Article

Assessment
2023, Vol. 30(5) 1590–1605

DOI: 10.1177/10731911221113571
journals.sagepub.com/home/asm
SAGE

Self-Rated Versus Clinician-Rated Assessment of Posttraumatic Stress Disorder: An Evaluation of Discrepancies Between the PTSD Checklist for *DSM-5* and the Clinician-Administered PTSD Scale for *DSM-5*

Lindsay B. Kramer[1,2], Sarah E. Whiteman[2], Jessica M. Petri[2,3], Elizabeth G. Spitzer[2,4], and Frank W. Weathers[2,4,5]

Abstract

Posttraumatic stress disorder (PTSD) is commonly assessed with self-rated or clinician-rated measures. Although scores from these assessment modalities are strongly associated, they are often discrepant for individual symptoms, total symptom severity, and diagnostic status. To date, no known studies have empirically identified the sources of these discrepancies. In the present study, we had three aims: (a) replicate previously identified discrepancies; (b) examine contribution of possible objective predictors of discrepancies, including negative response bias, random responding, conscientiousness, neuroticism, and verbal IQ; and (c) identify subjective sources of discrepancies through analysis of participant feedback. Trauma-exposed undergraduates ($N = 60$) were administered the PTSD Checklist for *DSM-5* (PCL-5), the Clinician-Administered PTSD Scale for *DSM-5* (CAPS-5), and other questionnaires. Interviewers identified discrepancies between corresponding PCL-5/CAPS-5 scores and asked participants to describe their attributions for discrepancies. Discrepancies, both dimensional and dichotomous, occurred at the item, cluster, and total score level. Objective predictors were weakly associated with discrepancies. The most commonly reported reasons for discrepancies were time-frame reminders, comprehension of symptoms, trauma-related attribution errors, increased awareness, and general errors. These findings help explain discordance between the PCL-5 and CAPS-5, and inform use and interpretation of these two widely used PTSD measures in clinical and research applications.

Keywords

posttraumatic stress disorder, assessment, PTSD Checklist for *DSM-5*, Clinician-Administered PTSD Scale for *DSM-5*, discrepancies

The development and psychometric evaluation of assessment measures for posttraumatic stress disorder (PTSD) has been a highly productive area of traumatic stress research since PTSD was introduced as a diagnostic category in the *Diagnostic and Statistical Manual of Mental Disorders–Third Edition* (*DSM-III*; American Psychological Association [APA], 1980). Numerous PTSD measures have been created, and many of them have been extensively validated (Bovin & Weathers, 2022; Weathers et al., 2014; Weathers & Keane, 1999). The most commonly used assessment methods for PTSD are self-rated questionnaires and clinician-rated interviews. Questionnaires such as the PTSD Checklist (PCL; Weathers et al., 1993) provide brief, inexpensive identification of individuals with PTSD in clinical and

[1]University of Cincinnati College of Medicine, OH, USA
[2]Auburn University, AL, USA
[3]Emory University School of Medicine, Atlanta, GA, USA
[4]VA Boston Healthcare System, Boston, MA, USA
[5]National Center for PTSD, Boston, MA, USA

Corresponding Author:
Lindsay B. Kramer, Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati College of Medicine, 3120 Burnet Avenue, Cincinnati, OH 45229, USA.
Email: kramerlb@ucmail.uc.edu

research settings. However, it is generally accepted that interviews such as the Clinician-Administered PTSD Scale (CAPS; Blake et al., 1990) are necessary to obtain an accurate diagnosis. The PCL and CAPS are the most widely used and extensively validated *DSM*-correspondent PTSD questionnaire and interview, respectively (Elhai et al., 2005; Weathers et al., 2001), and both were updated to correspond with changes to the PTSD diagnostic criteria in the *Diagnostic and Statistical Manual of Mental Disorders–Fifth Edition* (*DSM-5*; APA, 2013).

The PCL has been used for clinical screening and tracking symptom severity (Berlant & van Kammen, 2002), identification of PTSD in research studies (Hoge et al., 2008), and determination of epidemiological rates of PTSD (Erbes et al., 2007). The PCL identifies individuals with provisional PTSD, which is consistent with most questionnaires of psychopathology (Hopwood et al., 2008). Both the original PCL (McDonald & Calhoun, 2010; Weathers et al., 1993; Wilkins et al., 2011) and the updated *DSM-5* version (PCL-5; Weathers, Litz, et al., 2013) have demonstrated high internal consistency and test–retest reliability, as well as strong convergent and discriminant validity, structural validity, diagnostic utility, and sensitivity to clinical change (Blevins et al., 2015; Bovin et al., 2016; Wortmann et al., 2016). However, the PCL has several limitations. First, it assesses symptoms of PTSD, but does not assess trauma-relatedness of symptoms or functional impairment. Second, optimal PCL cutoff scores vary widely across studies, populations, and type of traumatic event (McDonald & Calhoun, 2010). For example, regarding the PCL-5, scores from 31 to 33 were found to be optimal in two samples of U.S. veterans (Bovin et al., 2016), while 34 was optimal in a U.K. veteran sample (Murphy et al., 2017), 23 was optimal in a Kurdish/Arab refugee sample (Ibrahim et al., 2018), and 41 was optimal in a first responder sample (Morrison et al., 2021). Third, responses to PCL items are answered through an individual's subjective understanding, awareness of their symptoms, and motivation to respond attentively and honestly, so scores are susceptible to various response biases. The PCL has good utility as a screening tool and as a measure of PTSD symptom severity, but researchers have cautioned against using it alone as a diagnostic tool (McDonald & Calhoun, 2010).

The CAPS is one of the most widely used and extensively validated structured interviews for PTSD diagnostic status and symptom severity. It assesses exposure to Criterion A traumatic event(s); PTSD symptom presence, frequency, and intensity; overall symptom severity; global distress; social and occupational impairment; and response validity. The CAPS also assesses trauma-relatedness of symptoms not inherently linked to the trauma (e.g., loss of interest or pleasure, detachment or estrangement, concentration difficulties), which allows individual symptoms to be specifically linked to the index event. This is an important feature as some PTSD symptoms overlap with other mental disorders such as depression. Therefore, clarifying that the symptoms are functionally related to the traumatic event ensures that the symptoms are part of a PTSD syndrome and not better accounted for by a different disorder.

CAPS items include multiple behaviorally anchored prompts that provide a detailed assessment of each symptom. Furthermore, raters are encouraged to rephrase prompts when respondents do not appear to understand what is being asked and clarify vague or insufficient responses. The CAPS (Elhai et al., 2005; Weathers et al., 1999, 2001) and the revised *DSM-5* version (CAPS-5; Weathers, Blake, et al., 2013) have demonstrated excellent psychometric properties, including high internal consistency, interrater reliability, and test–retest reliability, as well as strong convergent, discriminant, and structural validity (Weathers et al., 2018). Limitations of the CAPS include longer administration time and the need for clinical interviewers proficient in standardized administration and scoring.

The PCL and CAPS are both *DSM*-correspondent measures, and their scores have consistently been found to be strongly associated (e.g., Blanchard et al., 1996; Forbes et al., 2001). However, the correspondence is never perfect, and studies have found varying degrees of diagnostic discordance. Blanchard and colleagues (1996) examined the psychometric properties of the PCL using scores from the CAPS as the criterion. They found that correlations between PCL and CAPS items ranged widely (rs = .39–.79), but the correlation between PCL total severity score and CAPS total severity score was very high (r = .93). Regarding diagnostic utility, Blanchard and colleagues found that diagnostic efficiencies for PCL items in predicting the corresponding items on the CAPS were uniformly high (all .70 or higher). Furthermore, using the recommended PCL cutoff score of 50, they found a diagnostic efficiency of .83. Overall, their results indicate a high degree of correspondence between the PCL and CAPS for total severity and diagnostic status. However, there was substantial discordance for a few individual symptoms.

Other studies have compared the PCL and CAPS on their sensitivity to change as a function of treatment. Forbes et al. (2001) studied 97 male Vietnam veterans who completed the PCL and CAPS prior to treatment and at a 9-month follow-up. They examined change across the two time points, finding a 17.5% reduction in CAPS total severity, for a large effect size of .84, compared with an 8.4% reduction in PCL total severity, for a moderate effect size of .59. This suggests that, relative to the CAPS, the PCL is less sensitive to clinical change. Forbes and colleagues (2001) also found a low level of agreement between the PCL and the CAPS for several symptoms, including avoidance of reminders, flashbacks, sleep difficulties, upset by

reminders, and estrangement from others. Similar to Blanchard et al. (1996), they concluded that the PCL may not accurately assess the presence or severity of individual symptoms.

The PCL and CAPS have also been compared on their factor structure. Palmieri et al. (2007) evaluated the potential impact of assessment modality by comparing confirmatory factor analysis (CFA) results for the PCL and the CAPS in a sample of utility workers exposed to the World Trade Center Ground Zero site. The key finding was that the PCL and the CAPS displayed different latent structures. The CFA of the PCL supported a four-factor model with distinct reexperiencing, avoidance, dysphoria, and hyperarousal factors, while the CFA of the CAPS supported a four-factor model with distinct reexperiencing, avoidance, emotional numbing, and hyperarousal factors. Although these differences are small, they confirm that self-rated and clinician-rated measures can yield different latent models with different conceptual implications. On the contrary, Lee et al. (2019) found a similar structure across PCL and CAPS scores, with little evidence of a method effect.

To date, research comparing the updated PCL-5 and CAPS-5 suggests that the PCL/CAPS discordance may be even more salient in *DSM-5*. Newly added symptoms (i.e., blame, aggressive behavior, and risk-taking) and heavily revised symptoms (e.g., negative beliefs, inability to experience positive feelings) for *DSM-5* PTSD could be conceptually difficult for individuals to comprehend and accurately rate. Furthermore, Weathers and colleagues (2018) found PCL-5 and CAPS-5 total severity scores were highly correlated, but not to the degree expected for measures of the same construct. Bovin and colleagues (2016) found that PCL-5 scores of 31 to 33 were optimally efficient for predicting a CAPS-5 PTSD diagnosis, but even these best cutoffs resulted in substantial diagnostic discordance, with the CAPS-5 yielding a lower estimate of PTSD prevalence.

Several reasons may account for this discordance, including differences in instructions, format, content, and process (Monson et al., 2008). On the CAPS-5, individuals answer questions about frequency and intensity of individual PTSD symptoms, but on the PCL-5, they rate their subjective distress, or the degree to which they are "bothered by" each symptom. Individuals may lack awareness or have an exaggerated view of their covert PTSD symptoms (e.g., intrusive memories, avoidance of thoughts), which could lower reliability and validity when reporting these symptoms that are not behaviorally anchored (Foa et al., 2016). There are also no opportunities for multiple prompts or follow-up clarifications on the PCL-5 like on the CAPS-5 (Marmar et al., 2015). Furthermore, raters administering the CAPS-5 can prevent double-coding of symptoms (i.e., counting the same symptom episodes toward more than one symptom criterion) and use clinical judgment to rate symptoms according to the conceptual basis of PTSD symptoms and the clinical meaning of rating scale anchors. In contrast, the PCL-5 is based on participants' self-ratings, the accuracy of which is influenced by comprehension of items, possible consideration of non-PTSD or non-trauma-related distress in subjective ratings, and response bias.

Despite the importance of PCL-5/CAPS-5 concordance for diagnosis and treatment, no studies have attempted to empirically identify and test these potential sources of PCL-5/CAPS-5 discrepancies. In the present study, we utilized both quantitative and qualitative approaches to understand PCL-5/CAPS-5 discordance. First, we quantified PCL-5/CAPS-5 discrepancies at the item, cluster, total score, and diagnosis level. Furthermore, we examined discrepancies for both dimensional (e.g., 0–4) and dichotomous (present/absent) scores. We hypothesized that PCL-5 and CAPS-5 total scores would be strongly positively associated, but would also yield substantial levels of symptom-level and diagnostic-level discrepancies. In particular, we expected covert PTSD symptoms (i.e., intrusive memories, flashbacks, avoidance of thoughts and feelings) to be most discrepant.

Second, we examined potential objective predictors of discrepancies, including response validity indicators of negative response bias, inconsistency, and infrequency; personality traits of conscientiousness and neuroticism; and verbal IQ. As these predictors have not been examined previously for this purpose, their inclusion in the present study was exploratory. The rationale for examining response validity indicators is that infrequent and inconsistent responding, as well as the tendency to present oneself in an overly negative manner, are not only common sources of response bias, but also would be more likely on a questionnaire than on an interview, thereby creating discrepancy. The rationale for examining conscientiousness is that individuals high in this personality trait tend to be more cautious and thorough, and thus might be more likely to make a greater effort to read and comprehend questionnaire items, follow directions, and answer thoughtfully and honestly, all of which suggests a possible negative relationship between this trait and discrepancy. The rationale for examining neuroticism is that this trait has been found to be related to self-reporting more frequent and severe symptoms of illness (Costa & McCrae, 1985; Watson & Pennebaker, 1989). Furthermore, individuals high in neuroticism tend to be more introspective in nature and have also been associated with a greater tendency to interpret symptoms in a negative manner (Barsky & Klerman, 1983; Watson & Clark, 1984). These tendencies would likely have a more substantial impact on self-ratings on questionnaires than on clinician-ratings on structured interviews, thus leading to higher discrepancy. Last, the rationale for examining verbal IQ is that it could affect individuals' ability to read and comprehend questionnaire items. Wilkins et al. (2011) noted that the reading level of the PCL might be above the ability of some adults, so lower IQ might affect comprehension of PCL

items, especially some of the longer, more abstract symptoms, such as flashbacks or negative beliefs. Thus, it was hypothesized that verbal IQ would negatively predict discrepancy.

Third, we presented participants with their actual discrepant scores following PCL-5 and CAPS-5 administration and recorded their attributions about why the discrepancies might have occurred. We hypothesized that participants' attributions would generally reflect the main sources of discrepancy previously proposed by researchers, including statements that the CAPS-5 explicitly links symptoms to the index event, inquires in greater detail, improves comprehension of symptoms, and facilitates greater disclosure compared with the PCL-5. Overall, it is important to understand the concordance between the PCL-5 and CAPS-5 because both are measures of the same construct used in diagnosis and treatment planning decisions, but may be providing disparate information about an individual's PTSD symptoms.

Method

Participants and Procedure

Participants were undergraduate students ages 19 and older enrolled in a psychology course at a large public southeastern university. All procedures were approved by the university Institutional Review Board. In Part 1, students completed an online questionnaire battery consisting of an information letter, demographics questionnaire, the Life Events Checklist for *DSM-5* (LEC-5; Weathers, Blake, et al., 2013), the PCL-5, and the Personality Assessment Inventory (PAI; Morey, 1991) for research credit. Participants were eligible for Part 2 if they reported having experienced a *DSM-5* Criterion A traumatic event and endorsed moderate levels of current PTSD symptoms (total PCL-5 score $\geq$ 24). Criterion A status was determined by reviewing participants' responses on the LEC-5 and their written narrative description of their index event. Initially, syntax was created to provide an initial coding of Criterion A event exposure. Two doctoral students in clinical psychology then independently reviewed the syntax's coding and verified if the event met *DSM-5* Criterion A. Sixty of 200 eligible participants invited by email completed Part 2, which included an informed consent form, the Big Five Inventory (BFI; John & Srivastava, 1999), the Shipley–2; (Shipley et al., 2009), the CAPS-5, and an unstructured discrepancy interview. Individuals who chose to participate in Part 2 of the study were scheduled within 1 to 2 weeks following completion of Part 1. The only Part 1 data included in the analyses were the PAI data.

Part 2 lasted approximately 2 hours, and participants were compensated with additional research credit and a US$15 gift card. In Part 2, participants were first provided with informed consent. Participants then completed the PCL-5, referring to the same index event that had been identified in Part 1. Next, participants completed the BFI. The interviewer then administered the CAPS-5 in reference to the same event that had been previously identified. Afterward, participants completed the Shipley–2, while the interviewer reviewed PCL-5 and CAPS-5 responses to identify discrepancies. For the purposes of this study, a discrepancy was considered present when the corresponding CAPS-5 and PCL-5 item did not have the identical score (detailed method provided in "Data Analysis" section). We then conducted an audio-recorded interview in which participants commented on why they responded differently across the two PTSD measures. Finally, interviewers thanked and debriefed participants. Interviewers were doctoral students in clinical and counseling psychology, trained and supervised by the last author, a licensed clinical psychologist with expertise in the assessment of PTSD. Interviewers met regularly for discussion and reliability checks, for which they independently rated audiotaped interviews and discussed discrepant ratings.

The final sample (N = 60) was predominantly female (88.3%; n = 53) and ranged in age from 19 to 24 (M = 19.83; SD = 1.2) years. The racial breakdown was 81.7% European American/White (n = 49), 11.7% African American/Black (n = 7), 3.3% Asian American/Asian Origin (n = 2), 1.7% American Indian or Alaskan Native (n = 1), and 1.7% Other (n = 1). Traumatic event types included sexual assault (48.3%, n = 29); transportation accident (10.0%, n = 6); physical assault (15.0%, n = 9); unwanted sexual experience (6.7%, n = 4); sudden and violent death (8.3%, n = 5); serious accident at work, home, or during a recreational activity (5.0%, n = 3); exposure to toxic substance (3.3%, n = 2); and learning about suicide (3.3%, n = 2). In the final sample, the prevalence of *DSM-5* PTSD diagnosis based on the CAPS-5 was 51.7% (n = 31) and provisional *DSM-5* PTSD diagnosis based on PCL-5 was 51.7% (n = 31).

Measures

The LEC-5 is a self-report measure of traumatic event exposure consisting of 17 categories of traumatic stressors (e.g., natural disaster, fire or explosion, transportation accident, sexual assault). Respondents indicate the degree to which they have experienced each category of traumatic stressor by checking all response options that apply, including happened to me, witnessed it, learned about it, part of my job, not sure, or does not apply. In the current study, participants were asked to select their worst event from those they endorsed on the LEC-5 and provide a short written description used to verify Criterion A status. Previous versions of the LEC have been shown to be psychometrically sound in samples of college students and veterans (Gray et al., 2004).

The PCL-5 is a 20-item questionnaire that assesses *DSM-5* PTSD symptoms. For each symptom, respondents rate how much they were bothered by each symptom in the past month, with scores ranging from 0 (*not at all*) to 4 (*extremely*). PCL-5 scores have well-established reliability and validity (Blevins et al., 2015). In the present study, Cronbach's alpha for the full PCL-5 was .89.

The PAI is a 344-item self-report inventory developed to assess personality functioning and psychopathology. Participants respond to each item on a 4-point scale ranging from 1 (*not true at all*) to 4 (*very true*). The PAI has been extensively evaluated and has excellent psychometric properties (Morey, 1991, 2007). We examined three PAI scales: Negative Impression Management (NIM), which reflects a tendency to present oneself in an overly negative manner; Inconsistency (INC), which reflects a tendency to respond inconsistently to items; and Infrequency (INF), which reflects a tendency to respond to items in an idiosyncratic way.

The BFI is a 44-item questionnaire that assesses the Big Five personality domains of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. Respondents rate whether they demonstrate each particular trait on a 5-point scale, with scores ranging from 1 (*disagree strongly*) to 5 (*agree strongly*). Subscale scores are created by reverse scoring specified items, summing the ratings for the items on each subscale, and dividing by the total number of items to obtain a mean score. The BFI has demonstrated good internal consistency, test–retest reliability, and convergent validity with other personality measures (John & Srivastava, 1999). For the present study, we used the Conscientiousness and Neuroticism subscales, for which Cronbach's alphas were .81 and .86, respectively.

The CAPS-5 is a structured diagnostic interview for *DSM-5* PTSD that assesses the intensity and frequency of each PTSD symptom with behaviorally anchored prompts. Interviewers follow CAPS-5 scoring rules to combine intensity and frequency of symptoms and derive a symptom severity score on a 5-point scale: 0 (*absent*), 1 (*mild/subthreshold*), 2 (*moderate/threshold*), 3 (*severe/markedly elevated*), 4 (*extreme/incapacitating*). CAPS-5 diagnosis and total severity score have demonstrated strong reliability and validity (Weathers et al., 2018). In the present study, Cronbach's alpha for the full CAPS-5 was .87.

The Shipley–2 is a self-report measure of cognitive functioning and impairment that provides an estimate of verbal and non-verbal reasoning ability. Of its three subtests, only the 40-item vocabulary test was utilized in the present study to measure verbal ability. Respondents had 10 minutes to choose which of four listed words "means the same or nearly the same" as a specified target word. Correct responses were summed to create a total verbal ability score. The Shipley–2 has demonstrated good test–retest reliability and convergent validity with the Wechsler Adult Intelligence Scale–III (WAIS-III) and the Wonderlic Personnel Test (Shipley et al., 2009).

Data Analysis

There were no missing data for any variables in the present study. To address the first study aim, we documented the degree of discordance between the PCL-5 and the CAPS-5. First, using PCL-5 and CAPS-5 dimensional scores (i.e., item ratings 0–4), we calculated correlations between PCL-5 and CAPS-5 items, subscales (Reexperiencing [RE], Avoidance [AV], Negative Alterations in Cognition and Mood [NACM], Alterations in Arousal and Reactivity [AAR]), and total score. We also evaluated mean differences between PCL-5 and CAPS-5 items, subscales, and total score with paired *t*-tests. Due to the number of comparisons, we used the Benjamini and Hochberg (1995) procedure to control for the possibility of Type I error inflation (maintain at $\alpha = .05$), and calculated Cohen's *d* effect size of the difference (Cohen, 1988; $d = .20–.49$ small, $d = .50–.79$ medium, $d > .80$ large). We conducted these analyses separately for CAPS-5 Intensity and CAPS-5 Severity scores. This was because we anticipated that PCL-5 scores might be less discrepant from CAPS-5 Intensity scores than from CAPS-5 Severity scores, given that neither PCL-5 scores nor CAPS-5 Intensity scores are based on frequency of symptoms, whereas CAPS-5 Severity scores do incorporate frequency of symptoms. Although CAPS-5 Intensity ratings are not numerical, for the purpose of the analyses, CAPS-5 Intensity anchors were coded as 0 (*absent*), 1 (*minimal*), 2 (*clearly present*), 3 (*pronounced*), and 4 (*extreme*).

Second, we calculated dimensional discrepancy scores for each PTSD symptom by subtracting the CAPS-5 item score (0–4) from the corresponding PCL-5 item score (0–4), so that positive discrepancy scores indicated a higher PCL-5 score compared with the CAPS-5 score. We then summed the absolute values of item-level dimensional discrepancy scores to create discrepancy scores for total PTSD. We again conducted these analyses separately for CAPS-5 Intensity scores and CAPS-5 Severity scores. Using these dimensional discrepancy scores, we documented the frequency of any dimensional discrepancy, and the mean discrepancy between PCL-5 and CAPS-5 items, subscales, and overall PTSD. We conducted paired *t*-tests with the Benjamini and Hochberg (1995) correction, Cohen's *d* effect size, and McNemar tests to examine differences in the mean and frequency of dimensional discrepancies between PCL-5 scores and CAPS-5 Intensity and Severity scores at the item, subscale, and total scale levels.

Third, we evaluated dichotomous discrepancies. At the item level, a dichotomous discrepancy occurred when there was a disagreement between dichotomized (present/absent) scores on the PCL-5 and CAPS-5 Severity. A PCL-5 item was counted as a PTSD symptom if it was rated as 2 = *moderately* or higher,

and similarly, a CAPS-5 item Severity was counted as a PTSD symptom if it was rated as 2 = *moderate/threshold* or higher. At the diagnostic level, a discrepancy occurred when there was a disagreement between PCL-5 and CAPS-5 on PTSD diagnostic status. This was derived by using the dichotomized symptom scores and following the *DSM-5* PTSD diagnostic rule, which requires at least one RE symptom, one AV symptom, two NACM symptoms, and two AAR symptoms for a PTSD diagnosis. This produced four outcomes at both the item and diagnostic level: hits (PCL-5 present/CAPS-5 present), misses (PCL-5 absent/CAPS-5 present), false alarms (PCL-5 present/CAPS-5 absent), and correct rejections (PCL-5 absent/CAPS-5 absent). We calculated the frequency of dichotomous discrepancies (i.e., misses and false alarms), as well as chance-corrected PCL-5/CAPS-5 agreement at the item and diagnostic level using kappa coefficients.

We addressed the second study aim of evaluating potential objective predictors of discrepancy in three steps. First, we calculated correlations between the various objective predictors and dimensional discrepancy scores at the item, symptom cluster, and total score levels. Next, we combined the objective predictors in multiple regression to examine their aggregate ability to predict dimensional discrepancy scores. Third, we combined the objective predictors in logistic regression to examine their aggregate ability to predict dichotomous discrepancies. Given the number of regressions, we again used the Benjamini and Hochberg (1995) correction and reported the f^2 effect size for each prediction (Cohen 1988; f^2 = 0.02–0.14 small, 0.15–0.34 medium, and >0.35 large effect).

To address the third study aim, we conducted a thematic analysis to examine participant attributions for discrepancies, following Braun and Clarke's (2006) five phases of thematic analysis: (a) familiarization with the data and transcription of interview responses; (b) generation of codes for the data; (c) organization of codes into themes; (d) refinement of themes by either removal or combination; (e) finalization and denomination of themes. Overall, this process produces a thematic "map" of the entire dataset.

The first and second authors examined interview transcriptions and generated an initial list of codes with 30 potential themes. They refined this initial list to a final list of 21 themes, using a randomly selected sample of six interviews and consensus discussion. They then created a codebook with definitions and examples to guide final coding of all transcripts (see Supplementary Materials). Because of the range and complexity of responses, each PCL-5/CAPS-5 discrepant symptom was allowed to be coded with up to three different themes. The first author coded all 60 interviews to determine frequency and percentages of themes endorsed for each discrepant symptom between the PCL-5 and the CAPS-5 (see Supplemental Tables 7–9). As a reliability check, the second author coded a randomly selected 20 of the 60 interviews.

Results

Item-level descriptive statistics for the PCL-5 and CAPS-5 from Part 2 of the study are presented in Supplemental Tables 1 and 2.

Aim 1: Degree of Discrepancy Between the PCL-5 and CAPS-5

Correlations and mean differences between PCL-5 and CAPS-5 Intensity and Severity scores are displayed in Table 1. At the item level, the lowest correlations between the PCL-5 and both CAPS-5 Intensity and Severity items were for flashbacks, avoidance of internal and external reminders, hypervigilance, startle, and concentration. Conversely, the highest correlations were for nightmares, amnesia, inability to experience positive emotions, aggressive behavior, and reckless or self-destructive behavior. Paired *t*-tests controlling for Type I error revealed that mean PCL-5 scores were significantly higher than CAPS-5 scores for most items (11 items PCL-5 vs. CAPS-5 Intensity; 14 items PCL-5 vs. CAPS-5 Severity), all clusters, and total score, with variation in effect size from small to very large (Table 1; see Supplemental Tables 3 and 4 for additional information). This was particularly the case for PCL-5 versus CAPS-5 Severity. PCL-5 total scores were more discrepant from CAPS-5 Severity scores (11 points higher) than from CAPS-5 Intensity scores (9 points higher).

At the item level, the frequency of discrepancy occurrence between PCL-5 and CAPS-5 Intensity score and between PCL-5 and CAPS-5 Severity score are displayed in Table 2. For both of these types of discrepancy, 16 of the 20 items were discrepant more than half the time. Results of McNemar tests and paired *t*-tests with Benjamini and Hochberg (1995) correction indicated no significant differences for frequency or mean of any discrepancy between CAPS-5 Intensity and CAPS-5 Severity scores at the item, cluster, or total score level. These analyses indicate the frequencies of any discrepancy and mean discrepancy are highly similar for CAPS-5 Intensity and Severity scores. Therefore, for the remainder of the analyses, we used only CAPS-5 Severity scores.

Examination of item-level dichotomous discrepancies between PCL-5 and CAPS-5 revealed that negative beliefs, negative emotions, detachment or estrangement, and sleep problems had the highest prevalence of misses. Flashbacks, avoidance of external reminders, blame, hypervigilance, and startle had the highest prevalence of false alarms. Chance-corrected overall agreement between dichotomized PCL-5 and CAPS-5 scores ranged from κ = .24 to .64. Items with the lowest PCL-5/CAPS-5 agreement were flashbacks, avoidance of internal reminders, negative emotions, and startle. Items with the highest PCL-5/CAPS-5 agreement were nightmares, blame, inability to experience

Table 1. Correlations and Mean Differences Between PCL-5 and CAPS-5 Intensity and Severity Scores (N = 60).

	PCL-5		CAPS-5 Intensity		PCL-5 vs. CAPS-5 Intensity			CAPS-5 Severity		PCL-5 vs. CAPS-5 Severity		
Item	*M*	*SD*	*M*	*SD*	*r*	*t*	Cohen's *d*	*M*	*SD*	*r*	*t*	Cohen's *d*
Memories	1.78	1.09	1.53	1.02	.55**	1.93	0.25	1.45	0.98	.60**	2.77†	0.36
Dreams	1.27	1.27	1.20	1.23	.74**	0.57	0.07	1.08	1.12	.77**	1.70	0.22
Flashbacks	1.23	1.09	0.62	0.94	.43**	4.38†	0.57	0.48	0.75	.46**	5.80†	0.75
Cued distress	2.58	1.15	1.75	1.00	.50**	5.91†	0.76	1.57	0.85	.56**	8.01†	1.04
Cued physical reactions	2.22	1.21	1.68	1.00	.49**	3.67†	0.47	1.60	0.96	.57**	4.65†	0.60
Internal avoidance	2.80	1.18	1.92	1.01	.40**	5.67†	0.73	1.82	1.00	.34**	6.07†	0.78
External avoidance	2.48	1.32	1.27	1.19	.44**	7.09†	0.92	1.18	1.13	.47**	7.95†	1.03
Amnesia	1.63	1.54	0.90	1.30	.61**	4.46†	0.58	0.88	1.28	.62**	4.67†	0.60
Negative beliefs	1.70	1.32	1.72	1.33	.54**	−0.10	−0.01	1.65	1.29	.52**	0.30	0.04
Blame	2.13	1.48	1.43	1.33	.51**	3.89†	0.50	1.30	1.24	.55**	4.95†	0.64
Negative emotions	2.00	1.25	1.73	0.97	.48**	1.77	0.23	1.57	0.91	.54**	3.11†	0.40
Loss of interest	1.13	1.23	0.87	1.26	.47**	1.62	0.21	0.73	1.12	.51**	2.65†	0.34
Detachment	1.22	1.11	1.27	1.16	.47**	−0.33	−0.04	1.22	1.12	.51**	0.00	0.00
Anhedonia	1.13	1.19	0.92	1.06	.62**	1.69	0.22	0.88	1.04	.60**	1.93	0.25
Aggressive behavior	0.83	1.06	0.77	1.06	.64**	0.57	0.07	0.72	1.01	.60**	0.98	0.13
Reckless behavior	0.52	1.05	0.27	0.78	.76**	2.85†	0.37	0.23	0.67	.81**	3.43†	0.44
Hypervigilance	2.30	1.25	1.65	1.12	.43**	3.95†	0.51	1.53	1.08	.51**	5.08†	0.66
Startle	1.72	1.28	0.90	1.12	.41**	4.84†	0.62	0.83	1.03	.47**	5.66†	0.73
Concentration	1.72	1.28	1.07	1.15	.38**	3.73†	0.48	1.05	1.14	.38**	3.83†	0.49
Sleep	1.65	1.40	1.47	1.28	.58**	1.16	0.15	1.35	1.19	.59**	1.97	0.25
RE	9.08	4.58	6.78	3.65	.67**	5.17†	0.67	6.18	3.28	.76**	7.53†	0.97
AV	5.28	2.20	3.18	1.75	.52**	8.21†	1.06	3.00	1.67	.51**	9.00†	1.16
NACM	10.95	6.42	8.83	5.75	.63**	3.12†	0.40	8.23	5.43	.65**	4.17†	0.54
AR	8.73	4.64	6.12	3.82	.57**	5.07†	0.65	5.72	3.62	.60**	6.11†	0.79
Total score	34.05	14.34	24.92	12.22	.63**	6.10†	0.79	23.13	11.36	.69**	8.05†	1.04

Note. PCL-5 = PTSD Checklist for *DSM*-5; CAPS-5 = Clinician-Administered PTSD Scale for *DSM*-5; RE = Reexperiencing; AV = Avoidance; NACM = Negative Alterations in Cognition and Mood; AR = Arousal.

*$p < .05$; **$p < .001$; †Statistically significant *t*-test after applying Benjamini and Hochberg's (1995) procedure to hold the false discovery rate at $\alpha = .05$.

Table 2. Frequency and Mean Size of Discrepancy Between PCL-5 and CAPS-5 Intensity and Severity Scores (N = 60).

	CAPS-5 Intensity				CAPS-5 Severity			
Item	Frequency	%	*M*	*SD*	Frequency	%	*M*	*SD*
Memories	36	60.00	0.75	0.70	34	56.70	0.70	0.70
Dreams	28	46.70	0.57	0.70	26	43.30	0.52	0.68
Flashbacks	36	60.00	0.89	0.88	38	63.30	0.92	0.85
Cued distress	40	66.70	1.00	0.94	44	73.30	1.08	0.91
Cued physical reactions	44	73.30	0.97	0.78	43	71.70	0.92	0.77
Internal avoidance	41	68.30	1.12	0.99	43	71.70	1.22	1.03
External avoidance	40	66.70	1.32	1.23	41	68.30	1.33	1.23
Amnesia	31	55.00	0.93	1.13	33	55.00	0.92	1.12
Negative beliefs	34	56.70	0.88	0.90	36	60.00	0.92	0.89
Blame	42	70.00	1.17	1.03	42	70.00	1.17	1.01
Negative emotions	43	71.70	0.93	0.73	41	68.30	0.90	0.73
Loss of interest	31	51.70	0.83	0.99	35	58.30	0.83	0.91
Detachment	35	58.30	0.82	0.83	34	56.70	0.77	0.79
Anhedonia	26	43.30	0.58	0.83	25	41.70	0.58	0.85
Aggressive behavior	24	40.00	0.53	0.72	27	45.00	0.58	0.72
Reckless behavior	16	26.70	0.35	0.63	17	28.30	0.35	0.61
Hypervigilance	41	68.30	1.05	0.96	39	65.00	1.00	0.97
Startle	39	65.00	1.08	1.09	39	65.00	1.05	1.06
Concentration	40	66.70	1.08	1.03	41	68.30	1.10	1.02
Sleep	33	55.00	0.82	0.93	32	53.30	0.80	0.92
RE	58	96.70	4.17	2.18	58	96.70	4.13	2.24
AV	53	88.30	2.43	1.65	55	91.70	2.50	1.70
NACM	60	100.00	6.15	3.25	58	96.70	6.08	3.28
AR	58	96.70	4.92	2.67	59	98.30	4.88	2.57
Total Score	60	100.00	17.67	6.93	60	100.00	17.65	6.99

Note. PCL-5 = PTSD Checklist for *DSM-5*; CAPS-5 = Clinician-Administered PTSD Scale for *DSM-5;* RE = Reexperiencing; AV = Avoidance; NACM = Negative Alterations in Cognition and Mood; AR = Arousal.

positive emotions, and reckless behavior (see Table 3 for additional information).

Aim 2: Multiple Regression and Logistic Regression Analyses

There were few significant correlations between the various objective predictors and dimensional discrepancy scores for items and clusters (see Table 4). Results of the multiple regression analyses controlled for Type I error revealed the overall model was only significant for two PTSD symptoms. There was a large, significant overall effect for irritability and aggressive behavior, for which significant predictors were INF and Neuroticism. There was a large, significant overall effect for concentration, for which the significant predictors were Neuroticism and Verbal IQ. At the cluster level, the overall model demonstrated a large, significant effect for the AR cluster, for which the only significant predictor was Neuroticism. At the total discrepancy score level, the overall model was not significant. See Supplemental Table 5 for descriptive statistics for the objective predictors and Supplemental Table 6 for the multiple regression results.

Subsequently, the objective predictors were combined in logistic regression to examine their aggregate ability to predict dichotomous discrepancies at the item and diagnosis level (i.e., misses and false alarms). The overall model was not significant, and none of the predictors significantly predicted dichotomized discrepant items or diagnostic discrepancy score.

Aim 3: Qualitative Analyses

Following Braun and Clarke (2006), the first and second authors identified 21 themes of discrepancies, which were coded for frequency and percentage for each discrepant symptom between the PCL-5 and the CAPS-5 (see Supplemental Tables 7–9). We identified the five most frequently appearing themes: comprehension of symptoms, general errors, increased awareness, time-frame reminders, and trauma-related attribution error (see Table 5). The theme of comprehension of symptoms was coded if the participant

Table 3. Prevalence of PTSD Symptoms and Diagnosis Based on PCL-5 and CAPS-5 (N = 60).

	PCL-5 prevalence		CAPS-5 prevalence		Hits		Misses		False alarms		Correct rejections		Efficiency		
Item	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	κ (.5)
Memories	35	58.3	35	58.3	27	45.0	8	13.3	8	13.3	17	28.3	44	73.3	.45
Dreams	23	38.3	26	43.3	18	30.0	8	13.3	5	8.3	29	48.3	47	78.3	.56
Flashbacks	25	41.7	9	15.0	7	11.7	2	3.3	18	30.0	33	55.0	40	66.7	.25
Cued distress	49	81.7	38	63.3	37	61.7	1	1.7	12	20.0	10	16.7	47	78.3	.45
Cued physical reactions	40	66.7	40	66.7	33	55.0	7	11.7	7	11.7	13	21.7	46	76.7	.48
Internal avoidance	51	85.0	41	68.3	38	63.3	3	5.0	13	21.7	6	10.0	44	73.3	.25
External avoidance	47	78.3	28	46.7	27	45.0	1	1.7	20	33.3	12	20.0	39	65.0	.37
Amnesia	27	45.0	21	35.0	16	26.7	5	8.3	11	18.3	28	46.7	44	73.3	.45
Negative beliefs	31	51.7	36	60.0	24	40.0	12	20.0	7	11.7	17	28.3	41	68.3	.36
Blame	40	66.7	28	46.7	23	38.3	5	8.3	17	28.3	15	25.0	38	63.3	.57
Negative emotions	39	65.0	37	61.7	28	46.7	9	15.0	11	18.3	12	20.0	40	66.7	.28
Loss of interest	19	31.7	14	23.3	9	15.0	5	8.3	10	16.7	36	60.0	45	75.0	.38
Detachment	24	40.0	26	43.3	15	25.0	11	18.3	9	15.0	25	41.7	40	66.7	.32
Anhedonia	21	35.0	21	35.0	16	26.7	5	8.3	5	8.3	34	56.7	50	83.3	.63
Aggressive behavior	15	25.0	15	25.0	9	15.0	6	10.0	6	10.0	39	65.0	48	80.0	.47
Reckless behavior	6	10.0	4	6.7	3	5.0	1	1.7	3	5.0	53	88.3	56	93.3	.64
Hypervigilance	43	71.7	36	60.0	32	53.5	4	6.7	11	18.3	13	21.7	45	75.0	.45
Startle	33	55.0	21	35.0	18	30.0	3	5.0	15	25.0	24	40.0	42	70.0	.39
Concentration	32	53.3	24	40.0	17	28.3	7	11.7	15	25.0	21	35.0	38	63.3	.24
Sleep	29	48.3	33	55.0	24	40.0	9	15.0	5	8.3	22	36.7	46	76.7	.54
PTSD diagnosis	31	51.7	31	51.7	22	36.7	9	15.0	9	15.0	20	33.3	42	70.0	.37

Note. Symptoms and diagnosis are based on PCL-5 scores and CAPS-5 Severity scores dichotomized at 2 or above. Analyses are based on PCL-5 as the test and CAPS-5 as the criterion. Hits = PCL-5 Present/CAPS-5 Present; Misses = PCL-5 Absent/CAPS-5 Present; False Alarms = PCL-5 Present/CAPS-5 Absent; Correct Rejections = PCL-5 Absent/CAPS-5 Absent; Efficiency = Hits + Correct Rejections/Total; κ (.5) = quality of efficiency.

Table 4. Zero-Order Correlations for PCL-5 and CAPS-5 With Objective Predictor Variables (N = 60).

Item	PAI NIM	PAI INC	PAI INF	BFI CON	BFI NEURO	Verbal IQ
Memories	−.03	−.10	.07	.11	−.03	.12
Dreams	.14	.01	−.06	.05	.20	−.13
Flashbacks	−.13	−.08	.07	−.10	−.07	−.07
Cued distress	.01	.07	−.12	−.16	.11	−.05
Cued physical reactions	.42**	−.06	−.02	−.16	.04	−.25
Internal avoidance	−.15	.20	−.10	−.05	.17	.06
External avoidance	.17	−.03	−.06	−.15	.14	.11
Amnesia	.08	.13	−.17	−.14	.18	−.03
Negative beliefs	.02	.19	.11	.14	.00	.09
Blame	−.01	.13	.09	−.20	.11	−.03
Negative emotions	.03	.07	.25	−.07	.07	.14
Loss of interest	.03	.11	−.09	.15	.20	.10
Detachment	−.05	.14	−.06	.22	.21	−.03
Anhedonia	.27*	.02	.19	.04	.07	−.14
Aggressive behavior	−.02	.16	.35**	.11	.24	−.19
Reckless behavior	.36*	.11	−.08	−.15	.26*	−.10
Hypervigilance	.34*	.10	.11	−.29*	.19	−.08
Startle	.19	−.15	−.14	.03	.21	−.01
Concentration	.10	.14	.13	−.29*	.37**	.34**
Sleep	.08	.21	.08	−.32*	.27*	.10
RE	.13	−.06	−.03	−.07	.08	−.13
AV	.03	.10	−.10	−.14	.20	.12
NACM	.11	.22	.07	.02	.24	.02
AR	.36**	.18	.14	.33*	.52**	.07
Total Score	.23	.17	.05	−.17	.38*	.02

Note. RE = Reexperiencing AV = Avoidance; NACM = Negative Alterations in Cognition and Mood; AR = Arousal; PAI = Personality Assessment Inventory; NIM = Negative Impression Management; INC = Inconsistency; INF = Infrequency; BFI = Big Five Inventory; CON = Conscientiousness; NEURO = Neuroticism; Verbal IQ = Shipley–2.
*$p < .05$; **$p < .01$.

reported that they misunderstood or misinterpreted the question (e.g., "I did not understand the difference between intrusive memories and flashbacks. I answered it wrong because I was confused about what the question was asking for."). This theme was frequently reported for memories and flashback symptoms where participants self-endorsed more symptoms than experienced. The theme of general errors was coded if the participant reported making general errors such as not paying attention to the question, selecting the wrong answer, thinking about frequency, or forgetting a relevant experience. (e.g., "I don't remember marking that answer. I must not have read the question and I marked incorrectly."). This theme was frequently reported for cued distress, avoidance of internal and external reminders symptoms, with participants over-endorsing their symptoms on the PCL-5.

The theme of increased awareness was coded if the participant discussed the overall internal process that occurred while answering the assessment measure (e.g., "Having the time to think and talk about my symptoms helped me remember what I have been struggling with this past month."). This theme was frequently reported for memories, avoidance of internal and external reminders, and blame symptoms resulting in over-endorsement on the PCL-5. The theme of time-frame reminders was coded if the participant explained that they were not answering based on the past month or that they wanted reminders of the specific time frame (e.g., "When I answered the question, I was thinking about my overall distress. I did not remember to only answer based on the past month."). This theme was frequently reported for detachment and estrangement, startle, and concentration symptoms, which resulted in both under- and over-endorsement of symptoms on the PCL-5. Finally, the theme of trauma-related attribution error was coded if the participant interpreted their symptoms as an internal characteristic instead of recognizing the contribution of the traumatic event or misattributed symptoms to the trauma that were unrelated (e.g., "I have always had problems with sleeping, so I didn't realize the trauma had progressively changed my sleeping habits."). This theme was frequently reported for

Table 5. Frequencies of the Most Frequent Themes of Discrepancies Between the PCL-5 and the CAPS-5 by Item.

	Sx Comp		Gen Error		Increased awareness		Time Frame		TR Error		Sx Min		TR remind		Elaboration		Item content	
Items	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
B1	12	10.1	6	5.4	13	12.1	4	4.9	0	0	1	1.6	1	2.1	7	16.3	1	2.3
B2	5	4.2	5	4.5	4	3.4	1	1.2	0	0	3	4.8	0	0	1	2.3	0	0
B3	14	11.8	7	6.3	4	3.4	4	4.9	3	4.3	2	3.2	1	2.1	4	9.3	0	0
B4	7	5.9	8	7.1	6	5.6	4	4.9	3	4.3	5	7.9	1	2.1	5	11.6	2	4.1
B5	9	7.6	10	8.9	6	5.6	2	2.5	4	5.8	6	9.5	1	2.1	1	2.3	1	2.3
C1	5	4.2	11	9.8	8	7.5	5	6.2	3	4.3	6	9.5	0	0	2	4.1	2	4.1
C2	4	3.4	9	8.0	8	7.5	5	6.2	3	4.3	5	7.9	1	2.1	1	2.3	3	7.0
D1	5	4.2	5	4.5	4	3.4	3	3.7	0	0	3	4.8	0	0	2	4.1	5	11.6
D2	7	5.9	3	2.7	6	5.6	3	3.7	3	4.3	4	6.3	0	0	2	4.1	1	2.3
D3	4	3.4	7	6.3	10	9.3	4	4.9	0	0	4	6.3	1	2.1	1	2.3	14	32.6
D4	7	5.9	2	1.8	7	6.5	4	4.9	3	4.3	5	7.9	5	10.6	3	7.0	0	0
D5	6	5.0	3	2.7	3	2.8	5	6.2	4	5.8	4	6.3	2	4.3	3	7.0	0	0
D6	3	2.5	2	1.8	7	9.3	9	11.1	3	4.3	3	4.8	4	8.5	0	0	1	2.3
D7	4	3.4	3	2.7	3	2.8	5	6.2	3	4.3	1	1.6	3	6.4	0	0	1	2.3
E1	5	4.2	2	1.8	3	2.8	2	2.5	0	0	2	3.2	3	6.4	3	7.0	5	11.6
E2	4	3.4	4	3.6	2	1.9	2	2.5	2	2.9	0	0	0	0	2	4.1	0	0
E3	3	2.5	3	2.7	7	6.5	1	1.2	8	11.6	2	3.2	0	0	1	2.3	4	9.3
E4	5	4.2	8	7.1	3	2.8	6	7.4	5	7.2	0	0	6	12.8	3	7.0	1	2.3
E5	2	1.7	7	6.3	8	7.5	6	7.4	6	8.7	2	3.2	8	17.0	1	2.3	0	0
E6	4	3.4	3	2.7	2	1.9	3	3.7	10	14.5	2	3.2	9	19.1	0	0	1	2.3
Total freq.	119	—	112	—	107	—	81	—	69	—	63	—	47	—	43	—	43	—

Note. Sx Comp = Comprehension of Symptoms; Gen Error = General Errors; Time Frame = Time-frame Reminders; TR Error = Trauma-Related Attribution Error; Sx Min = Minimization of Symptoms; TR Remind = Trauma-Related Reminders; Elaboration = Opportunity for Elaboration; B1 = Memories; B2 = Dreams; B3 = Flashbacks; B4 = Cued distress; B5 = Cued physical reactions; C1 = Internal avoidance; C2 = External avoidance; D1 = Amnesia; D2 = Negative Beliefs; D3 = Blame; D4 = Negative emotions; D5 = Loss of interest; D6 = Detachment; D7 = Anhedonia; E1 = Aggressive behavior; E2 = Reckless behavior; E3 = Hypervigilance; E4 = Startle; E5 = Concentration; E6 = Sleep.

hypervigilance, concentration, and sleep symptoms, which resulted in both under- and over-endorsement of symptoms on the PCL-5.

Discussion

Self-rated PTSD questionnaires, such as the PCL-5, have consistently demonstrated strong psychometric properties (Blevins et al., 2015; Bovin et al., 2016; Wortmann et al., 2016). However, although questionnaires allow for efficient administration and scoring, researchers and clinicians have questioned the accuracy of self-rated measures compared with clinician-rated structured interviews. The present study is the first to empirically examine quantitative and qualitative reasons for discrepancies between a self-rated measure (PCL-5) and a clinician-rated measure (CAPS-5) in a trauma-exposed sample.

Consistent with our first aim, we identified a moderate degree of discrepancy between the PCL-5 and the CAPS-5. Similar to previous studies (e.g., Blanchard et al., 1996; Forbes et al., 2001), the PCL-5 and CAPS-5 scores at the item, cluster, and total score level were significantly and positively correlated, ranging from weak to very strong. It is possible that these correlations were attenuated due to restriction of range (i.e., due to selecting participants with at least moderate PCL-5 scores), but the wide range of PCL-5/CAPS-5 correlations provides evidence of discordance between the two measures. Providing further evidence of discordance, we found that mean PCL-5 scores were significantly higher than mean CAPS-5 scores for most items, cluster scores, and total score. While the frequency of discrepancy did not vary when comparing PCL-5 with CAPS-5 Intensity and CAPS-5 Severity, results indicated that the magnitude of PCL-5/CAPS-5 score discrepancy was higher for CAPS-5 Severity scores than CAPS-5 Intensity scores.

Next, item-level discrepancies occurred on more than half of the PTSD symptoms, the most discrepant being for symptoms of hypervigilance, exaggerated startle, internal avoidance, and external avoidance. This is consistent with Moshier and colleagues' (2018) network analysis that suggested participants may have had more difficulty responding differentially to similar symptoms (i.e., hypervigilance and exaggerated startle), whereas clinicians were able to distinguish between them. Avoidance symptoms may also be particularly difficult to rate accurately as participants may be less aware that they are engaging in avoidance behavior. In support of our hypothesis that covert symptoms may be more difficult to rate consistently, we identified PCL-5/CAPS-5 discrepancies among symptoms of cued distress, cued physical reactions, loss of interest, and blame.

When we examined dichotomous PCL-5/CAPS-5 discrepancies (i.e., misses, false alarms), we found considerable discrepancy between symptoms endorsed on PCL-5 versus CAPS-5. The most frequent misses (i.e., PCL-5 absent/CAPS-5 present) were for negative beliefs, negative emotions, detachment or estrangement, and sleep problems. One potential explanation is that respondents could have misunderstood the PCL-5 items. For example, for the negative belief item, respondents often do not understand the question on the CAPS-5 until they are prompted with specific examples. The most frequent false alarms (i.e., PCL-5 present/CAPS-5 absent) were for flashbacks, avoidance of external reminders, blame, hypervigilance, and startle. False alarms could suggest that respondents interpret normative experiences as clinically significant problems and endorse those items more readily on the PCL-5. For example, for the flashback item, this term is used broadly in popular culture to suggest a wide range of reexperiencing symptoms, and not just the more specific meaning of dissociative reexperiencing that is assessed on the CAPS-5. Items with the lowest PCL-5/CAPS-5 agreement were flashbacks, avoidance of internal reminders, negative emotions, and startle, while items with the highest agreement were nightmares, self-blame, inability to experience positive emotions, and reckless behavior. Some of the more discrepant symptoms, such as flashbacks, avoidance of internal reminders, and startle, are similar to previous studies' findings (e.g., Blanchard et al., 1996; Forbes et al., 2001).

Overall, four analytical methods of examining discrepancies revealed that influential PCL-5/CAPS-5 discrepancies, both dimensional and dichotomous, exist at the item, cluster, and total score levels in our sample. Against the CAPS-5 as the criterion, the PCL-5 generally had moderately high efficiencies for the presence of individual PTSD symptoms and PTSD diagnosis. However, only three items had efficiencies of 80% or higher, and the corresponding kappa coefficients were generally low to moderate. This suggests that although the PCL-5 and CAPS-5 overlap, they are not interchangeable for determining the presence of individual PTSD symptoms and PTSD diagnosis. While PCL-5 total score appears to be a useful proxy of likely PTSD diagnosis, the substantial item-level discrepancies identified highlight concerns about the diagnostic utility of individual items (e.g., Blanchard et al., 1996; Forbes et al., 2001) given the importance of accurately identifying present symptoms for clinical diagnosis and treatment decisions. These results should be interpreted also within the context that some discrepancy is expected given that the PCL-5 and CAPS-5 are designed for different purposes (e.g., screening probable PTSD vs. diagnosis). Consistent with Hopwood and colleagues' (2008) suggestion that symptoms prone to misinterpretation or requiring greater clinical inference are more amenable to assessment via interview method, we found that discrepancies were more frequent for complicated or nuanced PTSD symptoms.

Given anecdotal and theoretical explanations for discordance between self-rated and clinician-rated PTSD measures, we also empirically examined possible quantitative and qualitative reasons behind discrepancies. An exploratory multiple regression of six potential objective predictors of discrepancy (i.e., response validity indicators of negative impression management, inconsistency, infrequency; personality traits of conscientiousness and neuroticism; and verbal IQ) indicated that Neuroticism was the most predictive of discrepancy. Neuroticism predicted discrepant dimensional scores for symptoms of irritability and aggressive behavior and concentration problems as well as for the arousal and reactivity cluster. These results suggest that individuals reporting more neuroticism commonly experience unpleasant emotions (Raynor & Levine, 2009); tend to assess banal, everyday situations as threatening (Ebstrup et al., 2011); and therefore, may appraise and self-report certain symptoms as more severe than others, leading to greater discrepancies between their PCL-5 ratings and a clinician's ratings on the CAPS-5.

Other predictors of item-level dimensional discrepancy included the response validity indicator of INF, specifically for the symptom of irritability or aggressive behavior. This result suggests that a symptom with a more external manifestation could be influenced by response bias, consistent with concerns of random responding on symptom assessment (Weathers & Keane, 1999). We found that Verbal IQ was only a significant predictor of item-level dimensional discrepancy for concentration problems (associated with higher IQ score). Given the results of the thematic analysis which indicated many discrepancies for the concentration symptom were due to trauma-relatedness, it is possible that highly intelligent participants more readily noticed and endorsed problems in concentration on the PCL-5 that could have been due to various sources (e.g., academic stressors, general distress, or anxiety).

Response validity indicators, personality traits, and Verbal IQ did not significantly predict dichotomized discrepant items and diagnostic discrepancy scores. Therefore, none of the objective predictors adequately explained differences in whether participants and clinicians believed participants had a threshold symptom or diagnosis. Combined with the multiple regression results, this indicated that the objective predictors examined were more useful predictors of the magnitude of dimensional discrepancy (i.e., differences in self-rated vs. clinician-rated severity).

Our thematic analysis of participant attributions for PCL-5/CAPS-5 discrepancies identified 21 reasons for score discrepancies, the most frequently reported being comprehension of symptoms, general errors, increased awareness, time-frame reminders, and trauma-related attribution errors. In the interviews, participants often remarked that these discrepant responses were not intentionally inaccurate or dishonest. The most significant reason given for discrepancy was that participants struggled to comprehend

some of the symptoms when answering the PCL-5, such as differentiating between intrusive memories and flashbacks. This is consistent with prior literature on participants' overlapping experiences of vivid recollections and dissociative states (Marmar et al., 2015; Moshier et al., 2018) and the presence of psychopathology that affects an individual's ability to reflect on their symptoms (e.g., Enns et al., 2000). Our results also suggest that PTSD symptoms are possibly too conceptually complex for the respondent's understanding of the questions to reliably match what the researcher had in mind. This finding prompts the need to reword more conceptually complex symptoms (i.e., symptoms in the reexperiencing cluster) on self-report measures like the PCL-5 for easier respondent comprehension.

The second most prevalent reason for discrepancies was general errors such as participants reporting they forgot relevant experiences, selected the wrong answer, or did not read/hear the entire question. When completing the PCL-5, participants reported making judgments about the severity of their symptoms based on the frequency they occurred. This is problematic because the PCL-5 does not explicitly include frequency like the CAPS-5, and frequency alone is not a substitute for the intensity of distress the symptom causes (e.g., startling from noises daily, but without a reaction significant enough for anyone to notice or that causes impairment). Judgments about frequency are also reliant on sufficient memory and participant insight about symptoms and motivation to respond accurately. This finding supports the use of clinician-administered measures that allow for follow-up prompts, allow the participant to explain their distress, time-frame reminders, and separate coding of severity from frequency to achieve more accurate ratings.

Another pervasive theme was the ability to have more time to reflect on symptoms during a structured interview. Participants reported the more detailed exploration of their symptoms through discussion with a trained clinician provided opportunities to process the source of their distress and clarify their thoughts. Consistent with our quantitative findings and documented self-report difficulties (Enns et al., 2000), the lack of opportunity for this during questionnaire completion may increase the difficulty of accurately self-assessing and reporting more emotionally charged and complex symptoms of PTSD (i.e., memories, avoidance of internal and external reminders, and blame).

The last two reasons given for discrepancies are related to the influence of general distress on symptom reporting, which is a well-known theorized pitfall of self-rated assessment (Marmar et al., 2015) compared with structured interviews which tend to focus on the nature and pervasiveness of behaviors (Moshier et al., 2018). Participants reported their answers were often more based on their general distress, rather than reporting only the occurrence of symptoms from the past month and reporting symptoms only related to the traumatic event. Our qualitative findings indicated participants struggled to separate their PTSD-related distress from everyday struggles (e.g., concentration and sleep difficulties due to school, family, or job stressors in their overall subjective distress rating) and across time (e.g., including memorable symptoms from 6 months ago in their overall subjective distress rating). These themes became more frequent for the PTSD symptoms in the latter half of the assessment, suggesting that lengthier questionnaires may prompt more participant reliance on their non-specific distress. Therefore, the accuracy of PTSD diagnosis is improved by features of the CAPS-5 such as specific time-frame and trauma-relatedness prompts, behavioral examples of symptoms, and opportunities for trained clinicians to help respondents differentiate distress due to internal characteristics and other stressors from distress due to the traumatic event.

Given the current findings, clinicians who desire maximally interpretable PCL-5 scores may want to provide increased instruction to clients prior to their first completion of the PCL-5, such as reminding clients to read each question carefully before responding and that opportunities for clarification with the clinician are available. To prevent misunderstandings, clinicians may find it helpful to provide preemptive cues to the trauma-relatedness, distress, and time-frame components. To encourage attentive and specific responding, clinicians could also discuss the purpose of the PCL-5 with clients beforehand in greater detail.

Limitations of the present study include the use of a convenience sample of college students with limited diversity of age, race, and socioeconomic status and a cross-sectional research design without counterbalancing the administration of the CAPS-5 and PCL-5. We attempted to increase the generalizability of the results to clinical samples by including only trauma-exposed participants with moderate to severe PTSD symptoms. However, future studies should aim to replicate and extend our findings to more diverse and clinical samples. In addition, it is recommended that future studies collect information about whether participants have completed and/or are currently undergoing trauma-focused treatment. Examining treatment involvement as a predictor variable could potentially provide valuable information about whether individuals are more likely to have a lower number of self-reported versus clinician-rated discrepancies.

An important next step is to examine the potential reasons for discrepancies provided by participant qualitative responses. For example, future research studies could administer the PCL-5 with modifications such as more frequent time-frame reminders, inattentive responding and validity indices, or requirement of written examples of their symptoms that provide further context about participant insight, attention to time frame and trauma-relatedness, understanding of the item content, and potential sources of error. Another future direction could involve examining discrepancies between the PCL-5 and the CAPS-5 for individuals endorsing sub-threshold levels of PTSD. Finally, future

research should examine if assessment order has an effect on quantitative and qualitative PCL-5/CAPS-5 discrepancies.

This is the first known study to quantify and empirically examine sources of PCL-5/CAPS-5 discrepancies that may influence PTSD assessment and diagnosis. We presented evidence that the PCL-5 and the CAPS-5 are strongly associated, moderately concordant, but not interchangeable. Discrepancies were evident across individual symptoms, cluster scores, total score, and diagnosis, most often implicating symptoms that were conceptually complex, difficult to differentiate from one another, emotionally salient, or overlapping with general symptoms related to non-traumatic stress. Initial results indicate that elevated PCL-5 scores may be due to the influence of neuroticism, response bias, verbal IQ, self-rating errors, lack of opportunity to clarify symptoms, and difficulty differentiating one's experience of PTSD from general distress.

These findings are consistent with research showing that although the PCL-5 is an efficient, inexpensive screening tool comparable in some respects with the CAPS-5 (Lee et al., 2019), it yields higher prevalence estimates and severity ratings (Marmar et al., 2015) and is differentially sensitive to clinical change (see Forbes et al., 2001; Monson et al., 2008). The present study shows the PCL-5 is limited in representing the complexity of the PTSD symptom criteria, and the CAPS-5 is a favorable diagnostic tool because it allows for processing of respondents' symptoms, evaluating self-reported examples for conceptual fit, and using clinical judgment to make final ratings. The present study adds to our current knowledge of PCL-5/CAPS-5 discordance and suggests several ways to improve the measures' diagnostic and clinical utility.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Lindsay B. Kramer https://orcid.org/0000-0002-3881-0741
Sarah E. Whiteman https://orcid.org/0000-0003-1538-3991
Jessica M. Petri https://orcid.org/0000-0001-8755-5487

Supplemental Material

Supplemental material for this article is available online.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).

Barsky, A. I., & Klerman, G. L. (1983). Overview: Hypochondriasis, bodily complaints. and somatic styles. *American Journal of Psychiatry*, *140*, 273–283. https://doi.org/10.1176/ajp.140.3.273

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series A*, *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berlant, J., & van Kammen, D. P. (2002). Open-label topiramate as primary or adjunctive therapy in chronic civilian posttraumatic stress disorder: A preliminary report. *The Journal of Clinical Psychiatry*, *63*, 15–20. https://doi.org/10.4088/jcp.v63n0104

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Klauminzer, G., Charney, D. S., & Keane, T. M. (1990). A clinician rating scale for assessing current and lifetime PTSD: The CAPS-1. *The Behavior Therapist*, *13*, 187–188. https://doi.org/10.1037/t00072-000

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. https://doi.org/10.1016/0005-7967(96)00033-2

Blevins, C. A., Weathers, F. W., Davis, M. T., Witte, T. K., & Domino, J. L. (2015). The posttraumatic stress disorder checklist for *DSM-5* (PCL-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress*, *28*, 489–498. https://doi.org/10.1002/jts.22059

Bovin, M. J., Marx, B. P., Weathers, F. W., Gallagher, M. W., Rodriguez, P., Schnurr, P. P., & Keane, T. M. (2016). Psychometric properties of the PTSD Checklist for *Diagnostic and Statistical Manual of Mental Disorders–Fifth Edition* (PCL-5) in veterans. *Psychological Assessment*, *28*, 1379–1391. https://doi.org/10.1037/pas0000504

Bovin, M. J., & Weathers, F. W. (2022). Assessing PTSD symptoms. In J. G. Beck & D. M. Sloan (Eds.), *Handbook of traumatic stress disorders* (2nd ed., pp. 466–491). Oxford University Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*, 77–101. https://doi.org/10.1191%2F1478088706qp063oa

Cohen. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Costa, P. T., & McCrae, R. R. (1985). Hypochondriasis, neuroticism, and aging: When are somatic complaints unfounded? *American Psychologist*, *40*, 19–28. https://doi.org/10.1037/0003-066x.40.1.19

Ebstrup, J. F., Eplov, L. F., Pisinger, C., & Jørgensen, T. (2011). Association between the Five Factor personality traits and perceived stress: Is the effect mediated by general self-efficacy? *Anxiety, Stress & Coping*, *24*, 407–419. https://doi.org/10.1080/10615806.2010.540012

Elhai, J. D., Gray, M. J., Kashdan, T. B., & Franklin, C. L. (2005). Which instruments are most commonly used to assess traumatic event exposure and posttraumatic effects? A survey of traumatic stress professionals. *Journal of Traumatic Stress*, *18*, 541–545. https://doi.org/10.1002/jts.20062

Enns, M. W., Larsen, D. K., & Cox, B. J. (2000). Discrepancies between self and observer ratings of depression: The relation-

ship to demographic, clinical and personality variables. *Journal of Affective Disorders*, *60*, 33–41. https://doi.org/10.1016/s0165-0327(99)00156-1

Erbes, C., Westermeyer, J., Engdahl, B., & Johnsen, E. (2007). Post-Traumatic Stress Disorder and service utilization in a sample of service members from Iraq and Afghanistan. *Military Medicine*, *172*, 359–363. https://doi.org/10.7205/milmed.172.4.359

Foa, E. B., McLean, C. P., Zang, Y., Zhong, J., Rauch, S., Porter, K., Knowles, K., Powers, M. B., & Kauffman, B. Y. (2016). Psychometric properties of the Posttraumatic Stress Disorder Symptom Scale Interview for DSM-5 (PSSI-5). *Psychological Assessment*, *28*, 1159–1165. https://doi.org/10.1037/pas0000259

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. https://doi.org/10.1016/s0005-7967(00)00084-x

Gray, M. J., Litz, B. T., Hsu, J. L., & Lombardo, T. W. (2004). Psychometric properties of the life events checklist. *Assessment*, *11*, 330–341. https://doi.org/10.1177/1073191104269954

Hoge, C. W., McGurk, D., Thomas, J. L., Cox, A. L., Engel, C. C., & Castro, C. A. (2008). Mild traumatic brain injury in U.S. soldiers returning from Iraq. *The New England Journal of Medicine*, *358*, 453–463. https://doi.org/10.1056/nejmc086083

Hopwood, C. J., Morey, L. C., Grilo, C. M., Sanislow, C. A., McGlashan, T. H., Markowitz, J. C., & Skodol, A. E. (2008). A comparison of interview and self-report methods for the assessment of borderline personality disorder criteria. *Psychological Assessment*, *20*, 81–85. https://doi.org/10.1037/1040-3590.20.1.81

Ibrahim, H., Ertl, V., Catani, C., Ismail, A. A., & Neuner, F. (2018). The validity of the Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5) as screening instrument with Kurdish and Arab displaced populations living in the Kurdistan region of Iraq. *BMC Psychiatry*, *18*, 259. https://doi.org/10.1186/s12888-018-1839-z

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). Guilford Press.

Lee, D. J., Bovin, M. J., Weathers, F. W., Palmieri, P. A., Schnurr, P. P., Sloan, D. M., Keane, T. M., & Marx, B. P. (2019). Latent factor structure of *DSM-5* posttraumatic stress disorder: Evaluation of method variance and construct validity of novel symptom clusters. *Psychological Assessment*, *31*, 46–58. https://doi.org/10.1037/pas0000642

Marmar, C. R., Schlenger, W., Henn-Haase, C., Qian, M., Purchia, E., Li, M., Corry, N., Williams, C. S., Ho, C.-L., Horesh, D., Karstoft, K. I., Shalev, A., & Kulka, R. A. (2015). Course of posttraumatic stress disorder 40 years after the Vietnam War: Findings from the National Vietnam Veterans Longitudinal Study. *Journal of the American Medical Association Psychiatry*, *72*, 875–881. https://doi.org/10.1001/jamapsychiatry.2015.0803

McDonald, S., & Calhoun, P. (2010). The diagnostic accuracy of the PTSD Checklist: A critical review. *Clinical Psychology Review*, *30*, 976–987. https://doi.org/10.1016/j.cpr.2010.06.012

Monson, C., Gradus, J., Young-Xu, Y., Schnurr, P., Price, J., & Schumm, J. (2008). Change in posttraumatic stress disorder symptoms: Do clinicians and patients agree? *Psychological Assessment*, *20*, 131–138. https://doi.org/10.1037/e558112009-009

Morey, L. C. (1991). *The Personality Assessment Inventory Professional Manual*. Psychological Assessment Resources.

Morey, L. C. (2007). *The personality assessment inventory professional manual*. Psychological Assessment Resources.

Morrison, K., Su, S., Keck, M., & Beidel, D. C. (2021). Psychometric properties of the PCL-5 in a sample of first responders. *Journal of Anxiety Disorders*, *77*, 102339. https://doi.org/10.1016/j.janxdis.2020.102339

Moshier, S. J., Bovin, M. J., Gay, N. G., Wisco, B. E., Mitchell, K. S., Lee, D. J., Sloan, D. M., Weathers, F. W., Schnurr, P. P., Keane, T. M., & Marx, B. P. (2018). Examination of posttraumatic stress disorder symptom networks using clinician-rated and patient-rated data. *Journal of Abnormal Psychology*, *127*, 541–547. https://doi.org/10.1037/abn0000368

Murphy, D., Ross, J., Ashwick, R., Armour, C., & Busuttil, W. (2017). Exploring optimum cut-off scores to screen for probable posttraumatic stress disorder within a sample of UK treatment-seeking veterans. *European Journal of Psychotraumatology*, *8*(1), 1398001. https://doi.org/10.1080/20008198.2017.1398001

Palmieri, P. A., Weathers, F. W., Difede, J., & King, D. W. (2007). Confirmatory factor analysis of the PTSD Checklist and the Clinician-Administered PTSD Scale in disaster workers exposed to the World Trade Center Ground Zero. *Journal of Abnormal Psychology*, *116*, 329–341. https://doi.org/10.1037/0021-843x.116.2.329

Raynor, D. A., & Levine, H. (2009). Associations between the five-factor model of personality and health behaviors among college students. *Journal of American College Health*, *58*, 73–82. https://doi.org/10.3200/JACH.58.1.73-82

Shipley, W. C., Gruber, C. P., Martin, T. A., & Klein, A. M. (2009). *Shipley-2*. Western Psychological Services.

Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin*, *96*, 465–490. https://doi.org/10.1037/0033-2909.96.3.465

Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress, and distress: Exploring the central role of negative affectivity. *Psychological Review*, *96*, 234–254. https://doi.org/10.1037/0033-295x.96.2.234

Weathers, F. W., Blake, D. D., Schnurr, P. P., Kaloupek, D. G., Marx, B. P., & Keane, T. M. (2013). *Clinician-Administered PTSD Scale for DSM-5 (CAPS-5)*. National Center for PTSD.

Weathers, F. W., Bovin, M. J., Lee, D. J., Sloan, D. M., Schnurr, P. P., Kaloupek, D. G., Keane, T. M., & Marx, B. P. (2018). The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5): Development and initial psychometric evaluation in military veterans. *Psychological Assessment*, *30*, 383–395. https://doi.org/10.1037/pas0000486

Weathers, F. W., Keane, T., & Davidson, J. (2001). Clinician-Administered PTSD Scale: A review of the first ten years of research. *Depression and Anxiety*, *13*, 132–156. https://doi.org/10.1002/da.1029

Weathers, F. W., & Keane, T. M. (1999). Psychological assessment of traumatized adults. In P. A. Saigh & J. D. Bremner (Eds.), *Posttraumatic stress disorder: A comprehensive text* (pp. 219–247). Allyn & Bacon.

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., & Keane, T. M. (1993, October). The PTSD checklist (PCL): Reliability, validity, and diagnostic utility. Paper presented at the ninth annual conference of the ISTSS, San Antonio, TX, United States.

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). *The PTSD Checklist for DSM-5 (PCL-5)*. National Center for PTSD.

Weathers, F. W., Marx, B. P., Friedman, M. J., & Schnurr, P. P. (2014). Posttraumatic stress disorder in DSM-5: New criteria, new measures, and implications for assessment. *Psychological Injury and Law*, *7*, 93–107. https://doi.org/10.1007/s12207-014-9191-1

Weathers, F. W., Ruscio, A. M., & Keane, T. M. (1999). Psychometric properties of nine scoring rules for the Clinician-Administered Posttraumatic Stress Disorder Scale. *Psychological Assessment*, *11*, 124–133. https://doi.org/10.1037/1040-3590.11.2.124

Wilkins, K. C., Lang, A. J., & Norman, S. B. (2011). Synthesis of the psychometric properties of the PTSD checklist (PCL) military, civilian, and specific versions. *Depression and Anxiety*, *28*, 596–606. https://doi.org/10.1002/da.20837

Wortmann, J. H., Jordan, A. H., Weathers, F. W., Resick, P. A., Dondanville, K. A., Hall-Clark, B., Foa, E. B., Young-McCaughan, S., Yarvis, J. S., Hembree, E. A., Mintz, J., Peterson, A. L., & Litz, B. T. (2016). Psychometrics analysis of the PTSD Checklist-5 (PCL-5) among treatment-seeking military members. *Psychological Assessment*, *28*, 1392–1403. https://doi.org/10.1037/pas0000260