Published in final edited form as:

Clin Psychol Rev. 2013 February ; 33(1): 54–66. doi:10.1016/j.cpr.2012.10.004.

A Multi-Sample Confirmatory Factor Analysis of PTSD Symptoms: What Exactly Is Wrong with the DSM-IV Structure?

Grant N. Marshall, Terry L. Schell, and Jeremy N. V. Miles
RAND Corporation Santa Monica, CA

Abstract

Within the DSM-IV, PTSD symptoms are rationally classified as assessing one of three symptom domains: reexperiencing, avoidance/numbing, or hyperarousal. However, two alternative four-factor models have been advocated as superior to the DSM-IV framework, based on confirmatory factor analysis. In the Numbing model, symptoms of emotional numbing are differentiated from avoidance. In the Dysphoria model, several symptoms of numbing and hyperarousal are combined to form a factor purported to assess general psychological distress. Examination of these models, within 29 separate data sets, supports two conclusions. First, contrary to its conceptual underpinnings, the Dysphoria model differs empirically from the Numbing model solely in the correlation predicted between two hyperarousal symptoms; all other predicted correlations made by the two models are substantively identical. Second, when the factor analytic presumption of simple structure is relaxed to allow for potential presentation order effects, other plausible symptom structures emerge. In particular, the fit of the DSM-IV model improved dramatically and was a better fit to the data than either four-factor model. The ostensible inferiority of the DSM-IV model may be due to a methodological artifact stemming from the order in which symptoms are typically assessed. The provisional decision to revise the structure of PTSD symptoms in the DSM-5 in light of confirmatory factor analytic results may be misguided.

Keywords

Confirmatory factor analysis; posttraumatic stress disorder; symptom structure

Posttraumatic stress disorder (PTSD) has been the subject of considerable controversy, predating its formal codification in the Diagnostic and Statistical Manual of Mental Disorders-III (DSM-III; American Psychiatric Association, 1980). Areas of disagreement range from concerns about whether PTSD is a legitimate disorder as opposed to a sociopolitical construct (Scott, 1990; Summerfield, 2001; Young, 1995) to differences regarding the defining features of PTSD (Bovin & Marx, 2010; Spitzer, First, & Wakefield, 2007). One of the more challenging issues is the multiplicity of views regarding the optimal framework for conceptualizing the core symptoms of PTSD. Continued progression of knowledge regarding PTSD symptom structure is essential inasmuch as improvements in diagnosis, assessment, prevention, and treatment are contingent on a solid understanding the nature of PTSD. Yet, there is reason for concern that the pace of advancement in this area may have slowed.

As currently conceptualized, PTSD is an anxiety disorder exemplified by three clusters of symptoms that may follow from exposure to a traumatic life event (DSM-IV; American

Psychiatric Association, 2000). As described in the DSM-IV, these three symptom clusters are composed of seventeen symptoms and reflect the phenomena of *reexperiencing* (Criterion B), e.g., intrusive thoughts of the trauma, *avoidance and emotional numbing* (Criterion C), e.g., avoiding reminders of the trauma, and *hyperarousal* (Criterion D), e.g., hypervigilance.

Whereas the organizing framework represented in the DSM-IV was derived on the basis of expert clinical consensus, confirmatory factor analytic techniques have been used to examine the structure of PTSD symptoms (for reviews, see Elhai & Palmieri, 2011; Yufik & Simms, 2010). Based on model fit, two distinct four-factor models of PTSD symptom structure are now widely viewed as superior to the three-factor representation embodied in the DSM-IV. The first of these alternative models, proposed by King and colleagues (King, Leskin, King, & Weathers, 1998), posits the existence of four correlated factors. Two of these symptom clusters, i.e., reexperiencing and hyperarousal, map directly onto clusters proposed in the DSM-IV. The model departs from the DSM-IV framework, however, by hypothesizing that the Criterion C symptom cluster is composed of two separable symptom sets assessing avoidance and emotional numbing. Numerous confirmatory factor analytic studies have revealed this model—hereafter referred to as the Numbing model—to provide a very good fit to the data (e.g., Asmundson et al., 2000; Asmundson, Wright, McCreary, & Pedlar, 2003; DuHamel et al., 2004; King et al., 1998; Marshall, 2004; Palmieri & Fitzgerald, 2005; Palmieri, Marshall, & Schell, 2007a; Palmieri, Weathers, Difede, & King, 2007b; Schinka, Brown, Borenstein, & Mortimer, 2007).

A second four-factor model, proposed by Simms and colleagues (Simms, Watson, & Doebbeling, 2002), holds that PTSD symptoms can be differentiated into those assessing reexperiencing, avoidance, dysphoria, and hyperarousal. The first two of these dimensions, i.e., reexperiencing and avoidance, are identical in item content to those found in the Numbing model. This model differs from the Numbing model, however, in positing a Dysphoria factor which is thought to assess non-specific psychological distress characteristic of various disorders. The dysphoria dimension contains all symptoms construed in the Numbing model as measuring emotional numbing as well as three of the five symptoms regarded in the DSM-IV as assessing hyperarousal. The hyperarousal dimension is retained but is indicated by only the last two DSM-IV hyperarousal symptoms, i.e., *hypervigilance* (D4) and *exaggerated startle response* (D5). This conceptual framework, which we refer to as the Dysphoria model, has also been well-supported in confirmatory factor analytic investigations (e.g., Elhai, Gray, Docherty, Kashdan, & Kose, 2007; Elklit & Shevlin, 2007; Krause, Kaltman, Goodman, & Dutton, 2007; Palmieri et al., 2007b). Abbreviated symptom content and symptom-dimension alignments for the DSM-IV, Numbing, and Dysphoria models are depicted in Table 1.

At this point, neither four-factor alternative to the DSM-IV framework of PTSD symptom structure has emerged as clearly preferable, although a recent meta-analytic comparison of the two models determined that the Dysphoria model may provide a marginally better fit (Yufik & Simms, 2010). Despite the absence of consensus regarding an optimal representation of PTSD symptom structure, there is widespread agreement, based on factor analytic results, that the DSM-IV model provides an inadequate depiction of PTSD symptom structure. Moreover, this commonly-held conviction has prompted the frequent observation that the DSM should be reformulated to align the conceptualization of PTSD with confirmatory factor analytic findings (e.g., DuHamel et al., 2004; Hetzel-Riggen, 2009; King et al., 2008; Palmieri et al., 2007b; Simms et al., 2002).

The calls for revision appear to have been heeded. The provisional draft of the DSM-5 embraces certain aspects of the findings derived from confirmatory factor analysis

(American Psychiatric Association, 2010). At first glance, the integration of confirmatory factor analytic research results might be regarded as an unmitigated advance in the development of a more empirically-based psychiatric taxonomy. On closer inspection, however, it is not clear that the body of confirmatory factor analytic research findings is sufficiently compelling to have warranted revision of the forthcoming DSM-5.

The current study is a reanalysis of this literature, designed to investigate two limitations of confirmatory factor analytic research examining the structure of PTSD. First, factor analytic examinations of PTSD symptoms, just as structural equation modeling more generally (MacCallum, Wegener, Uchino, & Fabrigar, 1993), have tended to consider as plausible only a small set of potential models. Specifically, PTSD researchers have demonstrated a pronounced preference for models that conform to simple structure. The assumption that any pair of PTSD symptoms is correlated because the items share one, and only one, underlying cause simplifies model interpretation. However, simple structure is, at best, a heuristic aid, and may not adequately characterize psychological phenomena (Marsh, Hau, & Grayson, 2005). Near exclusive reliance on simple structure models, in the absence of clear and compelling justification, artificially constrains the scope of models that have come under examination and may have led to a biased understanding of the nature of PTSD.

In considering what additional influences on PTSD symptom structure might warrant departure from simple structure, it is noteworthy that PTSD researchers have focused almost exclusively on semantic content as the sole determinant of symptom covariation. Yet, reported PTSD symptoms may covary for reasons other than item content. In particular, the order in which a series of items is presented can create a context that influences covariation between items. Various theoretical accounts bear on the widespread empirical finding that responses to contiguous items may be more strongly correlated relative to responses to the same questions when placed further apart in an instrument (e.g., N. Schwarz & Sudman, 1992; N. Schwarz, 1999; Torangeau & Rasinski, 1988). For example, a considerable body of theory holds that people are motivated to maintain consistency (Festinger, 1957; Heider, 1958; Osgood & Tannenbaum, 1957; Bem, 1967). From this perspective, responses to successive questions might be answered in a manner that increases consistency with earlier responses. Other theories suggest cognitive mechanisms that could result in order effects. For example, an initial response may serve as a cognitive anchor that constrains the range of subsequent responses (Tversky & Kahneman, 1974). In addition, the process of answering an initial question may make information available that is primed for use in subsequent questions (e.g., Collins & Quillian, 1970; Schell, Klein & Babey, 1996).

To the degree that responses to queries about PTSD symptoms may vary as a function of the order in which items are presented, the failure to consider possible order effects in PTSD models may constitute a significant oversight. This omission is compounded inasmuch as PTSD symptoms are almost always presented to respondents in the sequence in which they are enumerated in the DSM. To model an order effect requires deviation from simple structure and, to our knowledge, has not been done in any of the published PTSD confirmatory factor analytic studies to date.

A second shortcoming of existing research is that the literature may rely too heavily on global model fit and visual inspection of path diagrams to draw theoretical inferences regarding PTSD symptom structure. Although path diagrams are quite useful for visualizing and communicating model assumptions, they have limited ability to facilitate understanding how and why any given model is superior to another in accounting for actual data. Indeed, path diagrams embodying substantially different causal assumptions can represent mathematically equivalent models (Hershberger, 2006; Lee & Hershberger, 1990; MacCallum et al., 1993; Raykov & Marcoulides, 2001). That is to say, the models make

exactly the same predictions about the empirical covariance matrix despite substantially different path diagrams. For example, given a data set in which the Dysphoria model fits better than the Numbing model, it would be common practice to use the Dysphoria model path diagram as a basis for inferring that symptom D1, i.e., *sleep disturbance*, should be interpreted as measuring dysphoria rather than hyperarousal. Yet, it is entirely possible that the Numbing model might actually be better at explaining correlations involving D1. The actual difference between models cannot be inferred easily from differences in the path diagrams.

Several theorists within the methodological literature have discussed the limitations of the standard practice in which researchers theoretically interpret the path diagram of the model with the best global fit (e.g., Bollen, 1989; Pearl, 1988). Although some methodologists have tried to work out new procedures that rely on local fit indices to derive theoretical conclusions (e.g., Shipley, 2000), these methods have not been adopted in any substantive literature. In addition, these methods do not address the challenge facing the PTSD researchers in which multiple plausible factor models fit the data similarly well. Of particular interest in the current analyses is whether the Dysphoria and Numbing models differ empirically in ways that are consistent with the conceptual distinction that has been made between the two models.

To improve the theoretical interpretation of empirical differences between models, this paper uses a novel descriptive approach to directly assess differences between two non-nested models. Specifically, we compute the difference between predictions made by each model for each element of the correlation matrix. In other words, we compare the correlations implied by the Numbing and Dysphoria models, on an element by element basis, to determine precisely where the two models make discrepant empirical predictions.¹ This strategy allows precise identification of differences in hypotheses implied by the models regardless of how those models are represented in path diagrams. In this way, the theoretical meaning of differences in aggregate model fit across two models can be grounded in a description of the specific data elements about which the models make different predictions.

This paper uses data sets from previously-published confirmatory factor analytic examinations of PTSD symptom structure to address two objectives. After first examining the fit of the DSM-IV, Numbing, and Dysphoria models across data sets, we sought: (1) to identify the specific correlations for which the Dysphoria and Numbing models make different predictions to guide the theoretical interpretation of differences in model fit; and (2) to examine the extent to which an alternative factor model, which relaxes the simple structure assumption to allow for possible order effects, provides a plausible account of the data.

Method

Identification of Eligible Studies

To examine the underlying structure of the DSM-IV PTSD symptoms, we reanalyzed data from studies that used confirmatory factor analysis to evaluate *DSM*-based PTSD measures. Eligible studies were identified through computerized database searches of MEDLINE and PsycINFO, using posttraumatic stress disorder, confirmatory factor analysis, and *symptom structure* as keywords. We confined ourselves to confirmatory factor analytic studies of traumatized adult samples that examined 17-item PTSD symptom severity measures that

¹This process assesses the empirical equivalence of different models rather than the fit of a given model to data.

mapped directly onto the 17 DSM-IV PTSD symptom criteria B, C, and D.² We also omitted studies for which complete data were available on only a selected subset of study participants (e.g., persons with qualifying, rather than skip-out, scores on key screening questions). Finally, we excluded samples with no reported history of trauma exposure.³ To ensure stable factor structures, we restricted our analyses to data sets that included a minimum of 200 participants (R. B. Kline, 2010).

To solicit data, we contacted via email the corresponding authors of 29 separate research articles published between January 2000 and October 2010 in English-language journals. We invited authors to send data in any of various forms (i.e., raw data, correlation, or covariance matrices) to increase likelihood of participation. From 27 of the 29 authors, we received 34 separate data matrices. Two of the 34 matrices were ultimately excluded as ineligible. Of the remaining 32 studies, three were omitted to address non-independent data stemming from administration of multiple PTSD measures to a single sample (i.e., Palmieri et al., 2007b; Scher, McCreary, Asmundson, & Resick, 2008) or to analyses conducted using non-independent samples (i.e., Simms et al., 2002).⁴ In all, the analytic sample consisted of 29 data sets. As part of our email query, we asked authors providing data to clarify the order in which PTSD items had been administered. With one exception (i.e., Simms et al., 2002), all items had been administered in the order in which they are found on their respective PTSD symptom inventories.

Description of Studies Analyzed

The studies contributing data sets for the current analyses are displayed in Table 2. As shown, data sets included responses from survivors of a range of trauma types including interpersonal violence, bereavement, life-threatening physical health conditions, military deployment, civil war, and terrorism. The two most common trauma types included warzone exposure (9 samples) and interpersonal violence (9 samples). Sample sizes ranged from $N = 208$ to $N = 15,603$. PTSD symptom severity assessments were derived from the Clinician-Administered PTSD Scale (Blake et al., 1995), the PTSD Symptom Checklist (Weathers et al., 1993), the Harvard Trauma Questionnaire (Mollica et al., 1992)⁵, the Posttraumatic Diagnostic Scale (PDS; Foa, Cashman, Jaycox, & Perry, 1997), the Posttraumatic Stress Disorder Questionnaire (Cross & McCanne, 2001), the Davidson Trauma Scale (DTS; Davidson et al., 1997), and the PTSD Symptom Scale (PSS; Foa, Riggs, Dancu, & Rothbaum, 1993). As shown in Table 2, for 22 of 29 data sets, PTSD symptom severity measures were administered in the order found in the DSM-IV. In five data sets, symptoms were administered in the order found in the DSM-III-R (American Psychiatric Association, 1987) which differs from the DSM-IV in the placement of a single item within the order.

²Although any study which administered an instrument assessing the 17 DSM-IV PTSD symptoms would be suitable for analysis, this study is limited to the specific investigations that form the basis of our current understanding about the structure of PTSD, i.e., published studies employing confirmatory factor analysis. In addition, prior efforts at reaching the larger literature consisting of all studies using any eligible measure of PTSD have yielded very poor response rates of less than 10% (i.e., Yufik & Simms, 2010).

³In particular, we omitted matrices from an investigation of college students (Hoyt & Yeater, 2010) and a study of non-deployed military personnel (Simms et al., 2002) for which no evidence of exposure to a potentially traumatic event had been provided.

⁴In their sample of recent survivors of sexual assault ($N = 216$), Scher et al. (2008) analyzed PTSD symptom severity information obtained using the Clinician-Administered PTSD Scale (CAPS; Blake et al., 1995) and the PTSD Symptom Scale (PSS; Foa, Riggs, Dancu, & Rothbaum, 1993). Similarly, Palmieri et al. (2007b) presented data from both the CAPS and the PTSD Symptom Checklist (PCL; Weathers et al., 1993) collected from a single sample of utility personnel who had worked at the World Trade Center in the aftermath of the 9/11 terrorist attacks. Finally, Simms et al. (2002) analyzed data from persons previously deployed to the Persian Gulf during the Gulf War ($N = 1896$) as well as a subgroup of individuals judged to have been exposed to a Criterion A event during deployment ($N = 419$). From these three studies, we randomly selected for analysis, the PSS data, the CAPS data, and the data from all previously deployed persons, respectively.

⁵The Harvard Trauma Questionnaire consists of 16, rather than 17, symptoms, and we did not routinely analyze studies that administered this instrument. The sole exception was the study by Palmieri et al. (2007a) in which the 16-item instrument had been expanded to 17 items by splitting the single avoidance item into separate items assessing avoidance of thoughts and feelings and avoidance of activities and places.

Two studies presented PTSD items in idiosyncratic orders inconsistent with both the DSM-III-R and the DSM-IV (i.e., Palmieri et al., 2007a; Simms et al., 2002).

Data Analysis

Because many investigators provided only correlation matrices, we analyzed correlation, instead of covariance, matrices.⁶ Data were examined using maximum likelihood (ML) estimation⁷, as implemented in Mplus 6.1 (Muthén & Muthén, 2008). Results were integrated across data sets using the MplusAutomation package (Hallquist, 2011) in R. Following the fit evaluation strategy validated by Hu and Bentler (1999), model fit was assessed using overall model chi-square, the comparative fit index (CFI; Bentler, 1990), and the standardized root mean square residual (SRMR; Bentler, 1995). These indices were augmented with Akaike's Information Criterion (AIC; Akaike, 1987) as a measure of parsimony which expresses model likelihood in a manner that can be compared across non-nested models. The causal assumptions underlying the AIC make it more appropriate than alternative indicators of parsimony such as the Bayesian Information Criterion (G. Schwartz, 1978; see Burnham & Andersen, 2004 or Gelman & Rubin, 1995, for discussion). In addition, differences in fit between nested models were assessed with chi-squared difference tests.⁸

We performed the analyses in three stages. As an initial analytic step, we used structural equation modeling to fit the DSM-IV, Numbing, and Dysphoria models to each data set individually and in the context of a single multiple group model. We present measures of model fit for individual data sets as well as for the aggregate, multi-group model which represents a sample-size weighted combination of the 29 data sets. Insofar as proponents of specific models of PTSD symptom structure have proposed no hypotheses regarding cross-group parameter equivalence, we posited only that models would be invariant with respect to configuration. In other words, we hypothesized that models would possess the same factor structure across different groups, and we imposed no other cross-group equality constraints. Thus, the specific factor loadings and item errors are permitted to vary across studies conducted on different populations, in different measurement contexts, and across scales with different response options.

Our approach contrasts with the strategy selected by Yufik and Simms (2010) in their recent meta-analysis of PTSD symptom structure. In particular, Yufik and Simms (2010) employed a two-stage meta-analytic procedure proposed by Cheung and Chan (2005). This two-stage procedure is predicated on the assumption that the covariance matrices are homogeneous across studies, allowing matrices to be aggregated into a single matrix for analysis. Chan and Cheung (2005) cautioned that their strategy is not appropriate when the homogeneity assumption is violated. For the 29 data sets, the covariance/correlation matrices were heterogeneous. In some data sets, for example, all correlations were uniformly high (e.g., r

⁶The literature has discussed the structure of PTSD symptoms as if that structure is independent of the sample variance. In other words, PTSD symptom structure is not posited to vary across samples with different symptom variances. Moreover, in these data, the variances of the 17 symptoms were very similar within each sample which lessens the likelihood that the results depend on the scale on which the variables are expressed. In addition, for those instances in which we received covariance or raw data matrices, we conducted sensitivity analyses. Essentially identical patterns emerged irrespective of whether correlations or covariances were analyzed. These factors converge to suggest that analysis of correlation matrices was appropriate.

⁷Given the typical distributions of PTSD symptoms, it would often be preferable to use robust estimation techniques rather than ordinary maximum likelihood estimation to improve the accuracy of parameter standard errors. However, robust estimation requires full raw data which was not available to us for most of the studies. As a sensitivity test, we re-estimated models using robust maximum likelihood in those instances in which we had raw data. No substantive differences emerged in results across estimation techniques, and we present only the ordinary maximum likelihood results.

⁸Obtained fit statistics may differ slightly from those reported in original articles due to slight discrepancies in sample sizes, and the use of different estimation methods.

= .70) whereas others showed more modest correlations (e.g., $r = .30$). The current analytic method does not require the assumption of homogeneity.⁹

In a second stage, we calculated the implied correlation matrices for both the Numbing and Dysphoria models for each of the 29 data sets. For each data set, we subtracted, element by element, the implied correlation matrix for the Dysphoria model from the implied correlation matrix of the Numbing model. This procedure allows identification of the specific correlations for which the two models make different predictions within each data set. In addition, model difference matrices can be averaged across data sets to identify the locations of discrepancies between the predictions made by the two models. Moreover, after inspecting the differences between implied correlation matrices for the Numbing and Dysphoria models, we identified the largest single source of divergence in the predictions made by each model and verified, using confirmatory factor analysis, that this individual element constituted the major difference between conceptual models.

Finally, we tested the fit of an additional model in which we relaxed the assumption of simple structure. Specifically, we tested the possibility that the order in which PTSD symptoms are presented on assessment instruments serves as another influence on the covariation among PTSD symptoms. In building this model, we were guided, as discussed above, by theories consistent with the possibility that each successive PTSD item has an influence on responses to the PTSD symptom which immediately follows (e.g., Collins & Quillian, 1970; Festinger, 1957; Osgood & Tannenbaum, 1957; Tversky & Kahneman, 1974).

Results

The Fit of Alternative Models

As an initial step, the fit of alternative models of PTSD symptom structure were examined. As shown in Table 3, nearly all models across all data sets (5 models x 29 data sets) had acceptable model fit as judged by the two-index criteria advocated by Hu and Bentler (1999). Specifically, most models either had an SRMR below .06 or a CFI above .95. In most cases, this was due to low SRMR values, indicating that the correlations implied by the models were very close to the actual correlations in absolute terms. Because most models have adequate model fit, we focus primarily on the relative differences in fit across models and data sets.

Aggregated across data sets, the Numbing and Dysphoria models fit better than the DSM-IV model on all fit indices as shown in Table 4. The magnitude of these differences was quite small, with an improvement in the CFI of less than 0.03 and an improvement in SRMR of less than 0.01. Yet, these differences were very stable across studies. The Numbing model was superior to the DSM-IV model in all 29 data sets according to the CFI and SRMR indices. As judged by the AIC, the Numbing model was superior to the DSM-IV model in 28 of 29 data. The Dysphoria model was superior to the DSM-IV model in 27 of 29 studies as determined by the CFI, the SRMR, and the AIC.

Comparison of the Numbing and Dysphoria Models

Comparing model fit to data—The Numbing and Dysphoria models fit the data equally well. Whereas the Numbing model was slightly superior when aggregated across all data,

⁹Substantial heterogeneity across studies was also observed by Yufik and Simms (2010) in their meta-analysis of PTSD symptom structure. Yufik and Simms (2010) chose to disregard caveats about the homogeneity assumption, combining across matrices “despite the evidence of matrix heterogeneity for the combined sample (p. 770).” Violation of the homogeneity assumption raises questions about the interpretability of the Yufik and Simms (2010) findings.

the difference was modest, i.e., only .002 on the CFI and .003 in the SRMR. Due to the large sample size ($N = 33384$), however, differences in χ^2 and AIC were substantial ($\Delta\chi^2 = \Delta AIC = 805$). In contrast, the Dysphoria model was preferred in a slight majority of the 29 individual data sets: 15 times as judged by the CFI and AIC indices, and 19 times as measured by the SRMR index. In addition, within some data sets, the magnitude of the difference across models was non-trivial. We investigated this discrepancy by conducting a sensitivity test to determine if aggregate results were unduly influenced by the study with an unusually large sample size ($N = 15603$). Deleting this data set from the aggregate analysis did not change the difference between the two models on any of the fit indices.

Comparing correlations predicted by the two models—To clarify the nature of the difference between the Numbing and Dysphoria models, implied correlation matrices for the two models were compared by subtracting the correlations predicted by the Dysphoria model from those of the Numbing model. This procedure allows identification of the specific locations in which the two models yield discrepant predictions, with positive values indicating instances in which the Numbing model predicted larger correlations relative to the Dysphoria model. As shown in Table 5, the two models yielded highly similar predictions for most correlations. In particular, when averaged across the 136 symptom pairs and 29 data sets, the mean difference in implied correlations equaled -0.001 ($SD = 0.025$). These small differences are consistent with the finding that global model fit indices (e.g., SRMR) tended to be quite similar for the two models.

Identification of differences in local fit—As shown in Table 5, despite the high similarity of the Numbing and Dysphoria models across the full correlation matrix, one substantial difference emerged with respect to local fit. Specifically, the two models yielded discrepant predictions for the correlation between (D4) *hypervigilance* and (D5) *exaggerated startle response*. The difference between model predictions for this correlation averaged -0.13 , and ranged from -0.03 to -0.29 across studies. This difference is more than 5 standard deviations from the mean of other differences in model predictions. Moreover, it is worth noting that the Dysphoria model always fit this particular correlation perfectly because the model operationalizes the hyperarousal factor using only these two indicators. Thus, this difference between the model predictions for the D4, D5 correlation necessarily reflects the misfit of the Numbing model relative to the Dysphoria model.

The six largest remaining model discrepancies, as shown in Table 5, also involved items D4 and D5. In particular, the correlations of items D4 or D5 with items D1, D2, and D3 were assigned higher values by the Numbing model in comparison with the Dysphoria model. These small positive differences are also best seen as resulting from the failure of the Numbing model to fit the D4, D5 element. Because the Numbing model underestimated the correlation between D4 and D5 relative to the Dysphoria model, it also overestimated the other matrix elements that depend on the factor loadings for D4 and D5.

To examine further the degree to which model discrepancies involved the predicted correlation between *hypervigilance* (D4) and *exaggerated startle response* (D5), we inspected the four largest differences in implied correlations for each data set. This strategy addresses the possibility that averaging across data sets might inadvertently obscure heterogeneity regarding the largest model difference in the individual data sets. As shown in Table 6, the predicted correlation between symptoms D4 and D5 was the single largest model discrepancy in 22 of 29 data sets. In one additional data set, the D4, D5 element was the second greatest model discrepancy and the largest negative discrepancy. In combination, these results converge in revealing that the Numbing and Dysphoria models make highly similar predictions for all elements except for the correlation between symptoms D4 and D5.

Comparison of the Numbing and Dysphoria models for all elements other than D4, D5—To better understand the importance of this finding, we fit an additional model that allows comparison of the fit between the Numbing and Dysphoria models for all elements other than D4, D5. This evaluation involved modifying the Numbing model to allow the residual terms for D4 and D5 to correlate freely. With the addition of this correlation, both the *Modified Numbing* and Dysphoria models each fit the D4, D5 element perfectly, and the two models can be compared to determine the extent to which the two models differ in their fit to all of the remaining elements. The Modified Numbing model can also be compared to the initial Numbing model to determine the extent to which model discrepancy in the predictions for D4, D5 actually constitute a significant source of misfit to the data for this model.¹⁰

Results indicated that the Modified Numbing model was a better fit than the initial Numbing model to the data, showing that the two models not only made discrepant predictions for the D4, D5 element but also that the D4, D5 element constituted a critical source of misfit of the Numbing model to the data. As shown in Table 3, the Modified Numbing model was significantly superior to the initial Numbing model in 26 out of 29 data sets. In 25 data sets, the difference was significant at the $p < .0001$ level; in an additional data set, the difference was significant at $p < .01$. In the three data sets in which this additional correlation was not statistically significant, the Modified Numbing model still constituted a superior fit to the initial Numbing model based on broader indices of model fit. As shown in Table 4, the Modified Numbing model also provided a dramatic improvement over the initial Numbing model in aggregate fit, $\Delta\chi^2(29) = 3349$.

To determine the degree to which the Numbing and Dysphoria models differed after accounting for the D4, D5 misfit by addition of the correlated error term, the Modified Numbing and Dysphoria models were compared. Results consistently indicate the superiority of the Modified Numbing model over the Dysphoria model. Although direct comparison of the Numbing and Dysphoria models via a χ^2 difference test is not possible because these models are not nested, they can be compared on measures of parsimony. As judged by the AIC, the Modified Numbing model was superior to the Dysphoria model in 26 of 29 data sets. Put another way, in 12 of the 15 studies (80%) in which the Dysphoria model fit better than the Numbing model, this advantage was reversed if the Numbing model was modified to fit the D4, D5 correlation. Moreover, when data were aggregated across studies, as shown in Table 4, the Modified Numbing model was substantially superior to the Dysphoria model, $\Delta\text{AIC} = 4133$. In all, analyses reveal that any advantage of the Dysphoria model over the Numbing model is substantially attributable to this single matrix element.

Contrary to what has been commonly assumed, the apparent advantage of the Dysphoria model over the Numbing model is not that it provides a better explanation of the correlations among the Dysphoria indicators C3-D3. The critical empirical distinction between models is that the Numbing model always underestimates the correlation between D4 and D5 whereas the Dysphoria model does not. Aside from this single data element, the symptom structure assumed by the Numbing model provides a better fit to the data than the structure assumed by the Dysphoria model.¹¹

¹⁰Two models can make discrepant predictions about a specific element without that element constituting a significant source of misfit for either model. So, it is not a foregone conclusion that allowing the D4 and D5 error terms to correlate in the Modified Numbing model would result in a significant improvement in model fit to actual data.

¹¹The Modified Numbing model is a useful tool to highlight how the Numbing and Dysphoria models differ, and it was used with this limited purpose in mind. The model is not necessarily a useful model for understanding PTSD. The best fitting model is not always the most appropriate if its elements cannot be interpreted.

Assessing the Fit of a Model that Included Response Order Effects

Building the order effect model—Most research on PTSD symptom structure has ignored the possibility of order effects despite substantial empirical and theoretical evidence to support the existence of such effects. To examine the extent to which a model that allows for presentation order effects provides a plausible account of the data, we compared the fit of the DSM-IV model which had been modified to include order effects to the standard DSM-IV model as well as to the Numbing and Dysphoria models. Specifically, in addition to the three latent constructs of re-experiencing, avoidance, and hyperarousal, each individual PTSD item was modeled as exerting a direct influence on the item that immediately followed. Thus, for items administered in the DSM-IV order, D5 was regressed onto D4, D4 was regressed onto D3, D3 was regressed onto D2, and so on. Due to slightly different item orders across data sets, the precise model specification varied.

The DSM-IV model was chosen for this purpose because neither the Numbing nor the Dysphoria model can be modified to incorporate an order effect in which the response to each item influences the response to the immediately subsequent item. The Numbing and Dysphoria models each contain one or more factors defined solely by two contiguous items. Each of these two-item factors formed from sequentially-administered items is fully redundant with the hypothesized order effect. In other words, because of the placement of these items in the symptom series, it is not possible to model simultaneously an order effect and a substantively meaningful two-item factor. With only two items comprising a factor, the two explanations are indistinguishable, and the model is not estimable.

Evaluating the fit of the order effects model—The Modified DSM-IV model proved dramatically superior to the standard DSM-IV model. As shown in Table 3, the Modified DSM-IV model constituted a statistically significant improvement over the DSM-IV model at the $p < .0001$ level in each of the 29 data sets. Similarly, as shown in Table 4, when aggregated across all studies, the results revealed the superiority of the Modified DSM-IV model, $\Delta\chi^2(464) = 17,531$, reflecting a chi-square improvement of 38 for each additional parameter. Based on the available data, it appears that more than half of the misfit for the DSM-IV model (as judged by model log-likelihood) is attributable to the tendency of the model to underestimate the correlations between the specific pairs of items that are presented sequentially. The evidence that order effects exist is not only statistically significant but also descriptively large and consistent across all data sets. The Modified DSM-IV model also provided a better fit than either of the four-factor models. As judged by the AIC, the Modified DSM-IV model was superior to the Numbing model in 28 of 29 data sets, and superior to the Dysphoria model in 27 of 29 data sets. In the aggregate data, the Modified DSM-IV model also had the lowest AIC and SRMR as well as the highest CFI.

Evaluating supplementary data relevant to order effects—In addition to the previously presented confirmatory factor analytic evidence which provides the primary findings consistent with an order effect, additional analyses also point to the possibility of an order effect. Specifically, we examined the extent to which naturally occurring variation in the order in which PTSD symptoms were presented in the 29 studies was associated with whether the D4, D5 data element emerged as the largest discrepancy in correlations predicted by the Numbing and Dysphoria models. Comparing the 23 data sets in which the D4, D5 element was the largest difference between models, 20 data sets (87%) had presented items in the DSM-IV order. By contrast, of the remaining 6 data sets in which some other matrix element best differentiated the models, only two (33%) were from studies that presented items in the DSM-IV item order, $\Delta\chi^2(1) = 7.47, p < .01$. These results show that variation in the order in which PTSD symptoms are presented is linked to differences in model discrepancies in predicted correlations. Specifically, when PTSD symptoms are

presented in the DSM-IV order, the Numbing model underestimates the D4, D5 correlation to a greater extent than when symptoms are presented in alternative orders. By contrast, the Dysphoria model always provides a better fit to this element because these two items form a 2-item factor.

Discussion

Extensive attention has been devoted to identifying the structure of PTSD symptoms using confirmatory factor analysis. In recent years, the main focus of research has expanded from documenting the fit of the DSM-IV model to determining which of two four-factor alternative conceptualizations of PTSD symptom structure should be preferred over the current DSM-IV framework (e.g., Yufik & Simms, 2010). This line of investigation has had considerable impact, undermining support for the existing DSM-IV structure and shaping provisional decisions concerning the criteria for PTSD in the forthcoming DSM-5 (American Psychiatric Association, 2010).

Our reanalysis of the existing confirmatory factor analytic literature supports two general conclusions. First, the nature of the empirical distinction between the Numbing and Dysphoria models appears to have been mischaracterized. Although the central feature of the Dysphoria model is ostensibly its differentiation of symptoms assessing general distress from PTSD-specific symptoms (Simms et al., 2002), the Numbing and Dysphoria models make nearly identical predictions regarding associations among all PTSD symptoms. The notable exception to this pattern involves the correlation between *hypervigilance* (D4) and *exaggerated startle response* (D5).

In those data sets in which the Dysphoria model fit better than the Numbing model, the advantage is due to the capacity of the Dysphoria model to explain the correlation between D4 and D5 (which is fit perfectly as a two-item factor) and not to the ability to account for correlations involving the other items that ostensibly measure dysphoria (i.e., C3 to D3). In other words, the most salient difference between the Numbing and Dysphoria models has nothing whatsoever to do with measurement of symptoms purported to assess dysphoria. Rather, the difference concerns two hyperarousal symptoms that are incidental to the Dysphoria model. These results, along with other recent findings (Marshall, Schell, & Miles, 2010; Miller et al., 2010) call into question the conceptual underpinnings of the Dysphoria model. After all, the Dysphoria model was intended to provide a better accounting of symptoms regarded as assessing non-specific distress and not merely to offer a better explanation of two hyperarousal symptoms.

Second, other models of PTSD symptom structure emerge as plausible representations when the inflexible presumption of simple structure is relaxed. In particular, a version of the DSM-IV model of PTSD—modified to allow for the possible influence of item presentation order—proved an excellent fit to the data and was generally superior to either four-factor alternative model. Thus, the widespread, but heretofore untested, assumption embedded in prominent models of PTSD symptom structure (i.e., that order effects play no role in producing the observed correlations among PTSD symptoms) appears inconsistent with the data. Indeed, the extent of model misfit that is plausibly attributed to an order effect is dramatically larger than the differences among the various factor models debated in the literature.

With few exceptions (e.g., Andrews, Joseph, Shevlin, & Troop, 2006; Miles, Marshall, & Schell, 2008), researchers have not considered *a priori* the possibility that useful models of PTSD symptom structure may fail to conform to simple structure. This preference for simple structure has resulted in a bias in favor of explaining all systematic relations among these

symptoms as indicating clinically meaningful features of PTSD. This predisposition is inconsistent with the long history of research that documents method effects in behavioral science research (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003; N. Schwartz, 1999; Tourangeau & Rasinski, 1988), and forecloses consideration of potentially meaningful models of symptom structure that may deviate from simple structure for reasons that are entirely justifiable based on prior research or theory.

The adverse impact of considering only a narrow range of PTSD models has been exacerbated in that most factor analytic studies have relied on PTSD instruments that present symptoms in the same order. Specifically, most studies administer PTSD measures which conform either to DSM-III-R or DSM-IV order, with these two sequences differing in the placement of only a single symptom. Moreover, the DSM-IV Numbing and Dysphoria factor models differ only with respect to where, within this common sequence, items are split into separate clusters. In other words, to the degree that the correlation between any two PTSD symptoms can be inflated by presenting these items sequentially, a bias has been incorporated into virtually every factor analytic study of PTSD symptom structure. As a result, it is possible that much of the debate about the structure of PTSD, as well as the proposed revisions to the definition of the disorder, has been driven by effects that may disappear if symptoms were asked in a different order.

Overall, our empirical results attest to the importance of distinguishing the structure of a specific measure of PTSD from the latent structure of PTSD as a clinical phenomenon. The two are commonly equated (e.g., Boelen, van den Hout, & van den Bout, 2008), and some have even suggested that it is unnecessary to draw a distinction between PTSD instruments and the constructs they measure (King, King, Orazam, & Palmieri, 2006). Yet, covariation among PTSD symptoms can be influenced by methodological factors in addition to the clinical phenomena of interest. Given that existing empirical work on PTSD symptom structure has been based on the implicit assumption that such methodological factors do not exist, researchers may be interpreting as clinically meaningful features of PTSD data that might be better attributed to specific measurement processes.

Model Selection for PTSD Research

Several factors may help to explain the widespread preference for simple structure models. Simple structure is familiar to researchers given its use in exploratory factor analysis wherein decisions must be made about a preferred factor rotation when an unlimited number of mathematically-equivalent solutions exist (Thurstone, 1947). Moreover, some deviations from simple structure (e.g., correlated errors) are viewed with suspicion because they may suggest a problematic, *a posteriori* strategy to improve inadequate models (MacCallum, Roznowski, & Necowitz, 1992). Irrespective of the reasons for favoring simple structure, departure from simple structure may be both justified and necessary to address shared method variance such as an order effect.

In this study, we considered a specific model of PTSD symptom structure that departed from simple structure. In addition to this particular alternative, however, other credible models that deviate from simple structure may also be well-fitting and useful for particular purposes. Depending on one's goals, models meriting attention include bi-factor models in which general and specific factors are simultaneously fit to the data (e.g., Andrews et al., 2006), and hybrid models of PTSD symptom structure in which, for example, the Numbing and Dysphoria models are simultaneously incorporated into a single model (e.g., Miles et al., 2008). Finally, models which posit that specific PTSD symptoms are not only shaped by trauma exposure but also by other PTSD symptoms over time appear useful for certain purposes (Schell, Marshall, & Jaycox, 2004; Marshall, Schell, Glynn, & Shetty, 2006).

The foregoing discussion of potentially useful models is consistent with the view that it may be advantageous for the field to recognize as quixotic the confirmatory factor analytic quest for a model that ‘best’ represents the structure of PTSD symptoms. Instead of viewing models as true or false, it may be preferable to focus on identifying models that are useful for specific purposes (Box, 1976; Box, 1979). As put succinctly by Box (1979), “all models are wrong but some are useful” (p. 202). It is not necessarily the case that any single model of PTSD symptom structure will emerge as the best for all objectives. Thus, attention should be devoted to determining the precise circumstances under which specific symptom structure models may prove useful.

The usefulness of a given model of PTSD symptom structure cannot be established solely by reference to the fit of confirmatory factor models. For this reason, the current review does not advocate that the field adopt a specific PTSD model on the basis of the relative model fits presented here. Model utility must be demonstrated empirically with respect to *external criteria* of scientific or clinical interest (Marshall et al., 2010; Miller et al., 2010). Thus, if one PTSD symptom model possesses demonstrated superiority over others in predicting response to therapeutic intervention, for example, then this evidence would attest to the value of this conceptualization. To this point, despite calls for greater attention to external evidence of validity (e.g., Marshall et al., 2010), model fit has been the primary means by which model superiority has been demonstrated.

Finally, these results pertain to the representation of PTSD in the forthcoming DSM-5. The American Psychiatric Association (2010) has issued a provisional pronouncement that the PTSD symptom structure framework will be revised in light of confirmatory factor analytic findings. Yet, the evidence base upon which this determination was made seems problematically incomplete. Researchers have almost invariably overlooked the possibility that non-substantive factors play a role in shaping the observed associations among PTSD symptoms. However, the current research indicates that order effects may inflate the correlations between pairs of PTSD symptoms that are presented sequentially. After accounting for the possibility of such methodological effects, the available data does not contain convincing evidence that either of the four factor models is superior to models that preserve the three substantive DSM-IV symptom clusters. Given that even scientifically defensible revisions to the DSM-5 may come at considerable cost (Frances, 2009), the decision to embrace a new conceptualization of PTSD symptom structure may be ill-advised.

Limitations

In considering these results, it is important to be mindful of the limitations of this study. First, this examination reanalyzed data from prior confirmatory factor analytic studies of PTSD symptom structure, and results may not be representative of all trauma-exposed individuals who develop PTSD symptoms. Similarly, to the extent that the analytic sample of studies used similar methods and instruments, this reanalysis will reflect any biases or errors that result from those methodological decisions. For example, our analyses were confined to studies using 17-item instruments that map directly onto the 17 symptoms found in the DSM-IV. Thus, the present results may not generalize to findings based on studies of instruments using a different number of symptoms.

In addition, although this reanalysis demonstrates that the DSM-IV model of PTSD, augmented to allow for order effects, provides a plausible account of most data sets examined here, future research is needed. In particular, although these results are strongly consistent with the existence of an order effect, these data do not rule out the possibility that another factor other than presentation order underlies the empirical observation that adjacent PTSD items are more strongly associated with one another than items placed further apart.

Similarly, while it is possible that some adjacent items, e.g., C1 and C2, or D4 and D5, may covary more strongly merely for substantive reasons (i.e., the presence of a theoretically interpretable latent variable), it is nonetheless the case that stronger associations for adjacent items are not limited to these two pairs of symptoms. The DSM-IV consistently underestimates the strength of the associations between almost all of the adjacent items, even while providing an excellent explanation of the associations between all non-adjacent items. Thus, it would appear that any explanation of the observation that all contiguous items are more strongly correlated—whether it is presentation order or some other explanation—must account for the systematic nature of this finding.

Future research that experimentally manipulates item order would help to directly address whether the observed results are best attributed to response order or to some other mechanism. Yet, it is important to note that experimental manipulation of item order does not solve the problem of how to measure PTSD in future studies. Even if response order was implicated, most researchers will neither want nor be able to administer PTSD items in multiple orders on a routine basis. The current analyses demonstrate an approach that allows for the possibility of modeling order effects within a conventional confirmatory factor analytic modeling framework.

Finally, the overall model fit results we present should not be used as the sole basis for drawing inferences regarding a preferred description of PTSD symptom structure. Although the modified DSM-IV model provides an excellent fit to the data, the Numbing and Dysphoria models are also adequate descriptions of the data. Moreover, a recent line of research indicates that a five-factor model of PTSD, in which items D1-D3 are differentiated from items D4-D5, might also provide a useful conceptualization of PTSD symptom structure (Elhai et al., 2011).¹² Although these results show that the DSM-IV model, when modified to allow for order effects, was the single best-fitting model that we tested, the current study should not be viewed as supporting the conclusion that this particular model be preferred over others. Ultimately, decisions about the best factor model for a given purpose should rest primarily on evidence supporting the validity and utility of the hypothesized latent constructs.

Conclusions

Despite the considerable energy devoted to examining the factor structure of PTSD symptoms, the scientific path the field has undertaken may not provide firm footing for DSM-5 revisions to the definition of the disorder. The current findings cast doubt on the claim that the DSM-IV model of PTSD must be changed to accommodate the findings from available factor analytic studies. Indeed, the only thing “wrong” with the DSM-IV model of PTSD is that it systematically underestimates the correlations between symptoms that are assessed sequentially when presented in the standard order. Insofar as non-clinical factors, rather than clinically meaningful features of PTSD, may be responsible for the apparent misfit of the DSM-IV model of PTSD symptom structure, the imminent revision of the DSM may be misguided.

Acknowledgments

This research was supported by grants R01MH56122, R01MH071636, and R01MH087657 from the National Institute of Mental Health and grant R01AA014246 from the National Institute on Alcohol Abuse and Alcoholism. The views expressed are ours and do not necessarily reflect those of the funders or RAND. We thank the many

¹²This recently-proposed five-factor model, which continues the trend toward drawing further distinctions based solely on the semantic content of PTSD symptoms, was not tested in the current analyses. This model, like the Numbing and Dysphoria models, cannot be modified to accommodate an order effect because certain factors are redundant with the order effect.

researchers who graciously shared their data with us as well as the thousands of trauma survivors who participated in these studies.

References

*References marked with an asterisk indicate studies included in the data analysis.

- Akaike H. Factor analysis and the AIC. *Psychometrika*. 1987; 52:317–332.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 3rd ed. Author; Washington, DC: 1980.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Revised 3rd ed. Author; Washington, DC: 1987.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed.. Author; Washington, DC: 1994.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. Revised 4th ed.. Author; Washington, DC: 2000.
- American Psychiatric Association. [November 11, 2011] DSM-5 development. Posttraumatic stress disorder. 2010. <http://www.dsm5.org/ProposedRevision/Pages/proposedrevision.aspx?rid=165#>
- American Psychiatric Association. [February 18, 2012] DSM-5 development. Proposed Draft Revisions to DSM Disorders and Criteria. 2011. <http://www.dsm5.org/proposedrevision/Pages/Default.aspx>
- American Psychiatric Association. [February 19, 2012] DSM-5 development. Frequently asked questions. 2011. <http://www.dsm5.org/about/Pages/faq.aspx#12>
- *. Andrews L, Joseph S, Shevlin M, Troop N. Confirmatory factor analysis of posttraumatic stress symptoms in emergency personnel: An examination of seven alternative models. *Personality and Individual Differences*. 2006; 41:213–224.
- Asmundson GJG, Frombach I, McQuaid J, Pedrelli P, Lenox R, Stein MB. Dimensionality of posttraumatic stress symptoms: A confirmatory factor analysis of DSM-IV symptom clusters and other symptom models. *Behaviour Research & Therapy*. 2000; 38:203–214. [PubMed: 10661004]
- Asmundson GJG, Wright KD, McCreary DR, Pedlar D. Post-traumatic stress disorder symptoms in United Nations peacekeepers: An examination of factor structure in peacekeepers with and without chronic pain. *Cognitive Behaviour Therapy*. 2003; 32:26–37. [PubMed: 16291532]
- *. Baschnagel JS, O'Connor RM, Colder CR, Hawk LW Jr. 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*. 2005; 18:677–684. [PubMed: 16382430]
- Bem DJ. Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*. 1967; 74:183–200. [PubMed: 5342882]
- Bentler PM. Comparative fit indexes in structural models. *Psychological Bulletin*. 1990; 107:238–246. [PubMed: 2320703]
- Bentler, PM. *EQS Structural Equations Program Manual*. Multivariate Software; Encino, CA: 1995.
- Blake DD, Weathers FW, Nagy LM, Kaloupek DG, Gusman FD, Charney DS, Keane TM. The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress*. 1995; 8:75–90. [PubMed: 7712061]
- *. Boelen PA, van den Hout MA, van den Bout J. The factor structure of posttraumatic stress disorder symptoms among bereaved individuals: A confirmatory factor analysis study. *Journal of Anxiety Disorders*. 2008; 22:1377–1383. [PubMed: 18342486]
- Bollen, KA. *Structural equations with latent variables*. Wiley; New York: 1989.
- Box GEP. Science and statistics. *Journal of the American Statistical Association*. 1976; 71:791–799.
- Box, GEP. Robustness in the strategy of scientific model building.. In: Launer, RL.; Wilkinson, GN., editors. *Robustness in statistics*. Academic Press; NY, NY: 1979. p. 201–236.
- Bovin MJ, Marx BP. The importance of the peritraumatic experience in defining traumatic stress. *Psychological Bulletin*. 2011; 137:47–67. [PubMed: 21090886]
- Burnham KP, Anderson DR. Multimodel Inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research* November. 2004; 33:261–304.

- Cheung MW, Chan W. Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods*. 2005; 10:40–64. [PubMed: 15810868]
- Collins AM, Quillian MR. Facilitating retrieval from semantic memory: The effect of repeating part of an inference. *Acta Psychologica*. 1970; 33:304–314.
- Cross MR, McCanne TR. Validation of a self-report measure of posttraumatic stress disorder in a sample of college-age women. *Journal of Traumatic Stress*. 2001; 14:135–147.
- *. Cuevas CA, Bollinger AR, Vielhauer MJ, Morgan EE, Sohler NL, Brief DJ, Miller AL, Keane TM. HIV/AIDS Cost Study: Construct validity and factor structure of the PTSD checklist in dually diagnosed HIV-seropositive adults. *Journal of Trauma Practice*. 2006; 5:29–51.
- *. Cyders MA, Burris JL, Carlson CR. Disaggregating the relationship between posttraumatic stress disorder symptom clusters and chronic orofacial pain: Implications for the prediction of health outcomes with PTSD symptom clusters. *Annals of Behavioral Medicine*. 2010; 1:1–12.
- Davidson JR, Book SW, Colket JT, Tupler LA, Roth S, David D, Hertzberg M, Mellman T, Beckham JC, Smith RD, Davison RM, Katz R, Feldman ME. Assessment of a new self-rating scale for post-traumatic stress disorder. *Psychological Medicine*. 1997; 27:153–160. [PubMed: 9122295]
- *. DuHamel KN, Ostrof J, Ashman T, Winkel G, Mundy EA, Keane TM, Morasco BJ, Vickberg SM, Hurley K, Burkhalter J, Chhabra R, Scigliano E, Papadopoulos E, Moskowitz C, Redd W. Construct validity of the posttraumatic stress disorder checklist in cancer survivors: Analyses based on two samples. *Psychological Assessment*. 2004; 16:255–266. [PubMed: 15456381]
- Elhai JD, Biehn TL, Aromour C, Klopper JJ, Frueh BC, Palmieri PA. Evidence for a unique PTSD construct represented by PTSD's D1-D3 symptoms. *Journal of Anxiety Disorders*. 2011; 25:240–345.
- *. Elhai JD, Engdahl RM, Palmieri PA, Naifeh JA, Schweinle A, Jacobs GA. Assessing posttraumatic stress disorder with or without reference to a single, worst traumatic event: Examining differences in factor structure. *Psychological Assessment*. 2009; 21:629–634. [PubMed: 19947796]
- *. Elhai JD, Gray MJ, Docherty AR, Kashdan TB, Kose S. Structural validity of the posttraumatic stress disorder checklist among college students with a trauma history. *Journal of Interpersonal Violence*. 2007; 22:1471–1478. [PubMed: 17925293]
- Elhai JD, Palmieri PA. The factor structure of posttraumatic stress disorder: A literature update, critique of methodology, and agenda for future research. *Journal of Anxiety Disorders*. 2011; 25:849–854. [PubMed: 21793239]
- *. Elhai JD, Palmieri P, Biehn TL, Frueh BC, Magruder KM. Posttraumatic stress disorder's frequency and intensity ratings are associated with factor structure differences in military veterans. *Psychological Assessment*. 2010; 22:723–728. [PubMed: 21038970]
- Elklit A, Shevlin M. The structure of PTSD symptoms: A test of alternative models using confirmatory factor analysis. *British Journal of Clinical Psychology*. 2007; 46:299–313. [PubMed: 17535523]
- *. Engdahl RM, Elhai JD, Richardson JD, Frueh BC. Comparing posttraumatic stress disorder's symptom structure between deployed and nondeployed veterans. *Psychological Assessment*. 2011; 23:1–6. [PubMed: 21171785]
- Festinger, L. A theory of cognitive dissonance. Row, Peterson; Evanston, IL: 1957.
- Feldman JM, Lynch JG. Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*. 1988; 73:421–435.
- Foa EB, Cashman L, Jaycox L, Perry K. The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. *Psychological Assessment*. 1997; 9:445–451.
- Foa EB, Riggs RS, Dancu CV, Rothbaum BO. Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *Journal of Traumatic Stress*. 1993; 6:459–473.
- Frances A. A warning sign on the road to *DSM-V*: Beware of its unintended consequences. *Psychiatric Times*. 2009; 26:1–4. <http://www.psychiatrictimes.com/dsm-5/content/article/10168/1425378>.
- Gelman A, Rubin DB. Avoiding model selection in Bayesian social research. *Sociological Methodology*. 1995; 25:165–173.
- Hallquist, M. [August 12, 2011] MplusAutomation: Automating Mplus model estimation and interpretation. R package version 0.4-3. 2011. <http://CRAN.R-project.org/package=MplusAutomation>

- Heider, F. The psychology of interpersonal relations. John Wiley & Sons; New York: 1958.
- Hershberger, SL. The problem of equivalent structural models.. In: Hancock, GR.; Mueller, RO., editors. Structural equation modeling: A second course. Information Age; Greenwich, CN: 2006. p. 13-41.
- *. Hetzel-Riggin MD. A test of structural invariance of posttraumatic stress symptoms in female survivors of sexual and/or physical abuse or assault. *Traumatology*. 2009; 15:46–59.
- Hoyt T, Yeater EA. Comparison of posttraumatic stress disorder symptom structure models in Hispanic and White college students. *Psychological Trauma: Theory, Research, Practice, and Policy*. 2010; 2:19–30.
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55.
- King LA, King DW, Orazem RJ, Palmieri PA. Research on the latent structure of PTSD. *PTSD Quarterly*. 2006; 17:1–7.
- King DW, Leskin GA, King LA, Weathers FW. Confirmatory factor analysis of the Clinician-Administered PTSD Scale: Evidence for the dimensionality of posttraumatic stress disorder. *Psychological Assessment*. 1998; 10:90–96.
- Kline, P. The handbook of psychological testing. Routledge; London, UK: 1993.
- Kline, RB. Principles and practice of structural equation modeling. The Guilford Press; New York, NY: 2010.
- *. Krause ED, Kaltman S, Goodman LA, Dutton MA. Longitudinal factor structure of posttraumatic stress symptoms related to intimate partner violence. *Psychological Assessment*. 2007; 19:165–175. [PubMed: 17563198]
- Lee S, Hershberger S. A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behavioral Research*. 1990; 25:313–334.
- *. Mansfield AJ, Williams J, Hourani LL, Babeu LA. Measurement invariance of posttraumatic stress disorder symptoms among U.S. military personnel. *Journal of Traumatic Stress*. 2010; 23:91–99. [PubMed: 20135678]
- MacCallum RC, Roznowski M, Necowitz LB. Model modification in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*. 1992; 111:490–504. [PubMed: 16250105]
- MacCallum RC, Wegener DT, Uchino BN, Fabrigar LR. The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*. 1993; 114:185–199. [PubMed: 8346326]
- Marsh, HW.; Hau, K-T.; Grayson, D. Goodness of fit in structural equation models.. In: Maydeu-Olivares, A.; McArdle, JJ., editors. *Contemporary Psychometrics: A Festschrift for Roderick P. McDonald*. Lawrence Erlbaum Associates; Mahwah, NJ: 2005. p. 225-340.
- *. Marshall GN. Posttraumatic Stress Disorder Symptom Checklist: Factor structure and English-Spanish measurement invariance. *Journal of Traumatic Stress*. 2004; 17:223–230. [PubMed: 15253094]
- Marshall GN, Schell TL, Glynn SM, Shetty V. The role of hyperarousal in the manifestation of posttraumatic psychological distress following injury. *Journal of Abnormal Psychology*. 2006; 115:624–628. [PubMed: 16866603]
- Marshall GN, Schell TL, Miles JNV. All PTSD symptoms are highly associated with general distress: Ramifications for the dysphoria symptom cluster. *Journal of Abnormal Psychology*. 2010; 119:126–135. [PubMed: 20141249]
- *. McDonald SD, Beckham JC, Morey R, Marx C, Tupler LA, Calhoun PS. Factorial invariance of posttraumatic stress disorder symptoms across three veteran samples. *Journal of Traumatic Stress*. 2008; 21:309–317. [PubMed: 18553409]
- Miles JNV, Marshall GN, Schell TL. Spanish and English versions of the PTSD Checklist–Civilian version (PCL-C): Testing for differential item functioning. *Journal of Traumatic Stress*. 2008; 21:369–376. [PubMed: 18720394]
- Miller MW, Wolf EJ, Harrington KM, Brown TA, Kaloupek DG, Keane TM. An evaluation of competing models for the structure of PTSD symptoms using external measures of comorbidity. *Journal of Traumatic Stress*. 2010; 23:631–638. [PubMed: 20839312]

- Mollica RF, Caspi-Yavin Y, Bollini P, Truong T, Tor S, Lavelle J. The Harvard Trauma Questionnaire: Validating a cross-cultural instrument for measuring torture, trauma and posttraumatic stress disorder in Indochinese refugees. *Journal of Nervous and Mental Disease*. 1992; 180:111–116. [PubMed: 1737972]
- *. Morina N, Böhme HF, Morina L, Asmundson GJ. The structure of post-traumatic stress symptoms in young survivors of war. *Psychiatry Research*. 2011; 186:306–309. [PubMed: 20728944]
- Muthén, B.; Muthén, L. Mplus 6.1. Author; Los Angeles, CA: 2008.
- *. Naifeh JA, Elhai JD, Kashdan TB, Grubaugh AL. The PTSD symptom scale's latent structure: An examination of trauma-exposed medical patients. *Journal of Anxiety Disorders*. 2008; 22:1355–1368. [PubMed: 18337058]
- *. Naifeh JA, Richardson JD, Del Ben KS, Elhai JD. Heterogeneity in the latent structure of PTSD symptoms among Canadian veterans. *Psychological Assessment*. 2010; 22:666–674. [PubMed: 20822279]
- Osgood, CE.; Suci, GJ.; Tannenbaum, PH. The measurement of meaning. The University of Illinois Press; Urbana, IL: 1957.
- *. Palmieri PA, Fitzgerald LF. Confirmatory factor analysis of posttraumatic stress symptoms in sexually harassed women. *Journal of Traumatic Stress*. 2005; 18:657–666. [PubMed: 16382424]
- *. Palmieri PA, Marshall GN, Schell TL. Confirmatory factor analysis of posttraumatic stress symptoms in Cambodian refugees. *Journal of Traumatic Stress*. 2007a; 20:207–216. [PubMed: 17427910]
- *. Palmieri PA, Weathers FW, Difede J, King DW. Confirmatory factor analysis of the PTSD Checklist and the Clinician-Administered PTSD Scale in disaster workers exposed to the World Trade Center Ground Zero. *Journal of Abnormal Psychology*. 2007b; 116:329–341. [PubMed: 17516765]
- Pearl, J. Probabilistic reasoning in intelligent systems: Networks of plausible inference. Kaufmann; San Mateo, CA: 1988.
- Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*. 2003; 88:879–903. [PubMed: 14516251]
- Raykov T, Marcoulides GA. Can there be infinitely many models equivalent to a given covariance structure model? *Structural Equation Modeling: A Multidisciplinary Journal*. 2001; 8:142–149.
- Schell TL, Klein SB, Babey S. Testing a hierarchical model of self-knowledge. *Psychological Science*. 1996; 7:170–173.
- Schell TL, Marshall GN, Jaycox LH. All symptoms are not created equal: The prominent role of hyperarousal in the natural course of posttraumatic psychological distress. *Journal of Abnormal Psychology*. 2004; 113:189–197. [PubMed: 15122939]
- *. Scher CD, McCreary DR, Asmundson GJ, Resick PA. 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*. 2008; 22:1137–1145. [PubMed: 18206346]
- Schwarz GE. Estimating the dimension of a model. *Annals of Statistics*. 1978; 6:461–464.
- Schwarz N. Self-reports: How the questions shape the answers. *American Psychologist*. 1999; 54:93–105.
- Schwarz, N.; Sudman, S. Order effects in social and psychological research. Springer-Verlag; New York, NY: 1992.
- Scott WJ. PTSD in DSM-III: A case in the politics of diagnosis and disease. *Social Problems*. 1990; 37:294–310.
- Shipley B. A new inferential test for path models based on directed acyclic graphs. *Structural Equation Modeling*. 2000; 7:206–218.
- *. Simms LJ, Watson D, Doebbeling BN. Confirmatory factor analyses of posttraumatic stress symptoms in deployed and nondeployed veterans of the Gulf War. *Journal of Abnormal Psychology*. 2002; 111:637–647. [PubMed: 12428777]
- Spitzer RL, First MB, Wakefield JC. Saving PTSD from itself in DSM-V. *Journal of Anxiety Disorders*. 2007; 21:233–241. [PubMed: 17141468]

- Summerfield D. The invention of post-traumatic stress disorder and the social usefulness of a psychiatric category. *British Medical Journal*. 2001; 322:95–98. [PubMed: 11154627]
- Thurstone, LL. *Multiple-factor analysis*. University of Chicago Press; Chicago: 1947.
- Tourangeau R, Rasinski KA. Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin*. 1988; 103:299–314.
- Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. *Science*. 1974; 185:1124–1131. [PubMed: 17835457]
- *. Ullman SE, Long SM. Factor structure of PTSD in a community sample of sexual assault survivors. *Journal of Trauma and Dissociation*. 2008; 9:507–524. [PubMed: 19042794]
- Weathers, FW.; Litz, BT.; Herman, DS.; Huska, JA.; Keane, TM. The PTSD Checklist (PCL): Reliability, validity, and diagnostic validity.. Paper presented at the Annual Meeting of the International Society for Traumatic Stress Studies; 1993, October 24-27;
- Young, A. *The harmony of illusions: Inventing post-traumatic stress disorder*. Princeton University Press; Princeton, NJ: 1995.
- Yufik T, Simms LJ. A meta-analytic investigation of the structure of posttraumatic stress disorder symptoms. *Journal of Abnormal Psychology*. 2010; 119:764–776. [PubMed: 21090877]

Table 1**Item Mappings for Alternative Models of PTSD symptoms**

DSM-IV PTSD symptom (abbreviated content)	Model		
	1	2	3
B1. Intrusive thoughts of trauma	R	R	R
B2. Recurrent dreams of trauma	R	R	R
B3. Flashbacks	R	R	R
B4. Emotional reactivity to trauma cues	R	R	R
B5. Physiological reactivity to trauma cues	R	R	R
C1. Avoiding thoughts of trauma	A, N	A	A
C2. Avoiding reminders of trauma	A, N	A	A
C3. Inability to recall aspects of trauma	A, N	N	D
C4. Loss of interest	A, N	N	D
C5. Detachment	A, N	N	D
C6. Restricted affect	A, N	N	D
C7. Sense of foreshortened future	A, N	N	D
D1. Sleep disturbance	H	H	D
D2. Irritability	H	H	D
D3. Difficulty concentrating	H	H	D
D4. Hypervigilance	H	H	H
D5. Exaggerated startle response	H	H	H

Note. PTSD = Posttraumatic Stress Disorder. Symptom clusters to which items contribute: R = Reexperiencing; A = Avoidance; N = Numbing; H = Hyperarousal; D = Dysphoria. Model 1 = DSM-IV; Model 2 = Numbing; Model 3 = Dysphoria. B1 through D5 correspond to symptom numbers as found in the DSM-IV.

Table 2

Characteristics of Data Sets Included in the Analyses

Study	Trauma	N ^a	Scale	Presentation Order
Andrews et al. (2006)	Emergency Responders	458	PSS	DSM-III-R
Baschnagel, O'Connor, Colder & Hawk (2005)	Terrorism	528	PDS	DSM-IV
Boelen et al. (2008)	Bereavement	347	PSS	DSM-IV
Cuevas et al. (2008)	HIV	208	PCL-C	DSM-IV
Cyders, Burris, & Carlson (2010)	Pain	215	PCL-C	DSM-IV
DuHamel et al. (2004)	Cancer	236	PCL-C	DSM-IV
Elhai et al. (2009)	Unspecified ^c	452	PCL-S	DSM-IV
Elhai, Gray, Docherty, Kashdan, & Kose (2007)	Unspecified ^c	215	PCL-S	DSM-IV
Elhai, Palmieri, Biehn, Frueh, & Magruder (2010)	Military	669	CAPS	DSM-IV
Engdahl, Elhai, Richardson, & Frueh (2011)	Military	1066	PCL-M	DSM-IV
Hetzel-Riggin (2009)	Abuse/Assault	2378	PTSD-Q	DSM-IV
Krause et al. (2007) Sample 1	IPV Service Seeking	405	PCL-C	DSM-IV
Krause et al. (2007) Sample 2	IPV	397	PCL-C	DSM-IV
Mansfield et al. (2010)	Military	15603	PCL-C	DSM-IV
Marshall (2004)	Community Violence	294	PCL-S	DSM-IV
McDonald et al. (2008) Sample 1	Military	320	DTS	DSM-III-R
McDonald et al. (2008) Sample 2	Military	313	DTS	DSM-III-R
McDonald et al. (2008) Sample 3	Military	814	DTS	DSM-III-R
Morina, Böhme, Morina, & Asmundson (2011)	Civil War (Civilians)	390	PSS	DSM-IV
Naifeh et al. (2008)	Unspecified ^b	252	PSS	DSM-IV
Naifeh et al. (2010)	Military	405	PCL-M	DSM-IV
Palmieri & Fitzgerald (2005)	Sexual Harassment	1218	PCL-C	DSM-IV
Palmieri et al. (2007a)	Refugees	527	HTQ	Nonstandard ^c
Palmieri et al. (2007b)	WTC Workers	3286	CAPS	DSM-IV
Scher et al. (2008) Sample 1	Sexual Assault	216	PSS	DSM-III-R
Scher et al. (2008) Sample 2	IPV	427	PDS	DSM-IV
Scher et al. (2008) Sample 3	Sexual Assault	280	CAPS	DSM-IV
Simms et al. (2002)	Military	1896	PCL-M	Nonstandard ^d
Ullman & Long (2008)	Sexual Assault	967	PDS	DSM-IV

Note. CAPS = Clinician-Administered PTSD Scale (Blake et al., 1995); PDS = Posttraumatic Diagnostic Scale (Cashman, Jaycox, & Perry, 1997); PSS = PTSD Symptom Scale (Foa, Riggs, Dancu, & Rothbaum, 1993); PCL = PTSD Symptom Checklist (Weathers et al., 1993); HTQ = Harvard Trauma Questionnaire (Mollica et al., 1992); DTS = Davidson Trauma Scale (Davidson et al., 1997); PTSD-Q = Posttraumatic Stress Disorder Questionnaire (Cross & McCann, 2001).

IPV = Intimate Partner Violence. PCL-M, PCL-C, and PCL-S correspond to military, civilian and specific versions of the PCL.

^a Minor discrepancies from sample sizes reported in original articles reflect differences in the data sets we received and/or differing approaches for handling missing data.

^b Insufficient information was provided by the authors to permit characterization of the nature of trauma exposure.

^c In the HTQ, PTSD symptoms are presented in an idiosyncratic order that does not conform to either the DSM-III-R or DSM-IV orders.

^d Although Simms et al. (2002) used the PCL-M, personal communication with the author revealed that the items had not been administered in the standard order.

Table 3

Fit Indices for Alternative Models in the Individual Datasets

Study	Fit Index	Model			
		DSM-IV	Numbing	Dysphoria	Modified DSM-IV
Andrews et al. (2006)	χ^2	602	504	566	463
	CFI	0.88	0.90	0.89	0.91
	SRMR	0.054	0.049	0.05	0.048
	AIC	10729	10638	10699	10599
Baschnagel et al. (2005)	χ^2	601	446	407	352
	CFI	0.84	0.89	0.90	0.92
	SRMR	0.059	0.048	0.047	0.044
	AIC	16106	15957	15918	15865
Boelen et al. (2008)	χ^2	266	233	230	209
	CFI	0.92	0.94	0.94	0.95
	SRMR	0.051	0.048	0.046	0.045
	AIC	13861	13835	13831	13812
Cuevas et al. (2008)	χ^2	269	206	195	185
	CFI	0.90	0.94	0.95	0.95
	SRMR	0.053	0.046	0.043	0.045
	AIC	8718	8661	8650	8641
Cyders et al. (2010)	χ^2	441	345	360	299
	CFI	0.90	0.93	0.93	0.94
	SRMR	0.054	0.044	0.042	0.044
	AIC	12259	12207	12222	12125
DuHamel et al. (2004)	χ^2	279	212	226	210
	CFI	0.87	0.92	0.91	0.92
	SRMR	0.058	0.051	0.051	0.051
	AIC	10411	10349	10364	10349

Study	Fit Index	Model				
		DSM-IV	Numbing	Dysphoria	Modified Numbing	Modified DSM-IV
Elhai et al. (2009)	χ^2	725	526	505	436	412
	CFI	0.88	0.92	0.93	0.94	0.94
	SRMR	0.056	0.042	0.042	0.042	0.042
	AIC	19466	19272	19251	19184	19184
Elhai et al. (2007)	χ^2	265	234	205	186	179
	CFI	0.92	0.94	0.95	0.96	0.96
	SRMR	0.050	0.045	0.041	0.041	0.043
	AIC	10089	10065	10036	10019	10036
Elhai et al. (2010)	χ^2	982	598	750	567	490
	CFI	0.94	0.97	0.96	0.97	0.97
	SRMR	0.03	0.024	0.026	0.024	0.02
	AIC	16882	16503	16655	16475	16422
Engdahl et al. (2011)	χ^2	785	644	560	596	430
	CFI	0.95	0.96	0.97	0.96	0.98
	SRMR	0.031	0.028	0.025	0.029	0.026
	AIC	44695	44559	44476	44514	44371
Hetzel-Riggin (2009)	χ^2	1572	1368	1624	1297	1069
	CFI	0.94	0.94	0.93	0.95	0.96
	SRMR	0.039	0.036	0.035	0.036	0.033
	AIC	93850	93652	93908	93583	93379
Krause, et al. (2007) Sample 1	χ^2	484	380	328	281	183
	CFI	0.92	0.94	0.95	0.96	0.98
	SRMR	0.046	0.041	0.037	0.039	0.029
	AIC	15613	15515	15463	15418	15344
Krause, et al. (2007) Sample 2	χ^2	383	309	296	270	214
	CFI	0.91	0.94	0.94	0.95	0.96
	SRMR	0.045	0.041	0.038	0.041	0.032

Study	Fit Index	Model			
		DSM-IV	Numbing	Dysphoria	Modified Numbing
	AIC	16388	16320	16307	16283
	χ^2	15821	10085	10374	8365
	CFI	0.92	0.95	0.95	0.96
Mansfield et al. (2010)	SRMR	0.037	0.03	0.029	0.031
	AIC	561451	555721	556010	554003
	χ^2	276	254	234	224
	CFI	0.93	0.94	0.95	0.95
Marshall (2004)	SRMR	0.047	0.044	0.042	0.042
	AIC	14419	14404	14384	14376
	χ^2	317	263	292	256
	CFI	0.92	0.94	0.93	0.94
McDonald et al. (2008) Sample 1	SRMR	0.056	0.05	0.052	0.049
	AIC	13337	13289	13318	13284
	χ^2	452	385	426	364
	CFI	0.93	0.94	0.93	0.95
McDonald et al. (2008) Sample 2	SRMR	0.045	0.039	0.04	0.039
	AIC	10778	10717	10758	10698
	χ^2	649	474	595	444
	CFI	0.92	0.95	0.93	0.95
McDonald et al. (2008) Sample 3	SRMR	0.052	0.045	0.049	0.044
	AIC	33238	33069	33190	33041
	χ^2	344	274	302	270
	CFI	0.93	0.95	0.94	0.95
Morina et al. (2010)	SRMR	0.043	0.038	0.041	0.038
	AIC	16051	15987	16015	15985
	χ^2	409	375	424	313
Naifeh et al. (2008)					258

Study	Fit Index	Model				
		DSM-IV	Numbing	Dysphoria	Modified Numbing	Modified DSM-IV
	CFI	0.91	0.92	0.90	0.94	0.95
	SRMR	0.048	0.046	0.051	0.053	0.040
	AIC	7384	7357	7406	7296	7265
Naifeh et al. (2010)	χ^2	409	343	314	308	232
	CFI	0.91	0.93	0.94	0.94	0.96
	SRMR	0.055	0.049	0.047	0.048	0.043
	AIC	20719	20659	20630	20626	20573
Palmieri, Weathers, et al. (2007)	χ^2	1677	1287	1384	1227	872
	CFI	0.91	0.93	0.93	0.94	0.96
	SRMR	0.041	0.038	0.035	0.037	0.03
	AIC	185455	185070	185168	185012	184682
Palmieri, Marshall & Schell (2007)	χ^2	467	416	441	414	322
	CFI	0.93	0.94	0.93	0.94	0.96
	SRMR	0.04	0.039	0.04	0.039	0.034
	AIC	17904	17860	17885	17859	17791
Palmieri & Fitzgerald (2005)	χ^2	1735	916	1228	792	838
	CFI	0.90	0.95	0.93	0.96	0.95
	SRMR	0.061	0.041	0.040	0.042	0.042
	AIC	58143	57330	57642	57208	57278
Scher et al. (2008) Sample 1	χ^2	207	201	188	176	156
	CFI	0.94	0.94	0.95	0.96	0.96
	SRMR	0.046	0.045	0.043	0.042	0.040
	AIC	9681	9682	9669	9659	9663
Scher et al. (2008) Sample 2	χ^2	329	294	231	210	200
	CFI	0.91	0.93	0.95	0.96	0.96
	SRMR	0.049	0.044	0.038	0.038	0.039
	AIC	18899	18870	18807	18788	18801

Study	Model					
	Fit Index	DSM-IV	Numbing	Dysphoria	Modified Numbing	Modified DSM-IV
Scher et al. (2008) Sample 3	χ^2	248	171	146	148	196
	CFI	0.89	0.95	0.97	0.97	0.92
	SRMR	0.060	0.048	0.044	0.046	0.053
	AIC	21592	21520	21495	21499	21571
Simms et al. (2002)	χ^2	698	577	493	504	409
	CFI	0.85	0.88	0.90	0.90	0.92
	SRMR	0.059	0.048	0.045	0.048	0.045
	AIC	15592	15477	15392	15406	15334
Ullman & Long (2008)	χ^2	1252	1066	867	670	456
	CFI	0.88	0.90	0.92	0.94	0.96
	SRMR	0.054	0.047	0.043	0.042	0.035
	AIC	38455	38275	38076	37881	37691

Note. When estimated in a single model, DSM-IV three-factor model $df=116$; Numbing model $df=113$; Dysphoria model $df=113$; Modified Numbing (with correlated residual between D4, D5) $df=112$; Modified DSM-IV (with sequential order effects) $df=100$. CFI = Comparative Fit Index (Bentler, 1990); SRMR = Standardized root mean square residual (Bentler, 1995); AIC = Akaike's Information Criterion (Akaike, 1987).

Table 4

Fit Indices for Alternative Models in a Multiple-Group Analysis

Fit Index	DSM-IV	Numbing	Dysphoria	Modified Numbing	Modified DSM-IV
df	3364	3277	3277	3248	2900
χ^2	32942	23383	24189	20034	15411
CFI	0.92	0.95	0.94	0.95	0.97
SRMR	0.043	0.036	0.039	0.036	0.031
AIC	1332163	1322816	1323621	1319488	1315560

Note. The Modified Numbing model includes correlated residual between D4 and D5; The Modified DSM-IV model includes sequential order effects. CFI = Comparative Fit Index (Bentler, 1990); SRMR = Standardized root mean square residual (Bentler, 1995); AIC = Akaike's Information Criteria (Akaike, 1987).

Table 5

Mean Differences in Correlations Implied by the Numbing and Dysphoria Models

	B1	B2	B3	B4	B5	C1	C2	C3	C4	C5	C6	C7	D1	D2	D3	D4
B2	.00															
B3	.00	.00														
B4	.00	.00	.00													
B5	.00	.00	.00	.00												
C1	.00	.00	.00	.00	.00											
C2	.00	.00	.00	.00	.00	.00										
C3	-.01	-.01	-.01	-.01	-.01	.00	.00									
C4	-.01	-.01	-.01	-.01	-.01	.01	.01	.01								
C5	.00	.00	.00	.00	.00	.01	.01	.02	.04							
C6	-.01	-.01	-.01	.00	-.01	.01	.01	.01	.03	.04						
C7	-.01	-.01	-.01	-.01	-.01	.00	.00	.01	.02	.03	.02					
D1	.02	.02	.02	.02	.02	.00	.00	-.03	-.03	-.03	-.03	-.04				
D2	.02	.02	.02	.02	.02	.00	.00	-.03	-.04	-.03	-.04	-.04	.03			
D3	.02	.02	.01	.02	.02	-.01	.00	-.03	-.04	-.04	-.04	-.05	.02	.02		
D4	-.01	-.01	-.01	-.01	-.01	-.02	-.02	-.01	.00	.00	.00	-.01	.06	.05	.06	
D5	-.02	-.02	-.02	-.02	-.02	-.02	-.02	-.01	.00	.00	.00	-.01	.06	.06	.06	-.13

Note. Item numbers correspond to PTSD symptoms B2 through D4 in the DSM-IV. Positive values indicate parameters wherein the Numbing model predicted larger values relative to the Dysphoria model. Negative coefficients correspond to instances in which the Numbing model produced smaller estimates in comparison to the Dysphoria model. Estimates greater than or equal to an absolute value of .05 are presented in boldface.

Table 6

Location and Size of the Four Largest Differences in Estimated Correlations

Sample	Largest Differences			
	1	2	3	4
Andrews et al. (2006)	-0.177 D4, D5	0.076 B5, D3	0.065 D1, D2	0.065 D2, D3
Baschnagel et al. (2005)	-0.165 D4, D5	0.99 D3, D4	0.93 D3, D5	0.88 D1, D4
Boelen et al. (2008)	-0.175 D4, D5	0.060 D3, D4	0.057 D3, D5	0.048 D2, D3
Cuevas et al. (2008)	-0.199 D4, D5	0.060 D2, D5	0.059 D1, D5	0.052 D3, D5
Cyders et al. (2010)	-0.178 D4, D5	0.084 D2, D5	0.077 D3, D5	0.072 D1, D5
DuHamel et al. (2004)	0.060 D3, D4	0.05 C5, C6	0.056 D1, D5	0.055 D2, D5
Elhai et al. (2009)	0.132 D4, D5	0.069 D3, D4	0.066 D2, D4	0.059 D3, D5
Elhai et al. (2007)	-0.263 D4, D5	0.070 C5, C6	0.067 C4, C6	0.063 C4, C5
Elhai et al. (2010)	0.058 D1, D4	-0.054 D4, D5	0.051 B4, D2	0.051 B1, D2
Engdahl et al. (2011)	-0.075 D4, D5	0.046 D2, D5	-0.045 B1, D4	-0.045 B5, D4
Hetzel-Riggin (2009)	-0.101 D4, D5	0.079 C5, C6	-0.075 C2, D5	0.072 C1, D5
Krause et al. (2007)	-0.200 D4, D5	0.090 D3, D4	0.085 D1, D4	0.080 D3, D5
Krause et al. (2007)	-0.188 D4, D5	0.075 D3, D5	0.067 D2, D5	0.061 D1, D5
Mansfield et al. (2010)	-0.123 D4, D5	0.069 D2, D4	0.064 D3, D4	0.064 D2, D5
Marshall (2004)	-0.148 D4, D5	0.071 D2, D5	0.067 D2, D4	0.054 D3, D4
McDonald et al. (2008)	0.096 D2, D5	0.082 D2, D4	0.080 D1, D5	0.074 D3, D5
McDonald et al. (2008)	0.068 D2, D4	0.065 D1, D4	0.060 C6, D3	0.050 D2, D5
McDonald et al. (2008)	0.086 C4, C5	-0.083 C7, D3	-0.082 C6, D3	0.081 D2, D5
Morina et al. (2011)	0.050 C5, C6	0.045 C4, C7	0.043 C7, D2	0.043 C7, D5
Naifeh et al. (2008)	-0.125 D4, D5	0.101 D3, D5	0.099 D1, D5	0.095 D3, D4
Naifeh et al. (2010)	-0.131 D4, D5	0.072 D1, D4	0.068 B1, D1	0.065 D1, D5
Palmieri et al. (2007a)	0.057 D1, D5	0.052 D2, D5	0.048 D3, D5	0.048 C7, D2
Palmieri et al. (2007b)	-0.119 D4, D5	0.065 D2, D5	0.057 D3, D5	0.051 D1, D5
Palmieri & Fitzgerald (2005)	-0.114 D4, D5	0.082 C4, C5	0.079 C5, C6	0.072 C4, C6
Sclier et al. (2008)	-0.165 D4, D5	0.078 D3, D4	0.075 D3, D5	0.054 C5, C6
Scher et al. (2008)	-0.138 D4, D5	0.114 D3, D5	0.113 D3, D4	0.099 D1, D5
Scler et al. (2008)	-0.172 D4, D5	0.132 D1, D5	0.122 D1, D4	0.106 B4, D1
Simms et al. (2002)	-0.178 D4, D5	0.041 D1, D5	0.039 D2, D5	0.035 D2, D5
Ullman & Long (2008)	-0.286 D4, D5	0.094 D3, D4	0.093 D2, D4	0.087 D3, D5

Note. Implied correlations for D4, D5 are presented in boldface.