Published in final edited form as:

J Anxiety Disord. 2009 January ; 23(1): 27–37. doi:10.1016/j.janxdis.2008.03.012.

Evaluating Symptom Expression as a Function of a Posttraumatic Stress Disorder Severity

Kathleen M. Palm¹, David R. Strong, and Laura MacPherson

Brown Medical School/Butler Hospital

Abstract

Little is known about the relative severity or typical sequence of Diagnostic and Statistical Manual (DSM-IV) symptoms of posttraumatic stress disorder (PTSD). Using data from the National Comorbidity Study – Replication (Kessler et al., 2004), the current study used a logistic item response model to assess the degree to which DSM-IV symptoms combine to define a primary construct underlying PTSD, to identify which symptoms are associated with greater severity of PTSD, and to determine whether the symptoms and symptom patterns are influenced by gender. Results suggested that PTSD symptoms can be combined to assess a single dimension of PTSD severity, providing support for a continuum of symptom severity. However, several DSM-IV symptoms provided overlapping information, potentially reducing the effectiveness of these symptoms in describing a broad range of PTSD. More precise assessment of PTSD severity may help improve the descriptive value of PTSD measures relationship to continuous measures of treatment outcomes, and ultimately inform more effective treatments.

Introduction

Few diagnostic categories have generated as much controversy as posttraumatic stress disorder (PTSD; Spitzer, First, & Wakefield, 2007). Indeed, debate as to the exact diagnostic criteria and central assumptions of PTSD have led to several significant revisions since it was originally included in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM; American Psychiatric Association [APA], 1980), with the list of symptoms increasing from 12 to 17 in the DSM-III-R (APA, 1987) and the clustering of symptoms changing over time. The varying lifetime prevalence estimates of PTSD found across studies and populations have also fueled debate regarding how to reliably characterize and assess PTSD [e.g., 15% of Vietnam veterans in one study (CDC, 1988) versus 30.9% in another (Kulka et al., 1990)]. Some researchers suggest that the manner in which PTSD symptoms are assessed (e.g., phrasing of questions; context demands; assessment method) may contribute to disagreements regarding prevalence estimates (e.g., Coyne & Thompson, 2007; McNally, 2007). As the next revision of the DSM is formulated, it is important that researchers continue to explore current operationalizations used to elicit the diagnostic classification of PTSD and evaluate the reliability and validity of current PTSD assessments.

The current DSM-IV-TR (APA, 2000) includes an operationalization of PTSD that organizes symptoms into three domains (i.e., re-experiencing, avoidance and numbing, hyperarousal).

¹Corresponding author: Kathleen M. Palm, Department of Psychology, Clark University, 950 Main St., Worcester, MA 01610. Email: kpalm@clarku.edu. Phone: (508) 793-7277. Fax: (508) 793-7265.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Absence of symptoms from one domain precludes attributing a high enough level of symptoms for assigning a diagnosis. The DSM-IV-TR (APA, 2000) does not assume each of the three symptom domains contributes equally to the PTSD syndrome, as the criteria domains differ both in number of potential symptoms and in threshold for determining a sufficient number of symptoms to decide if the criteria is met. Thus, among trauma-exposed individuals, a diagnosis of PTSD includes endorsement of at least one of five re-experiencing symptoms, three of seven avoidance and numbing symptoms, and two of five hyperarousal symptoms for more than one month, plus significant distress/impairment. If trauma-exposed individuals report anything less than the above symptoms, then they do not meet diagnostic criteria for PTSD.

Several researchers have questioned whether PTSD symptoms are best characterized using a binary classification of those with and without PTSD, or whether individuals above or below diagnostic threshold differ primarily in the degree of severity of resulting symptoms (Broman-Fulks et al., 2006; Forbes, Haslam, Williams, & Creamer, 2005; Gudmundsdottir & Beck, 2004; Ruscio, Ruscio, & Keane, 2002). However, little work has examined how the relative severity of individual PTSD symptoms relates to overall levels of PTSD either within or across symptom domains. Further, little is known about the relative severity of observed symptoms within a broad non-treatment seeking population. Identifying the kinds of symptoms that would begin to be observed near established diagnostic thresholds would increase the field's understanding of typical patterns of symptom responses and the level of phenotypic variability among individuals with PTSD. By using a single organizing dimension of PTSD, we can observe the co-occurrence of symptom reports among individuals with and without a PTSD diagnosis. This helps us to understand the kinds of symptoms observed and the degree of variability among individuals below diagnostic thresholds. This information can inform models of PTSD by suggesting symptom patterns that characterize levels of PTSD severity.

Factor Structure of PTSD

Attempts to validate the DSM-IV organization of symptoms within B, C, and D domains have been conducted using exploratory and confirmatory factor analytic methods. However, numerous factor analytic studies conducted across different trauma populations (e.g., male combat veterans, female sexual assault survivors, etc.) and assessment instruments [e.g., Clinically Administered PTSD Scale (Blake et al., 1995), Posttraumatic Stress Disorder Checklist (Weathers, Litz, Huska, & Keane, 1994), Mississippi PTSD Scale (Keane, Caddell, & Taylor, 1988)] have found differing solutions ranging from two (Buckley, Blanchard, & Hickling, 1998; Foa, Riggs, & Gershuny, 1995; Taylor, Kuch, Koch, Crockett, & Passey, 1998), three (Cordova, Studts, Hann, Jacobsen, & Andrykowski, 2000; Cox, Clara, & Enns, 2002; Thatcher & Krikorian, 2005), and four factors (Andrews, Joseph, Shevlin, & Troop, 2006; Asmundson et al., 2000; King, King, Fairbank, Keane, & Adams, 1998; McWilliams, Cox, & Asmundson, 2005; Palmieri & Fitzgerald, 2005; Simms, Watson, & Doebbeling, 2002). We would expect that if the DSM-IV organization of PTSD symptoms were robust, the organization of three domains would be supported across populations and different operationalizations of symptom indices. However, existing studies have not converged to support the DSM-IV based organization. These differing solutions may be a product of the varied assessment instruments or differential symptom endorsement across various trauma populations. Alternatively, symptoms within each domain might be linked more reliably by first understanding their relationship to the primary underlying construct of PTSD.

Although evidence suggests a full continuum of PTSD symptom severity and symptom counts provide an easy and intuitive index of PTSD severity, it remains unclear whether reliable individual differences can be detected using a count of DSM-IV based PTSD symptoms. It also is not clear how thresholds for determining the presence or absence of criteria relate to levels of the syndrome of PTSD (e.g., which symptoms are most likely to be the ones observed

among those above or below criteria thresholds?). Spitzer et al. (2007) proposed examining how PTSD symptoms specifically relate to the diagnosis of PTSD and that diagnostic thresholds should be generated empirically. However, to date no attempts have been made to link the likelihood of observing individual DSM-IV symptoms to an overall level of PTSD severity.

Item Response Analyses of PTSD

Given the potential for individual variability along a full continuum of severity, it is important to understand the relative severity of individual PTSD symptoms and which particular PTSD symptoms are likely to differentiate those above and below thresholds for a PTSD diagnosis. Methods based in item response theory (IRT; cf. Lord, 1980; Rasch, 1960) can greatly complement existing factor analytic studies by illustrating how individual responses change across different levels of PTSD severity. Further, item response models offer methods to evaluate the representativeness of specific items and the degree to which items define a common dimension.

To date, there have been two separate investigations using item response analyses to examine item functioning in assessments of PTSD. Conrad and colleagues (2004) examined the Mississippi PTSD scale using both true score theory and Rasch item response analyses in 803 male veteran inpatients. Conrad et al. found that when a single dimension of PTSD severity was examined, avoidance symptoms tended to characterize lower severity while reexperiencing symptoms were associated with greater PTSD severity. Betemps and colleagues (Betemps, Smith, Baker, & Rounds-Kugler, 2003) examined the performance of the CAPS (Blake et al., 1995), which was designed to approximate the 17 PTSD DSM-IV symptom criteria in a sample of 201 treatment-seeking veterans. They found that several items (e.g., “inability to recall important aspects of the trauma,” “acting or feeling as if the traumatic event were recurring”) did not discriminate well and thus provided little information about individual levels of PTSD. Further, it was observed that the CAPS items illustrated a ceiling effect such that selected items did not differentiate more severe levels of PTSD from moderate levels. Researchers using IRT procedures to examine the assessment of PTSD among veterans have produced initial evidence about the relative severity of PTSD symptoms and identified potential unreliability in rank-ordering veterans who report moderate to severe levels of PTSD. To date, these methodologies have yet to be applied directly to DSM-based interviews assessing PTSD symptoms among community samples where trauma events other than combat-related trauma may be a primary focus.

In the current study, we use methods based in IRT to examine the relative severity of each DSM-IV PTSD symptom, as assessed in the National Comorbidity Study- Replication (NCS-R), and each symptom’s ability to discriminate across a range of PTSD severity. The NCS-R has been used to assess estimates of mental health burden, to identify barriers to service, and to evaluate intervention targets on a national scale (Kessler et al., 2005; Kessler & Merikangas, 2004). The aim of the present study was to describe symptom expression as a function of PTSD severity and to determine whether the assessment of PTSD in the NCS-R provides reliable information about levels of PTSD severity beyond that provided by diagnostic classification. A secondary aim was to examine the kinds of symptoms likely to be observed near the threshold for a DSM-IV diagnosis of PTSD. Finally, given that several factor analytic studies of PTSD symptoms have focused exclusively on only men or women, we explored possible differential endorsement of PTSD symptoms in this sample across gender.

Method

Participants

The data included in the present analyses were from the NCS-R public-use set. As described elsewhere (Kessler et al., 2005), the NCS-R was a nationally representative face-to-face household survey conducted between February 2001 and April 2003. Questions regarding PTSD symptomology were included in Part II of the NCS-R interview, which assessed disorders that were considered to be secondary aims of the NCS-R. A total of 5,692 participants completed Part II of the assessment. The interview content was described to potential respondents, and verbal consent was obtained before conducting the interviews. A weighting procedure was used to adjust for differential probabilities of selection and non-response, and to adjust the sample to approximate the U.S. population on a number of sociodemographic characteristics (Kessler et al., 2004). The weighted prevalence estimate for lifetime PTSD in the NCS-R dataset was 6.8%. The NCS-R dataset includes sample weights to account for over-sampling of certain populations to ensure a representative sampling of the United States population based on the 2000 Decennial Census. These weights are not used in the current analyses, as our purpose was not to make generalizations about the US population but rather to examine the relationships of individual symptoms to an underlying continuum of PTSD and whether these differed across men and women.

In the PTSD module of the interview, respondents were asked about exposure to 27 different traumatic events (e.g., kidnapped, life-threatening automobile accident, rape). Participants who endorsed experiencing a traumatic event were asked about symptoms related to that event. If they reported more than one event, participants were asked about symptoms related to the worst event and an event that was randomly selected based on standardized procedures (WMH-CIDI; Kessler & Ustun, 2004). Interview questions used to operationalize PTSD symptoms are outlined in Table 1. A total of 4,985 endorsed at least one traumatic event and 1,946 met Criterion A (i.e., had been exposed to a traumatic event and their response involved intense fear, helplessness, and/or horror). Of the 1,946 who met Criterion A, 757 participants were excluded from the present analyses because these respondents did not endorse either of two items: “did you have any emotional problems after the event like upsetting memories or dreams, feeling emotionally distant or depressed, trouble sleeping or concentrating, or feeling jumpy or easily startled?” or “did any of these reactions last for 30 days or longer?” Thus, a total of 1,189 respondents was included in the present analyses. PTSD criteria requiring higher thresholds were assessed first; in other words, symptoms were assessed in the following order: criterion C, criterion B, and criterion D. The NCS-R interview used skip-outs after each criterion to determine whether the individual would still be eligible for a PTSD diagnosis. If an insufficient number of symptoms was endorsed to meet thresholds for any specific criterion, the subsequent criterion symptoms were not assessed and in this study, were presumed to be absent rather than missing. All 1,189 respondents received all criterion C symptom inquiries, 42 respondents did not receive any further symptom inquiries because they did not meet symptom thresholds for criterion C, and 65 respondents did not receive any criterion D symptoms because they did not meet symptom thresholds for Criterion B.

Of the 1,189 respondents, 604 (50.8%) met criteria for lifetime PTSD, 326 (27.4%) met criteria for PTSD in the past 12 months, and 160 (13.5%) met criteria in the past 30 days. The current sample had a mean age of 42.6 years ($SD=14.9$) and the majority was female (73.4%). Of the 1189 participants included in the present analyses, 227 (19.1%) reported seeing a psychiatrist, psychologist, social worker, counselor, or other mental health specialist in the past 12 months in the service utilization module of the NCS-R interview. In terms of lifetime treatment utilization, 738 (62.1%) indicated that they have ever participated in a counseling session (30 minutes or longer) with a professional, 617 (51.9%) indicated being on psychiatric medication

at some point, and 86 (7.2%) report ever being hospitalized for psychological or substance use problems.

Data analytic plan

Unidimensionality and local independence—Two assumptions of the item response models are 1) there is a primary underlying dimension of the latent construct (i.e., unidimensionality) and, 2) responses to a given symptom are independent from responses to other symptoms (i.e., local independence). Therefore, prior to evaluating individual item response probabilities, we evaluated whether we could assume that one primary latent construct was being measured and that responses to each symptom were largely independent of responses to other symptoms. Given the three symptom domains defined by the DSM-IV are used to index a general level of PTSD severity, we considered a single dimension underlying response to the 17 symptoms. However, given the 17 symptoms are placed *a priori* into three separate criterion domains, we also considered modeling each symptom domain as its own severity dimension (i.e., a three factor model). Finally, we evaluated a bifactor model (Holzinger & Swineford, 1937; Yung, Thissen, & McLeod, 1999). In the bifactor model, each symptom is allowed to load on the primary dimension of PTSD and to the dimension underlying its assigned criterion. Each criterion dimension is forced to be independent from the primary dimension of PTSD and from the two other criterion dimensions. Reise and Haviland (2005) suggest that the bifactor approach is a preferred method for examining whether a strong enough common factor underlies a complex construct such as PTSD rather than strictly comparing unidimensional versus multidimensional models using confirmatory factor analytic methods. When considering an IRT model, it is important to assess whether items continue to have reliable relationships after controlling for a shared relationship to a single common dimension of PTSD. If reliable relationships (e.g. strong loadings on the primary dimension) with the primary dimension remain after controlling for criterion related variability, a unidimensional IRT model for PTSD would be supported.

Fitting the unidimensional, three-factor, and bifactor model is useful in: a) guiding selection of an item response model (unidimensional vs. multidimensional) to characterize the relative severity of each PTSD symptom; b) determining whether consideration of the number of symptoms within each criterion increases reliability of assessments beyond a simple consideration of the total number of PTSD symptoms; and c) determining the impact of using a primary index of PTSD when multiple symptoms targeting a particular criterion are used (e.g. local dependence). The bifactor model has been used extensively in cognitive domain analyses (Gustafsson & Balke, 1993; Holzinger & Swineford, 1937) and increasingly has been useful in describing relationships among psychopathological and personality constructs (Patrick, Hicks, Nichol, & Krueger, 2007; Reise, Morizot, & Hays, 2007). We used confirmatory factor analysis of tetrachoric correlations to specify each of the three models in mPlus (Muthen & Muthen, 1998–2006). Robust Weighted Least Squares method was selected (Flora & Curran, 2004) and estimations allowed for the use of three fit indices in testing the fit of each of three models: the Comparative Fit Index (CFI: Bentler, 1990), the Tucker Lewis Index (TLI: Tucker & Lewis, 1973), and the root mean square error of approximation (RMSEA: Steiger, 1990). Cut-offs for model fit have been suggested to be $CFI \geq .96$, a $TLI \geq .95$, and the $RMSEA \leq .05$ (Yu, 2002). Each model was evaluated for improvements across each of these three indices. Given the present sample size and the sensitivity of the chi square statistics to sample size, chi squares difference tests were not used to compare models. We did observe univariate correlations between symptoms and overall PTSD severity and expected that the degree of correlation would mimic the factor loading and discrimination parameters estimated from multivariate model fitting.

Symptom-level estimates of problem severity—Item response models also evaluate patterns of symptom endorsement that result in each total score, and thereby provide an estimate of the likely ordering of symptom severity. This method allows an estimation of which symptom is likely to be endorsed at different levels of PTSD severity. Further, the two-parameter model can provide information about how well each item discriminates among individuals with PTSD symptoms. These estimates can be used to judge the effectiveness of each symptom and the amount of information about differences in PTSD severity contained in each symptom can be compared.

Gender influence on symptom endorsement—Use of an item response model allows us to compare independent estimates of the severity of symptoms across other variables (e.g., gender). The current comparison ensures that gender does not affect interpretation of total symptom counts. Estimation of differential item functioning (DIF) involves comparing analyses conducted separately within each group (Holland & Wainer, 1993). While there are many means to assess DIF, model-based assessment of DIF has the advantage of detecting DIF that arises from either differences in the severity of the symptom (threshold, or ‘b’ parameter) across women and men or in how the symptom discriminates among levels of PTSD (slope, or ‘a’ parameter) across women and men (Thissen, Steinberg, & Wainer, 1993). Model-based DIF utilize a likelihood-ratio test statistic to provide a significance test for the null hypothesis that the item parameters do not differ between two identified sub-groups (e.g. women and men). When evaluating each item, all other items are constrained to be the same across men and women, thus testing item differences while groups are anchored by the remaining items. We employed Version 2.0 of IRTLRDIF (Thissen, 2001) to complete DIF analyses.

Results

Unidimensionality and local independence

Results of the confirmatory models are presented in Table 2. These results suggest that fit indices from both the multidimensional model and the bifactor model were improved over the unidimensional model. Results from the three-factor model suggested substantial shared variability across the three factors, with an average correlation of .80 (SD = .06). Only the bifactor model produced fit indices that exceed suggested cutoffs across all three indices. When evaluating the need for a multidimensional model, we used the bifactor model to assess the relationship of each symptom with a primary dimension of PTSD after controlling for any relationships within each symptom domain. The symptoms’ relationship (e.g. factor loadings) to the primary dimension of PTSD across both models was quite similar, with the mean difference between loadings on a primary PTSD dimension across the two models being .014 (SD=.05). Thus, the bifactor model fits the data best, but controlling for criterion specific variability did not alter factor loadings on the primary dimension of PTSD substantially. Further, after controlling for loadings on the primary dimension, loadings within each criterion variable were reduced substantially relative to the multidimensional model. Our results suggest a unidimensional IRT model is unlikely to distort the item characteristics substantially and interpretation of item characteristics is unlikely to be biased due to these observed local dependencies. Although there was local dependence and a multidimensional model clearly fit the data better than a unidimensional model, after controlling for influences from multiple dimensions the items all clearly measured a strong common factor. Therefore, we selected an IRT model that captures variability along a single primary dimension of PTSD severity.

Symptom-level estimates of problem severity

In Table 3, order of the severity parameter estimates mirror the raw frequencies of each symptom, but the scale better reflects the relative magnitude of differences between symptoms of the continuum. Severity estimates ranged from the lowest level of PTSD severity of -1.28

(symptom C1) to the highest level of severity at .855 (symptom C7). Several of the PTSD symptoms were associated with similar levels of PTSD severity, suggesting an overlapping or clustering of symptoms on the continuum of PTSD severity. For example, symptoms B2, C6, and C4 had similar severity ratings between $-.232$ and $-.191$. Despite some overlap, the PTSD symptoms indexed a broad range of PTSD severity. The discrimination index provides information about how well each item discriminates among individuals with PTSD symptoms. Discrimination of an item corresponds to the slope and thus how rapidly the probability of endorsing an item changes across levels of the underlying latent PTSD construct, with higher discrimination values reflective of steeper slopes. The majority of items discriminated similarly, with symptoms C3 (.848) and B1 (2.390) suggesting the least and most discriminating items, respectively.

Figure 1 shows item characteristic curves (ICC) to visually illustrate differences in the severity and discrimination characteristics among the DSM-IV PTSD symptoms. Examination of these curves provides a visual representation of the relationship between the increased likelihood of a symptom (e.g., the y-axis) and individual level of PTSD symptom severity (e.g., the x-axis). A discrimination within levels of PTSD severity is made when the symptom becomes more likely to be present than not (i.e., when the ICC crosses .50). For example, item C1 was more likely to be endorsed at lower levels of PTSD severity and item C7 was more likely to be endorsed at higher levels of PTSD severity. As illustrated in Table 2, there was a relatively large gap between the first 15 symptoms with the highest severity of -0.047 and the two most severe symptoms with severities of 0.836 and 0.855 , suggesting that items in the current assessment have maximum reliability in capturing PTSD severity levels throughout the less severe end of the continuum.

Precision of the DSM-IV symptom index

Using the estimated measurement precision (i.e., information function from the model standard errors) of the DSM-IV PTSD symptoms, we can estimate the test information function (see Figure 2) to evaluate where on the continuum, the set of PTSD symptoms provide the most information or is most reliable in rank-ordering individuals. For reference, we calculated empirical Bayesian posteriori estimates for each level of PTSD among the 604 individuals with (PTSD +), and 585 individuals without a DSM-IV diagnosis (PTSD -). The average severity estimates for participants with PTSD was 0.55 ($SD = 0.65$) and -0.58 ($SD = 0.79$) for those without PTSD. Two vertical lines in Figure 2 position these group means with respect to the amount of information available from this set of DSM-IV symptoms. The 17 DSM-IV symptoms provide maximum information within a region of PTSD severity populated by those who do not meet the full DSM-IV diagnostic criteria. The amount of information decreases throughout the region populated by those with PTSD. Thus, the DSM-IV PTSD diagnostic classification rules place the diagnostic threshold beyond the reach of many symptoms. The lack of symptoms that characterize individuals above the diagnostic threshold may limit the ability to make further gradations within diagnosed groups. However, the PTSD symptoms do seem to capture variability within groups of individuals below a threshold for a diagnosis. For example, only 44 of the 585 participants reported no symptoms of PTSD and the median number of symptoms was 7.

Gender influence on symptom endorsement

Men and women indicated similar patterns of symptom endorsements with no significant DIF on 10 of the 17 symptoms ($p < .05$), although on average women had slightly higher PTSD severity than men (Cohen's $d = .25$). Women were more likely than men to report feeling distant emotionally ($b_{\text{men}} = -0.14$; $b_{\text{women}} = -0.45$; $DIF = .31$, $p < .05$) and feeling easily startled ($b_{\text{men}} = 0.18$; $b_{\text{women}} = -0.04$; $DIF = 0.22$, $p < .05$) at lower levels of PTSD. Conversely, men reported lack of plan for future ($b_{\text{men}} = 0.80$; $b_{\text{women}} = 1.20$; $DIF = -0.40$, $p < .05$), unwanted

memories ($b_{\text{men}} = -0.79$; $b_{\text{women}} = -0.56$; $\text{DIF} = -0.23, p < .05$), unpleasant dreams ($b_{\text{men}} = -0.22$; $b_{\text{women}} = 0.00$; $\text{DIF} = -0.22, p < .05$), and short-temper ($b_{\text{men}} = -0.03$; $b_{\text{women}} = 0.24$; $\text{DIF} = -0.23, p < .05$) at lower levels of PTSD than women. Finally, reports of experiencing flashbacks did not discriminate across levels of PTSD as well among men when compared to women ($a_{\text{men}} = 1.15$; $a_{\text{women}} = 1.74$; $\text{DIF} = .59, p < .05$).

Discussion

Findings

In the current study, we used an item response model to examine the symptoms that characterize a continuum of PTSD severity. Using data from the NCS-R (Kessler et al., 2005), we evaluated the degree to which PTSD symptoms were expressed as a function of a single primary continuum of PTSD severity and estimated the relative severity of the different symptoms. The PTSD symptoms sufficiently fit a unidimensional item response model, providing support for a continuum of symptom severity. Ordering of the PTSD severity estimates suggested that “trying not to think about the traumatic event” and “having an unwanted memory of the event” marked the lowest levels of severity. There was little difference in the level of PTSD severity between the point at which diagnostic criteria could first be met and endorsing up to seven additional symptoms. There were two symptoms, however, that marked the highest levels of PTSD severity: “no reason to plan for the future” and “inability to remember important parts of the traumatic event.” Interestingly, symptoms from the Criteria B, C, and D domains were represented across the severity continuum, suggesting symptoms from each of these domains are observed within both low and high levels of PTSD.

On average, women reported greater PTSD severity relative to men. After equating for level of PTSD severity, the likelihood of reporting specific symptoms was similar for men and women, with the exception that men were more likely to endorse avoidance (e.g., lack of plan for the future) and reexperiencing (e.g., unwanted memories and unpleasant dreams) symptoms at lower levels of PTSD severity relative to women. Further, women were more likely than men to report numbing (e.g., feeling emotionally distant) and hyperarousal (e.g., feeling easily startled) symptoms at lower levels of PTSD.

Although symptoms in the current study indexed a broad range of PTSD severity, several symptoms across Criteria B, C, and D mapped onto similar severity levels. The current findings suggest that items in the NCS-R interview capture a great deal of information, and potentially overlapping information, regarding low to moderate levels of PTSD severity. Items that are associated with similar severity levels provide limited incremental information about where a respondent falls along the continuum of PTSD severity given endorsement of one of those items. In contrast, there were few items that provided information regarding higher levels of PTSD severity. Similar to the findings of Betemps and colleagues (2003), who found a ceiling effect on the CAPS, the current findings may reflect a ceiling effect of the NCS-R interview. Thus, given the current assessment, it is difficult to differentiate between those who have moderate and severe PTSD symptomology, making it difficult to compare these individuals across other important dimensions such as co-occurring conditions, service utilization, and treatment efficacy. For example, the inability to reliably index the more severe end of the PTSD severity continuum could limit the ability to illustrate changes over time in intervention studies targeting severe PTSD populations. As also suggested by Betemps et al., future studies should seek to develop assessment items that reliably capture higher levels of PTSD severity.

Current findings differ from previous analyses of DSM-derived assessments of PTSD symptomology. As expected, there was a greater range of severity in the NCS-R community population compared to the veteran population examined by Betemps and colleagues (2003). Additionally, Betemps et al. found a somewhat different ordering of PTSD symptoms. In their

sample of male veterans with chronic PTSD, “feelings of detachment or estrangement” (C5) was associated with low levels of PTSD severity and “acting or feeling as if the traumatic event were recurring” (B3) was associated with more severe symptomology. In the current study, by indexing criteria C5 with ‘feeling emotionally distant,’ the severity of this symptom was reduced. The current study operationalization of criteria B3, with an emphasis on the report of ‘flashbacks,’ may also have reduced the severity of this symptom slightly.

Conrad and colleagues (2004) examined the item characteristics of the Mississippi PTSD Scale, which does not correspond directly with DSM-IV symptoms but examines some of the PTSD symptoms similarly. Consistent with findings of Conrad et al., we found that avoidance symptoms tended to characterize lower severity, while reexperiencing symptoms were associated with greater PTSD severity. In the sample of male veterans in Conrad et al.’s study, the highest severity item was “(l)ately, I have felt like killing myself,” which is similar to the highest severity item in the current study, “having no reason to plan for the future” (C7). However, in Conrad and colleagues’ analyses, the lowest severity items were, “I have a hard time expressing my feelings” and “(u)nexpected noises make me jump,” whereas “having trouble feeling normal feelings toward others” (C6) and feeling “jumpy or easily startled by ordinary noises” (D5) characterized mid-range levels of PTSD severity in the current study. It is possible that differences in phrasing for these items impacted the associated severity level. For example, in the NCS-R interview, symptom D5 was assessed with respect to “ordinary noises” compared to the Mississippi PTSD Scale which symptom D5 with respect to “unexpected noises.” In fact, it makes some intuitive sense that being startled by ordinary noises might be a more severe response than being startled by unexpected noises. Future research should evaluate range of severity estimates and the relative symptom severity in chronic versus acute populations to determine if these differences between studies are, in fact, due to population differences or other reasons such as the phrasing of items across instruments.

As anticipated, individuals meeting criteria for lifetime PTSD had relatively high severity estimates and endorsed more symptoms compared to respondents without a lifetime PTSD diagnosis. Although a diagnosis can be achieved with a minimum of six of the B, C, and D symptoms, those with lifetime PTSD generally scored in the highest ranges of a more continuous index of PTSD severity. Surprisingly, however, there was a relatively wide range of severity estimates for participants who failed to meet a full diagnosis of PTSD. There have been mixed findings regarding the clinical significance of sub-threshold PTSD (Breslau, Lucia, & Davis, 2004; Grubaugh et al., 2005; Schützwohl & Maercker, 1999; Stein, Walker, Hazen, & Forde, 1997). Current findings suggest there may be clinical utility in exploring the impact of PTSD symptoms among the subset of respondents who do not receive a full diagnosis. The ability to capture reliable variability even at low levels of PTSD and across symptom domains may suggest the possibility of using symptom counts from all domains simultaneously to index sub-threshold PTSD rather than requiring symptoms be expressed consistently across domains.

In the current study, women endorsed greater PTSD severity. Further, there were 7 symptoms that women and men endorsed differentially across levels of PTSD severity. Studies examining PTSD in the general population have consistently found higher rates of PTSD among women compared to men (Breslau, Davis, Andreski, Peterson, & Schultz, 1997; Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995; Stein, Walker, & Forde, 2000). Possible explanations for these gender differences have included differential rates of exposure to specific types of traumatic events (Kessler et al., 1995; Perkonig, Kessler, Storz, & Wittchen, 2000), differential responses to traumatic events (Breslau & Kessler, 2001), and biological differences (Yehuda, 1999). Additionally, differential symptom endorsement may be due to response bias (Gavranidou & Rosner, 2003; Peters, Issakidis, Slade, & Andrews, 2006). For example, Peters and colleagues (2006) suggested that men may be less likely than women to endorse particular symptoms (e.g., easily startled) that are interpreted as signs of weakness. Future studies should

continue to examine gender differences in symptom endorsement as these findings may have important implications for the assessment and treatment of PTSD.

Implications

Accumulating research, including the present study, suggests that a diagnostic system assessing the degree of PTSD symptomology (i.e., instead of a binary classification of presence/absence of diagnosis) may best capture variability among individuals exposed to a traumatic event using a full range of PTSD symptoms. Continuous indices of PTSD severity may be useful as supplements to extend research of the clinical correlates, associated co-morbidities, and clinical outcomes among individuals with a diagnosis of PTSD. Further, research examining the relationship between particular symptoms and how they relate to different levels of PTSD could inform the development of more precise assessments of PTSD symptomology and severity. For example, it appears that the DSM-IV PTSD symptoms as assessed in the NCS-R provides reasonable coverage at the low to moderate levels of the PTSD syndrome, however there may be a gap in symptoms that capture the range between of moderate to severe PTSD. Additionally, understanding where a patient falls on a continuum of PTSD severity could have important implications for appropriate treatment matching, such that someone with severe symptomology would receive more intensive treatment than someone endorsing symptoms in the less severe range of the severity continuum.

Given the cross sectional data and dichotomous rating for each symptom (present or absent), we could not examine whether changes in particular symptoms lead to improvement or worsening of other symptoms. The current IRT model does not imply a causal process, but an ordering process such that if a participant endorsed a higher severity symptom, it is expected that symptoms below that severity level would be endorsed. Examining symptom expression over time with a longitudinal IRT model would illustrate changes in the relative ordering of symptoms over time that could be examined in relation to changes in indices, such as whether or not a person continues to meet diagnostic criteria for PTSD at each time point. However, there is a significant possibility that severity of PTSD may be linked to influences beyond what can be explained by a single common factor. For example, there is a potential for a causal relationship among particular symptoms such that the development or exacerbation of one symptom may potentiate or inhibit others during the acute course of symptom development. In fact, some research suggests that symptoms characterizing increased arousal lead to the development of emotional numbing (Flack, Litz, Hsieh, Kaloupek, & Keane, 2000; Litz, Orsillo, Kaloupek, & Weathers, 2000; Taylor et al., 1998). Future research should prospectively examine PTSD symptom endorsement and the probability of observing specific symptom changes as a function of responses to other symptoms. In this way, researchers could illustrate how the probability of observing a specific symptom may increase dramatically if the respondent endorses a symptom of similar or greater severity.

Limitations

There are several limitations to the current study that merit discussion. First, current findings are limited by the nature of the assessment (i.e., assessment procedures, the phrasing of questions) used in the NCS-R. Notably, the NCS-R interview used skip-outs after each criterion to determine whether the next criterion would be assessed and, thus, these procedures may have influenced the structural findings regarding the dimensionality of PTSD. Although it is likely that individuals with low levels of PTSD may have been excluded from the present analyses due to the nature of the described events allowed in this survey, it is unlikely that these data would have changed the results since the majority of the symptoms seemed to assess a higher level of severity than these participants were likely to endorse. Further, we designated symptoms to be “absent” rather than “missing” if respondents were skipped-out of a section of the interview. Rather than assuming that responses were missing at random, we felt the

decision to presume an absence of the symptom was most in line with the assumptions of the survey method. This procedure may have decreased endorsement frequencies, as some of these respondents would have endorsed some of the symptoms if given the opportunity. Although the vast majority of respondents (85%) was administered all sections of the survey, the need to assign values is a limit of the study design.

Similarly, given that participants were required to endorse the screening item “did you have any emotional problems after the event like upsetting memories or dreams, feeling emotionally distant or depressed, trouble sleeping or concentrating, or feeling jumpy or easily startled?” prior to the assessment of Criteria B, C, and D symptomology, there was an increased likelihood of participants endorsing symptoms in the latter half of the assessment that were similar to the screening question (e.g., feeling emotionally distant or cut-off from people). Thus, the discriminability and relative severity indices of those particular items that were similar in content to the screening item may have been decreased as a function of the methodology used in the study rather than being a reflection of the item per se. Although these particular symptoms were represented across the continuum of severity and we did not observe a degree of local dependence that impacted our ability to estimate item characteristics, future research should examine item characteristics in assessments that are conducted in a manner such that the assessment of particular items is not dependent on previous responses.

There also may have been particular characteristics of the assessment items in the current study that impacted our findings. For example, phrasing of the questions in the NCS-R as well as population characteristics can impact the relative severity of each symptom dramatically. Each symptom of PTSD is an abstract construct and, thus, the relative ordering of these symptoms must be examined in the context of the current assessment. Thus, it is important that future research continue to examine the relative PTSD symptom severity across different questionnaires, interviews, and other means of assessment before making decisions about reliable assessment of the full PTSD syndrome. The current analyses are also limited by the lack of more detailed information regarding experiences related to the individuals’ traumatic event and associated symptoms. This cross-sectional study relied on lifetime retrospective reports and although an ordered pattern of symptom responses was observed, any attribution of the developmental progression of PTSD symptoms after experiencing a traumatic event requires longitudinal evaluation.

Finally, in the current study, we limited the analyses to examine the primary organization of DSM-IV PTSD symptoms along a latent unidimensional continuum. Since we were motivated primarily by an interest in organizing and understanding symptoms on a primary dimension of PTSD and the hierarchical-multidimensional model, in which a strong common factor was identified, was at least as reasonable as a multi-dimensional only model, we did not explore the relative fit of three versus four latent factors of PTSD. However, previous studies suggest that three (Cordova, Studts, Hann, Jacobsen, & Andrykowski, 2000; Cox, Clara, & Enns, 2002; Thatcher & Krikorian, 2005) or four factor-factor models (Andrews, Joseph, Shevlin, & Troop, 2006; Asmundson et al., 2000; King, King, Fairbank, Keane, & Adams, 1998; McWilliams, Cox, & Asmundson, 2005; Palmieri & Fitzgerald, 2005; Simms, Watson, & Doebbeling, 2002) may better approximate the symptoms of PTSD. Future research should consider multidimensional IRT models and whether they better characterize the dimensionality underlying PTSD symptomology over the more parsimonious unidimensional models.

Conclusions

To our knowledge, this is the first study to examine the relative severity of PTSD symptoms in a community sample. Future research should continue to explore the relative severity of PTSD symptoms as well as the ability to use available theoretical models to test placement of new symptoms within established hierarchies. Such efforts can fill gaps in the continuum of

PTSD severity and thus increase our understanding of how to adequately assess the full continuum of the PTSD syndrome. More precise assessment of PTSD severity may help inform the understanding of endophenotypes within those who experience PTSD, improve the descriptive value of PTSD measures' relationship to continuous measures of treatment outcomes, enhance predictive validity, and ultimately inform more effective treatments.

Acknowledgements

The National Comorbidity Survey - Replication (NCS-R) was supported by grant U01-MH60220 from the National Institute of Mental Health (NIMH), Bethesda, MD, with supplemental support from the National Institute of Drug Abuse (NIDA), Bethesda; the Substance Abuse and Mental Health Services Administration, Bethesda; grant 044708 from The Robert Wood Johnson Foundation, Princeton, NJ; and the John W. Alden Trust, Boston. Ronald C. Kessler, Principal Investigator.

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders-III. Vol. 3. Washington, DC: Author; 1980.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Vol. 3. Washington, DC: Author; 1987.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Vol. 4. Washington, DC: Author; 2000.
- Andrews L, Joseph S, Shevlin M, Troop N. Confirmatory factor analysis of posttraumatic stress symptoms in emergency personnel: An examination of seven alternative models. *Personality and Individual Differences* 2006;41:213–224.
- Asmundson GJ, Frombach I, McQuaid J, Pedrelli P, Lenox R, Stein MB. Dimensionality of posttraumatic stress symptoms: a confirmatory factor analysis of DSM-IV symptom clusters and other symptom models. *Behaviour Research and Therapy* 2000;38(2):203–214. [PubMed: 10661004]
- Bentler PM. Comparative fit indexes in structural models. *Psychological Bulletin* 1990;107(2):238–246. [PubMed: 2320703]
- Betemps EJ, Smith RM, Baker DG, Rounds-Kugler BA. Measurement precision of the clinician administered PTSD scale (CAPS): a RASCH model analysis. *Journal of Applied Measurement* 2003;4(1):59–69. [PubMed: 12700431]
- Blake DD, Weathers FW, Nagy LM, Kaloupek DG, Gusman FD, Charney DS, et al. The development of a Clinician-Administered PTSD Scale. *Journal of Traumatic Stress* 1995;8(1):75–90. [PubMed: 7712061]
- Breslau N, Davis GC, Andreski P, Peterson EL, Schultz LR. Sex differences in posttraumatic stress disorder. *Archives of General Psychiatry* 1997;54(11):1044–1048. [PubMed: 9366662]
- Breslau N, Kessler RC. The stressor criterion in DSM-IV posttraumatic stress disorder: an empirical investigation. *Biological Psychiatry* 2001;50(9):699–704. [PubMed: 11704077]
- Breslau N, Lucia VC, Davis GC. Partial PTSD versus full PTSD: an empirical examination of associated impairment. *Psychological Medicine* 2004;34(7):1205–1214. [PubMed: 15697047]
- Broman-Fulks JJ, Ruggiero KJ, Green BA, Kilpatrick DG, Danielson CK, Resnick HS, et al. Taxometric Investigation of PTSD: data from two nationally representative samples. *Behavior Therapy* 2006;37(4):364–380. [PubMed: 17071214]
- Buckley TC, Blanchard EB, Hickling EJ. A confirmatory factor analysis of posttraumatic stress symptoms. *Behaviour Research and Therapy* 1998;36(11):1091–1099. [PubMed: 9737061]
- Centers for Disease Control. Health status of Vietnam veterans. I. Psychosocial characteristics. The Centers for Disease Control Vietnam Experience Study. *Journal of the American Medical Association* 1988;259(18):2701–2707. [PubMed: 2833630]
- Conrad KJ, Wright BD, McKnight P, McFall M, Fontana A, Rosenheck R. Comparing traditional and Rasch analyses of the Mississippi PTSD Scale: revealing limitations of reverse-scored items. *Journal of Applied Measurement* 2004;5(1):15–30. [PubMed: 14757989]

- Cordova MJ, Studts JL, Hann DM, Jacobsen PB, Andrykowski MA. Symptom structure of PTSD following breast cancer. *Journal of Traumatic Stress* 2000;13(2):301–319. [PubMed: 10838677]
- Cox BJ, Clara IP, Enns MW. Posttraumatic stress disorder and the structure of common mental disorders. *Depression and Anxiety* 2002;15(4):168–171. [PubMed: 12112721]
- Coyne JC, Thompson R. Posttraumatic stress syndromes: Useful or negative heuristics? *Journal of Anxiety Disorders* 2007;21(2):223–229. [PubMed: 17084064]
- Flack WF Jr, Litz BT, Hsieh FY, Kaloupek DG, Keane TM. Predictors of emotional numbing, revisited: a replication and extension. *Journal of Traumatic Stress* 2000;13(4):611–618. [PubMed: 11109234]
- Flora DB, Curran PJ. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods* 2004;9(4):466–491. [PubMed: 15598100]
- Foa EB, Riggs DS, Gershuny BS. Arousal, numbing, and intrusion: symptom structure of PTSD following assault. *American Journal of Psychiatry* 1995;152(1):116–120. [PubMed: 7802101]
- Forbes D, Haslam N, Williams BJ, Creamer M. Testing the latent structure of posttraumatic stress disorder: a taxometric study of combat veterans. *Journal of Traumatic Stress* 2005;18(6):647–656. [PubMed: 16382434]
- Gavranidou M, Rosner R. The weaker sex? Gender and post-traumatic stress disorder. *Depression and Anxiety* 2003;17(3):130–139. [PubMed: 12768647]
- Grubaugh AL, Magruder KM, Waldrop AE, Elhai JD, Knapp RG, Frueh BC. Subthreshold PTSD in primary care: prevalence, psychiatric disorders, healthcare use, and functional status. *Journal of Nervous and Mental Disease* 2005;193(10):658–664. [PubMed: 16208161]
- Gudmundsdottir B, Beck JG. Understanding the pattern of PTSD symptomatology: a comparison of between versus within-group approaches. *Behaviour Research and Therapy* 2004;42(11):1367–1375. [PubMed: 15381444]
- Gustafsson J, Balke G. General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research* 1993;28:407–434.
- Holland, PW.; Wainer, H. *Differential item functioning*. Hillsdale, NJ: Lawrence-Erlbaum; 1993.
- Holzinger KJ, Swineford F. The bifactor method. *Psychometrika* 1937;2:41–54.
- Keane TM, Caddell JM, Taylor KL. Mississippi Scale for Combat-Related Posttraumatic Stress Disorder: three studies in reliability and validity. *Journal of Consulting and Clinical Psychology* 1988;56(1):85–90. [PubMed: 3346454]
- Kessler RC, Berglund P, Chiu WT, Demler O, Heeringa S, Hiripi E, et al. The US National Comorbidity Survey Replication (NCS-R): design and field procedures. *International Journal of Methods in Psychiatric Research* 2004;13(2):69–92. [PubMed: 15297905]
- Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* 2005;62(6):593–602. [PubMed: 15939837]
- Kessler RC, Merikangas KR. The National Comorbidity Survey Replication (NCS-R): background and aims. *International Journal of Methods in Psychiatric Research* 2004;13(2):60–68. [PubMed: 15297904]
- Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB. Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry* 1995;52(12):1048–1060. [PubMed: 7492257]
- Kessler RC, Ustun TB. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research* 2004;13(2):93–121. [PubMed: 15297906]
- King LA, King DW, Fairbank JA, Keane TM, Adams GA. Resilience-recovery factors in post-traumatic stress disorder among female and male Vietnam veterans: hardiness, postwar social support, and additional stressful life events. *Journal of Personality and Social Psychology* 1998;74(2):420–434. [PubMed: 9491585]
- Kulka, RA.; Schlenger, WE.; Fairbank, JA.; Hough, RL.; Jordan, BK.; Marmar, CR., et al. *Trauma and the Vietnam War generation: Report of findings from the National Vietnam Veterans Readjustment Study*. New York: Brunner/Mazel; 1990.
- Litz BT, Orsillo SM, Kaloupek D, Weathers F. Emotional processing in posttraumatic stress disorder. *Journal of Abnormal Psychology* 2000;109(1):26–39. [PubMed: 10740933]

- Lord, F. Applications of Item Response Theory to practical testing problems. Hillsdale, NJ: Erlbaum; 1980.
- McNally RJ. Can we solve the mysteries of the National Vietnam Veterans Readjustment Study? *Journal of Anxiety Disorders* 2007;21(2):192–200. [PubMed: 17045777]
- McWilliams LA, Cox BJ, Asmundson GJ. Symptom structure of posttraumatic stress disorder in a nationally representative sample. *Journal of Anxiety Disorders* 2005;19(6):626–641. [PubMed: 15927777]
- Muthen, LK.; Muthen, BO. Mplus user's guide. Vol. 4. Los Angeles, CA: Muthen & Muthen; 1998–2006.
- Palmieri PA, Fitzgerald LF. Confirmatory factor analysis of posttraumatic stress symptoms in sexually harassed women. *Journal of Traumatic Stress* 2005;18(6):657–666. [PubMed: 16382424]
- Patrick CJ, Hicks BM, Nichol PE, Krueger RF. A bifactor approach to modeling the structure of the Psychopathy Checklist-Revised. *Journal of Personality Disorders* 2007;2:118–141. [PubMed: 17492917]
- Perkonig A, Kessler RC, Storz S, Wittchen HU. Traumatic events and post-traumatic stress disorder in the community: prevalence, risk factors and comorbidity. *Acta Psychiatrica Scandinavica* 2000;101(1):46–59. [PubMed: 10674950]
- Peters L, Issakidis C, Slade T, Andrews G. Gender differences in the prevalence of DSM-IV and ICD-10 PTSD. *Psychological Medicine* 2006;36(1):81–89. [PubMed: 16356294]
- Rasch, G. Probabilistic models for some intelligence and attainment tests. Copenhagen: Denmark's Paedagogiske Institut; 1960.
- Reise SP, Haviland MG. Item response theory and the measurement of clinical change. *Journal of Personality Assessment* 2005;84(3):228–238. [PubMed: 15907159]
- Reise SP, Morizot J, Hays RD. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research* 2007;16:19–31. [PubMed: 17479357]
- Ruscio AM, Ruscio J, Keane TM. The latent structure of posttraumatic stress disorder: a taxometric investigation of reactions to extreme stress. *Journal of Abnormal Psychology* 2002;111(2):290–301. [PubMed: 12003450]
- Schutzwahl M, Maercker A. Effects of varying diagnostic criteria for posttraumatic stress disorder are endorsing the concept of partial PTSD. *Journal of Traumatic Stress* 1999;12(1):155–165. [PubMed: 10027149]
- Simms LJ, Watson D, Doebbeling BN. Confirmatory factor analyses of posttraumatic stress symptoms in deployed and nondeployed veterans of the Gulf War. *Journal of Abnormal Psychology* 2002;111(4):637–647. [PubMed: 12428777]
- Spitzer RL, First MB, Wakefield JC. Saving PTSD from itself in DSM-V. *Journal of Anxiety Disorders* 2007;21(2):233–241. [PubMed: 17141468]
- Steiger JH. Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research* 1990;25:173–180.
- Stein MB, Walker JR, Forde DR. Gender differences in susceptibility to posttraumatic stress disorder. *Behaviour Research and Therapy* 2000;38(6):619–628. [PubMed: 10846810]
- Stein MB, Walker JR, Hazen AL, Forde DR. Full and partial posttraumatic stress disorder: findings from a community survey. *American Journal of Psychiatry* 1997;154(8):1114–1119. [PubMed: 9247398]
- Taylor S, Kuch K, Koch WJ, Crockett DJ, Passey G. The structure of posttraumatic stress symptoms. *Journal of Abnormal Psychology* 1998;107(1):154–160. [PubMed: 9505048]
- Thatcher DL, Krikorian R. Exploratory factor analysis of two measures of posttraumatic stress disorder (PTSD) symptoms in a non-clinical sample of college students. *Journal of Anxiety Disorders* 2005;19(8):904–914. [PubMed: 16243638]
- Thissen D. IRTLRDIF v.2.0b: Software for the Computation of the Statistics Involved in Item Response Theory Likelihood-Ratio Tests for Differential Item Functioning. 2001 Unpublished manuscript
- Thissen, D.; Steinberg, L.; Wainer, H. Detection of differential item functioning using the parameters of item response models. In: Holland, PW.; Wainer, H., editors. *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1993. p. 67–113.

- Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 1973;38:1–10.
- Weathers, FW.; Litz, B.; Huska, J.; Keane, TM. PTSD Checklist (PCL). Boston, MA: National Center for PTSD - Behavioral Science Division; 1994.
- Yehuda R. Biological factors associated with susceptibility to posttraumatic stress disorder. *Canadian Journal of Psychiatry* 1999;44(1):34–39.
- Yu, CY. Unpublished Doctoral dissertation. University of California; Los Angeles: 2002. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes.
- Yung Y, Thissen D, McLeod LD. On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika* 1999;64:113–128.

Figure 1.
Item characteristic curve (ICC) for PTSD symptoms.

Figure 2.

Test information function for the 17 symptoms of Post Traumatic Stress Disorder. The vertical lines locate the average severity level for individuals with (PTSD +) and without (PTSD -) the DSM-IV diagnosis.

Table 1
NCS-R questions used to assess re-experiencing, avoidance, and hyperarousal symptomology

Symptom Cluster B: Re-experiencing

- Did you ever have repeated unwanted memories of (this experience) – that is, you kept remembering it even when you didn't want to?
 - Did you ever have repeated unpleasant dreams about (this experience)?
 - Did you have flashbacks – that is, suddenly act or feel as if (this experience) were happening all over again?
 - Did you get very upset when you were reminded of (this experience)?
 - When you were reminded of (this experience), did you ever have physical reactions like sweating, your heart racing, or feeling shaky?
 - In the weeks, months, or years after (this experience ended), did you try not to think about (it/what happened)?
-

Symptom Cluster C: Avoidance

- In the weeks, months, or years after (this experience ended), did you try not to think about (it/what happened)?
 - Did you purposely stay away from places, people or activities that reminded you of (this experience)?
 - Were you ever unable to remember some important parts of what happened?
 - Did you lose interest in doing things you used to enjoy?
 - Did you feel emotionally distant or cut-off from other people?
 - Did you have trouble feeling normal feelings like love, happiness, or warmth toward other people?
 - Did you feel you had no reason to plan for the future because you thought it would be cut short?
-

Symptom Cluster D: Hyperarousal

- During the time (this experience) affected you most, did you have trouble falling or staying asleep?
 - Were you more irritable or short-tempered than you usually are?
 - Did you have more trouble concentrating or keeping your mind on what you were doing?
 - Were you much more alert or watchful, even when there was no real need to be?
 - Were you more jumpy or easily startled by ordinary noises?
-

Kessler, R.C. (2008). ICPSR 4438: National Comorbidity Survey: Replication (NCS-R), 2001–2003.
<http://www.hcp.med.harvard.edu/ncs/replication.php>.

Table 2

Factor loadings of different confirmatory factor models

Item	Unidimensional		Multidimensional			Bifactor			
	1		1	2	3	G	g1	g2	g3
B1	0.81			0.85		0.77		0.53	
B2	0.67			0.71		0.64		0.36	
B3	0.68			0.72		0.67		0.23	
B4	0.75			0.79		0.75		0.18	
B5	0.70			0.74		0.72		0.07	
C1	0.52		0.56			0.57	-0.16		
C2	0.60		0.65			0.59	0.13		
C3	0.44		0.48			0.42	0.15		
C4	0.67		0.73			0.57	0.53		
C5	0.68		0.75			0.55	0.66		
C6	0.71		0.79			0.61	0.55		
C7	0.64		0.70			0.58	0.38		
D1	0.71				0.75	0.74			-0.02
D2	0.64				0.68	0.67			-0.06
D3	0.72				0.77	0.76			-0.15
D4	0.67				0.72	0.64			0.78
D5	0.73				0.78	0.70			0.39
CFI	.86		.91			.96			
TLI	.94		.96			.98			
RMSEA	.08		.06		.06	.04			

Note: G=general factor, g=group factor

Table 3

Symptom parameter estimates and symptom bias for PTSD symptoms in ascending order of severity

Symptom	Item Parameters		
	% Endorsement	Severity	Discrimination
C7. No reason plan for future	25.7	.855	1.502
C3. Unable to remember	33.2	.836	.848
D2. Irritable	50.8	-.047	1.461
B5. Physical reactions when reminded	51.5	-.049	1.831
B3. Flashbacks	51.4	-.073	1.695
D5. Jumpy/easily startled	54.7	-.154	1.820
C4. Lost interest in things used to enjoy	54.8	-.191	1.473
C6. Trouble feeling normal feelings	56.3	-.215	1.673
B2. Unpleasant dreams	56.6	-.232	1.576
C2. Avoid Reminders	57.2	-.280	1.305
D4. More alert/watchful	60.9	-.390	1.473
C5. Emotionally distant	63.6	-.489	1.453
D3. Trouble concentrating	66.7	-.507	1.878
B4. Got upset when reminded about	68.6	-.525	2.113
D1. Sleep problems	72.6	-.748	1.804
B1. Had unwanted memory	75.3	-.750	2.390
C1. Not think about	77.1	-1.280	1.056