School of Professional Psychology Faculty Publications

School of Professional Psychology

3-2014

Assessing Posttraumatic Stress in Military Service Members: Improving Efficiency and Accuracy

Caitlin L. Fissette

Douglas K. Snyder

Christina Balderrama-Durbin

Steve Balsis

Jeffrey A. Cigrang
Wright State University, jeffrey.cigrang@wright.edu

See next page for additional authors

Repository Citation

Authors

Caitlin L. Fissette, Douglas K. Snyder, Christina Balderrama-Durbin, Steve Balsis, Jeffrey A. Cigrang, G. Wayne Talcott, JoLyn Tatum, Monty Baker, Daniel Cassidy, Scott Sonnek, Richard E. Heyman, and Amy M. Smith Slep

Assessing Posttraumatic Stress in Military Service Members: Improving Efficiency and Accuracy

Caitlin L. Fissette, Douglas K. Snyder,
Christina Balderrama-Durbin, and Steve Balsis
Texas A&M University

Jeffrey Cigrang
Wright-Patterson Air Force Base, Dayton, Ohio

G. Wayne Talcott
University of Tennessee Health Science Center

JoLyn Tatum
Wright-Patterson Air Force Base, Dayton, Ohio

Monty Baker, Daniel Cassidy, and Scott Sonnek
Lackland Air Force Base, San Antonio, Texas

Richard E. Heyman and Amy M. Smith Slep
New York University

Posttraumatic stress disorder (PTSD) is assessed across many different populations and assessment contexts. However, measures of PTSD symptomatology often are not tailored to meet the needs and demands of these different populations and settings. In order to develop population- and context-specific measures of PTSD it is useful first to examine the item-level functioning of existing assessment methods. One such assessment measure is the 17-item PTSD Checklist–Military version (PCL-M; Weathers, Litz, Herman, Huska, & Keane, 1993). Although the PCL-M is widely used in both military and veteran health-care settings, it is limited by interpretations based on aggregate scores that ignore variability in item endorsement rates and relatedness to PTSD. Based on item response theory, this study conducted 2-parameter logistic analyses of the PCL-M in a sample of 196 service members returning from a yearlong, high-risk deployment to Iraq. Results confirmed substantial variability across items both in terms of their relatedness to PTSD and their likelihood of endorsement at any given level of PTSD. The test information curve for the full 17-item PCL-M peaked sharply at a value of $\theta = 0.71$, reflecting greatest information at approximately the 76th percentile level of underlying PTSD symptom levels in this sample. Implications of findings are discussed as they relate to identifying more efficient, accurate subsets of items tailored to military service members as well as other specific populations and evaluation contexts.

Keywords: posttraumatic stress disorder, PTSD Checklist, military, veterans, item response theory (IRT)

Over 2 million American men and women have served in Operations Enduring Freedom or Iraqi Freedom since September 11, 2001. This sustained mobilization has exacted an enormous toll on the service members deployed to these theaters of combat—of whom more than 800,000 have deployed multiple times (Sheppard, Malatras, & Israel, 2010). In a large study of help-seeking veterans returning from Iraq and Afghanistan, 37% received a mental health diagnosis—with the most prevalent diagnosis (22%) being posttraumatic stress disorder (PTSD; Seal et al., 2009). Many more returning service members have exhibited subthreshold symptoms of PTSD such that they endorse the exposure Criterion A and Criterion B symptom cluster and also meet diagnostic criteria for either the Criterion C or Criterion D symptom clusters but not both (Blanchard, Hickling, Taylor, Loos, & Gerardi, 1994). Reliably and efficiently assessing posttraumatic stress is critical to identifying those service members or veterans most in need of treatment.

According to the National Defense Authorization Act (2009), all service members of the Armed Forces deployed in connection with

a contingency operation are required to receive mental health assessments at four time points across the deployment cycle. At each time point, the service member is assessed for the presence of PTSD symptoms as well as other indicators of mental health functioning. PTSD assessment begins with the four-item Primary Care PTSD Screen (PC-PTSD; Prins et al., 2003), with two or more endorsed questions resulting in a follow-up assessment using the 17-item PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993). Several alternative versions of the PCL have been developed that address "a stressful military experience" for military personnel and veterans (PCL-M; Weathers et al., 1993), "a stressful experience from the past" for civilians (PCL-C; Weathers et al., 1993), or "the stressful event" for respondents identifying a specific traumatic event (PCL-S; Weathers et al., 1993); item content for the three versions is otherwise identical.

Although standardized assessment of PTSD can differ across settings, the PTSD Checklist–Military version (PCL-M; Weathers et al., 1993) continues to be one of the most widely used measures of posttraumatic stress in both military and Veterans Affairs (VA) health-care settings upon service members' return home. However, despite its widespread use and overall reliability and validity, the PCL has some notable shortcomings. Primary among these is the reliance upon an aggregate score in which all items receive the same weight (Bliese et al., 2008), neglecting important differences in both the prevalence and diagnostic efficiency among individual items. Indeed, prior research with the PCL using both item-total correlations and cluster score correlations (Blanchard, Jones-Alexander, Buckley, & Forneris, 1996; Lang & Stein, 2005) as well as examining item parameters (Bliese et al., 2008) has shown that items vary considerably in their discriminative information.

When conducting clinical assessment or rendering selection or intervention decisions based on the PCL's aggregate score, any combination of positively endorsed items is assumed to have equal predictive utility for evaluating PTSD regardless of whether these are high-prevalence symptoms relatively nonspecific to PTSD (e.g., trouble falling or staying asleep, feeling irritable) versus symptoms with lower prevalence but that are highly specific to PTSD (e.g., either trouble remembering or involuntarily reexperiencing a stressful military experience). However, it is likely that certain items would provide more information within a particular setting and that optimal item subsets would vary as a function of changes in context. Moreover, a cutoff score of 30 (see Lang, Laffaye, Satz, Dresselhaus, & Stein, 2003, and Yeager, Magruder, Knapp, Nicholas, & Frueh, 2007) could be obtained by ratings of mild or moderate impact on a large number of items endorsed at low levels of the PTSD latent construct or alternatively by ratings of more severe impact on a smaller number of items endorsed at high levels of the PTSD construct—presumably with quite different implications for intervention.

In addition, in some assessment contexts (e.g., when screening for PTSD during deployment to a theater of combat, or in comprehensive batteries assessing for a broad spectrum of mental and physical health concerns), inclusion of the full 17-item PCL may not be realistic. Instead, such settings would benefit from a subset of three or four items specifically selected for the prevalence and severity of PTSD symptoms in that population, and for various prevention or intervention purposes of the assessment targeting different PTSD thresholds. Consequently, assessment of posttraumatic stress in service members and veterans may be enhanced by

studying characteristics of individual items with respect to both their likelihood of endorsement at a given level or threshold of PTSD symptomatology and their degree of relatedness (or discriminative validity) to PTSD. Results of such an analysis could be used to adapt the PCL or tailor item selection for more efficient and accurate assessment of posttraumatic stress—targeting specific thresholds of this syndrome at varying levels depending on the setting and purpose (e.g., identifying subthreshold PTSD among active military for secondary prevention strategies prior to redeployment versus triaging impaired military personnel in-theater for crisis intervention).

Item response theory (IRT; Embretson & Reise, 2000) is ideally suited for examining item characteristics for purposes of adapting a measure to specific populations and settings. Specifically, using two-parameter logistic (2PL) analyses for each item, two defining characteristics can be identified, including a discrimination or slope parameter (a) and a difficulty parameter (b). The a parameter describes how closely an item relates to some latent construct (in this case, posttraumatic stress). The b parameter indicates the point on the latent construct at which the probability of endorsing the item is equal to 0.50. In combination, these two parameters enable one to select a subset of items having the highest discriminative potential at any given level of the targeted construct.

The current investigation examined item characteristics of the PCL in a sample of U.S. service members returning from a yearlong, high-risk deployment to Iraq. It was anticipated that, consistent with previous research, items would exhibit considerable variability with respect to both their rates of endorsement and their relatedness to PTSD. The goal of the study was to identify distinguishing item characteristics to facilitate tailoring more efficient assessment strategies for service members and veterans in future clinical and research applications.

Method

Participants

Participants were a cohort of 196 active-duty service members from a larger longitudinal investigation assessing a variety of risk and protective factors impacting U.S. Air Force (USAF) Security Forces across a yearlong deployment to Iraq. Two detachments of Airmen ($N = 318$) were tasked to train Iraqi police, a high-risk mission that required patrolling in communities with insurgent fighters; they deployed in two consecutive, 1-year deployment cycles during 2009 and 2010. Based on responses to a measure of deployment experiences described by Hoge and colleagues (2004), participants described high levels of exposure to combat-related stressful experiences during their time in Iraq. Nearly all (97%) knew someone who had been seriously injured or killed; over 70% had seen dead or seriously injured Americans, had witnessed extensive physical devastation and its impact on vulnerable citizens, or had experienced hostile reactions from civilians they were trying to help; and more than half had patrolled in areas with land mines, had aided in the removal of unexploded ordinances, or had been fired upon.

Following their deployment and return from Iraq, these Security Forces Airmen returned to their original bases scattered across the United States and other countries. At 6 to 9 months postdeployment, the Airmen were invited to return to Lackland Air Force

Base (AFB) in San Antonio, Texas, as part of this study to participate in focus-group discussions and complete follow-up measures—and 169 elected to do so. Reasons for the reduced sample at follow-up were that, at 6 or more months postdeployment, some of the Airmen had already separated from the military, a few chose not to participate, and some were not able to travel to the location of the follow-up assessment (Lackland AFB) despite provision of travel funds. To facilitate their participation, Airmen who could not attend the follow-up conference were invited to respond via a web-based survey, and an additional 35 participated via this method. Of the 204 Airmen participating in the follow-up assessment (85 at 6 months and 119 at 9 months postdeployment), eight failed to complete the PCL-M, resulting in a final study sample of 196 participants. These 196 Airmen did not differ from the larger sample of 318 Airmen assessed prior to deployment on any measure of demographic characteristics or predeployment measure of individual emotional or behavioral functioning, or intimate relationship functioning (all $ps > .50$).

Prior to completing measures at either pre- or postdeployment, the research team informed participants about the purpose of the study, the anonymity of their survey responses, and the volunteer nature of their participation. Study procedures were reviewed by, and conducted in full compliance with, the USAF Institutional Review Board. Airmen were provided with the study questionnaire inside a blank envelope and created a personal identifier known only to them, based on their mother's first and last initials, the calendar day on which they were born, and their hair color. No compensation was provided for completing study measures. However, approximately 98% of Airmen present at pre- and postdeployment assessment settings chose to participate.

Participants ranged in age from 19 to 46 years ($M = 25.4$, $SD = 5.7$); 181 (93%) were male, and 15 (7%) were female. The majority (67%) self-identified as Caucasian, 12% as Latino/a, 11% as African American, 7% as Asian, and 3% as "other." Officers constituted 4% of the sample, with the other pay grades distributed as follows: E1–E3 (junior enlisted): 24%, E4–E6 (midlevel enlisted or noncommissioned officers): 65%, and E7–E9 (senior noncommissioned officers): 7%. Most of the sample (77%) had deployed at least once previously, with 38% having had two or more prior deployments. Rates of PTSD as assessed by the PCL-M increased from 7% predeployment (6% mild to moderate, 1% severe) to 47% at postdeployment (26% mild to moderate, 21% severe) with a mean PCL-M score of 21.8 ($SD = 6.1$) prior to deployment and 35.4 ($SD = 16.1$) at postdeployment follow-up.

Measure

The PTSD Checklist (PCL; Weathers et al., 1993) was developed at the National Center for PTSD as a brief, self-report inventory for assessing the 17 symptoms of PTSD outlined in the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–IV;* American Psychiatric Association, 1994); items also correspond to the three clusters of PTSD: reexperiencing (Criterion B), avoidance/numbing (Criterion C), and hyperarousal (Criterion D). In the current study, the military version of this measure (PCL-M; Weathers et al., 1993) was used, asking respondents to consider the impact of their exposure to "stressful military experiences" and to rate each item regarding how much they had been "bothered by the problem in the past month" on a 5-point Likert scale ranging from

1 (*not at all*) to 5 (*extremely*), with scores ranging from 17 to 85. The PCL-M demonstrates excellent internal consistency ($\alpha = .96$) and test–retest reliability ($r = .96$; Weathers et al., 1993) and correlates highly with other standardized measures of PTSD (Forbes, Creamer, & Biddle, 2001). In the present study, for purposes of analysis using a two-parameter logistic model, participants' responses to each item were dichotomized in a manner consistent with that in previous literature (e.g., Weathers et al., 1993) such that omitted items or ratings of 1 or 2 (*not at all* or *a little bit*) were treated as nonendorsement, whereas ratings of 3, 4, or 5 (*moderate, quite a bit,* or *extremely*) were treated as endorsement of that item.

Results

As a preliminary analysis, confirmatory factor analysis (CFA) was used to determine whether the PCL-M demonstrated sufficient unidimensionality in this sample for further analysis using IRT. Evidence for unidimensionality is essential for meeting the two basic assumptions of IRT—that items assess a single underlying construct (e.g., posttraumatic stress) and that items are locally independent. When conducting IRT analysis, local independence can be assumed once unidimensionality has been established (Hambleton, Swaminathan, & Rogers, 1991). Evidence for unidimensionality was evaluated using Mplus Version 5.2 (Muthén & Muthén, 2007) with two goodness-of-fit indices: the Tucker–Lewis Index (TLI; Tucker & Lewis, 1973) and the comparative fit index (CFI; Bentler, 1990). Strong evidence for unidimensionality is obtained if the TLI and the CFI are both greater than .95. In the present sample, goodness-of-fit indices for the one-factor model suggested good unidimensionality, with TLI and CFI values of .98 and .97, respectively.

For comparison purposes, additional CFAs examined a three-factor model based upon *DSM–IV* criteria (American Psychiatric Association, 1994) as well as four-factor models proposed by King, Leskin, King, and Weathers (1998) and Simms, Watson, and Doebbeling (2002). Results provided support for these alternative models as well, with four-factor models obtaining modestly stronger support than alternative one- or three-factor models—TLI and CFI values of .98 and .98 for the three-factor *DSM–IV* model, and .99 and .99 for both the King et al. (1998) and Simms et al. (2002) four-factor models, respectively.

However, further inspection of a scree plot derived from principal components analysis of the PCL-M data provided additional support for a one-factor solution in this sample of USAF Security Forces following deployment. Specifically, the first factor obtained an eigenvalue of 8.42 and accounted for 50% of the explained variance, compared to eigenvalues of less than 1.46 and percentages of explained variance of less than 9% for subsequent factors. These results, combined with findings from the CFA, demonstrated sufficient unidimensionality of PCL-M data in this sample to proceed with IRT analysis.

Given preliminary evidence supporting the underlying assumptions of IRT in this sample, item parameters for the 17 items of the PCL-M were estimated using two-parameter logistic (2PL) analyses in Multilog Version 7 (Thissen, Chen, & Bock, 2003). Results of these analyses are presented in Table 1, listing for each item the discrimination or slope (*a*) parameter and the difficulty (*b*) param-

Table 1

Item Content and Parameters for 17 Items of the PCL-M

Item #	*a*	*b*
1. Repeated disturbing memories, thoughts, or images of a stressful military experience	2.63	0.67
2. Repeated, disturbing dreams of a stressful military experience	3.85	0.63
3. Suddenly acting or feeling as if a stressful military experience were happening again (as if you were reliving it)	4.58	1.05
4. Feeling very upset when something reminded you of a stressful military experience	5.47	0.65
5. Having physical reactions (e.g., heart pounding, trouble breathing, sweating) when something reminded you of a stressful military experience	3.64	0.68
6. Avoiding thinking about or talking about a stressful military experience or avoiding having feelings related to it	3.16	0.58
7. Avoiding activities or situations because they reminded you of a stressful military experience	4.13	0.84
8. Trouble remembering important parts of a stressful military experience	2.74	1.21
9. Loss of interest in activities that you used to enjoy	2.59	0.76
10. Feeling distant or cut off from other people	3.06	0.54
11. Feeling emotionally numb or being unable to have loving feelings for those close to you	2.26	0.68
12. Feeling as if your future will somehow be cut short	2.52	1.17
13. Trouble falling or staying asleep	2.69	−0.07
14. Feeling irritable or having angry outbursts	2.51	0.04
15. Having difficulty concentrating	2.61	0.59
16. Being watchful or on guard	2.06	0.25
17. Feeling jumpy or easily startled	1.77	0.56

Note. *n* = 196. PCL-M = PTSD Checklist–Military Version; *a* = item discrimination parameter; *b* = item difficulty-level parameter.

eter, along with item content. As anticipated, findings indicated substantial variability across items both in terms of their relatedness to PTSD and their likelihood of endorsement at any given level of PTSD. These differences and their implications are highlighted in Figures 1 and 2, in which item characteristic curves (ICCs) are plotted for each item with the latent construct PTSD (θ) represented on the *x*-axis and the probability of endorsing the item represented on the *y*-axis.

Figure 1 contrasts two items that, although having similar levels of relatedness to PTSD, have quite different likelihoods of being endorsed, depending on the underlying level of PTSD (θ) being experienced by the service member. Specifically, Item 13 ("trouble falling or staying asleep") has a 0.50 probability of being endorsed at low levels of PTSD (when θ = −0.07, representing the 48th percentile level), whereas Item 8 ("trouble remembering important parts of a stressful military experience") does not obtain a 0.50

probability of being endorsed until the level of underlying PTSD reaches a much higher level (i.e., when θ = 1.21, representing the 89th percentile level). Hence, if one wanted to select a subset of items that would distinguish service members experiencing the most acute levels of PTSD for crisis intervention while deployed to a combat environment, the three best items (with highest *b* levels) would include Items 8, 12, and 3 ("trouble remembering important parts of a stressful military experience," "feeling as if your future will somehow be cut short," and "suddenly acting or feeling as if a stressful military experience were happening again," respectively). By contrast, for screening at the lowest levels of PTSD for further assessment or prevention purposes—for example, prior to or following deployment—the three best items would include Items 13, 14, and 16 ("trouble falling or staying asleep," "feeling irritable or having angry outbursts," and "being watchful or on guard," respectively).
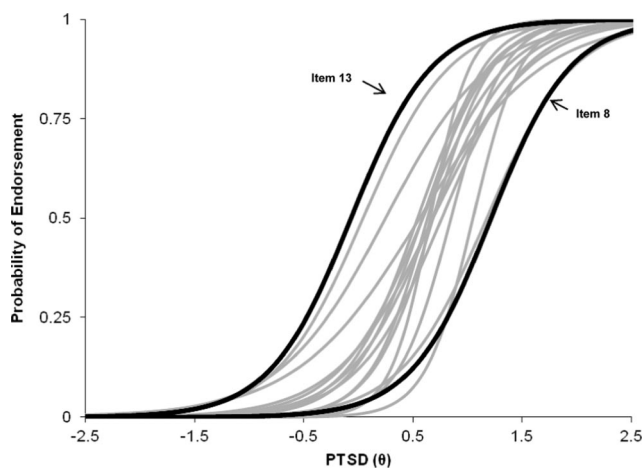

Figure 1. 17-item characteristic curves with bookend difficulty-level (*b*) curves highlighted. PTSD = posttraumatic stress disorder.
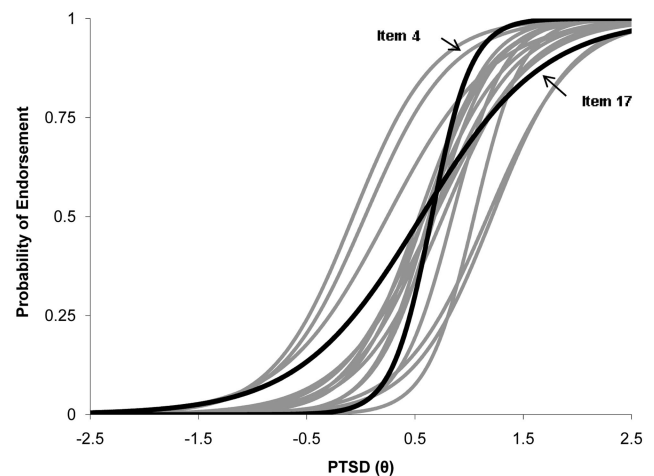

Figure 2. 17-item characteristic curves with bookend discrimination-level (*a*) curves highlighted. PTSD = posttraumatic stress disorder.

By comparison, Figure 2 contrasts two items that have similar likelihood of being endorsed at a given level of PTSD (in this case, a 0.50 probability at approximately the 73rd percentile level) but have sharply different levels of relatedness (or predictive validity) to PTSD. Specifically, Item 17 ("feeling jumpy or easily startled") has only modest relation to the underlying construct of PTSD ($a = 1.77$), compared with the strong relatedness to PTSD ($a = 5.47$) for Item 4 ("feeling very upset when something reminded you of a stressful military experience"). That is, if one desired specifically to target individuals in the top quartile of PTSD symptom severity (θ values approaching 0.675), the three most discriminating items at that level would include Items 4, 2, and 5 ("feeling upset when something reminded you of a stressful military experience," "repeated, disturbing dreams of a stressful military experience," and "having physical reactions . . . when something reminded you of a stressful military experience," respectively), whereas the three least discriminating items would be Items 17, 11, and 15 ("feeling jumpy or easily startled," "feeling emotionally numb or being unable to have loving feelings for those close to you," and "having difficulty concentrating," respectively).

Figure 3 presents the test information curve for the full 17-item PCL-M, reflecting both the overall discriminative capacity of this measure and that point along the PTSD latent construct at which maximum discrimination is provided. In this case, the curve peaks sharply at a value of $\theta = 0.71$, reflecting greatest information at approximately the 76th percentile level of underlying PTSD in this sample—with most of the discriminative information being provided between $\theta = -0.5$ (31st percentile level) and $\theta = 1.5$ (93rd percentile level).

Finally, to test for potential significant differences in item functioning between pre- and postdeployment data, we conducted IRT-based likelihood ratio differential item functioning (DIF) testing (Thissen, Steinberg, & Gerrard, 1986). This type of DIF testing involves statistically comparing IRT models with chi-square difference tests. These tests require first identifying anchor items to establish a common scale (to define the same latent variable) between the groups. We used a procedure adapted from Kim and Cohen (1995). Initial anchoring analyses and subsequent primary DIF analyses were conducted using a program developed by Thissen (2001), adopting a conservative studywide Bonferroni correction to reduce the chances of a false positive. For the analyses presented here, only one item was eligible for subsequent DIF testing, Item 12 ("feeling as if your future will somehow be cut short").

For the primary DIF analyses, we fitted the 2PL model with a and b parameters constrained equal for both groups, and with a and b parameters permitted to vary by group. The constraints significantly decreased model fit, and thus we identified evidence of DIF for this item, $\chi^2(2) = 26.0$, $p < .001$. We then conducted subsequent analyses to determine more specifically if the DIF was driven by the a parameters, the b parameters, or both. Specifically, we compared a model with a parameters constrained to be equal between groups but b parameters free to vary to a model that permitted both item parameters to vary between groups. We found no significance here, indicating no difference between the two groups' a parameters, $\chi^2(1) = 0.7$, $p = .40$. In other words, this item did not differ in its strength of relation to the latent PTSD continuum across groups. Then, we conducted a b DIF test conditional on equal slope parameters between the two groups. We found significance here, indicating a difference between the two groups' b parameters of 1.00 standard deviation, $\chi^2(1) = 25.3$, $p < .001$.

This finding indicates that Item 12 ("feeling as if your future will somehow be cut short") functioned differently when administered at pre- versus postdeployment. Specifically, the item was more readily endorsed by individuals predeployment, even when controlling for group mean differences. Conceptually, this differential functioning reflects that this item would likely have very different meaning to someone deploying for combat versus returning home to the relative safety of his or her community. Indeed, as noted earlier, only those service members experiencing the highest levels of PTSD symptomatology would be likely to endorse this item after returning stateside, as such an endorsement would indicate impaired functioning and distress. Conversely, such an endorsement prior to deployment may simply reflect an understandable concern elicited by the anticipation of entering a dangerous, unpredictable environment. In other words, when this item is endorsed at predeployment it may simply reflect that the service member is acknowledging that he or she is about to enter a war zone and thus has valid concerns about his or her future. Upon return, when facing a relatively safe future, only a few service members see their future bleakly or anticipate it being "cut short." These individuals very well may be experiencing significant PTSD. Stated yet another way, an endorsement of this item at
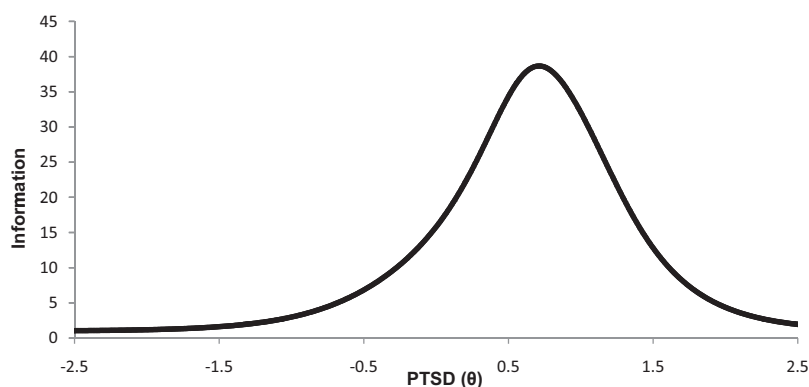

Figure 3. Test information curve for the full 17-item scale. PTSD = posttraumatic stress disorder.

predeployment can be viewed with relatively less alarm. Such an endorsement prior to deployment may reflect an understandable concern elicited by the anticipation of entering a dangerous, unpredictable environment.

Discussion

High rates of posttraumatic stress at both diagnostic and subthreshold levels among service members during and following deployment to a theater of combat operations compel both efficient and accurate assessment strategies for evaluating PTSD symptom severity throughout the deployment cycle. Identifying service members at all levels of PTSD intensity is particularly important given that veterans with subthreshold PTSD tend to be overlooked by clinicians and researchers despite their experiencing substantial trauma-related symptoms (Grubaugh et al., 2005). Differences in prevalence rates as well as differing objectives related to specific populations or evaluation contexts require the ability to tailor assessment strategies based on individual item characteristics. Although the PCL-M constitutes the most widely used measure of PTSD symptoms in both military and veteran health-care settings, aggregate scores based on the 17-item measure are not optimally matched to specific populations or evaluation contexts. As suggested by previous research and confirmed in the present study, items differ considerably with respect to their relatedness to PTSD and their likelihood of endorsement at a given level of an underlying PTSD construct. Additionally, any given aggregate score on the 17-item PCL remains ambiguous with respect to interpretation, because of multiple ways in which any given score may be obtained.

Moreover, in some evaluation contexts the full 17-item PCL may not be realistic to administer—for example, in comprehensive screenings examining a broad range of potential physical and mental disorders, or during interviews in primary care settings or by supervising officers in theaters of combat (in which only three to five items related to PTSD may be possible). In such situations, which subset of items comprises the optimal composite depends entirely on the objectives of the evaluation context. As apparent from Table 1 and explicated in the Results section, items best suited for identifying individuals at the highest levels of posttraumatic stress (with high levels of the b parameter) differ from those optimally suited to screen for individuals at the lowest levels. In theaters of combat where retaining the maximum number of deployed personnel is paramount, only those individuals experiencing acute levels of PTSD symptoms may be pulled from operations or reassigned to less critical functions. By contrast, when screening for potential PTSD-related difficulties following deployment, or screening for subthreshold PTSD among redeploying personnel, items more sensitive but less specific to PTSD (with low levels of the b parameter) may be desired. Using item parameters provided in Table 1, both clinicians and researchers could select a subset of items having high discrimination within a narrow range of PTSD severity (high levels of the a parameter within a narrow range of the b parameter), or the same number of items having lower discrimination (a) at any specific point but exhibiting moderate discrimination across a broader range of PTSD severity.

It is worth noting, in this context, that items on the PCL that correspond to the four items comprising the Primary Care PTSD Screen (PC-PTSD) developed by Prins et al. (2003) were not consistently those obtaining the highest discrimination (a parameter) values in the present study. For example, the item used by Prins et al. to capture hyperarousal ("being constantly on guard, watchful, or easily startled") had lower discrimination values on analogous PCL items among Airmen in this study than did items reflecting "feeling irritable or having angry outbursts" or "having difficulty concentrating." Similarly, the PC-PTSD item intended to capture reexperiencing ("having nightmares . . . or thoughts about it when you did not want to") had lower discrimination values on analogous PCL items in this study compared to PCL items in this sample reflecting "feeling very upset when something reminded you of a stressful military experience" or "suddenly acting or feeling as if a stressful military experience were happening again." The ideal subset of items for any specific assessment context will necessarily reflect both the item characteristics for representative samples in that setting and the specific purposes of the evaluation (e.g., the threshold level of PTSD being targeted by the assessment).

Item parameters for the PCL-M identified in this study were derived from a predominantly (93%) male sample of 196 service members following a yearlong deployment on a high-risk mission to Iraq. The composition of the sample did not permit comparison of parameters for women versus those for men, and hence the item characteristics obtained here may not generalize to women active-duty service members or veterans. Moreover, both the difficulty level of individual items and their relatedness to PTSD may differ when assessed at times other than postdeployment (e.g., predeployment or in theater). The rates of endorsement for specific items and their relation to PTSD might also be hypothesized to vary depending on the specialized training or experiences of specific personnel units contributing to varying levels of risk or resiliency to posttraumatic stress or other sequelae of exposure to combat-related stressors.

Similarly, item characteristics of the PCL may vary as a function of the type of stressor experienced—for example, when using civilian versions of the PCL to assess traumatic response to cancer (e.g., Andrykowski, Cordova, Studts, & Miller, 1998; DuHamel et al., 2004; Shelby, Golden-Kreutz, & Andersen, 2005; Smith, Redd, DuHamel, Vickberg, & Ricketts, 1999), motor vehicle accident (Blanchard et al., 1996), sexual harassment or assault (e.g., Blanchard et al., 1996; Palmieri & Fitzgerald, 2005), or intimate partner violence (e.g., Krause, Kaltman, Goodman, & Dutton, 2007). In these and similar applications, adopting IRT as a basis for examining item characteristics of the PCL may facilitate more efficient and accurate assessment by tailoring specific subsets of items optimally matched to the population, setting, and objectives of the evaluation context.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andrykowski, M. A., Cordova, M. J., Studts, J. L., & Miller, T. W. (1998). Posttraumatic stress disorder after treatment for breast cancer: Prevalence of diagnosis and use of the PTSD Checklist-Civilian Version (PCL-C) as a screening instrument. *Journal of Consulting and Clinical Psychology, 66,* 586–590. doi:10.1037/0022-006X.66.3.586

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246. doi:10.1037/0033-2909.107.2.238

Blanchard, E. B., Hickling, E. J., Taylor, A. E., Loos, W. R., & Gerardi, R. J. (1994). Psychological morbidity associated with motor vehicle accidents. *Behaviour Research and Therapy, 32,* 283–290. doi:10.1016/0005-7967(94)90123-6

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy, 34,* 669–673. doi:10.1016/0005-7967(96)00033-2

Bliese, P. D., Wright, K. M., Adler, A. B., Cabrera, O., Castro, C. A., & Hoge, C. W. (2008). Validating the Primary Care Posttraumatic Stress Disorder Screen and the Posttraumatic Stress Disorder Checklist with soldiers returning from combat. *Journal of Consulting and Clinical Psychology, 76,* 272–281. doi:10.1037/0022-006X.76.2.272

DuHamel, K. N., Ostroff, J., Ashman, T., Winkel, G., Mundy, E. A., Keane, T. M., . . . Redd, W. (2004). Construct validity of the Posttraumatic Stress Disorder Checklist in cancer survivors: Analyses based on two samples. *Psychological Assessment, 16,* 255–266. doi:10.1037/1040-3590.16.3.255

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Forbes, D., Creamer, M., & Biddle, D. (2001). The validity of the PTSD Checklist as a measure of symptomatic change in combat-related PTSD. *Behaviour Research and Therapy, 39,* 977–986. doi:10.1016/S0005-7967(00)00084-X

Grubaugh, A. L., Magruder, K. M., Waldorp, A. E., Elhai, J. D., Knapp, R. G., & Frueh, B. C. (2005). Subthreshold PTSD in primary care: Prevalence, psychiatric disorders, healthcare use, and functional status. *Journal of Nervous and Mental Disease, 193,* 658–664. doi:10.1097/01.nmd.0000180740.02644.ab

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Hoge, C. W., Castro, C. A., Messer, S. C., McGurk, D., Cotting, D. I., & Koffman, R. L. (2004). Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *New England Journal of Medicine, 351,* 13–22. doi:10.1056/NEJMoa040603

Kim, S., & Cohen, A. S. (1995). A comparison of Lord's chi-square, Raju's area measures, and the likelihood ratio test on detection of differential item functioning. *Applied Measurement in Education, 8,* 291–312. doi:10.1207/s15324818ame0804_2

King, D., Leskin, G., King, L., & Weathers, F. (1998). Confirmatory factor analysis of the clinician-administered PTSD scale: Evidence for the dimensionality of posttraumatic stress disorder. *Psychological Assessment, 10,* 90–96. doi:10.1037/1040-3590.10.2.90

Krause, E. D., Kaltman, S., Goodman, L. A., & Dutton, M. A. (2007). Longitudinal factor structure of posttraumatic stress symptoms related to intimate partner violence. *Psychological Assessment, 19,* 165–175. doi:10.1037/1040-3590.19.2.165

Lang, A. J., Laffaye, C., Satz, L. E., Dresselhaus, T. R., & Stein, M. B. (2003). Sensitivity and specificity of the PTSD Checklist in detecting PTSD in female veterans in primary care. *Journal of Traumatic Stress, 16,* 257–264. doi:10.1023/A:1023796007788

Lang, A. J., & Stein, M. B. (2005). An abbreviated PTSD Checklist for use as a screening instrument in primary care. *Behaviour Research and Therapy, 43,* 585–594. doi:10.1016/j.brat.2004.04.005

Muthén, L. K., & Muthén, B. O. (2007). *Mplus (Version 5.2)* [Computer software]. Los Angeles, CA: Muthén & Muthén.

National Defense Authorization Act for Fiscal Year 2010, Pub. L. No. 111–84, § 708, 123 Stat. 2197 (2009).

Palmieri, P. A., & Fitzgerald, L. F. (2005). Confirmatory factor analysis of posttraumatic stress symptoms in sexually harassed women. *Journal of Traumatic Stress, 18,* 657–666. doi:10.1002/jts.20074

Prins, A., Ouimette, P., Kimerling, R., Cameron, R. P., Hugelshofer, D. S., Shaw-Hegwer, J., . . . Sheikh, J. I. (2003). The Primary Care PTSD Screen (PC-PTSD): Development and operating characteristics. *Primary Care Psychiatry, 9,* 9–14. doi:10.1185/135525703125002360

Seal, K. H., Metzler, T. J., Gima, K. S., Bertenthal, D., Maguen, S., & Marmar, C. R. (2009). Trends and risk factors for mental health diagnoses among Iraq and Afghanistan veterans using Department of Veterans Affairs health care, 2002–2008. *American Journal of Public Health, 99,* 1651–1658. doi:10.2105/AJPH.2008.150284

Shelby, R. A., Golden-Kreutz, D. M., & Andersen, B. L. (2005). Mismatch of posttraumatic stress disorder (PTSD) symptoms and *DSM–IV* symptom clusters in a cancer sample: Exploratory factor analysis of the PTSD Checklist-Civilian Version. *Journal of Traumatic Stress, 18,* 347–357. doi:10.1002/jts.20033

Sheppard, S. C., Malatras, J. W., & Israel, A. C. (2010). The impact of deployment on U.S. military families. *American Psychologist, 65,* 599–609. doi:10.1037/a0020332

Simms, L. J., Watson, D., & Doebbeling, B. N. (2002). Confirmatory factor analyses of posttraumatic stress symptoms in deployed and nondeployed veterans of the Gulf War. *Journal of Abnormal Psychology, 111,* 637–647. doi:10.1037/0021-843X.111.4.637

Smith, M. Y., Redd, W., DuHamel, K., Vickberg, S. J., & Ricketts, P. (1999). Validation of the PTSD Checklist-Civilian Version in survivors of bone marrow transplantation. *Journal of Traumatic Stress, 12,* 485–499. doi:10.1023/A:1024719104351

Thissen, D. (2001). *IRTLRDIF v2.0b: Software for the computation of the statistics involved in item response theory likelihood ratio tests for differential item functioning*. Unpublished manuscript, University of North Carolina, L. L. Thurstone Psychometric Laboratory.

Thissen, D., Chen, W.-H., & Bock, R. D. (2003). *Multilog (Version 7)* [Computer software]. Lincolnwood, IN: Scientific Software.

Thissen, D., Steinberg, L., & Gerrard, M. (1986). Beyond group-mean differences: The concept of item bias. *Psychological Bulletin, 99,* 118–128. doi:10.1037/0033-2909.99.1.118

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38,* 1–10. doi:10.1007/BF02291170

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., & Keane, T. M. (1993). *The PTSD Checklist: Reliability, validity, and diagnostic utility*. Paper presented at the meeting of the International Society for Traumatic Stress Studies, San Antonio, TX.

Yeager, D. E., Magruder, K. M., Knapp, R. G., Nicholas, J. S., & Frueh, B. C. (2007). Performance characteristics of the Posttraumatic Stress Disorder Checklist and SPAN in Veterans Affairs primary care settings. *General Hospital Psychiatry, 29,* 294–301. doi:10.1016/j.genhosppsych.2007.03.004