The Structure of PTSD Among Two Cohorts of Returning Soldiers: Before, During, and Following Deployment to Iraq

Laura A. Meis  
Center for Chronic Disease Outcomes Research, Minneapolis, Minnesota; Minneapolis Veterans Affairs Health Care System, Minneapolis, Minnesota; and University of Minnesota Medical School

Christopher R. Erbes  
Minneapolis Veterans Affairs Health Care System, Minneapolis, Minnesota; and University of Minnesota Medical School

Matthew E. Kaler  
Minneapolis Veterans Affairs Health Care System, Minneapolis, Minnesota

Paul A. Arbisi  
Minneapolis Veterans Affairs Health Care System, Minneapolis, Minnesota; and University of Minnesota Medical School

Melissa A. Polusny  
Center for Chronic Disease Outcomes Research, Minneapolis, Minnesota; Minneapolis Veterans Affairs Health Care System, Minneapolis, Minnesota; and University of Minnesota Medical School

Evidence suggests either a four-factor emotional numbing or dysphoria model likely reflects the underlying structure of posttraumatic stress disorder (PTSD). Questions remain as to which of these structures best represents PTSD, how the structure changes with time, the applicability of models to returning veterans, and the validity of the symptom clusters. The present study addresses these questions among two longitudinal samples of National Guard soldiers assessed prior to, during, and following a combat deployment to Iraq. Findings support a four-factor intercorrelated dysphoria model of PTSD that remains stable across samples and time points. Differential associations were observed among PTSD symptom clusters over time and between symptom clusters and both depression and combat exposure, supporting important distinctions between symptom clusters.

**Keywords:** posttraumatic stress disorder, latent factor structure, longitudinal, returning veterans

According to the *Diagnostic and Statistical Manual, Fourth Edition (DSM–IV–TR)*, American Psychiatric Association, 2000), a posttraumatic stress disorder diagnosis (PTSD) requires an index traumatic event (Criterion A) and a requisite number of symptoms from each of three clusters: intrusive memories of the event (Criterion B), avoidance (Criterion C), and hyperarousal (Criterion D). However, considerable factor analytic evidence demonstrates these clusters fail to capture the underlying structure of PTSD across multiple populations, instruments, and cultures (e.g., King, Leskin, King, & Weathers, 1998; King, Orazem, Lauterbach, King, & Hebenstreit, 2009; McDonald et al., 2008; Naifeh, Elhai, Kashdan, & Grubaug, 2008; Palmieri, Marshall, & Schell, 2007). Two four-factor solutions, including either an emotional numbing (King et al., 1998) or dysphoria factor (Simms, Watson, & Doebling, 2002), receive the most support as alternative models (King, King, Orazem, & Palmieri, 2006).

These two models retain important conceptual differences. The King et al. (1998) model differentiates emotional numbing, an automatic muting of emotions in response to chronic elevated arousal (Foa, Riggs, & Gershuny, 1995; Foa, Zinbarg, & Rothbaum, 1992), from effortful or intentional avoidance of trauma cues (Asmundson, Stapleton, & Taylor, 2004), resulting in four

This research was supported by grants from the Minnesota Medical Foundation (Grant #3662-9227-06) and Department of Defense Con-
factors: intrusive memories, effortful avoidance, emotional numbing, and hyperarousal. Numbing, as distinct from avoidance, is less responsive to intervention, predicts poorer treatment response, and correlates more consistently with depression and attention than avoidance (Aasmunson et al., 2004). In numerous populations, the numbing model has received support over the dysphoria model (e.g., Naifeh et al., 2008; Palmieri & Fitzgerald, 2005; Palmieri, Marshall, et al., 2007; Schinka, Brown, Borenstein, & Mortimer, 2007; Suvak, Maguen, Litz, Silver, & Holman, 2008; Witteven et al., 2006), including recent samples of service members deployed to Iraq (Operation Iraqi Freedom, OIF) and Afghanistan (Operation Enduring Freedom, OEF; Mansfield, Williams, Hourani, & Babeu, 2010; McDonald et al., 2008).

Alternatively, the dysphoria model (Simms et al., 2002) includes a nonspecific suffering or dysphoria factor (e.g., difficulty experiencing emotion, loss of interest in activities, feeling distant from others, sleep problems, irritability, and impaired concentration) and trauma-specific PTSD-specific factors: intrusive memories, trauma-specific effortful avoidance, and PTSD-specific hyperarousal. The dysphoria model was originally framed within hierarchical models of anxiety and depression that grouped symptoms into either nonspecific distress symptoms (e.g., dysphoria) or disorder specific symptoms (e.g., reexperiencing, trauma specific avoidance, and hyperarousal; Brown, Chorpita, & Barlow, 1998; Clark & Watson, 1991; Mineka, Watson, & Clark, 1998). However, it is also consistent with a more modern, flexible model of anxiety and mood disorders that considers how each symptom both represents general distress and is specific to either depression or anxiety (Watson, 2009). The dysphoria model has also received support over the numbing model across multiple studies and populations (e.g., Armour & Shevlin, 2010; Boelen, van den Hout, & van den Bout, 2008; Baschnagel, O’Connor, Colder, & Hawk, 2005; Carragher, Mills, Slade, Teesson, & Silove, 2010; Elklit & Shevlin, 2007; Krause, Kaltman, Goodman, & Dutton, 2007; Olff, Sijbrandij, Opmeer, Carlier, & Gersons, 2009; Palmieri, Weathers, Difede, & King, 2007; Simms et al., 2002), including among OEF/OIF servicemembers (Pietrzak, Goldstein, Malley, Rivers, & Southwick, 2010). Greater cross-sectional associations are demonstrated between the dysphoria symptom cluster and measures of distress, including depression, generalized anxiety, anger, hopelessness, complicated grief, suicidal ideation, and severity of distress than between other symptom clusters and these variables (Boelen et al., 2008; Elklit & Shevlin, 2007; Grant, Beck, Marques, Palya, & Clapp, 2008; Naifeh, Richardson, Del Ben, & Elhai, 2010; Palmieri, Weathers, et al., 2007; Pietrzak et al., 2010; Simms et al., 2002). Additionally, lower cross-sectional associations are found between trauma exposure and dysphoria than between trauma exposure and other symptom clusters (Armour & Shevlin, 2010; Simms et al., 2002). However, the findings are not without contradiction, with recent work demonstrating distress evinces similar associations with dysphoria as the numbing model’s broader hyperarousal factor (Miller et al., 2010), and individual dysphoria items are no more strongly associated with general distress than other individual items assessing symptoms of PTSD (Marshall, Schell, & Miles, 2010).

To date, no clear consensus has emerged regarding the superiority of either the dysphoria or emotional numbing model. Several recent studies test more complex models by collapsing across factors. For example, one model allows symptoms D1-D3 (sleep problems, irritability/anger, and concentration problems) to cross-load on both dysphoria and arousal (Miles, Marshall, & Schell, 2008; Shevlin, McBride, Armour, & Adamson, 2009). While Shevlin et al. (2009) found that this model had a superior fit as compared to either the dysphoria or numbing models, all three models had excellent fit. Additionally, the cross-loaded items maintained notably larger factor loadings with dysphoria, suggesting that the model is largely consistent with the simple structure dysphoria model. A second hybrid model (Elhai et al., 2010) retains the reexperiencing, effortful avoidance, and trauma-specific hyperarousal clusters of the dysphoria model and divides the dysphoria symptoms into emotional numbing (C3-C7: amnesia for the event, reduced interest in activities, detachment from others, restricted range of affect, and foreshortened future) and items termed “dysphoric arousal” (D1-D3). While this five-factor model has a superior fit among female domestic violence victims, factor intercorrelations are especially pronounced for the two factors dividing the dysphoria items (numbing and dysphoric arousal, $r = .85$), suggesting that the differences between these symptom clusters may be inconsequential.

A handful of longitudinal studies have emerged from the literature (Buschnagel et al., 2005; King et al., 2009; Krause et al., 2007; Miles et al., 2008; Suvak et al., 2008) without clear support for either the numbing or dysphoria models and with mixed support for the longitudinal stability of the latent structure of PTSD (Buschnagel et al., 2005; King et al., 2009; Suvak et al., 2008). The failure to consistently find factor stability of PTSD models may be due to symptom clusters impacting one another differentially over time (Schell, Marshall, & Jaycox, 2004) and shifts in the structure of PTSD as the disorder develops, individuals recover, or symptoms fluctuate (Krause et al., 2007). Existing longitudinal studies are limited due to reliance on undergraduate samples with indirect exposure to trauma (i.e., 9/11 terrorist attacks; Baschnagel et al., 2005; Suvak et al., 2008), samples that are actively in treatment (that is, female intimate partner violence victims participating in a clinical treatment trial; Krause et al., 2007), the use of instruments lacking emotional numbing items (King et al., 2009), and the fact that they examine the structure of PTSD following a mass stressor (e.g., 9/11 terrorist attacks) without data prior to this exposure.

In spite of numerous efforts, the literature has yet to converge on a consistently supported factor solution for the structure of PTSD. Efforts to create hybrid models combining characteristics of the numbing and dysphoria models fail to add incrementally over existing, well-supported models of PTSD (i.e., dysphoria and numbing models). The lack of consensus emerging from the confirmatory factor analysis (CFA) literature regarding the structure of PTSD is likely due to variability in measures, methodology, samples, trauma exposure type, and the duration of time since traumatic exposure across samples studied. Longitudinal efforts are also hampered by the limitations of samples studied and instruments utilized. Research on the factor structure of PTSD in OEF/OIF veterans is reliant on a small group of studies utilizing cross-sectional designs (Mansfield et al., 2010; McDonald et al., 2008; Pietrzak et al., 2010). Examining OEF/OIF samples before, during, and after a combat deployment provides a unique avenue for addressing these limitations and allows for the study of PTSD in multiple contexts, including before a mass stressor (i.e., combat deployment), within the context of an ongoing stressor (i.e., during...
combat deployment), and some time later (i.e., postdeployment). In particular, the study of PTSD symptoms during a deployment may be important because many PTSD symptoms may be normative responses to the ongoing stressors of a combat deployment, contributing to an apparent instability in the structure of PTSD. In addition, the comparison of proposed PTSD models across groups and over time within a single study can provide insight into the superiority of a given model by demonstrating the model’s stability over the development and course of the disorder (King et al., 2006).

This study has three primary goals. The first is to examine the fit of previously supported structures for PTSD across two samples of combat-deployed OIF soldiers. Toward this end, we compared six previously supported models of PTSD, including two basic models for comparison: (1) a single factor and (2) the three-factor DSM–IV model; (3) the intercorrelated factors emotional numbing model (Simms et al., 2002). Additionally, it remains unclear whether the four-factor models are best represented by intercorrelated factors or a higher-order single PTSD factor (King et al., 2006). Higher-order models assume that an unidimensional aspect of PTSD is manifested by underlying symptom clusters with greater differentiation among clusters than an intercorrelated model (King et al., 1998). Both the (5) emotional numbing (Asmundson et al., 2000; King et al., 1998) and (6) dysphoria models (Krause et al., 2007; Palmieri & Fitzgerald, 2005) were tested as a unified higher-order syndrome of PTSD with four underlying symptom clusters.

Our second goal is to examine the stability of the factor structure of PTSD before, during, and after combat deployment. Consistent with previous literature, we expected to find support for either the emotional numbing (King et al., 1998) or dysphoria (Simms et al., 2002) models of PTSD and anticipated that this structure would remain invariant at both pre- and postdeployment and between samples. Since there are no studies in the literature addressing the stability of the structure of PTSD during a combat deployment, it is unclear whether the model fit will remain consistent at this time point.

Finally, our third goal is to examine the convergent and discriminant validity of symptom clusters longitudinally to contribute to the growing literature evaluating differences between PTSD symptom clusters (e.g., Miller et al., 2010; Simms et al., 2002). To accomplish this, we examined associations among symptom clusters over time and with additional correlates (i.e., depression and combat exposure).

Method

Participants and Procedures

Two samples were recruited from a US Army National Guard Brigade Combat Team (BCT) who conducted an extended deployment to OIF from March 2006 until July 2007. All participants provided informed consent. The study was approved by the Minnesota National Guard command and institutional review boards for the Department of Defense, University of Minnesota, and the Minneapolis Veterans Affairs Medical Center.

Sample 1. Sample 1 was drawn from 522 soldiers assessed on three occasions: approximately one month prior to deployment (Time 1), two to three months following return from deployment (Time 2), and 15 months following return from deployment (Time 3). At Time 1, during mobilization training, soldiers were invited to participate in a larger psychosocial survey assessing risk and resiliency among National Guard soldiers. Consistent with military regulations, no incentives were provided at Time 1 as soldiers were on active duty status. Predeployment surveys were administered in a group classroom setting using standardized conditions. Time 2 (N = 424; 81.2% of Time 1) and Time 3 surveys (N = 343; 80.9% of Time 2; 65.7% of Time 1) were administered via mail using repeated follow-up mailing procedures to maximize response rates and monetary incentives. Participants (n = 6 at Time 1; 1 at Time 2; and 3 at Time 3) who did not complete the PTSD symptom measure were omitted from final analyses. The final N for each wave was 516 at Time 1, 423 at Time 2, and 340 at Time 3.

Response bias. At Time 2, responders did not differ from nonresponders on gender, rank, ethnic minority status, predeployment PTSD symptoms, or predeployment protective/vulnerability factors. Nonresponders were less educated (M = 13.5 vs. M = 14.4 years of education), F(1, 513) = 16.26, p < .001; younger (M = 25.6 vs. M = 29.9 years old), F(1, 513) = 19.88, p < .001; more likely to be unmarried (69.4% vs. 51.2%; χ²(1) = 10.65, p = .001; and more likely to be enlisted (95.9% vs. 88.9%), χ²(1) = 4.43, p = .035, than responders. At Time 3, there were no significant differences between responders and nonresponders on gender, ethnicity, predeployment PTSD symptoms, or predeployment protective/vulnerability factors. Nonresponders were younger (M = 26.68 vs. 30.31 years old), F(1, 520) = 21.73, p < .001; less educated (M = 13.59 vs. 14.50 years of education), F(1, 513) = 25.13, p < .001; more likely to be unmarried (68.16% vs. 47.52%), χ²(1) = 20.20, p < .001; and more likely to be enlisted (95.53% vs. 87.46%), χ²(1) = 8.68, p = .003, than responders.

Sample 2. Sample 2 was assessed while deployed to Iraq (i.e., in the theater of operations) and again approximately 12 months following their return. Approximately 67% (N = 2,677) of the entire BCT completed surveys one month prior to redeployment home (Time 1) from the 16-month extended OIF deployment described above. Surveys were administered by the onsite investigator during large Group Warrior Transition briefings. Each soldier received a packet including a cover letter containing elements of informed consent, the survey, and an envelope. Participants were instructed to seal surveys (completed or uncompleted) in the envelope to ensure confidentiality and return them to the onsite investigator. No incentives were provided to soldiers for completing in-theater surveys due to their active-duty status. Participants were included in Sample 2 only if they did not participate in the Sample 1 cohort study (N = 2,449). Thirteen participants did not complete the PTSD Checklist (PCL: Weathers, Litz, Herman, Huska, & Keane, 1993), resulting in a total in-theater N of 2,436.

One year following return from deployment (Time 2), those who provided contact information and consent for future studies were surveyed via mail (N = 1,935 soldiers, 72.3% of the original sample), using the repeated mailings procedures. There were 953 participants who returned follow-up surveys (49.3%). Ten respondents did not complete the PCL and were not included in analyses (Time 2 final N = 943).

Response bias. Time 2 responders did not differ from nonresponders on gender, ethnicity, or PTSD symptoms. Nonresponders were less likely to be married, (37.0% vs. 49.0%),
\[ \chi^2(1) = 34.26, \ p < .001; \] more likely to be enlisted, (91.6% vs. 86.5%), \[ \chi^2(1) = 15.49, \ p < .001; \] and younger \( (M = 28.85 \text{ vs. } M = 31.47 \text{ years old}) \), \[ F(1, 2306) = 61.33, \ p < .001, \] than responders.

Demographic information is provided in Table 1. Nearly all participants reported experiences consistent with combat exposure. For example, 93.0% of Sample 1 and 98.3% of Sample 2 were exposed to hostile, incoming fire, and 90.4% of Sample 1 and 87.9% of Sample 2 engaged in combat patrols or missions. Additionally, a limited number of potentially traumatic events were assessed prior to deployment for Sample 1 using the Prior Stressors scale of the Deployment Risk and Resilience Inventory (DRRI; King, King, & Vogt, 2003; Vogt, Proctor, King, King, & Vasterling, 2008). We examined endorsement rates of items representing potential A1 events (items 1, 3, 8, 9, 14, and 15). High rates of exposure were endorsed (83.1% endorsed at least one event), including being physically injured by someone (60.8%), witnessing someone being assaulted or violently killed (31.3%), and experiencing a natural disaster (34.0%). Rates of potential trauma exposure were high and approached rates of lifetime trauma exposure reported in epidemiological studies (e.g., Breslau, Wilcox, Storr, Lucia, & Anthony, 2004).

### Measures

**PTSD symptoms.** The PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993) is a 17-item self-report instrument that was used to assess severity of the DSM-IV symptoms of PTSD. Item responses ranged from 1 (never) to 5 (extremely). The measure has high internal consistency and test–retest reliability as well as robust correlations with gold-standard semistructured diagnostic interviews and additional self-report measures of PTSD (Blanchard, Jones-Alexander, Buckley, & Forneris, 1996; Ruggiero, Del Ben, Scotti, & Rabalais, 2003). The Civilian Version (PCL-C) was administered prior to deployment (symptoms rated in response to “stressful experiences”), and the Military Version (PCL-M; Weathers, Huska, & Keane, 1991) was administered during and following deployment (symptoms rated in response to “stressful military experiences”).

**Depression.** Depression was assessed at each time point in each sample using the Beck Depression Inventory (BDI-II, Beck, Steer, & Brown, 1996), which assesses symptoms of depression across 21 items on a 4-point Likert scale. The psychometric properties and construct validity of the BDI-II are well established (Beck et al., 1996; Quilty, Zhang, & Bagby, 2010; Ward, 2006; Whisman, Perez, & Ramel, 2000).

**Combat exposure.** We utilized two scales (Combat Experiences Scale, CES; Aftermath of Battle scale) from the DRRI (King et al., 2003; Vogt et al., 2008). The CES is a 15-item scale assessing combat-related experiences. Items were assessed on a 5-point Likert scale, ranging from 1 (never) to 5 (daily or almost daily). The Aftermath of Battle scale is a 15-item scale assessing exposure to the consequences of combat, such as handling human remains, witnessing destroyed homes or villages, and exposure to severe wounds and disfigurement. Items were assessed dichotomously. Consistent with expectations, both scales are associated with neurocognitive deficits, PTSD, depression, physical symptoms, general poorer mental health functioning, and male gender (King et al., 2003; Vogt et al., 2008). The scales demonstrate adequate internal consistency (CES \( \alpha = .90 \); Aftermath: \( .86 \); Present sample: CES \( \alpha = .86-.84 \); Aftermath = .87–.86). The measures were assessed two to three months postdeployment (Time 2) for Sample 1 and 12 months postdeployment (Time 2) for Sample 2.

### Analyses and Results

#### Structure of PTSD

**Analyses.** The fit for each of the six models was evaluated for each sample and time point through a series of goodness-of-fit indices: comparative fit index (CFI; Bentler, 1990), Tucker-Lewis or nonnormed fit index (TLI; Bentler & Bonett, 1980), root mean square error of approximation (RMSEA; Steiger, 1990), and weighted root mean square residual (WMRR; Muthen & Muthen, 2001). Guidelines included: CFI and TLI > .90, RMSEA \( \leq .08 \), for acceptable fit; CFI and TLI > .95, RMSEA < .06, for desirable fit (Hu & Bentler, 1999). WRMR below or near 1.0 was considered satisfactory but has been relatively untested (Yu, 2002). As models were non-nested, decisions regarding model fit were determined by visual inspection (Muthen & Muthen, 2010).

We evaluated the degree to which the structure of PTSD was consistent across samples (i.e., measurement invariance) through multigroup confirmatory factor analyses (CFAs), comparing the third wave of Sample 1 (12 months following deployment return) and the second wave of Sample 2 (15 months following return), due to similarities in time since deployment. Invariance can be formally evaluated when the same model is supported across samples by comparing the fit of progressively restricted models. Similar to the series of steps performed by Krause et al. (2007), we first allowed all parameters across groups to vary freely. Then we estimated and compared three progressively restricted models, using chi-square difference tests. The first model evaluated configurural invariance: the degree to which the pattern of factor loadings remained consistent, allowing the size of loadings, factor intercorrelations, and residual variances to vary across groups.
Horn & McArdle, 1992). The second tested metric invariance: the degree to which the pattern of factor loadings and size of factor loadings were equivalent, allowing factor intercorrelations and residual variance to vary across groups. The third examined phi invariance: the degree to which the pattern of factor loadings, size of factor loadings, and intercorrelations among factors were equivalent across groups, allowing residuals to vary. Scalar or “strong” invariance (i.e., invariance of item thresholds) and invariance of the factor means were not evaluated, as we expected levels of PTSD symptoms may differ across samples (e.g., possibly due to differing levels of trauma exposure between samples 1 and 2 or subtle differences between time since deployment between the two groups). Full factorial invariance (i.e., further constraining the residual errors to be equal across groups) was also not tested, as support for invariant residual errors is seldom found in applied datasets (Horn & McArdle, 1992).

Lastly, when initial CFAs supported the same model overall time, longitudinal invariance was tested for each sample separately through repeated measures multi-group CFAs (McArdle, 2007). The same procedures and series of restrictions described above were applied. That is, progressively restrictive models were estimated across time points for Sample 1 and then Sample 2 to evaluate configural, metric, and phi invariance over time for each sample.

Non-normality and missing data. Initial review of the data revealed non-normality at the item-level data (i.e., skewness). As a result, the observed data were modeled as ordered categorical (i.e., ordinal). Factor analyses were computed on polychoric correlations (see Olsson, 1979), and robust weighted least squares estimation was employed (RWLS; Muthen & Muthen, 2010; see also Flora & Curran, 2004), specifically mean- and variance-adjusted weighted least squares estimations (WLSMV in Mplus; Muthen & Muthen, 2010). The only known simulation study on fit indices for CFAs with ordinal data, polychoric correlations, and RWLS estimation suggests the CFI best characterizes model fit (Yu, 2002). All models were evaluated using Mplus (version 6; Muthen & Muthen, 2010). Item-level PCL data were examined for their pattern of missingness. No item at any time point for either sample met or exceeded a 1% proportion of missingness. Further, missingness appeared to be evenly distributed across items for each sample at each time point. Thus, default pairwise procedures in Mplus v. 6 for WLSMV estimation were employed for dealing with missing data. This approach has been shown to produce consistent estimates when data are either MCAR (missing completely at random) or MAR (missing at random; Asparouhov & Muthen, 2010). Pairwise deletion was also used for correlational validity analyses.

Results. See Table 2 for PCL scores and descriptive statistics over and for each sample. For Sample 1, Time 1 (predeployment), rates of symptom endorsement (i.e., item response of either moderately, quite a bit, or extremely) ranged from 6.4% (B3: flashbacks/reliving the event) to 27.1% (C5: detached from others). For during and postdeployment samples, endorsement rates ranged from 8.2% (Sample 2, Time 1, C3: amnesia for the event) to 46.7% (Sample 1, Time 2, D4: hypervigilance). Table 3 presents fit statistics for tests on each sample at each time point for the six models tested. The dysphoria models achieved greatest support for fit across both samples and all time points (fit marginally superior to higher-order dysphoria model). The dysphoria model with correlated factors achieved adequate fit on the CFI and TLI in all samples at all time points, adequate fit on the RMSEA for Sample 1 (Time 1 and 2) and Sample 2 (Time 2), and adequate fit on the WRMR for Sample 1 (Time 1). Fit was comparatively lower, but still adequate, for the in-theater sample (Sample 2, Time 1). Standardized loadings ranged from .67 to .93, with particularly high and consistent loadings (i.e., above .75) for intrusions, trauma-specific avoidance, trauma-specific hyperarousal, and three dysphoria symptoms (detachment, restricted range of affect, and difficulty concentrating). Loadings remained high but more variable (.67 or greater) for the remaining dysphoria symptoms (reduced interest in activities, foreshortened future, sleep problems, and irritability/anger). Factor correlations ranged from .63 to .90 for each sample. See Table 4 for dysphoria model factor loadings and Table 5 for factor correlations, covariances, and variances. Strong cross-sectional factor intercorrelations were found, especially between intrusions and avoidance.

Cross-sample measurement invariance. Configural invariance was examined between Time 3 of Sample 1 (15 months postreturn) and Time 2 of Sample 2 (12 months postreturn) using a multigroup CFA of the dysphoria correlated model as described above. This model produced evidence of adequate fit on all fit indices, except the WRMR, $\chi^2(226) = 1117.00, p < .001$ (CFI = .98; TLI = .98; RMSEA = .08; WRMR = 1.77). A CFA was then conducted, constraining the factor loadings to be equal across samples (i.e., metric invariance) and, once again, showed evidence of satisfactory absolute model fit, except on the WRMR, $\chi^2(286) = 875.89, p < .001$ (CFI = .99; TLI = .98; RMSEA = .06; WRMR = 1.89). Additionally, a chi-square difference test, adapted for use in RWLS estimation (Asparouhov & Muthen, 2006), yielded a significant result, $\chi^2(13) = 33.61, p = .001$, indicating that the constraints reduced model fit when compared to the test of configural invariance. A final CFA was conducted with both constrained factor loadings and factor intercorrelations (i.e., phi invariance). The results again supported absolute fit with the exception of the WRMR, $\chi^2(245) = 856.66, p < .001$ (CFI = .99; TLI = .99; RMSEA = .04; WRMR = 1.93). The chi-square difference test, examining the difference in fit between metric and phi invariance models, was not significant, $\chi^2(6) = 6.38, p = .38$. Thus, model fit was not changed significantly when constraining

---

1 The cross-loaded dysphoria-numbing model (Shevlin et al., 2009) and the five-factor dysphoria-numbing model (Elhai et al., 2010) were also tested. Fit indices indicated similar or minor improvements in model fit for both these models over the other models tested. However, for the cross-loaded model, consistent with prior research (Shevlin et al., 2009), the only items distinguishing between the dysphoria and numbing models (D1-D3) largely loaded with dysphoria (i.e., .43–.73), with consistently lower loadings with hyperarousal (i.e., .10–.39), suggesting that the overall structure was largely consistent with the dysphoria model. Additionally, for the five-factor model, consistent with Elhai and colleagues’ (2010) findings, the correlation between the dysphoric arousal (D1-D3) and the numbing cluster (C3-C7) was quite high in each sample (i.e., .91, .91, .91, .85, and .91). Given previous theory and research supporting combining these symptoms into a dysphoria factor, it is unlikely these factors represent truly distinct constructs. Given the above findings, marginal differences in fit, and the reduced parsimony and interpretability of the two more complicated new models, we concluded these newer factor structures do not provide incremental information, interpretability, or clarity to prior models and proceeded with invariance tests and validity analyses with the simple structure dysphoria model.
the factor intercorrelations to be equal. To determine which factor loadings were responsible for the significant decline in fit for the test of metric invariance, post hoc analyses were conducted. We released constraints indicated by modification indices, while allowing the model to be nested within the less restrictive model. Results indicated that releasing the equality constraint on the loading for criterion C7 (sense of foreshortened future) diminished the chi-square difference test to nonsignificance.

**Stability of PTSD Structure Over Time**

**Results.** The consistency of support for the superior fit of the same model across all samples and time points provided initial support for the stability of PTSD’s structure. We then formally tested the degree to which the same construct was being assessed over time, through changes in underlying factor structure, using repeated-measures multigroup CFAs (McArdle, 2007) and the series of constraints described above. For each model in Sample 1, tests produced evidence of adequate fit on all of the fit indices but the WRMR: configural invariance, $\chi^2(339) = 1026.00, p < .001$ (CFI = .98; TLI = .97; RMSEA = .07; WRMR = 1.76); metric invariance, $\chi^2(365) = 896.59, p < .001$ (CFI = .98; TLI = .98; RMSEA = .06; WRMR = 1.95); and phi invariance, $\chi^2(377) = 755.13, p < .001$ (CFI = .99; TLI = .99; RMSEA = .05; WRMR = 2.34). In each case, more restrictive models resulted in

<table>
<thead>
<tr>
<th>Sample</th>
<th>Time Wave</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>α</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Predeployment</td>
<td>26.20</td>
<td>10.01</td>
<td>17–81</td>
<td>1.98</td>
<td>4.74</td>
<td>.92</td>
<td>.39*</td>
<td>.31*</td>
</tr>
<tr>
<td>2</td>
<td>2–3 Months Post</td>
<td>35.65</td>
<td>13.96</td>
<td>17–85</td>
<td>.75</td>
<td>-.07</td>
<td>.94</td>
<td>.66*</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>15 Months Post</td>
<td>35.76</td>
<td>14.89</td>
<td>17–80</td>
<td>.77</td>
<td>-.21</td>
<td>.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>In-Theater</td>
<td>31.53</td>
<td>12.40</td>
<td>17–85</td>
<td>1.14</td>
<td>1.11</td>
<td>.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>12 Months Post</td>
<td>34.21</td>
<td>14.66</td>
<td>17–85</td>
<td>1.03</td>
<td>.44</td>
<td>.95</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Post = postdeployment. *p < .001.

Table 2

**Descriptive Statistics for Total PTSD Symptom Severity Scores**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Time Wave</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>α</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Predeployment</td>
<td>26.20</td>
<td>10.01</td>
<td>17–81</td>
<td>1.98</td>
<td>4.74</td>
<td>.92</td>
<td>.39*</td>
<td>.31*</td>
</tr>
<tr>
<td>2</td>
<td>2–3 Months Post</td>
<td>35.65</td>
<td>13.96</td>
<td>17–85</td>
<td>.75</td>
<td>-.07</td>
<td>.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>15 Months Post</td>
<td>35.76</td>
<td>14.89</td>
<td>17–80</td>
<td>.77</td>
<td>-.21</td>
<td>.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>In-Theater</td>
<td>31.53</td>
<td>12.40</td>
<td>17–85</td>
<td>1.14</td>
<td>1.11</td>
<td>.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>12 Months Post</td>
<td>34.21</td>
<td>14.66</td>
<td>17–85</td>
<td>1.03</td>
<td>.44</td>
<td>.95</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Post = postdeployment. *p < .001.

Table 3

**Fit Indices for Item-Level CFAs of the PCL Across Samples and Time Points**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Time Wave</th>
<th>Model</th>
<th>$\chi^2(df)^*$</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
<th>WRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Predeployment</td>
<td>1 factor</td>
<td>703.58 (119)</td>
<td>0.92</td>
<td>0.91</td>
<td>0.10</td>
<td>1.6</td>
</tr>
<tr>
<td>2</td>
<td>2–3 Months Post</td>
<td>3 factor</td>
<td>496.52 (116)</td>
<td>0.95</td>
<td>0.94</td>
<td>0.08</td>
<td>1.29</td>
</tr>
<tr>
<td>3</td>
<td>15 Months Post</td>
<td>Numbing</td>
<td>306.83 (113)</td>
<td>0.97</td>
<td>0.97</td>
<td>0.06</td>
<td>0.97</td>
</tr>
<tr>
<td>2</td>
<td>In-Theater</td>
<td>Numbing (higher-order)</td>
<td>441.64 (115)</td>
<td>0.96</td>
<td>0.95</td>
<td>0.07</td>
<td>1.21</td>
</tr>
<tr>
<td>2</td>
<td>12 Months Post</td>
<td>Dysphoria</td>
<td>250.47 (113)</td>
<td>0.98</td>
<td>0.98</td>
<td>0.05</td>
<td>0.87</td>
</tr>
<tr>
<td>2</td>
<td>12 Months Post</td>
<td>Dysphoria (higher-order)</td>
<td>307.51 (115)</td>
<td>0.97</td>
<td>0.97</td>
<td>0.06</td>
<td>0.99</td>
</tr>
</tbody>
</table>

**Note.** CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean squared error of approximation; WRMR = weighted root mean square residual.

*All $\chi^2$ tests of model fit were significant at $p < .05.$
a significant reduction of model fit: configural versus metric, $\chi^2(26) = 56.77, p < .001$; metric versus phi, $\chi^2(12) = 38.63, p < .001$. Of note, while adding constraints to test metric and phi invariance reduced fit overall, reflecting the increasing strictness of models, three of four fit indices still reflected good fit in an absolute sense. Post hoc modifications indicated that, for the model testing metric invariance, releasing the equality constraints on the loadings for criteria C4 (diminished interest or pleasure in activities), C5 (detachment from others), C6 (restricted range of affect), D2 (anger and irritability), and D4 (hypervigilance) diminished the difference test to nonsignificance. For the model testing phi invariance, though, all equality constraints on the phi model required freeing (i.e., constraining any of the phi coefficients to equality across time points reduced fit).

For Sample 2, the test for configural invariance was supportive of fit on two of four indices, $\chi^2(226) = 3484.09, p < .001$ (CFI = .96; TLI = .95; RMSEA = .09; WRMR = 3.15). Both of the tests for metric and phi invariance achieved support for absolute fit on all indices except the WRMR; metric, $\chi^2(286) = 3474.90, p < .001$ (CFI = .96; TLI = .96; RMSEA = .08;
Convergent and Discriminant Validity of Symptom Clusters

Analyses. First, we examined convergent and discriminant validity using a multitrait-multiperiod matrix (Longley, Watson, & Noyes, 2005). We expected test-retest correlations (convergent validity) to be higher than correlations between different symptom clusters over time (discriminant validity), as retest correlations provide a standard for appraising the discriminant associations. However, we did not expect this pattern of associations with waves comparing pre- to postdeployment symptoms, particularly for trauma-specific symptoms (i.e., intrusions, avoidance, and hyperarousal), as the potential introduction of trauma exposure between assessment waves would likely make retest associations far less meaningful.

Second, we examined associations between symptom clusters and external correlates: depression and combat exposure. Consistent with previous cross-sectional research (e.g., Simms et al., 2002; Palmieri, Weathers, et al., 2007), we expected (1) stronger associations between depression and dysphoria symptoms over time than between depression and other symptom clusters and (2) stronger associations between combat exposure and intrusions over time (among waves following deployment onset) than with other symptom clusters. The sizes of the differences between correlations were compared using Steiger’s (1980) formula.

Results. The pattern of associations among symptom clusters over time was largely consistent with expectations (see Table 6). For pre- to postdeployment comparisons, with the exception of dysphoria (nonspecific to PTSD), test-retest correlations were similar to correlations between disparate symptom clusters over time. However, for assessments conducted after potential combat exposure, stability coefficients were higher than associations among differing symptom clusters over time. This pattern was less pronounced between assessments conducted in-theater (potentially prior to stabilization of trauma-specific symptoms and associated clinical impairment) and 12 months after return from deployment.

Correlations with external variables (depression and combat exposure) are displayed in Table 7. Again, the findings were largely consistent with expectations. For each assessment and wave, associations between dysphoria and depression over time were larger than associations between depression and other symptom clusters. In nearly every case, the difference in the size of these associations was significant. For combat exposure, the largest associations were found between severity of exposure and intrusive memories. Differences between these associations and those with other symptom clusters were typically significant when assessments were both conducted after combat exposure (vs. pre- to postdeployment). Differences with combat exposure were also more consistent and pronounced within 12 months of combat exposure (vs. 15 months postdeployment), suggesting that differential links between combat exposure and trauma-specific symptom clusters may be most pronounced within the first year following combat.

Discussion

We examined the superiority and stability of competing structural models of PTSD across two longitudinal samples of OIF

---

Table 6
Multitrait–Multioccasion Matrix of Symptom Clusters

<table>
<thead>
<tr>
<th></th>
<th>Sample 1 (n)</th>
<th>Sample 2 (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T2</td>
<td>T3</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>T1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrusions</td>
<td>.28 (419)</td>
<td>.26 (417)</td>
</tr>
<tr>
<td>Avoidance</td>
<td>.26 (419)</td>
<td>.29 (417)</td>
</tr>
<tr>
<td>Arousal</td>
<td>.20* (419)</td>
<td>.29 (417)</td>
</tr>
<tr>
<td>Dysphoria</td>
<td>.23* (419)</td>
<td>.28* (417)</td>
</tr>
<tr>
<td>T3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrusions</td>
<td>.62 (321)</td>
<td>.44* (320)</td>
</tr>
<tr>
<td>Avoidance</td>
<td>.48 (320)</td>
<td>.53 (319)</td>
</tr>
<tr>
<td>Arousal</td>
<td>.52* (321)</td>
<td>.33* (320)</td>
</tr>
<tr>
<td>Dysphoria</td>
<td>.46 (321)</td>
<td>.44* (320)</td>
</tr>
</tbody>
</table>

Note. Test–retest correlations are bolded along the diagonal. T1 = Time 1: Sample 1, predeployment; Sample 2, in-theater. T2 = Time 2: Sample 1, two months postdeployment. Sample 2, 12 months postdeployment. T3 = Time 3: Sample 1, 15 months postdeployment. * Value significantly different from test–retest correlation (Steiger, 1980), p < .05.
STRUCTURE OF PTSD AMONG RETURNING SOLDIERS

veterans that were followed over the course of a combat deployment. We found uniform support for the dysphoria model (Simms et al., 2002) and support for the stability of the this structure prior to, during, and following return from OIF deployment, suggesting that item content was interpreted similarly across the deployment cycle and the underlying construct remained stable. Adequate model fit was less consistent for all other models tested, including the numbing model (King et al., 1998). The findings are consistent with previous research supporting the dysphoria model across both nondeployed and deployed soldiers (Simms et al., 2002) and over time among treatment-seeking victims of intimate partner violence (Krause et al., 2007).

Support was somewhat less robust when constraining the size of item loadings and factor intercorrelations to be equal. Across samples, post hoc models suggested that C7 (sense of foreshortening future) was primarily responsible for significant declines in fit, suggesting that this item may be less consistently linked to dysphoria across samples, even when time since combat deployment is roughly equivalent (approximately 12–15 months postdeployment). For longitudinal models, a greater breadth of symptoms contributed to reductions in fit when adding model constraints (e.g., dysphoria and hypervigilance items for both samples; reexperiencing items for Sample 2). In fact, in the most restrictive models restraining any specific symptom significantly reduced model fit. Reduced fit with these constraints may reflect structural changes due to increases in the absolute value or variability of PTSD symptom severity scores, changes in measurement error with time, and/or changes in relations between individual symptoms following combat exposure. However, for each sample and time point, standardized parameter estimates for the dysphoria model do not appear to be dramatically different in valence or magnitude. Thus, these tests of increasingly strict invariance are likely reflecting very subtle differences in structure.

The intercorrelated dysphoria model demonstrated a modestly superior fit over the higher-order dysphoria model. This suggests that PTSD is not best represented by a larger latent construct with independent manifest indicators. Among our samples, PTSD is best represented by separate and distinct but moderately to highly correlated symptom clusters that, when grouped together, form PTSD. The findings support the distinction between each of these symptom clusters, may support the existence of PTSD subtypes created through differing combinations of symptoms (King et al., 1998), and replicate prior work supporting intercorrelated over higher-order models (Boelen et al., 2008; Elklit & Shevlin, 2007; Krause et al., 2007; Palmieri, Weathers, et al., 2007).

The dysphoria model evinced poorer comparative fit for the in-theater sample than for other assessment waves. This may reflect the unique context of a combat-deployment environment. During deployment, soldiers may have experiences that are ostensibly consistent with descriptions of PTSD symptoms but fail to reflect the underlying construct of PTSD. For example, sleeping difficulties and irritability may result from harsh environmental conditions (e.g., uncomfortable sleeping quarters) and realistic, even adaptive, awareness of potential danger and the need for preparedness. Soldiers completed assessments in Iraq near the end of their deployment. Consequently, it is possible that they were experiencing positive feelings related to their impending return and this is reflected in the reduced fit of the dysphoria model. However, average PCL scores reflect elevated distress levels compared to predeployment, suggesting that this alternative interpretation is unlikely.

Of note, improvements in model fit for the dysphoria over the numbing model were small. The field of items assessing emotional numbing (5 items) and dysphoria (8 items) are limited and overlapping (5 items are shared; all numbing items are also dysphoria items). While our findings provide clear support for a dysphoria

### Table 7

**Discriminant Validity: Correlations With Symptom Cluster Scales for Sample 1**

<table>
<thead>
<tr>
<th>Time</th>
<th>Cluster</th>
<th>Sample 1 (n)</th>
<th>Sample 2 (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1 BDI</td>
<td>T2 BDI</td>
<td>T3 BDI</td>
</tr>
<tr>
<td>T1</td>
<td>.26 (.414)</td>
<td>.23 (.418)</td>
<td>.14 (.417)</td>
</tr>
<tr>
<td></td>
<td>.23 (.336)</td>
<td>.18 (.418)</td>
<td>.11 (.417)</td>
</tr>
<tr>
<td></td>
<td>.24 (.414)</td>
<td>.22 (.418)</td>
<td>.11 (.417)</td>
</tr>
<tr>
<td></td>
<td>.41 (.514)</td>
<td>.34 (.418)</td>
<td>.12 (.417)</td>
</tr>
<tr>
<td>T2</td>
<td>.21 (.418)</td>
<td>.42 (.418)</td>
<td>.42 (.417)</td>
</tr>
<tr>
<td></td>
<td>.57 (.414)</td>
<td>.40 (.320)</td>
<td>.36 (.417)</td>
</tr>
<tr>
<td></td>
<td>.49 (.414)</td>
<td>.38 (.418)</td>
<td>.32 (.417)</td>
</tr>
<tr>
<td></td>
<td>.34 (.418)</td>
<td>.59 (.320)</td>
<td>.29 (.418)</td>
</tr>
<tr>
<td>T3</td>
<td>.14 (.335)</td>
<td>.34 (.320)</td>
<td>.40 (.320)</td>
</tr>
</tbody>
</table>

*Note.* The largest correlation between a symptom cluster and the variable of interest is bolded for each wave. T1 = Time 1: Sample 1, predeployment; Sample 2, in-theater. T2 = Time 2: Sample 1, two months postdeployment. Sample 2, 12 months postdeployment. T3 = Time 3: Sample 1, 15 months postdeployment; BDI = Beck Depression Inventory; CES = Combat Events Scale; Aftermath = Aftermath of Battle.

*Value significantly different from correlation with intrusions (Steiger, 1980), *p* < .05. **Value significantly different from correlation with intrusions (Steiger, 1980), *p* < .05.
factor, existing measures of PTSD symptoms are insufficient to rule out the simultaneous presence of an emotional numbing component that is not adequately assessed by existing measures of PTSD due to overlap with dysphoria items. Developing instruments with items that provide better assessments and differentiate emotional numbing from dysphoria may aid in determining if both components are unique and necessary to understanding the construct of PTSD. Future studies will need to develop and make use of a broader pool of nonoverlapping items.

**Validity of Symptom Clusters**

Strong cross-sectional factor intercorrelations were found, especially between intrusions and avoidance. These strong associations are consistent with prior research (e.g., Krause et al., 2007; Palmieri, Marshall, et al., 2007; Pietrzak et al., 2010). However, prior theory, extensive prior factor analysis, and our own examination of discriminant validity across symptom clusters, suggest important differences between symptom clusters. Strong correlations may be due to a causal link (e.g., intrusive memories causing symptoms of dysphoria) or feedback loop among trauma-specific symptom clusters (e.g., intrusions lead to avoidance and avoidance to further intrusions).

Our findings examining differential associations with symptom clusters, both cross-sectionally and longitudinally, were largely consistent with proposed theories related to the pathogenesis of PTSD. While the overall structure of PTSD remains consistent over time (i.e., symptoms vary together in clusters), correlations with symptom cluster scores vary with deployment. Specifically, following deployment onset, each trauma-specific symptom cluster (intrusions, avoidance, and hyperarousal) becomes more strongly and consistently predictive of itself (i.e., higher retest correlations) than of other symptom clusters. This suggests that when a sample is exposed to a mass event with probable trauma exposure, trauma-specific symptom clusters increase in their stability, specificity, and distinctiveness, providing evidence of discriminant validity. Regardless of deployment, the dysphoria cluster better predicted subsequent or prior assessments of dysphoria than it predicted other trauma-specific symptom clusters. The stability of dysphoria, even from before to after combat deployment, supports the nontrauma-specific nature of this symptom cluster.

Relations between symptom clusters and external variables provided additional evidence of the dysphoria model’s validity. Consistent with prior cross-sectional findings (e.g., Simms et al., 2002), depression evinced stronger associations with dysphoria symptoms than other symptom clusters, both cross-sectionally and over time, even prior to deployment. This finding reinforces that the dysphoria cluster is highly saturated with broad general distress (Elklin, Armour, & Shevlin, 2009; Grant et al., 2008; Simms et al., 2002; Palmieri, Weathers, et al., 2007). Combat and aftermath of battle exposure were more strongly associated with intrusions than other symptom clusters, including avoidance, supporting symptom clusters’ convergent validities and highlighting important differences between clusters. While less pronounced, intrusive memories and hyperarousal prior to deployment were also more predictive of subsequent combat exposure than other symptom clusters. This may indicate that those with more severe PTSD-specific symptoms prior to deployment are at greater risk for more severe exposure to combat and its aftermath or for having more salient memories of the severity of these exposures.

Watson (2009) proposed a quadripartite model, which classifies individual symptoms of anxiety and depressive disorders along two dimensions: distress level and level of specificity to either anxiety or depression. Based on a review of the literature, he concluded that the dysphoria cluster of PTSD is relatively more saturated with distress/negative affect than other symptom clusters. The remaining symptom clusters (intrusions, avoidance, and hyperarousal) correlate comparatively with both depression and anxiety, suggesting they, and likely PTSD as a construct, lie between depression and anxiety rather than specifically within either category of disorders. Watson’s model allows for distress and specificity to vary independently and acknowledges the centrality of nonspecific distress symptoms (e.g., dysphoria) to the core of psychopathology. Our findings support this framework as well as prior hierarchical models of PTSD (Brown et al., 1998; Clark & Watson, 1991; Mineka et al., 1998).

**Limitations**

While the homogeneity of each of our samples allows for reducing additional measurement noise by holding the participant pool constant, it also limits the generalizability of our findings. Specifically, our samples were comprised of mostly male, White, National Guard soldiers from the Midwest. The findings may not generalize to more ethnically diverse samples, women, or other branches of the armed forces. Additionally, despite explicit instructions that surveys were voluntary and confidential and the use of sealed envelopes, soldiers may have responded differently due to concerns that responses would not be held private, would negatively impact their ability to serve, and/or would negatively impact their larger military career. Third, while not a limitation per se, it is important to highlight that the variability, extent, and timing of any trauma exposure among participants before and during deployment means that comparisons between time periods cannot be interpreted as differences before and after exposure to any traumatic event. Additionally, as is the case with any longitudinal study, differences between respondents and nonrespondents may bias our findings and limit their application to underrepresented groups. As identified above, while the PCL is a widely used measure of PTSD symptoms, additional research is needed using structured diagnostic interviews for PTSD, a broader pool of items separately assessing emotional numbing and dysphoria symptoms, and additional items representing those factors with few items within the dysphoria model (hyperarousal, 2 items; trauma-specific avoidance, 2 items). Comparisons of self-report and interview methods of assessing PTSD over time are required to establish the impact of method variance on these findings. While active debate over the utility and necessity of Criterion A2 (presence of fear, helplessness, or horror at the time of the event) is ongoing, neither criterion A2 nor primary trauma linked to symptoms of PTSD were assessed and diagnostic information was not available.

Different administration methods were employed (predeployment: classroom; in-theater: mail; after deployment: mail). This may have influenced the results; however, analyses do not indicate any differences which may be attributable to the administration method. Finally, prior to deployment, respondents were instructed...
to report symptoms in response to stressful experiences, while after the deployment onset, respondents were instructed to report symptoms in response to stressful military experiences. While both instructions are fairly nonspecific and may be interpreted to refer to a host of events, differences in instructions may have influenced findings. Of note, prior research has demonstrated that the factor structure of PTSD is largely invariant among those with prior trauma histories, regardless of whether participants were asked to respond to a specific trauma or stressful experiences generally (Elhai et al., 2009), and there are no indications among our findings that the latent structure of PTSD or associations with symptom clusters were influenced by this change in instructions.

References
Marshall, G. N., Schell, T. L., & Miles, J. N. V. (2010). All PTSD symptoms are highly associated with general distress: Ramifications for


Received August 22, 2010
Revision received April 1, 2011
Accepted April 6, 2011