode (blackout) *QV* *(specify)*_____ __ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	**F** ____ **I** ____

29. derealization

Frequency	**Intensity**	**Past month**
In the past month have there been times when things going on around you seemed unreal or very strange or unfamiliar? **[IF NO:] (What about times when people you knew suddenly seemed unfamiliar?)** What was that like? How often? **[IF NOT CLEAR:] (Was it due to an illness or the**	**How strong was (DEREALIZATION)? How long did it last? What did you do while this was happening?** (Did other people notice your behavior? What did they say?)	**F** ____ **I** ____

effects of drugs or alcohol?) When did you first start feeling that way? **(After the [EVENT]?)** 0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day *Description/Examples*	0. No derealization 1. Mild, slight derealization 2. Moderate, definite but transient derealization 3. Severe, considerable derealization, marked confusion about what is real, may persist for several hours 4. Extreme, profound derealization, dramatic loss of sense of reality or familiarity *QV* *(specify)*_____ __ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	

30. depersonalization

Frequency	**Intensity**	**Past month**
In the past month have there been times when you felt as if you were outside of your body, watching yourself as if you were another person? **[IF NO:] (What about times when your body felt strange or unfamiliar to you, as if it had changed in some way?)** What was that like? How often? **[IF NOT CLEAR:] (Was it due to an illness or the effects of drugs or alcohol?)** When did you first start feeling that way? **(After the [EVENT]?)** 0. Never 1. Once or twice 2. Once or twice per week 3. Several times per week 4. Daily or almost every day	**How strong was (DEPERSONALIZATION)? How long did it last? What did you do while this was happening?** (Did other people notice your behavior? What did they say?) 0. No depersonalization 1. Mild, slight depersonalization 2. Moderate, definite but transient depersonalization 3. Severe, considerable depersonalization, marked of detachment from self, may persist for several hours 4. Extreme, profound depersonalization, dramatic loss of sense of detachment from self	**F** ____ **I** ____

Description/Examples	*QV* *(specify)*_____ __ **Trauma-related?** 1 definite 2 probable 3 unlikely Current _____ Lifetime _____	

APPENDIX L: Debriefing Form

Debriefing Form

The purpose of this research was to compare how well different formats of PTSD instruments agree with one another for diagnostic purposes. You were selected for the research because you indicated experiencing some level of stress in response to an event in your life. People who were asked to continue their participation varied widely in the nature of their responses – some acknowledged severe symptoms while others were almost asymptomatic.

Please keep in mind that all information collected during this research project is confidential. Your identifying information (e.g.-name) will be removed from the file in order to protect your privacy. Your data will be assigned a research ID number based on how many participants have already completed the study. The research data will be stored in a locked office and in a password protected computer at the VAAAHCS. Data will be encrypted to provide additional protection. To prevent any potential negative consequences to you, and information gathered during the study will **not** be included in your medical records unless you report risk of harm to self or others. Data will be retained for 7 years after the last publication from the data set. Patient identifiers connected to research ID numbers will be included in a file also secured at the VA that is stored in a locked cabinet separate from the rest of the study data and destroyed at the same interval as the study data.

Sometimes discussing stressful events can be distressing and cause a person to remember troubling events. Persons often become tearful or upset when responding to questions like those that you answered today. If you are feeling upset, please tell the interviewer. There is no rush to leave, if you need a few minutes to regain your composure, please stay until you feel better.

If you find that you continue to have difficulty managing your emotions after you leave this session or believe you may be a danger to yourself or others, professional help is available to you.

Veterans should contact their primary provider at the VA. Veterans can also access triage services at the Mental Health Clinic. The phone number is 734-213-6998. If you need help when this center is closed, please contact 911 emergency services for mental health assistance.

Above all, please feel free to contact Dr. Sheila Rauch at (734) 769-7100 x6040 or Dr. Dean Lauterbach at (734) 487-0785 if you are having any difficulties as a result of this study.

While we do not expect many individuals to develop symptoms that warrant further care, you should be aware that there are many treatment options available to you and that it is not unusual to feel down for a while after discussing a traumatic event.