Measurement Invariance of Posttraumatic Stress Disorder Symptoms Across Three Civilian Trauma Types

by

Benjamin Hammond Carter

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 8, 2012

Keywords: PTSD, CFA, measurement invariance,
PCL, factor structure

Approved by

Frank Weathers, Chair, Professor of Psychology
Christopher Correia, Associate Professor of Psychology
Alejandro Lazarte, Associate Professor of Psychology
Richard Mattson, Assistant Professor of Psychology

Abstract

The factor structure of posttraumatic stress disorder (PTSD) remains the subject of intense investigation. The DSM three-factor conceptualization of PTSD has not been empirically supported; rather, two four-factor models of PTSD (King, Leskin, King, & Weathers, 1998; Simms, Watson, & Doebbeling, 2002) have garnered the majority of support from confirmatory factor analytic (CFA) studies. Recently, interest has turned to examination of the generalizability of these well-supported models across diverse samples. Termed factorial or measurement invariance, these studies answer the question of whether PTSD maintains the same factor structure, and can be measured equivalently, across samples. However, no studies have examined PTSD's measurement invariance across distinct, homogeneous samples exposed to different trauma types. The current study examined the factor structure and measurement invariance of PTSD in, and across, three groups of trauma-exposed college students (N = 854) using the PTSD Checklist—Specific Version (PCL-S; Weathers, Litz, Herman, Juska, & Keane, 1993). Participants were grouped according to self-reported direct exposure to one of three distinct trauma types: motor vehicle accidents (MVA), sexual assault (SA), and sudden unexpected death of a loved one (SUD). Five models were tested using within-groups CFA, and three models showed adequate fit in each trauma group (Elhai, Biehn, Armour, Klopper, Frueh, & Palmieri, 2011; King et al., 1998; Simms et al., 2002). Multiple-group CFA showed that factor loadings were equivalent across groups for the Elhai et al. (2011) and the King et al. (1998) models, but not for the Simms et al. (2002) model. However, intercepts differed between groups in all three models. These findings suggest that PTSD symptoms as measured by the PCL-S cannot be compared across individuals who have experienced different types of traumatic events, as observed scores on the PCL-S may not correspond to levels of the latent factors equivalently across trauma types. Implications and limitations of the current results are discussed.

Acknowledgments

Many thanks to Dr. Frank Weathers for his expertise, wisdom, support, patience, and friendship over the last five years. His ideas have been interesting and inspiring, and his optimism has been a key ingredient in all of my work. Thank you also to Rich Mattson and Chris Correia who were not only committee members but also clinical supervisors, instructors, mentors, and friends. I also thank Tracy Witte for her time, expertise, and interest in this project. Of course I owe eternal gratitude to my wife, Jana Carter, and three children, Caitlyn, Jackson, & Lauren. Your sacrifices are appreciated and you are the reason that the hard work is all worthwhile. I greatly appreciate the graduate students and undergraduate research assistants who contributed their time and abilities, as well as the many participants who shared their personal experiences. Finally, thank you to my parents, Don & Judy Carter, who have always encouraged me and been examples for me.

Table of Contents

List of Tables

List of Abbreviations

PTSD	Posttraumatic Stress Disorder
DSM	Diagnostic and Statistical Manual
EFA	Exploratory Factor Analysis
CFA	Confirmatory Factor Analysis
PCL	PTSD Checklist
LEC	Life Events Checklist
MVA	Motor Vehicle Accident
SA	Sexual Assault
SUD	Sudden Unexpected Death of a Loved One
CFI	Comparative Fit Index
TLI	Tucker-Lewis Index
SRMR	Standardized Root Mean Square Residual
RMSEA	Root Mean Square Error of Approximation
BIC	Bayesian Information Criterion

Introduction

Posttraumatic stress disorder (PTSD) was introduced as a diagnostic category in the *Diagnostic and Statistical Manual, 3rd edition* (*DSM-III*; American Psychiatric Association [APA], 1980) in formal recognition of the characteristic stress-response syndrome that may develop following exposure to catastrophic life events. PTSD has been studied extensively over the last thirty years and a substantial body of research has emerged supporting its construct validity in the areas of phenomenology, etiology, assessment, and intervention (Friedman et al., 2007). Nonetheless, PTSD has generated considerable controversy and debate (Rosen & Lilienfeld, 2008) and many key issues regarding the disorder have yet to be resolved.

From its inception, PTSD has been conceptualized as a complex disorder with numerous symptoms grouped in related but distinct clusters (Wilson, 2004). Thus, one crucial area of investigation involves two closely related questions: What are the core symptoms of PTSD and how are they clustered? The first question has to do with the descriptive psychopathology of the disorder and articulating the defining features that distinguish it from other disorders. The second question has to do with discovering the latent structure of PTSD, i.e., the underlying factors that give rise to the various PTSD symptoms. Structural validity evidence is essential, first because it is an important aspect of construct validity (American Educational Research Association, 1999), and second because each underlying factor may operate through different mechanisms and thus represent different paths to the disorder as well as predict different outcomes (Cattell, 1978). Knowledge of the structure of PTSD may help refine its diagnostic criteria, inform its appropriate placement within the current psychiatric nosology, and contribute to the identification of its unique aspects that can then be targeted in prevention and treatment programs.

Such questions regarding structural validity evidence are best answered through factor analysis. Exploratory factor analysis (EFA) is a largely atheoretical descriptive technique that attempts to explain large covariance matrices in terms of a few latent factors (Fabrigar, Wegener, MacCallum, & Strahan, 1999). It is used in an attempt to identify a set of factors that is responsible for the covariation among the observed variables, that is, a set of hypothetical constructs believed to represent distinct causal mechanisms (Cattell, 1978). While EFA is appropriate for identifying the possible latent structure of a construct, it is not suitable for examining the fit of hypothesized structures or comparing the fit of competing structures. For this reason, focus has shifted to confirmatory factor analysis (CFA) studies in which models are specified a priori and tested against alternative models to determine which most accurately reflects the relationship between symptoms.

Structural Studies of PTSD

The current PTSD diagnostic criteria in the *Diagnostic and Statistical Manual, Fourth Edition, Text Revision* (*DSM-IV-TR*; APA, 2000) include 17 symptoms grouped in three symptom clusters (see Table 1): (1) persistent reexperiencing of the traumatic event (Criterion B); (2) avoidance of event-related stimuli and numbing of general responsiveness (Criterion C); and (3) hyperarousal (Criterion D). These symptom clusters were derived rationally rather than empirically (Brett, Spitzer, & Williams, 1988), and the validity of the factor structure they imply has been challenged. Several EFA and CFA studies have suggested alternative models to explain the relationships among the 17 PTSD symptoms (e.g., Asmundson, Frombach, McQuaid, Pedrelli, Lenox, & Stein, 2000; Elhai, Biehn, Armour, Klopper, Frueh, & Palmieri, 2011; King, Leskin, King, & Weathers, 1998; Simms, Watson, & Doebbling, 2002). Alternative conceptualizations have included factor structures with between two (e.g., Taylor et al., 1998) and five factors (e.g., Elhai et al., 2011), and have been evaluated in a variety of populations using a variety of PTSD assessment instruments (King, King, Orazem, & Palmieri, 2006; Yufik & Simms, 2010). To date, cumulative evidence supports two competing four-factor models. Nonetheless, findings continue to vary across studies, and there remains uncertainty as to which model most accurately represents the symptom structure of PTSD. This variability is likely attributable to methodological differences across studies with respect to sampling (e.g., trauma type, gender, sample size, PTSD prevalence), assessment instruments (e.g., self-report vs. structured interview, DSM-correspondent vs. non-DSM-correspondent measures), and analytic strategies (e.g., choice of models to test, estimation method). In the next section, studies finding support for alternative models of PTSD are briefly summarized. Because the literature has moved beyond EFA and focused almost exclusively on CFA to evaluate competing structural models of PTSD, only CFA findings are included in the review.

Two-factor models.

Although some prominent etiological models of PTSD imply a two-factor model, (e.g., Foa, Zinbarg, & Rothbaum, 1992; Horowitz, Wilner, & Alvarez, 1979), only six CFA studies have supported a two-factor solution, and they vary considerably with respect to the nature of the two factors evaluated. Two studies agreed on a model with an intrusions and an avoidance factor (Shevlin, Hunt, & Robbins, 2000; van der Ploeg et al., 2004), while two other studies supported a model with an intrusion/avoidance factor and a hyperarousal/numbing factor (Asmundson, Wright, McCreary, & Pedlar, 2003; Buckley, Blanchard, & Hickling, 1998). A fifth study found support for a model with two factors, labeled depression/avoidance and anxiety/arousal (Maes et al., 1998). Recently, Synder, Elhai, North, and Heaney (2009) proposed a unique two-factor model with factors labeled trauma/dysphoria and self-dysfunction. Importantly, of the studies finding support for two-factor models, only one specified and tested multiple alternative models that included more than two factors (Asmundson et al., 2003). Similarly, only two studies used measures with items that correspond directly to *DSM-IV-TR* PTSD symptom criteria (Asmundson et al., 2003; Buckley et al., 1998). Thus, support for a two-factor structure of PTSD is minimal.

Three-factor models.

A similarly limited number of published studies have provided support for three-factor models of PTSD, and the proposed three-factor models are equally diverse. Three studies found some support for a three-factor model that resembles the *DSM-IV-TR* conceptualization, with re-experiencing, avoidance/numbing, and hyperarousal factors (Cordova, Studts, Hann, Jacobsen, & Andrykowski, 2000; Foy, Wood, King, King, & Resnick, 1997; Taft et al., 2007). Another study found support for a model with intrusion/active avoidance, numbing/passive avoidance, and arousal factors (Anthony, Lonigan, & Hecht, 1999). Finally, Lancaster, Melka, and Rodriguez (2009) found support for a three-factor model with intrusion/avoidance, dysphoria, and hyperarousal factors. Each of these studies tested multiple models, though only three employed DSM-correspondent measures (Cordova et al., 2000; Lancaster et al., 2009; Taft et al., 2007). Thus, as with the two-factor models, there is only limited support for three-factor models, and even then, some of these do not replicate the *DSM-IV-TR* PTSD symptom clusters.

Four-factor models.

Most CFA studies of PTSD have identified one of two four-factor models as the best fit: the so-called numbing model, originally described by King, Leskin, King, and Weathers (1998), and the so-called dysphoria model originally described by Simms, Watson, and Doebbling (2002). The primary difference between the King et al. (1998) numbing model and the *DSM-IV-TR* three-factor model is that the numbing model separates the *DSM-IV-TR* avoidance and numbing cluster into two distinct factors labeled avoidance (C1-C2) and emotional numbing (C3-C7). This is a reflection of evidence that avoidance and numbing are differentially related to measures of psychopathology and post-treatment outcomes, and that they differentially predict poor treatment response (see Asmundson, Stapleton, & Taylor, 2004 for a review).

Using the CAPS in a sample of 524 treatment-seeking male military veterans, King and colleagues (1998) tested four models of PTSD: a four-factor first-order model, a two-factor higher-order model, a single-factor higher-order model, and a single-factor first-order model. Goodness-of-fit indices suggested that the intercorrelated four-factor model (reexperiencing, effortful avoidance, emotional numbing, and hyperarousal) provided the best fit for the data. The numbing model has been supported in other CFA studies using different measures in a wide variety of populations, including primary medical care patients (Asmundson et al., 2000; Grubaugh, Long, Elhai, Frueh, & Magruder, 2010; Naifeh, Elhai, Kashdan, & Grubaugh, 2008), emergency responders (Andrews, Joseph, Shevlin, & Troop, 2006), UN peacekeepers (Asmundson, Wright, McCreary, & Pedlar, 2003), cancer survivors (DuHamel et al., 2004), survivors of community violence (Marshall, 2004), war veterans (Amdur & Liberzon, 2001; McDonald, Beckham, Morey, Marx, Tupler, & Calhoun, 2008), sexually harassed women (Palmieri & Fitzgerald, 2005), war refugees (Palmieri, Marshall, & Schell, 2007; Rasmussen, Smith, & Keller, 2007; Sack, Seeley, & Clarke, 1997), college students (Elhai, Gray, Docherty, Kashdan, & Kose, 2007), elderly survivors of hurricanes (Schinka, Brown, Borenstein, & Mortimer, 2007), individuals indirectly exposed to the 9/11 attacks (Suvak, Maguen, Litz, Silver, & Holman, 2008), and nationally representative samples (Cox, Mota, Clara, & Asmundson, 2008; McWilliams, Cox, & Asmundson, 2005).

In contrast to the numbing model, Simms and colleagues (2002) found support for a first-order, four-factor model that involved combining symptoms D1 through D3 from the hyperarousal factor with items from the numbing factor (C3-C7) to create what they called the dysphoria factor. Simms and colleagues made the argument that several of PTSD's symptoms are examples of general emotional distress common to other anxiety and mood disorders, and that these symptoms cohere together to form a factor that is common across internalizing disorders, often referred to as negative emotionality or negative affectivity (Brown & Barlow, 2009; Clark & Watson, 2006; Miller, 2003; Watson, 2005). By combining these items, this model separates PTSD-specific items (e.g., intrusion, avoidance, and hyperarousal factors) from more general items (dysphoria factor), and is consistent with structural research that has attempted to explain the comorbidity across anxiety and mood disorders (e.g., Clark & Watson, 2006; Miller, 2003; Mineka, Watson, & Clark, 1998).

Using the PCL in a sample of 3,695 Gulf War veterans, Simms and colleagues (2002) tested six models of PTSD, including the four-factor numbing model and the four-factor dysphoria model described above. The dysphoria model consistently fit the data better than any of the other models, though the difference between the dysphoria and numbing models was small. The authors cross-validated their findings in a sample of non-deployed Gulf War era veterans and obtained identical results. Like the numbing model, the dysphoria model has been supported in multiple samples including rape and sexual assault victims (Armour & Shevlin, 2010; Elklit, Armour, & Shevlin, 2010; Ullman & Long, 2008), New York undergraduates following the 9/11 terrorist attacks (Baschnagel, O'Conner, Colder, & Hawk, 2005), bereaved individuals (Armour & Shevlin, 2010; Elklit, Armour, & Shevlin, 2010; Boelen, van den Hout, van den Bout, 2008), motor vehicle accident survivors (Elklit & Shevlin, 2007; Grant, Beck, Marques, Palyo, & Clapp, 2008), treatment-seeking victims of intimate partner violence (Krause, Kaltman, Goodman, & Dutton, 2007), nationally representative epidemiological samples (Carragher, Mills, Slade, Teesson, & Silove, 2010), combat veterans (Pietrzak, Goldstein, Malley, Rivers, & Southwick, 2010), and civilian trauma survivors (Olff, Sijbrandij, Opmeer, Carlier, & Gersons, 2009).

Five-factor models.

Only three CFA studies have found support for a five-factor model in comparison with other, more parsimonious models. Witteveen and colleagues (2006) tested eight alternative models using the Self-Rating Inventory for Posttraumatic Stress Disorder (SRIP; Hovens, van der Ploeg, Bramsen, Klaarenbeek, Schreuder, & Rivero, 1994) in a sample of 1,168 police officers and fire fighters involved in the 1992 airline crash in Amsterdam. They found the most support for a five-factor model, though some of the reported fit statistics were below traditional cutoffs for acceptable model fit. Their model, including intrusion, avoidance, hyperarousal, emotional numbing, and sleep disturbance factors, was also supported by Morina and colleagues (2010) who used the Impact of Event Scale—Revised (IES-R; Weiss & Marmar, 1997) in 1662 residents of the former Yugoslavia who experienced potentially traumatic events during the Yugoslav Wars (1991-1995). Of note, the Witteveen et al. (2006) study also utilized the IES in the same sample and instead found most support for a four-factor model originally proposed by

Amdur and Liberzon (2001) that added a sleep disturbance factor to intrusions, avoidance, and numbing. Fit indices were also moderate at best for the four-factor IES model.

Most recently, Elhai and colleagues (2011) created a new five-factor model by specifying the three symptoms that differentiate between the numbing and dysphoria four-factor models (D1-D3) to load on their own factor. Elhai and colleagues argued that symptoms D1-D3 are conceptually different from both hyperarousal and dysphoria. They cite Watson's (2005) conclusion that the D1-D3 symptoms involve general distress or dysphoria, while the D4-D5 symptoms involve anxious arousal that is consistent with fear-based disorders. They also suggest that D1-D3 differ somewhat from dysphoria because the symptoms involve restlessness and agitation, as opposed to the numbing of responsiveness seen in the other dysphoria symptoms. In support of this, Shevlin, McBride, Armour, and Adamson (2009) found that while D1-D3 load on both the hyperarousal and dysphoria factors, they load only weakly and do not appear to be clear indicators of either factor. Finally, Elhai and colleagues note that Simms and colleagues (2002) made two changes to the numbing model simultaneously (separated D1-D3 from the hyperarousal factor and combined them with the numbing symptoms), thereby making it impossible to know which modification improved model fit. The five-factor model was proposed to determine whether separating D1-D3 from both the hyperarousal and dysphoria factors is empirically supported.

Using the PTSD Symptom Scale-Self Report (PSS; Foa, Riggs, Dancu, & Rothbaum, 1993) in a sample of 252 female victims of domestic violence, the five-factor model fit significantly better than the four-factor numbing and dysphoria models. However, because the sample size was small, and because it has yet to be validated in other samples, the status of the five-factor model as a proposed structure of PTSD remains tentative. Nonetheless, it represents a possible solution for resolving the conflict between the numbing and dysphoria models and merits further evaluation.

In the only published meta-analysis of structural studies of PTSD, Yufik and Simms (2010) obtained correlation matrices from 40 studies that used DSM-correspondent measures of PTSD. Aggregating the correlation matrices allowed a test of the most prominent models of PTSD in a large (*N* = 14,827) and diverse sample, thereby overcoming the limited generalizability of previous studies. Though both the numbing and dysphoria four-factor models showed good fit to the aggregated data, the dysphoria model was superior across all sample types. Furthermore, the authors concluded that PTSD structure did not appear to vary as a function of the measures used or of trauma types assessed. However, they did not formally test for invariance in the traditional multiple groups CFA framework; rather, they separated the samples into homogeneous groups and observed that the dysphoria model fit slightly better in each of the groups (Yufik & Simms, 2010). In the combined sample, the difference in fit between the numbing and dysphoria models was small, and the models were equivalent on three of the six fit indices. In addition, the authors noted that certain trauma types (e.g., motor vehicle accidents, terrorism exposure, sexual assault) were underrepresented in their meta-analysis and should be further explored.

Review of the various CFA studies reported above shows a clear consensus of support for two four-factor models. However, efforts to reconcile the differences between the two models (e.g., Shevlin et al., 2009) have yet to provide clear evidence of the superiority of one model over the other. Differentiating between the two is difficult as it is not possible to statistically evaluate differences in model fit between non-nested models. Furthermore, the differences in fit between the two models have most often been slight. Establishing that one model more consistently fits the data well in diverse samples, using different assessment methods and measures, may contribute valuable information to efforts to distinguish between the two models. Efforts to determine such generalizability are described below.

Measurement Invariance of PTSD Models

Once a theoretically and empirically strong model is identified and replicated using CFA, the model is then tested across multiple samples to determine if the factor structure generalizes across sample differences such as trauma type, age, ethnicity, and gender, as well as across differences in measurement. Using the same measure in multiple samples, one can determine whether the symptoms (represented by the items of the measure) relate to one another and to the latent factors in the same way across samples. Such tests of equivalence across samples are referred to as tests of factorial invariance, a subset of measurement invariance (Byrne, 1998; Gregorich, 2006; Meredith, 1993; Vandenberg & Lance, 2000). When measurement invariance is found, individuals from different subpopulations who are at the same level of a latent construct should receive the same observed score on an indicator of the construct. Where measurement invariance does not hold, comparisons across subpopulations are likely to be misleading and group differences on a construct may be artifactual. Measurement invariance can take many forms, such as demonstrating that items cluster together in a similar way across samples or that items are equally reliable across samples. However, to demonstrate basic measurement invariance, it is only necessary to show that factor loadings (regression slopes) and item intercepts are equivalent across samples. When such equivalence is demonstrated, comparisons of group means can be made with confidence that the observed scores represent real levels of the latent construct.

Establishing equivalence across samples has increasingly become the focus of study in the field of traumatic stress. Studies of measurement invariance can help determine which of the two four-factor models most accurately represents the latent structure of PTSD. If one model generalizes better across samples, it may be preferred to other models. Recently, the structural stability and generalizability of previously supported models has been tested across a variety of samples. Most commonly, these studies have focused on three broad domains: invariance across ethnic, racial, or geographic subsamples; invariance over time; and, invariance across assessment methods or measures. In these investigations, researchers test increasingly stricter levels of invariance, or equality. The levels of invariance within this hierarchy of tests have been labeled in various ways. The least restrictive level of invariance specifies that the pattern of indicator-factor relationships is the same across samples and has been referred to as configural, pattern, or equal form invariance (Brown, 2006; Meredith, 1993; Horn & McArdle, 1992). A stricter test specifies equivalence of factor loadings for like items across groups, and is referred to as metric, weak factorial, or equal factor loadings invariance. Scalar, strong factorial, or equal intercepts invariance refers to the next level of equivalence, and requires both equal factor loadings and item intercepts across groups. Finally, specifying equality of like items' unique variances across groups is referred to as strict factorial invariance or invariant uniquenesses (Vandenberg & Lance, 2000). While tests of configural, metric, and scalar invariance are common in the PTSD structural literature, tests of strong and strict factorial invariance are less common.

These tests of invariance are hierarchical; therefore, previous, more conservative requirements must be met prior to testing more rigorous standards. However, when invariance is not found, it is recommended that partial invariance be tested to determine the source of non-invariance (Byrne, Shavelson, & Muthén, 1989). This is done by systematically releasing non-invariant parameter estimates from constraint and observing the change in model fit compared to the baseline model. Once the non-invariant parameters are identified and released, the next level of invariance can be tested using the modified partial invariance model if at least one item per factor is invariant, in addition to the marker indicator (Byrne et al., 1989). Unfortunately, most studies have discontinued tests of measurement invariance without examining partial invariance, possibly due to concern about over-fitting the model to the sample, or to inflation of Type I error rates with the multiple hypothesis tests that are required. It may also be due to a lack of guidelines in the literature directing partial invariance analysis. In any case, the lack of partial invariance analysis is unfortunate because the specific source of non-invariance can remain hidden when partial invariance is not examined, leaving consumers uninformed regarding the specific properties of measurement instruments.

The majority of measurement invariance studies have examined the performance of a single model of PTSD across two or more groups or subgroups, typically selecting either the numbing or dysphoria models for evaluation. Though some studies have examined the invariance of alternative models (e.g., Anthony et al., 1999; Anthony et al., 2005; King et al., 2009; Wang et al., 2011; Witteveen et al., 2006), this review will focus on the sizeable majority of studies that have explored one of the two four-factor models.

Measurement invariance of the numbing model.

Most analyses of the numbing model have found support for metric invariance and have failed to find support for scalar invariance. In two samples of Spanish- or English-speaking survivors of community violence, Marshall (2004) found evidence of metric and partial scalar factorial invariance of the numbing model using two versions of the PCL. In that study, only one item of the PCL (B4) was identified as non-equivalent at the scalar invariance level between the two samples. Suvak and colleagues (2008) assessed PTSD in a national probability sample immediately following the 9/11 attacks, and again approximately 2-months, 6-months, and 1-year after the attacks. The numbing model demonstrated metric invariance across the four time-points. Comparing deployed vs. non-deployed veterans, Mansfield and colleagues (2010) found evidence of metric but not scalar invariance. McDonald and colleagues (2008) found support for metric invariance of the four-factor numbing model of PTSD when comparing Vietnam-era with post-Vietnam-era treatment-seeking veterans, but did not find metric invariance when comparing treatment-seeking with non-treatment-seeking veterans. Hoyt and Yeater (2010) compared Hispanic and White college students without assessing for trauma history or requiring PCL-C responses to reference a specific traumatic event. They found metric invariance for both the numbing and dysphoria models, as well as for an alternative four-factor model (Smith et al., 1999).

Armour and colleagues (2011) compared war-exposed Bosnian children who endorsed PTSD Criteria A1 and A2 with those who endorsed only A1, and found a lack of invariance beyond the configural level. Similarly, Norris and colleagues (2001) found support for configural but not metric invariance when they tested the fit of the numbing model in Mexican and American survivors of two hurricanes.

Measurement invariance of the dysphoria model.

Results from the examination of measurement invariance of the dysphoria model have been mixed, with a broad range of reported levels of invariance. Simms and colleagues (2002) found support for metric invariance when they evaluated the consistency of their four-factor dysphoria model in a sample of U.S. military veterans deployed during the Gulf War and a demographically equivalent sample of veterans who were not deployed. In five subsamples of victims of sexual or physical abuse or assault, Hetzel-Riggin (2009) found evidence for strict factorial invariance for the dysphoria model across groups. In a sample of female victims of intimate partner violence (IPV), Krause and colleagues (2007) found metric and phi (equal factor intercorrelations) invariance of the dysphoria model when the participants were assessed again one year later. They also cross-validated the model by testing it in both a treatment-seeking group of IPV victims and in a sample of women not seeking treatment but who screened positive for IPV. The model proved to be very stable (metric and phi invariance) across settings as well as over time.

Baschnagel and colleagues (2005) tested the fit of the dysphoria model among Western New York undergraduates at both 1-month and 3-months following the 9/11 terrorist attacks. The authors found configural but not metric invariance across the two assessment time points. However, the only item that was non-equivalent over time was the amnesia item (C3), and this item has consistently loaded only weakly on avoidance or numbing factors in previous factor analytic research (Foa et al., 1995; King et al., 1998; Lancaster et al., 2009; Palmieri et al., 2005; Simms et al., 2002). Comparing deployed vs. non-deployed Canadian military veterans, Engdahl and colleagues (2011) found a lack of invariance for the dysphoria model on all structural parameters, with the exception of item intercepts, which they tested despite finding a lack of invariance at the configural and metric levels. Ullman and Long (2008) also did not find evidence of configural or metric invariance of the dysphoria model when they compared groups by race (Caucasian vs. African-America) and education level (high school or less vs. some college or beyond).

Studies comparing assessment methods.

Though not a direct test of measurement invariance, Palmieri and colleagues (2007) examined the role of assessment instruments in determining the factor structure of PTSD. They used both the PCL and the CAPS in a sample of 2,960 utility workers who worked at the World Trade Center Ground Zero site following the 9/11 attacks. The authors tested five models and found nearly equivalent support for the numbing and dysphoria models. When the PCL data were used in the factor analysis, the data suggested a slightly better fit to the dysphoria model. However, when the CAPS data were used, the data suggested a slightly better fit to the numbing model.

Other studies have also evaluated assessment methods in a more formal measurement invariance approach. Elhai and colleagues (2009) manipulated the instructions on a self-report PTSD measure to form three groups of undergraduate research participants. One random subset of participants was instructed to rate symptoms of PTSD in reference to their single worst traumatic event, whereas another was instructed to rate symptoms in reference to their overall trauma history in general. A third group of non-trauma-exposed participants was instructed to rate PTSD symptoms globally, based on any stressful event. The authors found dimensional, but not configural, invariance across the three groups. While the best-fitting models for each group were all four-factor models, the numbing model fit the trauma-general and non-trauma exposed groups best. The dysphoria model was the best fit in the trauma-specific group. However, the differences in model fit were small and it was not possible to determine if those differences were statistically significant.

A subsequent study compared PTSD symptoms when assessed in terms of frequency vs. intensity (Elhai, Palmieri, Biehn, Frueh, & Magruder, 2010). The authors reported significant variability in CFA results when frequency vs. intensity data was used, finding invariance only in the items associated with the avoidance factor of the four-factor numbing model (Elhai et al., 2010). Finally, Elhai and colleagues (2011) merged the response options of two commonly-used self-report measures of PTSD, the PCL and PSS, and manipulated what instruction set and item prompts participants received. This manipulation represented an attempt to examine the role that a measure's items and instructions play in the latent structure of the construct. They reported that the dysphoria model fit best in both the PCL and PSS groups, but showed only metric invariance.

In summary, there is moderate support for metric invariance of the numbing model across a variety of samples, though support for scalar invariance is lacking. There is also support for metric invariance of the dsyphoria model, however, findings from dysphoria model studies have been more variable. Overall, as most measurement invariance studies of the numbing and dsyphoria models have found at least metric invariance, it appears that the basic relationships between variables are consistent, and that the latent constructs predict item indicators equivalently across groups. Where there is a lack of invariance, it has typically been between groups in which the severity of trauma exposure has been distinctly different (e.g., deployed vs. non-deployed veterans; Armour et al., 2011; Elhai et al., 2009; Engdahl et al., 2011; McDonald et al., 2008). This pattern of non-invariance suggests that the structure of PTSD may differ based on characteristics of the precipitating event (e.g., severity, event type). Unfortunately, none of the previous studies systematically examined invariance across well-defined types of traumatic events. Given that PTSD is assumed to be an equivalent syndrome regardless of the type of traumatic event that precipitated it, this leaves open the key question of whether the structure of PTSD is invariant across trauma types.

Present Study

In line with the recent emphasis on evaluating the generalizability of proposed PTSD models, the primary purpose of the present study was to examine the invariance of the structure of PTSD across trauma types. Contrary to the assumption that PTSD is the same syndrome across trauma types, recent research has suggested that different traumatic events may be associated with differences in PTSD symptom severity and symptom profiles (Elhai, Frueh, Gold, Gold, & Hamner, 2000; Green, Krupnick, Stockton, Goodman, & Corcoran, 2001; Kelley, Weathers, McDevitt-Murphy, Eakin, & Flood, 2009; Kirz, Drescher, Gussman, Klein, & Schwartz, 2001; Wilson, Smith, & Johnson, 1985). For example, Armour and Shevlin (2010) included 10 trauma types as covariates in a multiple causes multiple indicators model (MIMIC) and regressed each of the latent factors of the dysphoria model onto the trauma covariates. The results indicated that direct combat experience, witnessing the severe injury or death of another, rape, and sexual molestation were significantly associated with endorsement of re-experiencing symptoms. Rape, sexual molestation, and physical assault were all significantly associated with avoidance symptoms; and rape, physical assault, being threatened with a weapon, and being kidnapped or held hostage were all significantly associated with symptoms of arousal. None of the trauma types were associated significantly with the dysphoria symptoms (Armour & Shevlin, 2010). It should be noted that the authors did not establish measurement invariance across the trauma types, leaving the possibility that the different associations between trauma types and latent means are due to the indicators not performing in the same way across trauma types. Nonetheless, it may be that the mechanisms underlying PTSD differ somewhat across trauma types, especially as traumatic events differ on the degree of interpersonal closeness involved. Factor analysis can identify how the relationships between symptoms, and the latent factors that explain them, differ across trauma types. Examining these relationships in the context of measurement invariance allows a finer examination of any potential differences, and is essential in order to make comparisons across groups. Further, in order to justify collapsing individual trauma types into heterogeneous groups, it is crucial to demonstrate that existing measures of PTSD symptoms perform similarly across trauma types.

To examine differences in the factor structure of PTSD across trauma types, three conceptually distinct trauma types were chosen for comparison: motor vehicle accident (MVA); sexual assault (SA); and sudden, unexpected death of a loved one (SUD). Motor vehicle accident was included because it involves classical fear conditioning without the interpersonal aspects of both SA and SUD. Sexual assault was included because it is regarded as a prototypical precipitant of PTSD (Kessler et al., 1995), and because sexual assault victims have one of the highest rates of lifetime PTSD (Breslau et al., 1998). Finally, SUD was included because it is distinct from MVA and SA in that it typically does not involve perpetration or a direct physical threat to the survivor. Also, with its high prevalence and moderate conditional risk for PTSD, it has been identified as "the single most important trauma as a cause of PTSD" (Breslau et al., 1998; p. 628).

Goals of the Present Study

The broad aim of this study is to provide structural validity evidence for the PTSD construct using CFA. Specifically, previously proposed models of PTSD were tested in three groups representing three distinct trauma types to determine which model best represented the data in all three groups. Invariance across trauma types of each model was tested with the following goals: (1) to examine whether the latent structure of PTSD is equivalent across trauma types; (2) to determine whether scores on a measure of PTSD can be meaningfully compared across groups; and, (3) to contribute to the ongoing debate regarding the superiority of one of the four-factor models widely supported in the literature.

Method

Participants and Procedures

Participants were 2161 undergraduates attending a large public university in the southeastern United States. Data were collected in five studies, over four years, involving assessment of PTSD. All studies were approved by the university institutional review board. Studies were advertised on the university's password-protected website that lists ongoing research opportunities open to student participation. Participants self-identified as having experienced "stressful life events." In three of the five studies, participants completed a paper and pencil research packet that included an informed consent statement, a brief demographics form, a research version of the Life Events Checklist (LEC; Blake et al., 2004), the specific version of the PCL (PCL-S; Weathers, Litz, Herman, Huska, & Keane, 1993), and other measures of mood, anxiety, and personality. On the LEC, participants identified and briefly described their worst stressor, which served as the index event for inquiry on the PCL. In the remaining two studies, participants completed the same measures in a computer lab using online survey software (Qualtrics).

Only female participants were included, reducing eligible participants to 1761. One hundred and five (6%) participants were eliminated because they did not complete the LEC or PCL, and 474 (29%) were eliminated because their index event did not meet the PTSD Criterion A1 threshold. These restrictions reduced the sample to 1182, all of whom reported exposure to at least one event that met the Criterion A1 requirement, as determined by responses on the LEC and as coded by a research team directed by the second author. The research team developed an extensive coding system for evaluating whether reported events met *DSM-IV-TR* PTSD (APA, 2000) Criterion A1, and for classifying index events into specific event type categories (see Kelley et al., 2009). Decisions regarding Criterion A exposure and event type were based on the LEC and included a narrative description of the index event. Interrater reliability for the three event types used in the current study was very high. In a subset of the full sample (n = 860), percentage agreement was 95.8% and Cohen's kappa was .90 (p < .001). All rating discrepancies were resolved by discussion among the research team until consensus ratings for all participants were achieved.

Next, participants were selected if their index event was classified as one of the three event types used in the current study and did not overlap with the two other event types. This reduced the sample to 875. Finally, for the MVA and SA groups, only participants who directly experienced their index event, as opposed to witnessing or learning about it, were included. For the SUD group, participants were included if they directly experienced (n = 139), witnessed (n = 76), or learned about the event (n = 127). This resulted in a final sample of 854, including 253 MVA, 259 SA, and 342 SUD. Ethnic composition was 79.9% White (n = 683), 18.4% African American (n = 126), 1.5% Multi-racial (n = 13), 1.4% Asian or Pacific Islander (n = 12), 1.0% other (n = 9), <1% Hispanic (n = 8), and <1% Native American or Alaskan Native (n = 3). Mean age was 20.17 (SD = 2.36). The three groups differed significantly on age F(2,850) = 5.704, p = .003, with the mean age in the SA group (20.54) significantly higher than that of the MVA group (19.85) t(510) = 3.38, p = .001, but not significantly different from the SUD group (20.11) t(598) = 2.00, p = .052. Chi-square analysis revealed a difference in marital status across groups χ^2 (8) =18.47, p = .018; a greater proportion of the SA group were married, which may be related to this group's higher average age. There were no significant differences in ethnicity or work status across groups.

Finally, the PTSD prevalence in each group was examined. A dichotomous PTSD diagnosis was derived from the PCL-S by considering items rated as 3 or higher as an endorsed symptom, then following the *DSM-IV-TR* diagnostic rule of one B symptom, three C symptoms, and two D symptoms (Weathers, 2008). The groups differed significantly in the prevalence of PTSD, χ^2 (2, N = 854) = 65.35, $p < .001$. Sexual assault had the highest percentage of PTSD (37%), followed by SUD (15%) and MVA (11%). Total PCL-S scores averaged 30.43 (SD = 11.15) for the MVA group, 41.05 (SD = 13.48) for the SA group, and 32.69 (SD = 12.62) for the SUD group. These differences were significant F(2, 851) = 52.67 $p < .001$, with post hoc t-tests suggesting that total PCL-S scores in the SA group differed significantly from those of the MVA (t (510) = 9.69, $p < .001$) and SUD (t (599) = 7.80, $p < .001$) groups. Using a PCL total score of 44 as a cutoff for probable PTSD, 15% (n = 38) of the MVA group screened positive for PTSD, while 41% (n = 106) of the SA group and 18% (n = 60) of the SUD group screened positive.

Measures

The LEC is the self-report assessment portion of the Clinician-Administered PTSD Scale (CAPS; Blake et al., 1990; Weathers, Keane, & Davidson, 2001). The LEC consists of 17 items, including 16 items assessing exposure to specific categories of traumatic events (e.g., natural disaster, sexual assault) and one item labeled "other." Respondents indicate their lifetime exposure to each of the event categories by checking one or more of five options: happened to me, witnessed it, learned about it, not sure, and does not apply. The LEC has good temporal stability and convergent validity as a screening measure for trauma exposure (Gray, Litz, Hsu, & Lombardo, 2004).

In the present study, an extended research version of the LEC was used. After completing the standard LEC, participants identified their worst event, then responded to four items assessing Criterion A1, two items assessing Criterion A2, and one item assessing the number of times the participant had experienced an event comparable to the worst event. Next, participants wrote a brief narrative describing their worst event. Finally, they completed the PCL-S, rating their PTSD symptoms with respect to their worst event.

The PCL-S is a 17-item self-report measure that assesses each of the 17 *DSM-IV-TR* symptoms of PTSD (Weathers et al., 1993). Respondents identify a specific traumatic event, then indicate how much they were bothered by each symptom in the past month, using a 5-point scale (1 = not at all to 5 = extremely). The PCL-S has been used extensively in a wide variety of trauma populations and has excellent psychometric properties (Blanchard, Jones-Alexander, Buckley, & Forneris, 1996; Ruggiero, Del Ben, Scotti, & Rabalais, 2003; Weathers, 2008).

Data Analytic Strategy

A series of confirmatory factor analytic models were specified and estimated in Mplus version 6.0 (Muthén and Muthén, 2010; see Table 1). Based on convention established in previous CFA studies of PTSD, the following five models were selected for evaluation in each group separately: (1) the two-factor model reported in Asmundson and colleagues (2003), with Intrusions/Avoidance and Hyperarousal/Numbing (two-factor model); (2) the *DSM-IV-TR* three-factor model with Reexperiencing, Avoidance/Numbing, and Hyperarousal (three-factor model); (3) the King et al. (1998) four-factor numbing model with Reexperiencing, Avoidance, Numbing, and Hyperarousal (numbing model); (4) the Simms et al. (2002) four-factor dysphoria model with Intrusions, Avoidance, Dysphoria, Hyperarousal (dysphoria model); and, (5) a five-factor model with Reexperiencing, Avoidance, Numbing, Dysphoric Arousal, and Anxious Arousal (five-factor model; Elhai et al., 2011). In our initial CFA models, we handled scale dependency by fixing the variance of each latent variable to 1 and freely estimating the factor loadings. All factors were allowed to correlate and, for most models, no correlated errors or crossloadings were specified. Error covariances were estimated when theoretically defensible and empirically supported in all three groups based on modification indices (Brown, 2006). Given that modification indices can be inflated when sample sizes are large (Brown, 2006), we also examined the expected parameter change (EPC) for any proposed modification, which indicates the magnitude by which a specific parameter is expected to change (Kaplan; 1989, 1990).

Descriptive statistics for all three samples are provided for each PCL-S item in Table 2. For each sample, all available data were used in estimation, with missing data handled by direct maximum likelihood (ML) which is generally regarded as the best method for handling missing data in CFA (Brown, 2006). The proportion of complete data for each pair of variables was well above the recommended cutoff of 50% for the use of direct ML estimation in Mplus (range: .98 to 1.00; Muthén & Muthén, 1998-2010). Preliminary examination of PCL-S items showed univariate and multivariate nonnormality, as evidenced by examination of histograms for each item, as well as Mahalanobis distance values and Mardia's (1970) coefficients. Because PCL-S items were non-normally distributed, confirmatory factor analyses were conducted using *robust* maximum likelihood estimation (MLR; Yuan and Bentler, 2000) with the Yuan-Bentler scaled chi-square (Y-Bχ^2; Yuan & Bentler, 1999). Robust maximum likelihood estimates standard errors under conditions of multivariate nonnormality (Muthén & Muthén, 1998-2010). The Y-Bχ^2 is analogous to the Satorra-Bentler scaled chi-square (S-Bχ^2; Satorra and Bentler, 1994) in that it allows calculation of chi-square-dependent fit statistics that are robust to nonnormality. Based on a minimum sample size of 253 (MVA) and a minimum of 109 degrees of freedom (the five-factor model), it is estimated that power to detect close fit exceeded .95 across analyses (MacCallum, Browne, & Sugawara, 1996).

Model fit was evaluated using multiple complementary goodness-of-fit indices (Brown, 2006). Two indices of absolute fit, chi-square (χ^2) and standardized root mean square residual (SRMR), evaluate how similar the correlations observed in the input matrix are to the correlations predicted by the model. The comparative fit index (CFI; Bentler, 1990) and the Tucker-Lewis index (TLI; Tucker & Lewis, 1973) are indices of comparative fit in that they evaluate the fit of the specified model in relation to a more restricted baseline model in which the covariances of the input indicators are fixed to zero. The root mean square error of approximation (RMSEA; Steiger & Lind, 1980) is a parsimony corrected index in that it incorporates a penalty function for poor model parsimony (Brown, 2006). Finally, the Bayesian information criterion (BIC; Schwarz, 1978) allows comparison of non-nested models and also accounts for sample size and model complexity. Adequate model fit is indicated by CFI and TLI estimates of .90 or better ($\geq$ .95 indicates good model fit), SRMR estimates of .10 or lower ($\leq$ .08 indicates good model fit), and RMSEA estimates of .08 or lower ($\leq$ .06 indicates good model fit; Browne & Cudeck, 1993; Hu & Bentler, 1999). Lower BIC values indicate better model fit, with a 10-point difference suggesting a 150:1 likelihood ($p < .05$) that the model with the lower value fits best (Raftery, 1995). The chi-square statistic is the Yuan-Bentler scaled chi-square (Yuan & Bentler, 1999), which is a chi-square test of overall model fit for continuous non-normal outcomes that involves dividing the standard chi-square value by a scaling correction factor to adjust for non-normality.

Based on examination of the above goodness-of-fit indices, the models that showed the best fit across samples were selected for measurement invariance testing. Following established guidelines (Brown, 2006; Meredith, 1993; Meredith & Teresi, 2006), multiple-group CFAs (MGCFA) were run to test for invariance across groups in the following sequence: configural, metric, and scalar invariance. Measurement invariance testing is hierarchical in that the assumption of invariance must hold for the previous step prior to testing subsequent, more stringent assumptions. Thus, the models were run sequentially until invariance was no longer supported. Multiple groups were specified, with equality constraints specified to model increasingly stricter forms of measurement invariance, generating nested models for which model fit could be compared empirically. Equality across groups was tested in a pairwise manner (i.e. MVA with SA, MVA with SUD, SA with SUD) with equality constraints imposed on the appropriate parameters, such that parameters were freely estimated within the group, but constrained to be equal to the corresponding parameters in the comparison group. Though even stricter forms of measurement invariance can be tested (e.g., strict invariance), they are rare in applied research and are often considered to be overly restrictive and not central to measurement invariance evaluation (e.g., Bentler, 1995; Brown, 2006; Byrne, 1998). Therefore, measurement invariance analysis in this paper was stopped after the scalar invariance stage. For MGCFA analyses, scale dependency was handled by fixing a marker indicator to 1 with marker indicators selected based on R^2 values.

The primary test used for model comparison is the chi-square difference test where a non-significant increase in the chi-square value suggests invariance. However, chi-square difference testing is affected by sample size and the unequal sample sizes between the three groups used in MGCFA could bias results (Brown, 2006). Thus, the difference between the CFI values of the nested and comparison model was also evaluated, as simulation studies suggest that the performance of this fit index for MGCFA is superior to that of chi-square as it appears to be more robust to differences in sample size and model complexity (Cheung & Rensvold, 2002). A change in CFI of .01 or less is consistent with the presence of invariance.

Where non-invariance was found, the procedures for examining partial invariance set forth by Byrne, Shavelson, & Muthén (1989) were followed. Examination of modification indices, standardized residuals, and unstandardized factor loadings helped to identify sources of strain in the model. Where appropriate, one parameter at a time was selected to be freely estimated across groups while the other parameters remained constrained to equality. This process was repeated until the modified model was found to be invariant to the baseline model on the basis of the chi-square difference test and change in CFI, or until no further suggested modifications appeared.

Results

CFA

Results of the confirmatory factor analyses are presented in Table 3. Of the five models initially tested, the five-factor model showed the best overall fit in each sample as evidenced by the lowest BIC scores. In addition, both four-factor models demonstrated acceptable fit in indices of absolute and parsimony-corrected fit (SRMR and RMSEA respectively). Sources of strain were examined for each of the models that showed evidence of acceptable fit in all three samples. On the basis of prominent and consistent modification indices, as well as expected parameter change, a single error covariance was freed for estimation in each of the four-factor models.

For the numbing model, modification indices consistently (i.e. across all three groups) and strongly suggested allowing the error variances of PCL-S items 16 (*Hypervigilance*) and 17 (*Exaggerated startle response*) to correlate (MI range: 36.78 to 91.17; Completely Standardized Expected Parameter Change [EPC] range: .47 to .71). The error covariance of items 16 and 17 was the only modification suggested in all three samples, and the only modification of such magnitude. In addition to this empirical support, it was determined that theoretical support existed for allowing this error covariance as the items appear to share similarity in their wording that is not shared by the other items loading on the same factor. In addition, these items are the sole indicators of the Hyperarousal factor as arranged in the Simms et al., (2002) dysphoria model, suggesting that they share unique variance relative to the remaining hyperarousal symptoms. A prominent source of strain was also observed in the dysphoria model as estimating the error covariance of PCL-S items 10 (*Detachment*) and 11 (*Restricted affect*) was consistently suggested in modification indices in all three samples. It was the most prominent recommended modification in the MVA sample and among the most prominent suggested modifications in the SA and SUD samples (MI range: 12.94 to 78.98; EPC range: .28 to .75). As with the numbing model, the error covariance of items 10 and 11 was the only modification suggested in all three samples and the only modification of such magnitude. Items 10 and 11 also appear to share unique variance relative to the other items loading on the dysphoria factor as they both reference blunted emotions in an interpersonal context. The five-factor model did not have any consistent or prominent modification indices and thus was not modified in any way.

Following modification, each of the models appeared to fit well across the three samples (see Table 3). Based on BIC scores, the modified dysphoria model provided the best fit in the MVA sample (difference in BIC score >10), while the five-factor model fit the SA sample best (difference in BIC score > 10). The modified numbing model and five-factor model fit the SUD sample approximately equivalently (difference in BIC score = 5), both of which fit better than the dysphoria model (difference in BIC score > 10). The χ^2 values of the best fitting models were all statistically significant, suggesting model strain. However, χ^2 is sensitive to sample size and often suggests rejecting the model in large samples due to only minor discrepancies between the sample and implied covariance matrices (Brown, 2006; Tanaka, 1987). The χ^2/df ratio was less than two for each of the three best fitting models (values < 2.0 indicate a good fitting model). Overall, fit was poorest in the MVA sample, with some fit indices falling below established cutoffs for the modified numbing and five-factor models.

Parameter estimates for each of the best fitting models—including standardized and unstandardized factor loadings, standard errors, and factor correlations—are presented in Tables 4-6. In all three samples, all items loaded significantly on their specified latent factor (all *p*s < .001). For the modified numbing model, all completely standardized factor loadings ranged from .42 to .92, with the exception of PCL-S item 8 (*Inability to recall aspects of trauma*) which consistently loaded weakly across samples (from .20 to .32). For the modified dysphoria model, all completely standardized factor loadings ranged from .49 to .88, again with the exception of item 8 which loaded from .27 to .31 across samples. Completely standardized factor loadings ranged from .42 to .92 for the five-factor model, with loadings for item 8 between .20 and .32. R-square values for item 8 ranged from .041 to .104 in the modified numbing model, from .071 to .098 in the modified dysphoria model, and from .041 to .104 in the five-factor model, suggesting that the latent factors account for very little of the true variance of this item. These values were statistically significant at the $p = .05$ level in only the SA and SUD samples. As an index of reliability, the poor R-square values suggest that only a small portion of the variance of item 8 can be considered "true score" variance. Factor correlations ranged from .35 to .86 in the numbing model, from .47 to .86 in the dysphoria model, and from .41 to .86 in the five-factor model. The highest factor correlations were always between the reexperiencing and avoidance factors.

Measurement Invariance

Because both modified four-factor models and the five-factor model showed adequate fit across samples, each was maintained for examination of measurement invariance.

Numbing model.

Support for configural invariance was found as the modified numbing model fit adequately in each group (see Table 7). Support for metric invariance was also found across samples as evidenced by non-significant χ^2 difference tests and change in CFI (ΔCFI) of less than .01. Support for scalar invariance was not found as the χ^2 difference test was significant at the $p < .001$ level in each of the three pairwise comparisons. At this point, partial invariance at the scalar invariance level was examined by freeing one indicator intercept at a time based on modification indices and observing the χ^2 difference test values.

For the pairwise comparison of the MVA and SA samples, indicator intercepts were freed from equality (to be freely estimated) in a sequential fashion until no further modifications were suggested. This resulted in the free estimation of intercepts for items 12 and 3 in each sample while the remaining intercepts remained constrained to equality across the samples. Evidence for partial invariance was equivocal as the χ^2 difference test was significant at the $p < .05$ level and CFI and TLI values did not meet the cutoff for adequate fit, but the ΔCFI was less than .01 and the SRMR and RMSEA values were in the adequate range. Additional intercepts were not released as there were no further modifications suggested.

Similarly, after releasing the intercepts of items 16, 17, 8, 5, and 3, the pairwise comparison of the MVA and SUD samples generated a significant χ^2 difference test at the $p < .05$ level, but the ΔCFI was less than .01 and the remaining fit indices were all within the adequate range. Evidence for partial scalar invariance was found for the SA and SUD comparison when the intercepts for items 8, 5, 12, 15, 16, and 17 were freely estimated.

Dysphoria model.

Support for configural invariance was also found as the modified dysphoria model fit adequately in each group (see Table 8). Support for metric invariance was found only for the SA and SUD sample comparison. Support was not found for scalar invariance in the SA and SUD comparison. Indicator intercepts were freed sequentially until support for partial scalar invariance was found, resulting in the free estimation of items 8, 5, 11, and 12 across groups.

Because full metric invariance was not found MVA and SA comparison and in the MVA and SUD comparison, individual factor loadings were examined as potential sources of strain. Freeing individual loadings one at a time, as suggested by modification indices and examination of unstandardized factor loadings, resulted in partial invariance for the MVA and SA comparison when items 11, 10, and 9 were freed. For the MVA and SUD comparison, partial invariance required releasing the factor loadings of items 11, 10, and 9 to be freely estimated in each sample. Examination of scalar invariance proceeded with the partial invariance models (i.e., the factor loadings identified above remained freely estimated in the equal intercepts model) but failed to find support for equal intercepts in either comparison.

Five-factor model.

Support for configural and metric invariance was found for the five-factor model (see Table 9). Support for equal intercepts was not found in any of the three comparisons. Partial invariance analysis proceeded by releasing from equality each of the indicator intercepts identified by modification indices until no further intercepts were suggested. For the partial scalar invariance comparison of the MVA and SA samples, indicator intercepts for items 12 and 3 were released from equality. Again, evidence for partial invariance was equivocal as the S-B χ^2 difference test was significant at the $p < .05$ level, yet ΔCFI was less than .010 and the remaining fit indices were adequate or better. For the comparison of MVA and SUD, indicator intercepts for items 8, 5, and 3 were released, again resulting in equivocal support for partial invariance. Finally, the SA and SUD comparison followed the same pattern. After intercepts were released for items 8, 5, and 12, the S-B χ^2 difference test remained significant at the $p < .05$ level, while ΔCFI was less than .010 and the remaining fit indices were adequate or better.

Discussion

To date, this is the only study to examine differences in PTSD factor structure across distinct, strictly-defined, homogeneous trauma types. Previous studies examined differences in PTSD symptom profiles across trauma types (e.g., Armour & Shevlin, 2010; Kelley et al., 2009), but have not focused on the differences in the latent structure. Initial CFA results provided support for the limited equivalence of three proposed models of the symptom structure of PTSD (Elhai et al., 2011; King et al., 1998; Simms et al., 2002). The best fit overall across groups was found with a five-factor model (Elhai et al., 2011) that to this point has not been cross-validated. The previously well-supported four-factor numbing (King et al., 1998) and dysphoria models (Simms et al., 2002) also fit the data adequately, but each showed a clear and consistent source of strain. Because modification indices consistently suggested a single, prominent modification in each model, and because both suggested modifications also appeared theoretically justified, one error covariance was freely estimated in each model. This minimal model modification is consistent with previous CFA research that has allowed correlated errors (e.g., Baschnagel et al., 2005; Marshall et al., 2004; Norris et al., 2001). There were no prominent or consistently suggested modifications for the five-factor model, thus the two modified four-factor models and the five-factor model were compared with each other throughout the remainder of the analyses.

Though the five-factor model most consistently fit across samples, it was not always the absolute best-fitting model. The modified dysphoria model fit the MVA sample best, while the five-factor model was the best fit in the SA sample, and the modified numbing model and five-factor model were the best fitting models in the SUD sample. All item indicators loaded significantly on their respective factors and all completely standardized factor loadings, with the exception of item 8, were "salient" or reasonably good indicators of the latent factors. PCL-S item 8 (C3) has consistently loaded only weakly onto its designated factor in previous research (Foa et al., 1995; King et al., 1998; Lancaster et al., 2009; Palmieri et al., 2005; Simms et al., 2002) and the current results are consistent with that trend. The weak loading of item 8 and its related low R-square value suggests that the latent factor associated with this item accounts for only a small proportion of its variance. In other words, much of the variance of this item is error variance, suggesting poor reliability. This item has frequently been identified as a having poor psychometric properties and should be examined for revision in the newest edition of the DSM.

In each of the models, correlations between latent factors were in the moderate to high range. The highest factor correlation was consistently between the Reexperiencing and Avoidance factors in the MVA sample and exceeded commonly used cutoffs for determining if factors represent distinct constructs (e.g., .85). However, such high factor correlations are common in previous PTSD CFA research (e.g., Hoyt & Yeater, 2010; Krause et al., 2007; Palmieri et al., 2007), and there is theoretical justification for assuming that the constructs of reexperiencing and avoidance are highly related. Avoidance has been proposed as a coping strategy for preventing the distressing emotions that accompany reminders and memories of the event (Horowitz, 1982; McFarlane, 1992), and therefore can be expected to correlated strongly with reexperiencing symptoms.

The measurement invariance analyses were conducted to examine the consistency of PTSD's factor structure across samples, and to determine to what extent the PCL-S measurement properties generalize across samples (Vandenberg & Lance, 2000). Measurement invariance analyses showed that for all three models, the basic relationships between variables were consistent across samples (configural invariance). That is, the number of factors and the pattern of indicator-factor relationships remained the same in all three samples. In the modified numbing model and the five-factor model, factor loadings were equivalent across samples (metric invariance), suggesting that the indicators have comparable relationships to the latent factors in each of the samples tested. In other words, a unit change in the latent factor should be associated with an equivalent change in the observed indicators. However, this was not the case with the modified dysphoria model. Comparison of the MVA sample with the SA and SUD samples revealed that indicators consistently loaded more weakly on their respective factors in the MVA sample than they did in the other two samples. In the MVA sample, items 9 (*Loss of interest*), 10 (*Detachment*), and 11 (*Restricted affect*) were particularly problematic as their unstandardized factor loadings varied the most between the MVA sample and the other two samples. This suggests that these items may function differently for individuals who have experienced a motor vehicle accident compared with those who experienced a sexual assault or sudden unexpected death of a loved one. Factor loadings were equivalent between the SA and the SUD samples.

In contrast to the support for metric invariance, support for scalar invariance intercepts was not found for any of the three models. In the absence of scalar invariance, the same level of the latent factor may correspond with different observed scores across groups, and equal scores on the observed measure may correspond with different levels of the latent factor. Where scalar invariance is not found, factor means cannot be compared across groups. As Horn and McArdle (1992; p.117) stated: “If there is no evidence indicating the presence…of measurement invariance…findings of differences between individuals and groups cannot be unambiguously interpreted.” Previous research has established that different trauma types are associated with different PTSD profiles (e.g., Kelley et al., 2009). However, only three studies of measurement invariance of PTSD models have found support for scalar invariance, and each has examined a different four-factor model (Hetzel-Riggin et al., 2009; Marshall, 2004; Wang et al., 2011). Thus, there is limited justification for the comparison of PTSD factor scores across groups and conclusions based on such comparisons should be tempered.

When a significant omnibus test of invariance suggests non-invariance, fit diagnostics can be utilized to identify the specific parameters responsible for the non-invariance. The source(s) of the lack of invariance of item intercepts were explored through strategies suggested by Byrne, Shavelson, & Muthén (1989). Unequivocal support for partial invariance of item intercepts was found only for the SA and SUD comparison within the modified numbing and modified dysphoria models. However, this partial invariance was achieved only after releasing at least five indicator intercepts to be estimated. The most common problematic items were 6, 5, 8, 12, and 3 in order of the frequency that they were identified as having non-invariant intercepts across groups. This preponderance of unequal intercepts (and the unequal factor loadings in the modified dysphoria model) again contraindicates group mean comparisons. Though measurement invariance analysis can continue when partial invariance is found, it was determined that there were too many non-invariant parameters to justify any interpretation of findings of strict invariance (Brown, 2006).

Though the five-factor model showed promising fit across all groups, it remains in need of further construct validation before it can be compared on the same level as the two four-factor models. One potential shortcoming of the five-factor model is the tendency of measurement models to show improvement in fit as they become more complex. That is, there is a general preference for more parsimonious models as complex models run the risk of becoming over-fit to idiosyncrasies of the original sample. Future studies should explore relationships of the five-factor model with external measures of psychopathology to determine if each factor is indeed distinct. Structural equation modeling (SEM) could be used to regress external measures of psychopathology onto each of the five latent factors. Correlational analyses with external measures would also be useful to determine if the dysphoric and anxious arousal factors show evidence of discriminant validity.

Each of the three models shown to fit the data in this study is closely related. The numbing and dysphoria models differ only in their assignment of PCL-S items 13-15 to either the hyperarousal factor (King et al., 1998) or the dysphoria factor (Simms et al., 2002). The five-factor model differs from the numbing model only in its division of the five hyperarousal items into two distinct factors (Elhai et al., 2011). This degree of overlap likely explains why CFA studies continue to find report mixed findings, some in support of the numbing model, others in support of the dysphoria model, and most finding similar fit between models (e.g., Palmieri et al., 2007). Thus, progress in identifying the most accurate representation of PTSD's latent structure may rely on studies of the generalizability of proposed models (King et al., 2006). Demonstrating that a particular model is invariant across samples and over time provides validity evidence for the construct as proposed. In this study, despite similar fit across groups in the initial CFA, the numbing model and the five-factor model achieved the highest level of invariance across trauma types. Given that the five-factor model lacks cross-validation in other samples and external validation and is less parsimonious than the numbing model, the present study appears to support the four-factor numbing model as the most accurate representation of the latent structure of PTSD.

This study is also consistent with previous measurement invariance studies that have largely failed to find support for scalar invariance of the numbing and dysphoria models (e.g., Anthony et al., 1995; Hoyt & Yeater, 2010; Mansfield et al., 2010; Suvak et al., 2008). This consistent finding is evidence that measures of PTSD do not function equivalently across groups and should not be compared without careful consideration. Examination of partial scalar invariance across studies might identify those items that should not be included in cross-group comparisons. Unfortunately, other studies have not examined partial invariance at the item indicator level.

Most importantly, as understanding of the latent structure of PTSD improves, so does understanding of treatment selection and course, identification of risk factors, and accuracy of measurement. For example, as formulated in *DSM-IV-TR*, the avoidance and numbing items are combined in a single factor. However, Asmundson and colleagues (2004) summarized a number of findings that would have been obscured given the assumption that avoidance and numbing symptoms share a unified latent factor. For instance, high emotional numbing at baseline is predictive of poorer response to cognitive-behavioral treatments, whereas a high level of avoidance is not (Taylor, Fedoroff, Koch, Thorardson, Fecteau, & Nicki, 2001). Similarly, emotional numbing correlates higher with depression than does avoidance (Asmundson, Stein, & McCreary, 2002), and exposure therapy outperforms eye movement desensitization and reprocessing for symptoms of avoidance, but not for symptoms of emotional numbing (Taylor, Thordarson, Maxfield, Redoroff, Lovell, & Ogrodniczuk, 2003). It is possible that accurate placement of the D1-D3 symptoms in latent models will spur further advancement in the understanding of the assessment, treatment, and prediction of PTSD, especially as additional research explores correlates of the distinct factors.

Results of the current study are qualified by several methodological limitations. First, the use of a non-clinical sample of trauma-exposed undergraduate participants may limit its generalizability to clinical populations with higher levels of PTSD. Nonetheless, PTSD diagnostic status and PCL-S total scores met or exceeded those of other CFA studies in undergraduate (e.g., Baschnagel et al., 2005; Elhai et al., 2009; Hoyt & Yeater, 2010; Lancaster et al., 2009) and non-clinical populations (e.g., Asmundson et al., 2000; Marshall, Schell, & Miles, 2010). PCL-S scores also met or exceeded those reported in more severely trauma-exposed samples (e.g., Elhai et al., 2010; Engdahl et al., 2011; Palmieri et al., 2007). Though the majority of participants did not meet full criteria for PTSD, previous taxometric research suggests that PTSD is a dimensional construct, and thereby permits evaluation across the full range of symptom severity (Ruscio, Ruscio, & Keane, 2002). Another limitation is found in the self-selected nature of the sample, a characteristic that may prevent comparison to the broad population of civilian trauma survivors. The sample was also predominantly White and entirely female, potentially limiting generalizability to more diverse populations and male participants. Finally, the trauma groups were intentionally designed to be non-overlapping in order to make comparisons across trauma types possible. This tactic may have created atypical samples, as risk of PTSD is increased with exposure to multiple traumatic events (Breslau, Chilcoat, Kessler, & Davis, 1999).

Inherent in each of the three models tested are problems with underidentified factors. Each specifies only two indicators for the avoidance factor, and the dysphoria and five-factor models both propose a hyperarousal (or anxious arousal) factor with only two indicators. Ideally, each factor would include more indicators to ensure that the domain is fully measured, and that it is stable (Brown, 2006). However, this study was designed to examine existing models of PTSD, not to propose alternative models. Future research should explore additional indicators for the avoidance and hyperarousal factors. An additional limitation of the current study is the reliance on a self-report measure of PTSD rather than including a diagnostic interview. Many studies have identified method idiosyncrasies that may influence CFA results (e.g., Elhai et al., 2010, Palmieri et al., 2007; Scher et al., 2008; Witteveen, 2006), and it is possible that the findings are specific to the structure and stability of the PCL rather than to PTSD in general. However, the numbing model has been supported in multiple samples using multiple measures, including the gold-standard structured clinical interview for PTSD (e.g., CAPS; King et al., 1998; Palmieri et al., 2007). Although the PCL-S is the gold-standard self-report measure of PTSD and has excellent psychometric properties, structured clinical interviews can provide for a more valid assessment through follow-up questions and probes that help clarify responses. As Palmieri et al. (2007) note, self-reporting of PTSD symptoms may be hindered if participants do not realize that they are engaging in avoidance behavior, do not recognize that they are emotionally numb, or do not find their hyperarousal behaviors to be unusual. Therefore, future research should explore the factor structure of PTSD across similarly distinct trauma groups utilizing structured clinical interviews for assessment.

Structural studies of this sort are an important contribution to the ongoing validation of the PTSD construct. Moreover, this study is in line with recent efforts to explore the generalizability of PTSD's factor structure across settings, time, trauma types, and demographic differences. Establishing that a model of PTSD is invariant across samples is a prerequisite for comparing group means and generally evaluating the claim that PTSD is an equivalent syndrome. This study is consistent with the majority of measurement invariance investigations of PTSD which have failed to find the necessary equivalence of parameters in order to justify the direct comparison of PTSD measurement across groups. Although these findings need to be replicated across other trauma types and using other self-report and interview assessment methods, the current study serves to alert the field of traumatic stress that PTSD may not be an equivalent syndrome across trauma types. These findings apply only to the current conceptualization of PTSD and to extant measures of PTSD. With the upcoming revision of the DSM, including the likely addition of three PTSD symptoms and the reorganization into four clusters, these matters of measurement invariance will need to be revisited with the revised criteria and associated measurement tools.

References

Amdur, R. L., & Liberzon, I. (2001). The structure of posttraumatic stress disorder symptoms in combat veterans: a confirmatory factor analysis of the Impact of Event Scale. *Journal of Anxiety Disorders, 15*, 345-357.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA/APA/NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

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

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text Revision). Washington, DC: Author.

Andrews, L., Joseph, S., Shevlin, M., & Troop, N. (2006). Confirmatory factor analysis of posttraumatic stress symptoms in emergency personnel: an examination of seven alternative models. *Personality and Individual Differences, 41,* 213-224.

Anthony, J. L., Lonigan, C. J., & Hecht, S. A. (1999). Dimensionality of posttraumatic stress disorder symptoms in children exposed to disaster: results from confirmatory factor analyses. *Journal of Abnormal Psychology*, *108,* 326-336.

Anthony, J. L., Lonigan, C. J., Vernberg, E. M., La Greca, A. M., Silverman, W. K., & Prinstein, M. J. (2005). Multisample Cross-Validation of a Model of Childhood Posttraumatic Stress Disorder Symptomatology. *Journal of Traumatic Stress*, *18*, 667-676.

Armour, C., Layne, C. M., Naifeh, J. A., Shevlin, M., Durakovic-Belko, E., Djapo, N., Pynoos, R. S., & Elhai, J. D. (2011). Assessing the factor structure of posttraumatic stress disorder symptoms in war-exposed youths with and without criterion A2 endorsement. *Journal of Anxiety Disorders, 25*, 80-87.

Armour, C. & Shevlin, M. (2010). Testing the dimensionality of PTSD and the specificity of the dysphoria factor. *Journal of Loss and Trauma: International Perspectives on Stress & Coping*, *15*, 11-27.

Asmundson, G. J. G., Frombach, I., McQuaid, J., Pedrelli, P., Lenox, R., & Stein, M. B. (2000). Dimensionality of posttraumatic stress symptoms: a confirmatory factor analysis of DSM-IV symptom clusters and other symptom models. *Behaviour Research and Therapy, 38,* 203-214.

Asmundson, G. G., Stein, M. B., & McCreary, D. R. (2002). Posttraumatic stress disorder symptoms influence health status of deployed peacekeepers and nondeployed military personnel. *Journal of Nervous and Mental Disease, 190*, 807-815.

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

Baschnagel, J. S., O’Conner, 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.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246.

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Charney, D. S., & Keane, T. M. (2004). Clinician-Administered PTSD Scale. Los Angeles: Western Psychological Services.

Blake, D. D., Weathers, F. W., Nagy, L., Kaloupek, D. G., Klauminzer, G., Charney, D. S. (1990). A clinician rating scale for assessing current and lifetime PTSD: The CAPS-1. *The Behavior Therapist*, *18,* 187-188.

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

Boelen, P. A., van den Hout, M. A., & van den Bout, J. (2008). The factor structure of posttraumatic stress disorder symptoms among bereaved individuals: A confirmatory factor analysis study. *Journal of Anxiety Disorders*, *22*, 1377-1383.

Breslau, N., Chilcoat, H. D., Kessler, R. C., & Davis, G. C. (1999). Previous exposure to trauma and PTSD effects of subsequent trauma: Results from the Detroit Area Survey of Trauma. *The American Journal of Psychiatry*, *156*, 902-907.

Breslau, N., Kessler, R. C., Chilcoat, H. D., Schultz, L. R., Davis, G. C., & Andreski, P. (1998). Trauma and posttraumatic stress disorder in the community: The 1996 Detroit Area Survey of Trauma. *Archives of General Psychiatry, 55,* 626-632.

Brett, E. A., Spitzer, R. L., & Williams, J. B. (1988). DSM-III-R criteria for posttraumatic stress disorder. *American Journal of Psychiatry, 145,* 1232-1236.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Brown, T. A., & Barlow, D. H. (2009). A proposal for a dimensional classification system based on the shared features of the DSM-IV anxiety and mood disorders: Implications for assessment and treatment. *Psychological Assessment, 21*, 256–271.

Browne, M. W., & Cudeck, R. (1993). Alternate ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*, 456-466.

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

Carragher, N., Mills, K., Slade, T., Teesson, M., & Silove, D. (2010). Factor structure of posttraumatic stress disorder symptoms in the Australian general population. *Journal of Anxiety Disorders*, *24*, 520-527.

Cattell, R. B. (1978). *The scientific use of factor analysis in the behavioral and life sciences.* New York: Plenum Press.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233-255.

Clark, L. A., & Watson, D. (2006). Distress and fear disorders: An alternative empirically based taxonomy of the “mood” and “anxiety” disorders. *British Journal of Psychiatry, 189,* 481–483.

Cordova, M. J., Studts, J. L., Hann, D. M., Jacobsen, P. B., & Andrykowski, M. A. (2000). Symptom structure of PTSD following breast cancer. *Journal of Traumatic Stress, 13,* 301-319.

Cox, B. J., Moata, N., Clara, I., & Asmundson, G. J. G. (2008). The symptom structure of posttraumatic stress disorder in the National Comorbidity Replication Survey. *Journal of Anxiety Disorders, 22*, 1523-1528.

DuHamel, K. N., Ostrof, J., Ashman, T., Winkel, G., Mundy, E. A., Keane, T. M., et al. (2004). Construct Validity of the Posttraumatic Stress Disorder Checklist in Cancer Survivors: Analyses Based on Two Samples. *Psychological Assessment*, *16*, 255-266.

Elhai, J. D., Biehn, T. L., Armour, C., Klopper, J. J., Frueh, B. C., & Palmieri, P. A. (2011). Evidence for a unique PTSD construct represented by PTSD’s D1-D3 symptoms. *Journal of Anxiety Disorders*, *25,* 340-345.

Elhai, J. D., Engdahl, R. M., Palmieri, P. A, Naifeh, J. A., Schweinle, A., & Jacobs, G. A. (2009). Assessing posttraumatic stress disorder with or without reference to a single, worst traumatic event: Examining differences in factor structure. *Psychological Assessment*, *21*, 629-634.

Elhai, J. D., Frueh, C. B., Gold, P. B., Gold, S. N., & Hamner, M. B. (2000). Clinical presentations of posttraumatic stress disorder across trauma populations: A comparison of MMPI-2 profiles of combat veterans and adult survivors of child sexual abuse. *Journal of Nervous and Mental Disease*, *188*, 708-713.

Elhai, J. D., Gray, M. J., Docherty, A. R., Kashdan, T. B., & Kose, S. (2007). Structural validity of the Posttraumatic Stress Disorder Checklist among college students with a trauma history. *Journal of Interpersonal Violence, 22,* 1471-1478.

Elhai, J. D., Palmieri, P. A., Biehn, T. L., Frueh, B., & Magruder, K. M. (2010). Posttraumatic stress disorder's frequency and intensity ratings are associated with factor structure differences in military veterans. *Psychological Assessment*, *22*, 723-728.

Elklit, A., Armour, C., & Shevlin, M. (2010). Testing alternative factor models of PTSD and the robustness of the dysphoria factor. *Journal of Anxiety Disorders*, *24*, 147-154.

Elklit, A., & Shevlin, M. (2007). The structure of PTSD symptoms: a test of alternative models using confirmatory factor analysis. *British Journal of Clinical Psychology, 46,* 299-313.

Engdahl, R. M., Elhai, J. D., Richardson, J. D., & Frueh, B. C. (2011). Comparing posttraumatic stress disorder's symptom structure between deployed and nondeployed veterans. *Psychological Assessment*, 23, 1-6.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4,* 272-299.

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.

Foa, E. B., Riggs, D. S., & Gershuny, B. S. (1995). Arousal, numbing, and intrusion: Symptom structure of PTSD following assault. *American Journal of Psychiatry, 152,* 116-120.

Foa, E. B., Zinbarg, R., & Rothbaum, B. O. (1992). Uncontrollability and unpredictability in post-traumatic stress disorder: An animal model. *Psychological Bulletin*, *112*, 218-238.

Foy, D. W., Wood, J. L., King, D. W., King, L. A., & Resnick, H. S. (1997). Los Angeles Symptom Checklist: Psychometric evidence with an adolescent sample. *Assessment, 4*, 377-384.

Friedman, M. J., Keane, T. M., & Resick, P. A. (2007). *Handbook of PTSD: Science and practice*. New York, NY: Guilford Press.

Grant, D. M., Beck, J. G., Marques, L., Palyo, S. A., & Clapp, J. D. (2008). The structure of distress following trauma: Posttraumatic stress disorder, major depressive disorder, and generalized anxiety disorder. *Journal of Abnormal Psychology*, *117*, 662-672.

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

Green, B. L., Krupnick, J. L., Stockton, P., Goodman, L., Corcoran, C., & Petty, R. (2001). Psychological outcomes associated with traumatic loss in a sample of young women. *The American Behavioral Scientist, 44*, 817-837.

Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care, 44*, S78–S94.

Grubaugh, A. L., Long, M. E., Elhai, J. D., Frueh, B., & Magruder, K. M. (2010). An examination of the construct validity of posttraumatic stress disorder with veterans using a revised criterion set. *Behaviour Research and Therapy*, *48*, 909-914.

Hetzel-Riggin, M. (2009). A test of structural invariance of posttraumatic stress symptoms in female survivors of sexual and/or physical abuse or assault. *Traumatology*, *15*, 46-59.

Horn, J. L. & McArdle, J. J. (1992). A practical guide to measurement invariance in aging research. *Experimental Aging Research, 18,* 117-144.

Horowitz, M. (1982). Stress response syndromes and their treatment. In L. Goldberger, & S. Breznitz (Eds.), *Handbook of stress: theoretical and clinical aspects* (pp. 711–732). New York: The Free Press.

Horowitz, M. J., Wilner, N., & Alvarez, W. (1979). Impact of Event Scale: A measure of subjective stress. *Psychosomatic Medicine*, *41*, 209-218.

Hovens, J. E., van der Ploeg, H. M., Bramsen, I., Klaarenbeek, M. T., Schreuder, J. N., & Rivero, V. V. (1994). The development of the Self-Rating Inventory for Posttraumatic Stress Disorder. *Acta Psychiatrica Scandinavica, 90,* 172-183.

Hoyt, T., & Yeater, E. A. (2010). Comparison of posttraumatic stress disorder symptom structure models in Hispanic and White college students. *Psychological Trauma: Theory, Research, Practice, and Policy*, *2*, 19-30.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Kaplan, D. (1989). Model modification in covariance structure analysis: Application of the expected parameter change statistic. *Multivariate Behavioral Research, 24,* 285-305.

Kaplan, D. (1990). Evaluating and modifying covariance structure models: A review and recommendation. *Multivariate Behavioral Research, 25,* 137-155.

Kelley, L. P., Weathers, F. W., McDevitt-Murphy, M. E., Eakin, D. E., & Flood, A. M. (2009). A comparison of PTSD symptom patterns in three types of civilian trauma. *Journal of Traumatic Stress*, *22*, 227-235.

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

King, L. A., King, D. W., Orazem, R. J., & Palmieri, P. A. (2006). Research on the latent structure of PTSD. *PTSD Research Quarterly, 17,* 1-7.

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.

King, D. W., Orazem, R. J., Lauterbach, D., King, L. A., Hebenstreit, C. L., & Shalev, A. Y. (2009). Factor structure of posttraumatic stress disorder as measured by the Impact of Event Scale–Revised: Stability across cultures and time. *Psychological Trauma: Theory, Research, Practice, and Policy*, *1*, 173-187.

Kirz, J. L., Drescher, K. D., Klein, J. L., Gusman, F. D., & Schwartz, M. F. (2001). MMPI-2 assessment of differential post-traumatic stress disorder patterns in combat veterans and sexual assault victims. *Journal of Interpersonal Violence, 16,* 619-639.

Krause, E. D., Kaltman, S., Goodman, L. A., & Dutton, M. A. (2007). Longitudinal factor structure of posttraumatic stress symptoms related to intimate partner violence. *Psychological Assessment, 19,* 165-175.

Lancaster, S. L., Melka, S. E., & Rodriguez, B. F. (2009). A factor analytic comparison of five models of PTSD symptoms. *Journal of Anxiety Disorders*, *23*, 269-274.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*, 130-149.

Maes, M., Delmeire, L., Schotte, C., Janca, A., Creten, T., Mylle, J., et al. (1998). The two-factorial symptom structure of post-traumatic stress disorder: Depression—avoidance and arousal—anxiety. *Psychiatry Research, 81*, 195-210.

Mansfield, A. J., Williams, J., Hourani, L. L., & Babeu, L. A. (2010). Measurement invariance of posttraumatic stress disorder symptoms among U.S. military personnel. *Journal of Traumatic Stress*, *23*, 91-99.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519-530.

Marshall, G. N. (2004). Posttraumatic stress disorder symptom checklist: factor structure and English–Spanish measurement invariance. *Journal of Traumatic Stress, 17,* 223-230.

Marshall, G. N., Schell, T. L., & Miles, J. V. (2010). All PTSD symptoms are highly associated with general distress: Ramifications for the dysphoria symptom cluster. *Journal of Abnormal Psychology*, *119*, 126-135.

McDonald, S. D., Beckham, J. C., Morey, R., Marx, C., Tupler, L. A., & Calhoun, P. S. (2008). Factorial invariance of posttraumatic stress disorder symptoms across three veteran samples. *Journal of Traumatic Stress, 21,* 309-317.

McFarlane, A. C. (1992). Avoidance and intrusion in posttraumatic stress disorder. *Journal of Nervous and Mental Disease,180,* 439–445.

McWilliams, L. A., Cox, B. J., & Asmundson, G. J. G. (2005). Symptom structure of posttraumatic stress disorder in a nationally representative sample. *Journal of Anxiety Disorders, 19,* 626-641.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525-543.

Meredith, W., & Teresi, J. A. (2006). An Essay on Measurement and Factorial Invariance. *Medical Care, 44,* S69-S77.

Miller, M. W. (2003). Personality and the etiology and expression of PTSD: A three-factor model perspective. *Clinical Psychology: Science and Practice*, *10*, 373–393.

Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49*, 377–412.

Morina, N., Böhme, H. F., Ajdukovic, D., Bogic, M., Franciskovic, T., Galeazzi, G. M., & ... Priebe, S. (2010). The structure of post-traumatic stress symptoms in survivors of war: Confirmatory factor analyses of the Impact of Event Scale—Revised. *Journal of Anxiety Disorders*, *24*, 606-611.

Muthén, L. K., & Muthén, B. O. (1998-2007). *Mplus User's Guide. Fifth Edition.* Los Angeles, CA: Muthén & Muthén.

Naifeh, J. A., Elhai, J. D., Kashdan, T. B., & Grubaugh, A. L. (2008). The PTSD Symptom Scale's latent structure: An examination of trauma-exposed medical patients. *Journal of Anxiety Disorders*, *22*, 1355-1368.

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

Olff, M., Sijbrandij, M., Opmeer, B. C., Carlier, I. V. E., & Gersons, B. P. R. (2009). The structure of acute posttraumatic stress symptoms: ‘Reexperiencing’, Active Avoidance’, ‘Dysphoria’, and ‘Hyperarousal’. *Journal of Anxiety Disorders, 23,* 656-659.

Palmieri, P. A, & Fitzgerald, L. F. (2005). Confirmatory factor analysis of posttraumatic stress symptoms in sexually harassed women. *Journal of Traumatic Stress, 18,* 657–666.

Palmieri, P. A., Marshall, G. N., & Schell, T. L. (2007). Confirmatory factor analysis of posttraumatic stress symptoms in Cambodian refugees. *Journal of Traumatic Stress, 20,* 207-216.

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 Centre Ground Zero. *Journal of Abnormal Psychology, 116,* 329-341.

Pietrzak, R. H., Goldstein, M. B., Malley, J. C., Rivers, A. J., & Southwick, S. M. (2010). Structure of posttraumatic stress disorder symptoms and psychosocial functioning in Veterans of Operations Enduring Freedom and Iraqi Freedom. *Psychiatry Research, 178,* 323-329.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methods, 25,* 111-163.

Rasmussen, A., Smith, H., & Keller, A. S. (2007). Factor structure of PTSD symptoms among West and Central African refugees. *Journal of Traumatic Stress, 20,* 271-280.

Rosen, G. M., & Lilienfeld, S. O. (2008). Posttraumatic stress disorder: An empirical evaluation of core assumptions. *Clinical Psychology Review, 28,* 837-868.

Ruggiero, K. J., Del Ben, K., Scotti, J. R., & Rabalais, A. E. (2003). Psychometric properties of the PTSD-Checklist, civilian version. *Journal of Traumatic Stress, 16,* 495-502.

Ruscio, A. M., Ruscio, J., & Keane, T. M. (2002). The latent structure of posttraumatic stress disorder: A taxometric investigation of reactions to extreme stress. *Journal of Abnormal Psychology, 111,* 290-301.

Sack, W. H., Seeley, J. R., & Clarke, G. N. (1997). Does PTSD transcend cultural barriers? A study from the Khmer Adolescent Refugee Project. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 49-54.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye, C. C. Clogg, A. von Eye, C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399-419). Thousand Oaks, CA US: Sage.

Scher, C. D., McCreary, D. R., Asmundson, G. J. G., & Resick, P. A. (2008). The structure of post-traumatic stress disorder symptoms in three female trauma samples: A comparison of interview and self-report measures. *Journal of Anxiety Disorders, 22*, 1137-1145.

Schinka, J. A., Brown, L. M., Borenstein, A. R., & Mortimer, J. A. (2007). Confirmatory factor analysis of the PTSD checklist in the elderly. *Journal of Traumatic Stress, 20,* 281-289.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*, 461-464.

Shevlin, M., Hunt, N., & Robbins, I. (2000). A confirmatory factor analysis of the Impact of Event Scale using a sample of World War Two and Korean War veterans. *Psychological Assessment*, *12,* 414-417.

Shevlin, M., McBride, O., Armour, C., & Adamson, G. (2009). Reconciling the differences between the King et al. (1998) and Simms et al. (2002) factor models of PTSD. *Journal of Anxiety Disorders*, *23*, 995-1001.

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

Smith, M. Y., Redd, W., DuHamel, K., Vickberg, S. J., & Ricketts, P. (1999). Validation of the PTSD Checklist—Civilian Version in survivors of bone marrow transplantation. *Journal of Traumatic Stress*, *12*, 485-499.

Steiger, J. H., & Lind, J. M. (1980, June). *Statistically based tests for the number of common factors*. Paper presented at the meeting of the Psychometric Society, Iowa City, IA.

Suvak, M., Maguen, S., Litz, B. T., Silver, R., & Holman, E. A. (2008). Indirect exposure to the September 11 terrorist attacks: Does symptom structure resemble PTSD?. *Journal of Traumatic Stress*, *21*, 30-39.

Taft, C. T., Kaloupek, D. G., Schumm, J. A., Marshall, A. D., Panuzio, J., King, D. W., & Keane, T. M. (2007). Posttraumatic stress disorder symptoms, physiological reactivity, alcohol problems, and aggression among military veterans. *Journal of Abnormal Psychology, 116*, 498-507.

Tanaka, J. S. (1987). 'How big is big enough?': Sample size and goodness of fit in structural equation models with latent variables. *Child Development*, *58*, 134-146.

Taylor, S., Fedoroff, I. C., Koch, W. J., Thordarson, D. S., Fecteau, G., & Nicki, R. M. (2001). Posttraumatic stress disorder arising after road traffic collisions: Patterns of response to cognitive–behavior therapy. *Journal of Consulting and Clinical Psychology*, *69,* 541-551.

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.

Taylor, S., Thordarson, D. S., Maxfield, L., Fedoroff, I. C., Lovell, K., & Ogrodniczuk, J. (2003). Comparative efficacy, speed, and adverse effects of three PTSD treatments: Exposure therapy, EMDR, and relaxation training. *Journal of Consulting and Clinical Psychology, 71*, 330-338.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1-10.

Ullman, S. E., & Long, S. M. (2008). Factor structure of PTSD in a community sample of sexual assault survivors. *Journal of Trauma & Dissociation, 9,* 507-524.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*, 4-69.

van der Ploeg, E. M., Mooren, T. T. M., Kleber, R. J., van der Velden, P. G., & Brom, D. (2004). Construct validation of the Dutch version of the Impact of Events Scale. *Psychological Assessment*, *16*, 16-26.

Wang, L., Zhang, J., Shi, Z., Zhou, M., Huang, D., & Liu, P. (2011). Confirmatory factor analysis of posttraumatic stress symptoms assessed by the Impact of Event Scale—Revised in Chinese earthquake victims: Examining factor structure and its stability across sex. *Journal of Anxiety Disorders, 25*, 369-375.

Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology, 114*, 522–536.

Weathers, F. W. (2008). PTSD Checklist. In G. Reyes, J. D. Elhai, & J. D. Ford (Eds.), Encyclopedia of psychological trauma (pp. 491-494). Hoboken, NJ: Wiley.

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 and Anxiety, 13*, 132-156.

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

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

Wilson, J. P. (2004). PTSD and Complex PTSD: Symptoms, Syndromes, and Diagnoses. In: J. P. Wilson, & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD (2nd ed.)* (pp. 7-44). New York: Guilford Press.

Wilson, J. P., Smith, W. K., & Johnson, S. K. (1985). A comparative analysis of PTSD among various survivor groups. In C. R. Figley (Ed.), Trauma and its wake: Volume 1 (pp. 142-172). New York: Bruner/Mazel.

Witteveen, A. B., van der Ploeg, E. M., Bramsen, I., Huizink, A. C., Slottje, P., Smid, T., et al. (2006). Dimensionality of the posttraumatic stress response among police officers and fire fighters: An evaluation of two self-report scales. *Psychiatry Research*, *141*, 213-228.

Yuan, K., & Bentler, P. M. (1999). F tests for mean and covariance structure analysis. *Journal of Educational and Behavioral Statistics*, *24*, 225-243.

Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. In M. E. Sobel & M. P. Becker (Eds.), *Sociological Methodology 2000* (pp. 165-200). Washington, DC: American Sociological Association.

Yufik, T. & Simms, L. J. (2010). A meta-analytic investigation of the structure of posttraumatic stress disorder symptoms. *Journal of Abnormal Psychology, 119*, 764-776.

Appendix: Tables

Table 1
Item Mapping for CFA Models

PCL-S Item	DSM-IV PTSD Symptoms		Model 1	2	3	3a*	4	4a†	5
1	B1	Intrusive thoughts of trauma	I,A	R	R	R	I	I	R
2	B2	Recurrent dreams of trauma	I,A	R	R	R	I	I	R
3	B3	Flashbacks	I,A	R	R	R	I	I	R
4	B4	Emotional reactivity to trauma cues	I,A	R	R	R	I	I	R
5	B5	Physiological reactivity to trauma cues	I,A	R	R	R	I	I	R
6	C1	Avoiding thoughts of trauma	I,A	A	A	A	A	A	A
7	C2	Avoiding reminders of trauma	I,A	A	A	A	A	A	A
8	C3	Inability to recall aspects of trauma	N,H	A	N	N	D	D	N
9	C4	Loss of interest	N,H	A	N	N	D	D	N
10	C5	Detachment	N,H	A	N	N	D	D	N
11	C6	Restricted affect	N,H	A	N	N	D	D	N
12	C7	Sense of foreshortened future	N,H	A	N	N	D	D	N
13	D1	Sleep disturbance	N,H	H	H	H	D	D	DA
14	D2	Irritability	N,H	H	H	H	D	D	DA
15	D3	Difficulty concentrating	N,H	H	H	H	D	D	DA
16	D4	Hypervigilance	N,H	H	H	H	H	H	AA
17	D5	Exaggerated startle response	N,H	H	H	H	H	H	AA

Note. PTSD = posttraumatic stress disorder. I = Intrusions; R = Reexperiencing; A = Avoidance; N = Emotional Numbing; H = Hyperarousal; D = Dysphoria; DA = Dysphoric Arousal; AA = Anxious Arousal.
* = Model 3 was modified allowing estimation of an error covariance between PCL16 and PCL17.
† = Model 4 was modified allowing estimation of an error covariance between PCL10 and PCL11.

Table 2
Descriptive Statistics of PCL Items by Sample

PCL-S Item	DSM-IV PTSD Symptom	MVA Sample (*n* = 253)					SA Sample (*n* = 259)					SUD Sample (*n* = 342)				
		M	SD	SE	Skew	Kurt	M	SD	SE	Skew	Kurt	M	SD	SE	Skew	Kurt
1	B1	2.39	1.08	.068	0.55	-0.43	2.67	1.16	.072	.40	-0.68	2.59	1.22	.066	0.34	-0.94
2	B2	1.63	0.96	.060	1.48	1.27	2.05	1.19	.074	0.93	-0.17	1.95	1.13	.061	1.10	0.43
3	B3	1.82	1.04	.066	1.09	0.25	1.85	1.11	.069	1.33	1.01	1.71	1.06	.057	1.52	1.62
4	B4	2.49	1.32	.083	0.56	-0.86	3.10	1.26	.078	-0.05	-1.09	2.96	1.22	.066	0.11	-0.99
5	B5	2.07	1.22	.077	0.89	-0.33	2.42	1.29	.080	0.52	-0.87	1.89	1.19	.064	1.16	0.25
6	C1	1.87	1.16	.073	1.28	0.72	3.52	1.31	.081	-0.43	-1.00	2.57	1.39	.075	0.37	-1.19
7	C2	1.97	1.17	.074	1.10	0.28	2.74	1.33	.083	0.16	-1.17	2.14	1.30	.071	0.87	-0.47
8	C3	2.00	1.35	.085	1.12	-0.07	2.40	1.40	.087	0.54	-1.05	1.58	1.00	.054	1.74	2.12
9	C4	1.26	0.63	.039	2.77	8.52	1.90	1.11	.069	1.15	0.46	1.50	0.97	.052	2.10	3.79
10	C5	1.23	0.66	.041	3.50	13.27	2.28	1.31	.081	0.80	-0.48	1.73	1.11	.060	1.59	1.70
11	C6	1.19	0.61	.038	3.97	17.14	2.23	1.33	.083	0.77	-0.64	1.51	0.99	.053	2.07	3.64
12	C7	1.48	0.97	.061	2.28	4.64	1.69	1.14	.071	1.65	1.68	1.64	1.08	.058	1.74	2.15
13	D1	1.61	1.08	.068	1.79	2.31	2.49	1.40	.087	0.45	-1.14	1.93	1.24	.067	1.23	0.40
14	D2	1.44	0.87	.055	2.21	4.59	2.22	1.26	.079	0.72	-0.63	1.70	1.08	.058	1.52	1.41
15	D3	1.64	1.04	.066	1.76	2.52	2.47	1.36	0.84	0.55	-0.92	2.02	1.24	.067	1.08	0.07
16	D4	2.41	1.38	.087	0.62	-0.87	2.64	1.39	.086	0.25	-1.28	1.67	1.04	.056	1.54	1.59
17	D5	1.94	1.13	.071	1.17	0.64	2.44	1.32	.082	0.52	-0.90	1.59	1.02	.055	1.88	2.88

Table 3

Goodness-of-Fit Statistics for Confirmatory Factor Analyses

Model	df	Y-B χ^2	CFI	TLI	SRMR	RMSEA (90% CI)	BIC
MVA (*N* = 253)							
Two-factor	118	383.69	0.790	0.758	0.087	0.094 (0.084-0.105)	11006.76
Three-factor	116	433.21	0.750	0.707	0.123	0.104 (0.094-0.115)	11037.53
Four-factor numbing	113	321.87	0.835	0.802	0.103	0.085 (0.075-0.097)	10926.94
Four-factor numbing*	112	281.44	0.866	0.838	0.103	0.077 (0.066-0.089)	10876.27
Four-factor dysphoria	113	308.44	0.846	0.814	0.072	0.083 (0.072-0.094)	10920.79
Four-factor dysphoria†	112	220.19	0.915	0.896	0.065	0.062 (0.050-0.074)	10799.72
Five-factor	109	240.43	0.896	0.871	0.089	0.069 (0.057-0.081)	10837.07
SA (*N* = 259)							
Two-factor	118	388.96	0.830	0.804	0.073	0.094 (0.084-0.105)	13235.60
Three-factor	116	379.53	0.834	0.806	0.077	0.094 (0.083-0.104)	13230.40
Four-factor numbing	113	317.80	0.871	0.845	0.065	0.084 (0.073-0.095)	13176.85
Four-factor numbing*	112	214.24	0.936	0.922	0.064	0.059 (0.047-0.071)	13060.22
Four-factor dysphoria	113	238.13	0.921	0.905	0.052	0.065 (0.054-0.077)	13083.18
Four-factor dysphoria†	112	224.40	0.929	0.914	0.051	0.062 (0.050-0.074)	13071.75
Five-factor	109	186.66	0.951	0.939	0.046	0.052 (0.038-0.065)	13043.88
SUD (*N* = 342)							
Two-factor	118	348.99	0.877	0.858	0.062	0.076 (0.067-0.085)	15743.45
Three-factor	116	342.63	0.879	0.859	0.065	0.076 (0.066-0.085)	15741.68
Four-factor numbing	113	263.51	0.920	0.904	0.052	0.063 (0.053-0.072)	15644.79
Four-factor numbing*	112	202.91	0.952	0.941	0.049	0.049 (0.038-0.059)	15563.27
Four-factor dysphoria	113	241.03	0.932	0.918	0.049	0.058 (0.047-0.068)	15614.48
Four-factor dysphoria†	112	229.67	0.937	0.924	0.048	0.055 (0.045-0.066)	15602.07
Five-factor	109	194.71	0.954	0.943	0.045	0.048 (0.037-0.059)	15568.61

Note. All Y-B χ^2 were significant at $p < .001$. Y-B χ^2 = Yuan-Bentler scaled chi-square; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; SRMR = Standardized Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation; BIC = Bayesian Information Criterion.

* = Four-factor numbing model was run again allowing error variances of items 16 and 17 to correlate.

† = Four-factor dysphoria model was run again allowing error variances of items 10 and 11 to correlate.

Table 4

*Factor Loadings and Factor Relationships of the Four-Factor Numbing Model**

PCL-S Item	Reexperiencing			Avoidance			Numbing			Hyperarousal		
	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD
						Factor Loadings[a]						
1	.74/.80	.73/.85	.79/.97									
	(.060)	(.060)	(.049)									
2	.63/.61	.73/.87	.70/.78									
	(.073)	(.076)	(.065)									
3	.65/.68	.71/.78	.72/.76									
	(.071)	(.079)	(.068)									
4	.79/1.04	.70/.87	.76/.93									
	(.065)	(.061)	(.045)									
5	.72/.88	.78/1.01	.71/.84									
	(.071)	(.062)	(.065)									
6				.60/.70	.59/.77	.65/.91						
				(.084)	(.084)	(.067)						
7				.68/.80	.78/1.04	.83/1.09						
				(.091)	(.075)	(.071)						
8							.20/.27	.32/.45	.29/.29			
							(.124)	(.093)	(.072)			
9							.51/.32	.72/.80	.76/.73			
							(.075)	(.073)	(.075)			
10							.92/.60	.83/1.08	.84/.93			
							(.093)	(.072)	(.065)			
11							.83/.51	.78/1.03	.73/.72			
							(.094)	(.069)	(.074)			
12							.42/.41	.65/.74	.60/.65			
							(.11)	(.082)	(.080)			
13										.81/.87	.68/.95	.76/.94
										(.088)	(.079)	(.072)
14										.79/.69	.70/.87	.79/.85
										(.080)	(.075)	(.068)
15										.88/.91	.86/1.16	.83/1.03
										(.083)	(.063)	(.058)
16										.41/.57	.46/.64	.52/.54
										(.099)	(.091)	(.077)
17										.48/.55	.51/.67	.63/.64
										(.090)	(.088)	(.077)

(*Table 4 continues*)

(*Table 4 continued*)

	Factor Correlations[b]											
	Reexperiencing			Avoidance			Numbing			Arousal		
	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD
Reexp.	1.00	1.00	1.00									
Avoidance	.86	.72	.72	1.00	1.00	1.00						
Numbing	.45	.61	.61	.34	.57	.61	1.00	1.00	1.00			
Arousal	.68	.65	.74	.51	.51	.58	.71	.79	.85	1.00	1.00	1.00

	Error Covariances		
	MVA	SA	SUD
16 with 17	.45	.63	.49

Note. MVA = Motor Vehicle Accident; SA = Sexual Assault; SUD = Sudden Unexpected Death.

* With covariance between PCL-S items 16 and 17 estimated.

[a] Completely standardized followed by unstandardized loadings (and standard errors); all loadings significant at $p < .05$.

[b] Values along the diagonal are factor variances set to one for purposes of scaling the latent variable; correlations all significant at $p < .05$.

Table 5

*Factor Loadings and Factor Relationships of the Four-Factor Dysphoria Model**

PCL-S	Intrusions			Avoidance			Dysphoria			Hyperarousal		
Item	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD
					Factor Loadings[a]							
1	.74/.79 (.059)	.73/.84 (.061)	.79/.96 (.049)									
2	.63/.60 (.070)	.72/.86 (.076)	.69/.78 (.065)									
3	.66/.69 (.070)	.71/.78 (.078)	.72/.77 (.069)									
4	.79/1.04 (.063)	.70/.88 (.061)	.75/.92 (.045)									
5	.73/.89 (.069)	.79/1.02 (.061)	.71/.84 (.065)									
6				.59/.69 (.088)	.60/.78 (.086)	.66/.91 (.066)						
7				.69/.81 (.094)	.77/1.02 (.076)	.83/1.08 (.071)						
8							.28/.38 (.114)	.31/.44 (.093)	.27/.27 (.072)			
9							.49/.31 (.065)	.68/.75 (.073)	.72/.70 (.080)			
10							.64/.42 (.084)	.74/.96 (.080)	.77/.85 (.074)			
11							.58/.35 (.082)	.70/.92 (.074)	.65/.64 (.075)			
12							.49/.47 (.095)	.65/.74 (.080)	.58/.62 (.077)			
13							.81/.87 (.088)	.61/.86 (.093)	.73/.91 (.074)			
14							.80/.70 (.080)	.68/.86 (.073)	.78/.84 (.068)			
15							.88/.91 (.082)	.81/1.09 (.066)	.81/1.01 (.057)			
16										.75/1.02 (.084)	.83/1.15 (.065)	.76/.79 (.072)
17										.76/.85 (.082)	.87/1.14 (.069)	.86/.88 (.077)

(*Table 5 continues*)

(*Table 5 continued*)

	Factor Correlations[b]											
	Reexperiencing			Avoidance			Numbing			Arousal		
	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD
Intrusions	1.00	1.00	1.00									
Avoidance	.85	.73	.72	1.00	1.00	1.00						
Dysphoria	.67	.65	.71	.52	.56	.62	1.00	1.00	1.00			
Arousal	.77	.62	.63	.65	.50	.47	.57	.53	.67	1.00	1.00	1.00

	Error Covariances		
	MVA	SA	SUD
10 with 11	.65	.30	.26

Note. MVA = Motor Vehicle Accident; SA = Sexual Assault; SUD = Sudden Unexpected Death.

* With covariance between PCL-S items 10 and 11 estimated.

[a] Completely standardized followed by unstandardized loadings (and standard errors); all loadings significant at p < .05.

[b] Values along the diagonal are factor variances set to one for purposes of scaling the latent variable; correlations all significant at p < .05.

Table 6

Factor Loadings and Factor Relationships of the Five-Factor Model

PCL	RE			AV			NU			DA			AA		
Item	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD
							Factor Loadings[a]								
1	.74/.79 (.059)	.73/.85 (.061)	.79/.96 (.049)												
2	.63/.60 (.070)	.72/.86 (.076)	.69/.78 (.065)												
3	.66/.69 (.070)	.71/.78 (.079)	.72/.76 (.069)												
4	.79/1.04 (.063)	.70/.88 (.061)	.76/.92 (.045)												
5	.73/.89 (.069)	.79/1.02 (.061)	.71/.84 (.065)												
6				.59/.68 (.087)	.60/.78 (.085)	.65/.91 (.067)									
7				.70/.82 (.094)	.77/1.02 (.076)	.84/1.09 (.071)									
8							.20/.27 (.124)	.32/.45 (.093)	.29/.29 (.072)						
9							.51/.32 (.075)	.72/.80 (.073)	.76/.74 (.074)						
10							.92/.60 (.094)	.83/1.08 (.072)	.84/.92 (.065)						
11							.83/.51 (.094)	.78/1.03 (.069)	.73/.72 (.074)						
12							.42/.41 (.110)	.65/.74 (.082)	.60/.65 (.079)						
13										.80/.86 (.088)	.67/.94 (.083)	.76/.94 (.073)			
14										.80/.70 (.079)	.71/.89 (.075)	.79/.85 (.068)			
15										.90/.93 (.081)	.87/1.18 (.064)	.84/1.04 (.059)			
16													.75/1.02 (.085)	.83/1.15 (.065)	.75/.78 (.072)
17													.76/.85 (.082)	.87/1.14 (.069)	.87/.89 (.078)

(*Table 6 continues*)

(*Table 6 continues*)

	Factor Correlations[b]														
PCL	RE			AV			NU			DA			AA		
Item	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD	MVA	SA	SUD
RE	1.00	1.00	1.00												
AV	.86	.74	.72	1.00	1.00	1.00									
NU	.45	.61	.62	.35	.58	.61	1.00	1.00	1.00						
DA	.65	.60	.72	.48	.49	.57	.71	.78	.86	1.00	1.00	1.00			
AA	.77	.63	.63	.65	.50	.47	.41	.48	.56	.54	.52	.70	1.00	1.00	1.00

Note. RE = Reexperiencing; AV = Avoidance; NU = Numbing; DA = Dysphoric Arousal; AA = Anxious Arousal; MVA = Motor Vehicle Accident; SA = Sexual Assault; SUD = Sudden Unexpected Death.

[a] Completely standardized followed by unstandardized loadings (and standard errors); all loadings significant at $p < .05$.

[b] Values along the diagonal are factor variances set to one for purposes of scaling the latent variable; correlations all significant at $p < .05$.

Table 7

Goodness-of-Fit Statistics for MGCFA and Partial Invariance (Numbing Model)

	χ^2	*df*	S-B χ^2_{diff}	Δdf	CFI	ΔCFI	TLI	SRMR	RMSEA
MVA vs SA									
Configural	501.86**	224	--	--	0.902	--	0.880	0.086	0.070 (0.061-0.078)
Metric	511.74**	237	14.65	13	0.903	0.001	0.888	0.090	0.067 (0.059-0.075)
Scalar	601.47**	250	94.38**	13	0.875	0.028	0.865	0.104	0.074 (0.067-0.082)
Partial [12, 3]	530.66**	247	18.82*	10	0.899	0.004	0.889	0.093	0.067 (0.059-0.075)
MVA vs SUD									
Configural	484.32**	224	--	--	0.917	--	0.899	0.077	0.063 (0.055-0.070)
Metric	496.34**	237	15.98	13	0.917	0.000	0.905	0.082	0.061 (0.053-0.068)
Scalar	662.20**	250	221.40**	13	0.869	0.048	0.857	0.104	0.074 (0.068-0.081)
Partial [16, 17, 8, 5, 3]	510.87**	244	14.67*	7	0.915	0.002	0.905	0.084	0.061 (0.053-0.068)
SA vs SUD									
Configural	416.10**	224	--	--	0.945	--	0.933	0.056	0.053 (0.045-0.061)
Metric	431.67**	237	15.19	13	0.944	0.001	0.936	0.059	0.052 (0.044-0.060)
Scalar	571.26**	250	166.34**	13	0.908	0.036	0.900	0.072	0.065 (0.058-0.072)
Partial [8, 5, 12, 15, 16, 17]	443.40**	243	11.92	6	0.943	0.001	0.936	0.059	0.052 (0.045-0.060)

Note. SB χ^2_{diff} = Satorra-Bentler Scaled χ^2 difference. The item intercepts listed in brackets were freely estimated while the remaining intercepts were constrained to equality.

**p* < .05 ***p* < .001

Table 8

Goodness-of-Fit Statistics for MGCFA and Partial Invariance (Dysphoria Model)

	χ^2	*df*	S-B χ^2_{diff}	Δdf	CFI	ΔCFI	TLI	SRMR	RMSEA
MVA vs SA									
Configural	444.15**	224	--	--	0.922	--	0.905	0.058	0.062 (0.053-0.070)
Metric	499.88**	237	51.36**	13	0.907	.015	0.893	0.081	0.066 (0.058-0.074)
Partial [11, 10, 9]	460.47**	233	16.35	9	0.919	.003	0.906	0.062	0.062 (0.053-0.070)
Scalar	724.72**	250	339.88**	17	0.832	.090	0.817	0.182	0.086 (0.079-0.093)
MVA vs SUD									
Configural	449.75**	224	--	--	0.928	--	0.913	0.056	0.058 (0.050-0.066)
Metric	499.56**	237	45.05**	13	0.916	.012	0.904	0.073	0.061 (0.054-0.068)
Partial [11, 10, 9]	465.17**	234	15.56	10	0.926	.002	0.914	0.061	0.058 (0.050-0.065)
Scalar	675.12**	251	287.81**	17	0.865	.061	0.853	0.095	0.075 (0.069-0.082)
SA vs SUD									
Configural	454.57**	224	--	--	0.934	--	0.920	0.049	0.059 (0.051-0.066)
Metric	472.67**	237	18.09	13	0.933	0.001	0.923	0.055	0.058 (0.050-0.065)
Scalar	576.31**	250	119.81**	13	0.907	0.026	0.899	0.062	0.066 (0.059-0.073)
Partial [8, 5, 11, 12]	482.60**	245	8.75	8	0.932	0.001	0.925	0.055	0.057 (0.049-0.064)

Note. SB χ^2_{diff} = Satorra-Bentler Scaled χ^2 difference. The factor loadings and item intercepts listed in brackets were freely estimated while the remaining factor loadings and intercepts were constrained to equality.

**p* < .05 ***p* < .001

Table 9

Goodness-of-Fit Statistics for MGCFA and Partial Invariance (Five-Factor Model)

	χ^2	*df*	S-B χ^2_{diff}	Δdf	CFI	ΔCFI	TLI	SRMR	RMSEA
MVA vs SA									
Configural	432.28**	218	--	--	0.924	--	0.905	0.071	0.062 (0.053-0.070)
Metric	444.61**	230	15.02	12	0.924	0.000	0.910	0.075	0.060 (0.052-0.069)
Scalar	535.81**	242	95.46**	12	0.896	0.028	0.883	0.090	0.069 (0.061-0.077)
Partial [12, 3]	465.52**	239	21.40*	9	0.920	0.004	0.909	0.077	0.061 (0.053-0.069)
MVA vs SUD									
Configural	435.28**	218	--	--	0.931	--	0.914	0.067	0.058 (0.058-0.066)
Metric	452.33**	230	18.61	12	0.929	0.002	0.916	0.072	0.057 (0.049-0.065)
Scalar	563.24**	242	139.36**	12	0.898	0.033	0.885	0.082	0.067 (0.060-0.074)
Partial [8, 5, 3]	466.48**	237	14.25*	7	0.927	0.002	0.916	0.075	0.057 (0.049-0.065)
SA vs SUD									
Configural	382.12**	218	--	--	0.953	--	0.941	0.046	0.050 (0.042-0.058)
Metric	398.04**	230	15.84	12	0.952	0.001	0.943	0.049	0.049 (0.041-0.057)
Scalar	499.66**	242	117.81**	12	0.926	0.026	0.917	0.058	0.060 (0.052-0.067)
Partial [8, 5, 12]	417.24**	238	20.21*	8	0.949	0.003	0.941	0.050	0.050 (0.042-0.058)

Note. SB χ^2_{diff} = Satorra-Bentler Scaled χ^2 difference. The item intercepts listed in brackets were freely estimated while the remaining intercepts were constrained to equality.

**p* < .05 ***p* < .001