Self-Rated versus Clinician-Rated Assessment of Posttraumatic Stress Disorder: An Evaluation of Diagnostic Discrepancies between the PTSD Checklist for *DSM-5* (PCL-5) and the Clinician-Administered PTSD Scale for *DSM-5* (CAPS-5)

by

Lindsay Brooke Kramer

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

Auburn, Alabama
August 8, 2020

Keywords: Posttraumatic stress disorder, assessment, PTSD Checklist for *DSM-5*, Clinician-Administered PTSD Scale for *DSM-5*, discrepancies

Approved by

Frank Weathers, Chair, Professor of Psychology
Christopher Correia, Professor of Psychology
Daniel Svyantek, Professor of Psychology
Tracy Witte, Associate Professor of Psychology

Abstract

Posttraumatic stress disorder (PTSD) is commonly assessed with self-rated or clinician-rated measures. Self-rated measures are more widely used, but clinician-rated measures are considered the gold standard for diagnosis. Although scores from these distinct PTSD assessment methods are strongly associated, they are often discrepant for individual symptoms, total symptom severity, and PTSD diagnostic status. To date, no known studies have empirically identified the sources of these discrepancies. Accordingly, the present study had three aims: (a) replicate previously identified discrepancies; (b) examine the contribution of several objective predictors of discrepancies, including negative response bias, careless or random responding, conscientiousness, neuroticism, and verbal IQ; and (c) identify subjective sources of discrepancies through qualitative 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 several other self-rated measures. Participants then provided open-ended feedback regarding their attributions for their discrepant symptom scores on the PCL-5 versus CAPS-5. Results indicated that the most discrepant symptoms were cued physical reactions, avoiding internal reminders, blame, loss of interest, and hypervigilance. Further, in general, objective predictors were only weakly associated with discrepancies, and for total discrepancy, neuroticism was the only significant predictor. Last, qualitative analyses revealed that 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 elucidate the nature and sources of discordance between the PCL-5 and the CAPS-5, and will inform the use and interpretation of these measures in a wide range of clinical and research applications.

Acknowledgments

I am incredibly grateful to everyone who has been with me every single step of the way. First and foremost, I am so appreciative of my advisor, Dr. Frank Weathers. Thank you for helping me become a better researcher, writer, and clinician. Thank you so much for all of your guidance and support. It has been an honor learning from you these past five years. I would also like to thank the other faculty members at Auburn University, and especially, my other committee members, Dr. Chris Correia, Dr. Tracy Witte, and Dr. Dan Svyantek. I have had the privilege of having these individuals as professors, as clinical supervisors, and as research mentors. Overall, my successes and academic achievements would not have been possible without my committee members.

Much of this research was completed thanks to my lab members and friends, Sarah Whiteman, Jess Petri, Liz Spitzer, and Jessica Townsdin. I could never have completed this extensive research project without their assistance. From making research packets, emailing potential participants, helping with data collection, and assisting with analyses, I am forever indebted to each of them for their help in making this project successful.

My biggest fear when I began graduate school wasn't that it would be mentally and emotionally challenging. It was that I wouldn't find a core group of friends. I am so happy to admit that I was wrong. I have made so many incredible, intelligent, and wonderful friends since moving to Auburn. Thank you for the wine nights, Bachelor Mondays, and self-care days. Thank you for always encouraging me and lending an ear when I needed help. And to my

friends outside of graduate school, I cannot begin to express my gratitude for all of the support you have provided.

Lastly, but definitely most importantly, to my family. My parents and my brother are the best cheerleaders you will ever meet. I was raised to dream big, but even I did not think I could get to this point in my education. However, with my family's unwavering support, love, and encouragement, I have pursued my dreams, and I am eternally grateful.

Graduate school can be extremely challenging. However, I was never alone. Thank you to everyone who believed in me, who helped me become the person I am today, and who will continue to follow me along this journey.

Table of Contents

List of Tables

Self-Rated versus Clinician-Rated Assessment of Posttraumatic Stress Disorder: An Evaluation of Diagnostic Discrepancies between the PTSD Checklist for *DSM-5* (PCL-5) and the Clinician-Administered PTSD Scale for *DSM-5* (CAPS-5)

The development and psychometric evaluation of assessment measures for posttraumatic stress disorder (PTSD) has been one of the most productive areas 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, 1980; Reardon, Brief, Miller, & Keane; 2014; Weathers & Keane, 1999; Weathers, Marx, Friedman, & Schnurr, 2014). Dozens of measures are now available to assess PTSD symptom severity and diagnostic status, many of which have been extensively validated. The most commonly used assessment methods for PTSD are self-rated questionnaires and clinician-rated diagnostic interviews, each of which has distinct advantages and disadvantages.

Questionnaires are the most commonly used PTSD measures (Elhai, Gray, Kashdan, & Franklin, 2005). The most widely used and well-validated ones include the PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993), the PTSD Symptom Scale-Self Report (PSS-SR; Foa, Riggs, Dancu, & Rothbaum, 1993), the Davidson Trauma Scale (DTS; Davidson et al., 1997), and the original and revised versions of the Impact of Event Scale (IES; Horowitz, Wilner, & Alvarez, 1979; IES-R; Weiss & Marmar, 1997). Questionnaires provide a brief and inexpensive, yet effective, way of identifying individuals with PTSD in both clinical and research settings. They require minimal training to administer and score, and typically are straightforward to interpret (Paulhus & Vazire, 2007). Also, questionnaires allow individuals to consider their symptoms privately, without being influenced by an interviewer, and some respondents may prefer this response format. In clinical settings, questionnaires can be used to

screen for possible PTSD and to efficiently track change in symptom severity over the course of treatment. In research settings, questionnaires can be used to screen for potential participants, estimate prevalence of PTSD, classify participants for between-groups analyses, and provide a continuous measure of PTSD for various correlational analyses (Reardon, Brief, Miller, & Keane; 2014; Weathers & Keane, 1999).

PTSD questionnaires also have several disadvantages. First, most PTSD questionnaires do not explicitly link individual symptoms to the traumatic event. This is a problem because many symptoms of PTSD reflect nonspecific distress and may overlap with other disorders such as depression and anxiety disorders (Simms, Watson, & Doebbeling, 2002). For example, many PTSD symptoms, such as emotional numbing and the hyperarousal symptoms, are not inherently linked to a traumatic event, so an explicit link must be established to determine if such symptoms should be considered part of the PTSD syndrome (Brewin, 2005; Weathers & Keane, 1999). Second, the use of questionnaires assumes that respondents can read the items, comprehend the meaning of the items and rating scale anchors, have sufficient memory and insight to provide accurate retrospective reports, and are motivated to respond openly and honestly. These assumptions may not always be justified in PTSD, particularly because some PTSD symptoms are conceptually complex and not easily understood by respondents (Weathers et al., 2014). Psychopathology can also impact an individual's ability to accurately report on their symptomatology (e.g., Enns, Larsen, & Cox, 2000).

Third, social desirability and other demand characteristics can impact report of symptomatology (Frisch, 1998), particularly on questionnaires. For example, individuals may be reluctant to report significant mental health problems due to a fear of being stigmatized or being denied opportunity to advance in the military (Hoge et al., 2004; Marmar et al., 2015).

Conversely, individuals with PTSD may be reluctant to report a decrease in symptoms due to a fear of losing their disability status (McNally & Frueh, 2013) or losing their identity as a trauma survivor. Therefore, response bias, which includes "faking good," "faking bad," or random responding, can distort scores on self-rated PTSD measures (Weathers & Keane, 1999). Empirical studies demonstrate an increased prevalence of defensiveness, symptom exaggeration, and malingering when individuals self-report PTSD symptoms (e.g., Frueh et al., 2003; Hickling, Blanchard, Mundy, & Galovski, 2002). Finally, questionnaires typically yield high diagnostic base rates, and thus may "overdiagnose" (Hopwood, Morey, Edelen, Shea, & Grilo, 2008; Hunt & Andrews, 1992).

Because of these various disadvantages of self-rated measures, it is generally accepted that clinician-rated structured interviews are necessary to obtain an accurate diagnosis of a psychological disorder (e.g., Resnick, Kilpatrick, & Lipovsky, 1991). There are several well-validated, dedicated PTSD interviews, including the Clinician-Administered PTSD Scale (CAPS; Blake et al., 1990), the PTSD Symptom Scale Interview (PSS-I; Foa, Riggs, Dancu, & Rothbaum, 1993), and the Structured Interview for PTSD (SI-PTSD; Davidson, Smith, & Kudler, 1989). Structured interviews allow for a more detailed exploration of symptomatology because respondents are able to discuss their experiences with a clinician. Clinicians understand the conceptual basis for each symptom, which allows them to determine if respondents accurately understand a symptom, explain or clarify as needed, determine whether responses fit the symptom criterion, and ask follow-up questions to obtain sufficient information to make a valid rating. Structured interviews also provide the opportunity for clinicians to gather additional information through behavioral observations. Last, final ratings on structured interviews are based on clinical judgment, and thus represent the respondent's answers interpreted through the

lens of the clinician's expertise.

Although structured interviews are considered the most valid assessment method, they have several disadvantages. First, interviews take substantially longer than questionnaires to administer, placing greater time burden on respondents and interviewers. Second, interviews require clinicians with sufficient training in differential diagnosis, use of structured interviews, and proficiency in administering and scoring a specific interview (Weathers & Keane, 1999). Third, clinician bias can potentially impact assessment of PTSD symptom severity. For example, clinicians may vary in their understanding of certain symptom criteria and thus may differ in whether they consider responses as qualifying for a given symptom. Also, clinicians may vary in their use of a rating scale and thus differ in the severity ratings they assign for symptoms. Fourth, as with questionnaires, interviews may also be susceptible to various forms of response bias. Clinical judgment might mitigate this somewhat, but it can be very difficult for clinicians to discern inaccurate responses. For example, Hickling and colleagues (2002) recruited trained professional actors to simulate PTSD and had them covertly present as routine referrals to an ongoing PTSD treatment study. Actors were assessed with the CAPS, administered by carefully trained clinicians with 1.5 to 5.5 years of experience. None of the actors were detected initially, although three were detected once clinicians were informed of the deception and asked to guess which individuals were simulators.

The PCL and the CAPS

Although numerous PTSD measures exist, the focus of the current study is on the PCL and CAPS, since they are the most widely used and extensively validated *DSM*-correspondent PTSD questionnaire and interview, respectively (Elhai, Gray, Kashdan, & Franklin, 2005; Weathers, Keane, & Davidson, 2001). The PCL was originally developed with a sample of

predominantly Vietnam and Gulf Veterans (Weathers et al., 1993). The PCL has been used for clinical screening and tracking symptom severity (e.g., Berlant & van Kammen, 2002), as a diagnostic tool to identify PTSD in research studies (e.g., Hoge et al., 2008), and to determine rates of PTSD in large epidemiologic research studies (Erbes, Westermeyer, Engdahl, & Johnsen, 2007).

PCL scores have demonstrated excellent internal consistency, test-retest reliability, and diagnostic utility (e.g., Keen et al., 2008; McDonald & Calhoun, 2010; Weathers et al., 1993; Wilkins, Lang, and Norman 2011). Additionally, the PCL has demonstrated convergent validity by correlating highly with total scores of other self-report measures of PTSD, as well as total scores from structured PTSD diagnostic interviews (e.g., Keen et al., 2008; Weathers et al., 1993). McDonald and Calhoun (2010) reviewed 18 studies of the PCL and concluded that it had strong diagnostic accuracy.

However, there are several limitations to the PCL. First, the PCL only assesses the symptoms of PTSD, and does not examine trauma relatedness of symptoms or functional impairment. Second, optimal cutoff scores vary widely across studies, depending on the population and trauma type. Therefore, no single cutoff score is appropriate for all populations. Third, although McDonald and Calhoun (2010) concluded that the PCL has utility as a screening test for PTSD, they cautioned against using it as a diagnostic tool, suggesting that it should be used in conjunction with structured interviews when determining a diagnosis of PTSD.

Developed in 1990 at the National Center for PTSD, the CAPS has become one of the most widely used structured PTSD interviews and is a widely accepted criterion measure for assessing PTSD diagnostic status and symptom severity. Initially validated on combat veterans, the CAPS has been extensively validated in a wide range of trauma populations, such as rape and

torture victims, motor vehicle accidents, and survivors of severe human suffering (Weathers, Keane, & Davidson, 2001). The CAPS was designed as a comprehensive, *DSM*-correspondent structured interview for PTSD, with several unique features. It includes items assessing Criterion A, core and associated symptoms, and global distress, social and occupational impairment, response validity, and overall severity. Further, the CAPS assesses both the frequency and intensity of individual PTSD symptoms with behaviorally anchored prompts and rating scales. Multiple follow-up prompts for each symptom allow interviewers to clarify any ambiguity or confusion respondents might experience. Finally, unlike other PTSD interviews, the CAPS directly assesses the trauma relatedness of individual symptoms that are not specifically linked to a traumatic event.

The CAPS has been widely used in clinical, research, and forensic settings, and has been shown to have excellent psychometric properties, including high internal consistency, interrater reliability, and test-retest reliability, as well as strong convergent and discriminant validity, diagnostic utility, structural validity, and sensitivity to clinical change (Elhai, Gray, Kashdan, & Franklin, 2005; Weathers, Keane, & Davidson, 2001; Weathers, Ruscio, & Keane, 1999). Limitation of the CAPS include longer administration time compared to other PTSD interviews and the need for more extensive training to become proficient at standard administration and scoring.

Previous Research on the Relationship between the PCL and the CAPS

The PCL and CAPS are both *DSM*-correspondent measures, so, not surprisingly, they have consistently been found to be strongly associated (e.g., Blanchard, Jones-Alexander, Buckley, & Forneris, 1996; Forbes, Creamer, & Biddle, 2001; Monson et al., 2008; Palmieri, Weathers, Difede, & King, 2007). Despite this, severity scores on these two measures are never

perfectly correlated, and diagnostic scores typically yield at least some degree of diagnostic discordance. Since each of the measures is frequently used as a primary indicator of both PTSD symptom severity and diagnostic status, it is important to examine the discordance between them.

A number of reasons have been proposed to account for this discordance, including differences in instructions, format, content, and process (e.g., Monson et al., 2008). First, the CAPS inquires about the frequency and intensity of individual PTSD symptoms, whereas the PCL asks respondents to rate their subjective distress, or the degree to which they are "bothered by" each PTSD symptom. Second, CAPS interviewers ask multiple prompts per symptom and can clarify as needed, whereas PCL items consist of a single stem and there is no opportunity for follow-up or clarification (Marmar et al., 2015). Third, clinicians can prevent double-counting of symptoms, i.e., counting the same symptom episodes toward more than one symptom criterion. For the self-rated PCL, however, this rule is too complex and nuanced to include in the instructions, so there is nothing preventing respondents from double-counting symptom episodes and thereby potentially inflating their overall severity score. Last, CAPS ratings are based on the clinical judgment of a trained interviewer who understands the conceptual basis of PTSD symptoms and the clinical meaning of the rating scale anchors. In contrast, the PCL is based on participants' self-ratings, which may be inaccurate due to not comprehending at least some PTSD symptoms, not anchoring ratings to clinical significance, or the influence of one or more types of response bias.

Although no previous study has empirically identified the sources of discrepancies between the PCL and CAPS, a number of studies have examined the degree of correspondence between the two measures across a range of traumatized samples. Blanchard, Jones-Alexander,

Buckley, and Forneris (1996) examined the psychometric properties of the PCL in a sample of 40 motor vehicle accident and sexual assault victims, using scores from the CAPS as the criteria. They found that correlations between PCL and CAPS items ranged widely (*r*s = .386 to .788), but that the correlation between PCL total severity score and CAPS total severity score was very high (*r*= .929). The lowest correlations were for the hypervigilance and the psychogenic amnesia items(*r*s= .386 and .479, respectively). 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 .700 or higher). Further, using the recommended PCL cutoff score of 50, they found a diagnostic efficiency of .825. Overall, their results indicate a high degree of correspondence between the PCL and CAPS for total severity and diagnostic status. However, correspondence was well short of perfect, and 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, Creamer, and Biddle (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 to 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.

It may be that participants underreported their symptom improvement on the PCL. Consistent with that possibility, in their meta-analysis on this issue, Van Etten & Taylor (1998) found that greater effect size changes do normally occur with clinician assessment measures compared to patient ratings. However, without empirical evidence identifying the specific

sources of this discordance, the specific reason for this underreporting can only be hypothesized. Forbes and colleagues 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.

Monson et al. (2008) also examined the concordance between the PCL and the CAPS over the course of treatment in two randomized clinical trials with male Vietnam veterans. They found significant positive associations between the two measures in total symptom severity score and in symptom clusters. In contrast to Forbes et al. (2001), however, Monson et al. found that the PCL indicated greater change as a result of treatment than did the CAPS. Specifically, for every standard deviation change in total PCL score, there was only an .82 change in total CAPS score. However, when comparing clinically significant change on the two measures, there were no differences in the percentages of agreement between the PCL and the CAPS in both improvement and exacerbation. Therefore, this study found that both measures were generally consistent in detecting change.

The PCL and CAPS have also been compared on their factor structure. Palmieri, Weathers, Difede, and King (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. Palmieri and colleagues (2007) suggested that their results may be due to the fact that the CAPS allows for clarification with follow-up questions and probes. Additionally, they suggested that individuals may not realize they are engaging in avoidance or numbing behaviors and may focus on more dysphoric and distress symptoms in self-rating their symptoms.

It is important to note that the previous information is derived from studies that utilized the original versions of the PCL and the CAPS. The diagnostic criteria for PTSD were significantly revised for *DSM–5* (APA, 2013). Due to these changes, *DSM*-correspondent assessment measures of PTSD, including the PCL and the CAPS, were revised to incorporate changes to the diagnostic criteria. However, the issue of discrepancy between the PCL and CAPS remains. In fact, discrepancies may even be more likely for *DSM-5*, because the symptoms that were added (blame, aggressive behavior, and risk-taking) and those that were heavily revised(e.g., negative beliefs, inability to experience positive feelings) are among the most conceptually difficult for individuals to comprehend and accurately rate.

Similar to the original version of the PCL, the PTSD Checklist for *DSM-5 (PCL-5;* Weathers, Litz, Keane, Palmieri, Marx, & Schnurr, 2013) has demonstrated good internal consistency, test-retest reliability, and convergent and discriminant validity, structural validity, diagnostic utility, and sensitivity to clinical change (Bovin et al., 2016; Blevins, Weathers, Davis, Witte, & Domino, 2015; Wortmann et al., 2016). Additionally, the Clinician-Administered PTSD Scale for *DSM-5* (CAPS-5; Weathers, Blake, et al., 2013) diagnosis has demonstrated good interrater reliability and test-retest reliability, and CAPS-5 total severity score has demonstrated good internal consistency, interrater reliability, test–retest reliability, and

convergent and discriminant validity (Weathers et al., 2017).

To date, several studies have compared the PCL-5 and CAPS-5. Marmar et al. (2015) used the CAPS-5 as the primary diagnostic measure in the National Vietnam Veterans Longitudinal Study (NVVLS). They found that the PCL-5, using a cutoff score of 38, showed high concordance with the CAPS-5 in signal detection analyses, with a sensitivity of 0.83, specificity of 0.95, and correct classification rate (CCR) of 95%. However, they found that the CAPS-5 yielded a lower estimate of PTSD prevalence. This finding confirms that discrepancies occur between the PCL-5 and CAPS-5, and that clinician ratings yield lower prevalence estimates than do self-ratings. Marmar et al. suggested that this effect may be due in part to revised probes on the CAPS-5 that might improve differentiation of milder from more clinically significant symptoms.

Weathers et al. (2017) examined the correlations between CAPS-5 scores and the PCL-5 for evidence of convergent validity. CAPS-5 total severity score was found to be strongly correlated with the PCL-5 (r = .66). Using a subset these data, Bovin and colleagues (2016) examined the psychometric properties of the PCL-5 in a sample of veterans. They conducted signal detection analyses (Kraemer, 1992) to examine the diagnostic utility of PCL-5 scores for predicting PTSD diagnostic status on the CAPS-5. Analyses revealed that scores of 31 to 33 was optimally efficient for predicting a CAPS-5 diagnosis. However, the resulting kappa was only .58, which indicates substantial diagnostic discordance. Additionally, similar to Marmar et al (2015), the CAPS-5 yielded a lower estimate of PTSD prevalence. Therefore, the cumulative evidence from these studies suggests that, although correspondence between the PCL-5 and CAPS-5 is strong, discrepancies still exist between the assessments.

Lastly, similar to the Palmieri et al. (2007) study, Lee and colleagues (2019) examined

the factor structure of PTSD symptoms using both the PCL-5 and the CAPS-5 to determine the impact of assessment method. They found that fit for the two measures was similar across the various models evaluated, and concluded that there was little evidence for a method effect. However, they noted that specific items (i.e., cued distress, avoidance of thoughts, and avoidance of external reminders) varied in salient loadings, indicating some notable discrepancy between the PCL-5 and CAPS-5 at least for certain items.

Present Study

Although correspondence between the PCL and CAPS is generally high, the correspondence is never perfect, and discrepancies, especially between individual symptoms, are common. Investigators have speculated about possible reasons for these discrepancies, but no studies have attempted to empirically identify their sources. To address this gap in the literature, the present study aimed to utilize both quantitative and qualitative research to understand the discordance. For the quantitative approach, several potential objective predictors that might plausibly account for discordance were examined, including response validity indicators of negative response bias, inconsistency, and infrequency; personality traits of conscientiousness and neuroticism; and verbal IQ. For the qualitative approach, participants were presented with their actual discrepant scores on the PCL-5 and CAPS-5 and then asked, in an unstructured interview format, to provide feedback about why they think their responses were discrepant.

Accordingly, this study had three primary aims. The first aim was to document the degree of discordance between *DSM-5* versions of the PCL and CAPS. Previous research indicates that discrepancies will occur, but as a first step for the present study it was important to (a) quantify discrepancies; (b) examine them at the level of individual PTSD symptoms, symptom clusters, and diagnostic status; and (c) examine them for dimensional and dichotomous

(present/absent) scores, all based on the revised *DSM-5* PTSD symptom criteria. It was hypothesized that PCL-5 and CAPS-5 scores would be strongly positively associated, but would also yield substantial levels of diagnostic and symptom discrepancies. Specifically, it was hypothesized that covert PTSD symptoms (i.e., intrusive memories, flashbacks, avoidance of thoughts and feelings) would be the most discrepant. This was proposed because individuals display lower reliability when reporting symptoms that are not behaviorally anchored (e.g., Foa et al., 2016). In addition, individuals may struggle to understand the differences between intrusive memories and flashback symptoms (Marmar et al., 2015). Weathers and colleagues (2017) also suggested that individuals may be less aware of their avoidance, especially of internal cues (i.e., thoughts and feelings) and are therefore less likely to report these symptoms on self-report measures compared to interviews.

The second aim was to determine the extent to which potential objective predictors could account for discordance between the PCL-5 and CAPS-5. As noted above, predictors examined included the response validity indicators of negative response bias, infrequency, and inconsistency; personality traits of conscientiousness and neuroticism; and verbal IQ. Since these predictors have not been examined previously for this purpose, their inclusion in the present study is 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. Specifically, infrequency refers to responding to items in an idiosyncratic way, whether due to reading difficulties, random responding, or failure to follow test instructions, whereas inconsistency refers to failing to attend consistently or appropriately to item content due to carelessness, reading difficulties, or

confusion.

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). Further, 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, Lang, and Norman (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.

The third aim was to elicit qualitative feedback from participants regarding their attributions for discrepant symptoms between the PCL-5 and the CAPS-5. Although to date no known study has asked participants to explain their discrepant responses, it was hypothesized that participants' most commonly reported attributions would generally reflect the main sources of discrepancy previously proposed by researchers, including statements that compared to the

PCL-5, the CAPS-5 explicitly links symptoms to the index event, inquires in greater detail, improves comprehension of symptoms, and facilitates greater disclosure.

Method

Participants and Procedure

The study consisted of two separate sessions. In Part 1, undergraduate students 19 and older enrolled in a psychology course at Auburn University were invited to complete an online survey related to "a very stressful life event" and were compensated with extra credit. Participants indicated consent by selecting that they agree to participate in the study. They then responded to several questionnaires, including the extended version of the Life Events Checklist for *DSM-5* (LEC-5; Weathers et al., 2013), PCL-5, and the Personality Assessment Inventory (PAI; Morey, 1991). The questionnaires took approximately 90 minutes to complete.

Some participants who completed Part I of the study were invited to participate in Part 2. Those who chose to participate in Part 2 were emailed and scheduled within one to two weeks following completion of Part 1. Participants were considered eligible for Part 2 if their self-identified worst event on the LEC-5 met *DSM-5* Criterion A and they had a score of 24 or higher on the PCL-5, indicating at least moderate levels of PTSD symptom severity. This choice of a cutoff score was intended to ensure that participants were reporting high enough levels of PTSD symptoms so that sufficient discrepancies would be found. It also establishes generalizability to research and clinical settings. However, the disadvantage of using a cutoff score is that it creates a restriction of range in both PCL-5 and CAPS-5 scores, which could lead to lower correlations between the two measures. Criterion A status was determined by reviewing participants' responses on the LEC-5 and their written narrative description of their index event. Sixty participants completed Part 2 and were included in analyses.

Part 2 was conducted in therapy rooms in Auburn University Psychological Services Center. Participants signed an informed consent form, completed the PCL-5 and the Big Five Inventory (BFI; John & Srivastava, 1999), and were administered the CAPS-5. Interviewers were doctoral students in clinical and counseling psychology, trained and supervised by a licensed clinical psychologist with expertise in the assessment of PTSD. Throughout the study, interviewers met periodically for reliability checks, for which they independently rated audiotaped interviews and discussed discrepant ratings.

After completion of the CAPS-5, participants had ten minutes to complete the Shipley-2 (Shipley, Gruber, Martin, & Klein, 2009). While they were filling out this measure, the interviewer compared their responses on the CAPS-5 and PCL-5. After identifying all discrepancies between the CAPS-5 and PCL-5, the interviewer conducted an open-ended interview, asking participants to comment on why they responded differently across the two PTSD measures. The discrepancy interview was audio-recorded to assist with qualitative analysis.

Part 2 took approximately two hours. Participants were provided with a copy of the debriefing form. They were then assigned extra credit for completion and received a $15 gift card. The University's Institutional Review Board approved the study protocol in October 2017.

Part 1 Measures

Demographics – A custom demographics questionnaire was used to assess sex, age, race, ethnicity, income, parent relationship status (e.g., never married, married, divorced) relationship status (e.g., single, married), student status (e.g., part-time, full-time), and work status (e.g., part-time, full-time).

The *Life Events Checklist for DSM-5* (LEC-5; Weathers et al., 2013) was used to identify Criterion A status. The LEC-5 is a self-rated measure of trauma 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*. Of the events participants experienced, they are then asked to select the one event that bothered them the most and to briefly describe it. The narrative description of this index event was used to verify Criterion A status. Index events were coded as meeting *DSM-5* Criterion A if participants endorsed their worst event on the LEC-5 as *having happened to me directly* or *witnessed it.* Additionally, index events were coded as meeting *DSM-5* Criterion A if participants also endorsed one of the following: *my life was in danger, someone else's life was in danger,* or *if it involved sexual violence*. Index events were also coded as *DSM-5* Criterion A if participants endorsed *having learned about it happening to a close family member or close friend* and *if it involved accident or violence or sexual violence*. Previous versions of the LEC have been shown to be psychometrically sound in a variety of samples (Gray, Litz, Hsu, & Lombardo, 2004).

The *PTSD Checklist for DSM-5 (PCL-5;* Weathers, Litz, Keane, Palmieri, Marx, & Schnurr, 2013) is a 20-item questionnaire designed to assess *DSM–5* symptoms of PTSD. For each symptom, respondents provide a severity rating ranging from 0 to 4 that indicates how much they were bothered by each symptom (0= *not at all* to 4= *extremely*) in the past month. Scores range from 0 to 80, and higher scores indicate more severe PTSD symptoms. Symptom clusters include re-experiencing, avoidance, negative alterations of cognitions and moods, and arousal and reactivity. PCL-5 scores have demonstrated strong internal consistency ($\alpha = .94$),

test-retest reliability (*r* = .82), and convergent (*rs* = .74 to .85) and discriminant (*rs* = .31 to .60) validity (Blevins, Weathers, Davis, Witte, & Domino, 2015). In the present study, Cronbach's alpha was .90.

The *Personality Assessment Inventory* (PAI; Morey, 1991) is a 344-item self-report inventory developed to assess personality functioning and psychopathology. There are four validity, 11 clinical, five treatment consideration, and two interpersonal scales. Instructions ask participants to respond to each item on a four-point scale ranging from 1, *not true at all*, to 4, *very true*. The PAI has been evaluated extensively and has been shown to have excellent psychometric characteristics (Morey, 1991; Morey, 2007), including in trauma-exposed college students similar to the sample for the present study (McDevitt-Murphy, Weathers, Flood, Eakin, & Benson, 2007). PAI scales used in the present study were Negative Impression Management (NIM), which measures the tendency to present oneself in an overly negative manner; Inconsistency (INC), which measures the tendency to respond inconsistently to items; and Infrequency (INF), which measures the tendency to respond to items in an idiosyncratic way.

Part 2 Measures

The *PTSD Checklist for DSM-5 (PCL-5;* Weathers, Litz, Keane, Palmieri, Marx, & Schnurr, 2013) was again used to assess the *DSM–5* symptoms of PTSD. Cronbach's alpha for this administration was .89.

The *Big Five Inventory* (BFI; John & Srivastava, 1999) is a 44-item questionnaire that assesses the Big Five personality domains. The five subscales are Extraversion (8 items), Agreeableness (9 items), Conscientiousness (9 items), Neuroticism (8 items), and Openness (10 items). For the present study, only the Conscientiousness and Neuroticism subscales were examined. Conscientiousness measures the tendency to be cautious, thorough, and organized.

Neuroticism measures the tendency to be prone to negative emotions, stress, and emotional instability. All items consist of short phrases that are based on prototypical trait adjectives related to each construct and are rated on a 5-point scale (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. John and Srivastava (1999) reported alpha coefficients from .75 to .80 for subscales and 3-month test-retest reliabilities from .80 to .90. Additionally, the BFI has been shown to have good convergent validity with other Big 5 personality measures. In the present study, Cronbach's alphas for the conscientious and neuroticism subscales were .81 and .86, respectively.

The *Clinician-Administered Posttraumatic Stress Disorder Scale for DSM-5* (CAPS-5; Weathers, Blake, et al., 2013) is a structured diagnostic interview for *DSM-5* PTSD. Interviewers assess the intensity and frequency of each PTSD symptom using multiple behaviorally anchored prompts. Using predetermined scoring rules, they then combine intensity ratings and frequency information into a single 5-point (0 to 4) symptom severity score, for which 0 = *absent*, 1= *mild/subthreshold*, 2 = *moderate/threshold*, 3 = *severe/markedly elevated*, and 4 = *extreme/incapacitating*. Item scores of 2 or higher are considered symptoms endorsed and are counted toward a PTSD diagnosis. PTSD diagnostic status is determined by following the *DSM-5* diagnostic rule, which requires at least one reexperiencing symptom, one avoidance symptom, two symptoms of negative alterations in cognition and mood, and two symptoms of hyperarousal, as well as a duration of at least one month and clinically significant distress or impairment. CAPS-5 diagnosis has demonstrated strong interrater reliability (α = .78 to 1.00) and test-retest reliability (α= .83). CAPS-5 total severity score has demonstrated high internal consistency (α= .88) and interrater reliability (ICC = .91), good test–retest reliability (ICC =

.78), and good convergent and discriminant validity (Weathers et al., 2017). Cronbach's alpha for CAPS-5 full scale was .87.

The *Shipley-2* (Shipley, Gruber, Martin, & Klein, 2009*)* is a self-report measure of cognitive functioning and impairment that provides a good estimation of verbal and non-verbal reasoning ability. It comprises three subtests: vocabulary, abstraction, and block patterns. In the present study, only the vocabulary test was utilized to provide a measure of verbal ability. The 40-item vocabulary test requires participants to choose which of four listed words ''means the same or nearly the same'' as a specified target word. There is a 10-min time limit. The Shipley-2 was scored by summing participant's correct responses on the test. Concurrent validation evidence shows that the Shipley-2 correlates strongly with the WAIS-III (r=.45 to .87) and the Wonderlic Personnel Test (r=.47 to .64). Additionally, test- retest reliability ranges from r = .87 to .94 (Shipley et al., 2009).

Data Analytic Plan

The primary outcome for the present study was the extent of discrepancy between the PCL-5 and the CAPS-5. For dimensional scores, i.e., scores based on the 0-4 rating scales for PCL-5 and CAPS-5 items, discrepancy scores were created for each PTSD symptom by subtracting the CAPS-5 item score for a given symptom from the corresponding PCL-5 score and taking the absolute value. Discrepancy scores for PTSD symptoms B1(memories) to B5 (cued distress) were summed to create the Reexperiencing Cluster discrepancy score. Discrepancy scores for PTSD symptoms C1 (avoiding internal reminders) and C2 (avoiding external reminders) were summed to create the Avoidance Cluster discrepancy score. Discrepancy scores for PTSD symptoms D1 (amnesia) to D7 (numbing) were summed to create the Negative Alterations in Cognition and Mood Cluster discrepancy score. Discrepancy scores

for PTSD symptoms E1 (irritability or aggressive behavior) to E6 (sleep) were summed to create the Arousal and Reactivity Cluster discrepancy score. Discrepancy scores for all twenty PTSD symptoms were summed to create total PTSD discrepancy score.

For dichotomous (present/absent) scores, PCL-5 and CAPS-5 items were dichotomized such that PCL-5 items rated as *2 = Moderately* or higher, and CAPS-5 item Severity scores rated as *2 = Moderate/threshold* or higher, were considered as symptoms endorsed or present. PCL-5 and CAPS-5 items rated as 0 or 1 were considered absent. In addition, a dichotomous diagnostic variable was computed for both the PCL-5 and CAPS-5 by using dichotomous symptom scores and following the *DSM-5* PTSD diagnostic rule, which requires one reexperiencing symptom, one avoidance symptom, two symptoms of negative alterations in cognition and mood, and two hyperarousal symptoms. Discrepancies for dichotomized items and diagnosis were combinations of PCL-5 present/CAPS-5 absent and PCL-5 absent/CAPS-5 present. The dimensional discrepancy scores were utilized as the dependent variables in the multiple regression analyses, while the dichotomous discrepancy scores were utilized as the dependent variables in the logistic regression analyses.

Importantly, the CAPS-5 and PCL-5 differ in terms of the aspect of symptom severity on which ratings are based, and this difference might be an important source of discrepancy. Specifically, CAPS-5 ratings are based on symptom intensity and severity (with severity being derived from a combination of intensity and frequency), whereas the PCL-5 ratings are based essentially on subjective distress, without explicit consideration of frequency. Thus, PCL-5 ratings arguably are conceptually more similar to CAPS-5 Intensity ratings than they are to CAPS-5 Severity scores, and therefore might be more strongly correlated with and less discrepant from CAPS-5 Intensity scores relative to CAPS-5 Severity scores. Accordingly, for

dimensional scores, two distinct sets of discrepancy scores were calculated: one between CAPS-5 Intensity scores and PCL-5 scores, and the other between CAPS-5 Severity scores and PCL-5 scores. Although CAPS-5 Intensity ratings are not numerical, for the purpose of the analyses, CAPS-5 Intensity anchors were coded the following values: 0 = *Absent*, 1= *Minimal*, 2 = *Clearly Present*, 3 = *Pronounced*, and 4 = *Extreme*.

The first aim of documenting the degree of discordance between the PCL-5 and the CAPS-5 was addressed in four steps, the first three involving dimensional scores, and the last one involving dichotomous scores. First, correlations were calculated between PCL-5 and CAPS-5 items, symptom clusters, and total score. Based on Blanchard et al. (1996) and Forbes et al. (2001), these correlations were expected to be moderately strong, generally in the .50-.70 range, although possibly somewhat lower because of the restriction of range due to selecting only participants with PCL-5 scores at least in the moderate range. Second, mean differences between PCL-5 and CAPS-5 items, symptom clusters, and total score were evaluated with paired *t*-tests. In keeping with previous findings that self-ratings tend to be higher than clinician ratings, mean PCL-5 scores were expected to be higher than CAPS-5 scores for most items, all four symptom cluster scores, and total severity score.

Third, the frequency of any discrepancy and the mean discrepancy were calculated between PCL-5 and CAPS-5 items, symptom clusters, and overall PTSD. Differences between discrepancy levels for CAPS-5 Intensity versus Severity were evaluated with McNemar tests and paired *t*-tests. Fourth, examining dichotomous scores, the PCL-5 and CAPS-5 were compared on the prevalence of individual PTSD symptoms and PTSD diagnosis. Using the PCL-5 as the test and the CAPS-5 as the criterion, the outcomes of 2 x 2 tables (hits, misses, false alarms, correct rejection) were calculated and evaluated with kappa coefficients. Discrepancies were either

misses (absent on the PCL-5 but present on the CAPS-5) or false alarms (present on the PCL-5 but absent on the CAPS-5).

The second aim of evaluating potential objective predictors of discrepancy was addressed in three steps. First, correlations were calculated between the various objective predictors and dimensional discrepancy scores for individual items, symptom clusters, and total PTSD. Next, the objective predictors were combined in multiple regression to examine their aggregate ability to predict these same discrepancy scores. Third, the objective predictors were combined in logistic regression to examine their aggregate ability to predict discrepancies between dichotomized items and diagnosis (i.e., combinations of PCL-5 present/CAPS-5 absent and PCL-5 absent/CAPS-5 present).

Last, qualitative analysis was conducted to examine the feedback provided from the participants regarding the discrepancies. Thematic analysis, a type of qualitative analysis, was employed. This is a method for identifying, analyzing, and reporting patterns or themes within data. The overall process starts when the researcher looks for patterns of meaning and the endpoint is the reporting of the content and meaning of themes in the data. Braun and Clark (2006) described five phases of thematic analysis. These phases were followed for the present study.

The first phase of thematic analysis is familiarization of the data, which includes transcribing interview responses. The second phase involves generation of codes of data. The third phase involves organizing the different codes into themes. The fourth phase consists of refining the themes by either removing a theme if there is not enough codes to support the theme or combining several themes into one overall theme. Phase five involves defining, further

refining, and naming the themes. Overall, this process produces a thematic "map" of the entire dataset.

Results

A total of 2,539 number of individuals completed Part 1 of the study. There were three inclusion criteria for Part 2 of study. First, participants had to indicate that they wanted to be contacted. Second, participants had to report an index event that met *DSM-5* Criterion A. Third, participants had to report a PCL-5 total score greater than or equal to 24. Two hundred individuals met all three criteria and were contacted to participate in Part 2, sixty of whom completed Part 2.

The final sample was predominantly female (88.3%; $n = 53$), and the ethnic 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 ($n = 1$). Age ranged from 19 to 24 ($M = 19.83$; $SD = 1.2$). Event types for the *DSM-5* Criterion A group 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 suicide (3.3%, $n = 2$). Item-level descriptive statistics for the PCL-5 and CAPS-5 from Part 2 of the study are presented in Appendix A, Tables A1 and A2.

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

The first aim was to identify the degree of discrepancy between the PCL-5 and CAPS-5. First, correlations between PCL-5 and CAPS-5 items, symptom clusters, and total severity score were computed to quantify discordance between the two measures. As shown in Table 1,

correlations between PCL-5 and CAPS-5 Intensity scores ranged from .38 to .76 for individual items, and from .52 to .67 for the four symptom clusters, with a correlation of .63 for total score. Correlations between PCL-5 and CAPS-5 Severity scores were comparable but generally slightly higher, ranging from .34 to .81 for items, and from .51 to .76 for clusters, with a correlation of .69 for total score. The lowest correlations for PCL-5 items with both CAPS-5 Intensity and Severity items were found for B3 (flashbacks), C1 and C2 (avoidance of internal and external reminders), E3 (hypervigilance), E4 (startle), and E5 (concentration), whereas the highest were found for B2 (nightmares), D1 (amnesia), D7 (inability to experience positive emotions), E1 (aggressive behavior), and E2 (reckless or self-destructive behavior).

Second, as shown in Table 1, mean differences between PCL-5 and CAPS-5 items, symptom clusters, and total score were evaluated with paired *t*-tests. Similar to previous findings (e.g., Bovin et al., 2016; Forbes et al. 2001; Marmar et al., 2015; Parker-Guilbert, Leifker, Sippel, & Marshall, 2014), mean PCL-5 scores were significantly higher than CAPS-5 scores for most items, all four symptom cluster scores, and total score. This was particularly the case for PCL-5 versus CAPS-5 Severity scores, indicating that PCL-5 scores are more discrepant from CAPS-5 Severity scores than from CAPS-5 Intensity scores. Notably, PCL-5 total score was nearly 11 points higher than CAPS-5 Severity total score, and a little over 9 points higher than CAPS-5 Intensity total score.

Third, as shown in Tables 2 and 3, discrepancies between PCL-5 and CAPS-5 scores were common. Table 2 presents the raw discrepancies between PCL-5 items and CAPS-5 Intensity scores, and Table 3 presents the raw discrepancies between PCL-5 items and CAPS-5 Severity scores. The pattern of discrepancies was similar for both CAPS-5 Intensity and Severity scores. First, for most symptoms, the modal discrepancy was 0, and most other discrepancies

were +/- 1, indicating a fairly high level of correspondence between PCL-5 and CAPS-5 scores. Second, consistent with the higher mean scores for PCL-5 in Table 1, the pattern of discrepancies in Tables 2 and 3 indicate that PCL-5 scores were generally higher than CAPS-5 scores, in that most discrepancies were in the positive range. Further, there were relatively few discrepancies of -2, almost none of -3, and none of -4, indicating that it was rarely the case that the CAPS-5 item score was more than a point higher than the PCL-5 item score.

As shown in Table 4, for individual items, the frequency of any discrepancy between the PCL-5 score and CAPS-5 Intensity score ranged from 26.7% for E2 (reckless behavior) to 73.3% for B5 (cued physical reactions), and 16 of 20 items were discrepant more than half the time. In addition, considering absolute values of discrepancies, mean discrepancy ranged from .35 for E2 to 1.32 for C2 (avoidance of external reminders). Similarly, the frequency of any discrepancy between the PCL-5 score and CAPS-5 Severity score ranged from 28.3% for E2 to 73.3% for B4 (cued distress), and again 16 of 20 items were discrepant more than half the time. Mean discrepancy ranged from .35 for E2 to 1.33 for C2. Next, differences between CAPS-5 Intensity and Severity scores with respect to frequency of any discrepancy and mean discrepancy were examined with McNemar tests and paired *t*-tests. No significant differences were found for frequency of any discrepancy, and only one significant difference was found for mean discrepancy, specifically for C1 (avoidance of internal reminders), for which the discrepancy for CAPS-5 Severity was higher than for CAPS-5 Intensity. These analyses indicate the frequency of any discrepancy and mean discrepancy are highly similar for CAPS-5 Intensity and Severity scores. Therefore, for the remainder of the analyses, only CAPS-5 Severity scores were used, since these are the final and only numerical rating on the CAPS-5, and thus the only CAPS-5 scores used and reported in the literature.

Fourth, as shown in Table 5, discrepancies between PCL-5 and CAPS-5 dichotomous scores were examined based on misses (absent on the PCL-5 but present on the CAPS-5) and false alarms (present on the PCL-5 but absent on the CAPS-5). D2 (negative beliefs), D4 (negative emotions), D6 (detachment or estrangement), and E6 (sleep problems) were found to have the highest prevalence of misses. B3 (flashbacks), C2 (avoidance of external reminders), D3 (blame), E3 (hypervigilance), and E4 (startle) were found to have the highest prevalence of false alarms. Kappa coefficients, indicating chance-corrected overall agreement between the PCL-5 and CAPS-5, ranged from .24 to .64. Items with the lowest kappas, and thus lowest agreement, were B3 (flashbacks), C1 (avoidance of internal reminders), D4 (negative emotions), and E4 (startle). Items with the highest kappas, and thus highest agreement, were B2 (nightmares), D3 (blame), D7 (inability to experience positive emotions), and E2 (reckless behavior).

Second Aim: Multiple Regression and Logistic Regression Analyses

The second aim was to evaluate potential objective predictors of PCL-5/CAPS-5 discrepancies, including negative impression management, inconsistency, and infrequency; personality traits of conscientiousness and neuroticism; and verbal IQ. Descriptive statistics for the objective predictors are presented in Table 6. First, correlations were calculated between the predictors and PTSD items. As shown in Table 7, there were relatively few significant correlations between the various objective predictors and discrepancy scores for items, clusters, and total score, and only Neuroticism was significantly correlated with discrepancy score for total score.

Second, multiple regression analyses were run to examine the aggregate ability of the objective predictors to predict discrepancy scores. The overall model was non-significant for all

but four symptoms: cued distress, aggressive behavior, hypervigilance, and concentration difficulties. The overall model was significant for cued distress, $F(6, 50) = 2.90$, $p = .02$, with NIM, $\beta = .46$, $t = 3.35$, $p < .001$, and verbal IQ, $\beta = -.27$, $t = -2.08$, $p = .04$, significantly associated with the symptom. The overall model was significant for irritability and aggressive behavior, $F(6, 50) = 3.67$, $p < .001$, with INF, $\beta = 30$, $t = 2.34$, $p = .02$ and Neuroticism, $\beta = .50$, $t = 3.58$, $p < .001$, significantly associated with the symptom. The overall model was significantly associated with hypervigilance, $F(6, 50) = 2.44$, $p < .04$, with NIM, $\beta = .35$, $t = 2.48$, $p = .02$, significantly associated with the symptom. Last, the overall model was significantly associated with concentration, $F(6, 50) = 4.18$, $p < .00$, with Neuroticism, $\beta = .36$, $t = 2.62$, $p = .01$, and Verbal IQ, $\beta = .40$, $t = 3.31$, $p < .001$, significantly associated with the symptom.

Examining the four clusters, the overall model was only significant for the arousal and reactivity cluster, $F(6, 50) = 6.92$, $p < .001$, with neuroticism, $\beta = .46$, $t = 3.73$, $p < .001$, as the only significant predictor of this cluster. Last, the overall model was significant for total discrepancy score, $F(6, 50) = 2.61$, $p = .02$, with neuroticism, $\beta = .35$, $t = 2.41$, $p = .02$, as the only significant predictor of this cluster. See Table 8 for the full breakdown of multiple regression analyses results.

Third, the objective predictors were combined in logistic regression to examine their aggregate ability to predict discrepant dichotomized items and diagnostic score (i.e., combinations of PCL-5 present/CAPS-5 absent and PCL-5 absent/CAPS-5 present). Analyses revealed that none of the predictors significantly predicted dichotomized discrepant items or diagnostic discrepancy score. Therefore, these results will not be presented.

Third Aim: Qualitative Analyses

Two doctoral students in clinical psychology performed the initial coding procedures. Following the steps of Braun and Clarke (2006), these researchers first examined transcriptions of each interview conducted. An initial list of codes were created that were then transformed into a list of approximately 30 potential themes. 10% of the interviews were selected to help narrow down the list of themes. During this stage, the researchers were able to refine the themes to come up with a final list of 21 themes of discrepancies. A codebook with definitions of codes and example quotations derived from consensus discussions and used to guide final coding of all transcripts can be found in Appendix B. Once the list of themes was finalized, the first researcher went through each interview to determine frequency and percentages of themes endorsed for each discrepant system between the PCL-5 and the CAPS-5 (see Tables 9-11). The second researcher coded a random subset (approximately 33%) of the interviews. Interrater reliability was excellent (κ = .78). Because of the range of responses, each discrepant symptom could have been coded with up to three different themes. Only the top five themes (i.e., time-frame reminders, comprehension of symptoms, trauma-related attribution error, increased awareness, and general errors) are discussed in detail to assist with brevity.

The theme of time-frame reminders was coded if participant explained that they were not answering based on the past month or that they wanted reminders of the specific timeframe. An example is: "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 found to be frequently reported for detachment and estrangement, startle, and concentration symptoms.

The theme of comprehension of symptoms was coded if the participant reported that they misunderstood or misinterpreted the question. An example is: “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.” The theme was found to be frequently reported for memories and flashbacks symptoms.

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. An example is: “I have always had problems with sleeping so I didn’t realize the trauma had progressively changed my sleeping habits.” This theme was found to be frequently reported for hypervigilance, concentration, and sleep symptoms.

The theme of increased awareness was coded if the participant discussed the overall internal process that occurred while answering the assessment measure. An example is: “Having the time to think and talk about my symptoms helped me understand what I am actually going through. Thinking it through helped me to process and organize my thoughts and helped me remember what I have been struggling with this past month.” The theme was found to be frequently reported for memories, avoidance of internal and external reminders and blame symptoms.

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.. An example is: “I don’t remember marking that answer. I must not have read the question and I marked incorrectly.” The theme was found to be frequently reported for cued distress, avoidance of internal and external reminders symptoms.

Discussion

Self-rated measures, such as the PTSD Checklist for *DSM-5* (PCL-5), have consistently demonstrated strong psychometric properties (Bovin et al., 2016; Blevins, Weathers, Davis, Witte, & Domino, 2015; Wortmann et al., 2016). However, although self-rated measures allow for efficient administration, researchers and clinicians have questioned the accuracy of self-rated measures compared to structured interviews. The present study examined the discrepancies between a self-rated measure (i.e., PCL-5) and a clinician-rated measure (i.e., CAPS-5) in a trauma-exposed sample. This is the first known study to empirically examine the quantitative and qualitative reasons behind discrepancies.

The first aim of the study was to identify the degree of discrepancy between the PCL-5 and the CAPS-5. Discrepancies were addressed in four steps, the first three involving dimensional scores, and the last one involving dichotomous scores. First, as previous studies have determined (e.g., Blanchard et al., 1996; Forbes, 2001), the PCL-5 total score and all specific symptoms significantly and positively correlated with the CAPS-5 total score and symptoms, respectively. The lowest correlations for PCL-5 items with both CAPS-5 Intensity and Severity items were found for B3 (flashbacks), C1 and C2 (avoidance of internal and external reminders), E3 (hypervigilance), E4 (startle), and E5 (concentration), whereas the highest were found for B2 (nightmares), D1 (amnesia), D7 (inability to experience positive emotions), E1 (aggressive behavior), and E2 (reckless or self-destructive behavior).

Although highly associated with each other, the correlations were never close to perfect. It is important to note that the correlations are possibly somewhat lower because of the restriction of range due to selecting only participants with PCL-5 scores at least in the moderate

range. However, the wide range of correlations between the PCL-5 and the CAPS-5 provides evidence of discordance between the two measures.

Second, and in continued support of discrepancies between the measures, mean PCL-5 scores were significantly higher than CAPS-5 scores for most items, all four symptom clusters, and total score. This was particularly the case for PCL-5 versus CAPS-5 Severity scores, indicating that PCL-5 scores are more discrepant from CAPS-5 Severity scores than from CAPS-5 Intensity scores.

Third, item-level discrepancies occurred on more than half of the PTSD symptoms. It was found that symptom pairs (e.g., hypervigilance and exaggerated startle, and avoidance of internal and external reminders) were among the more discrepant findings. This finding corresponds with Moshier and colleagues (2019) who evaluated assessment modality using network analysis. As they suggested, it is possible that participants double-coded these symptom pairs, while clinicians were able to distinguish between them. Furthermore, when considering avoidant symptoms, it is possible that participants are less aware that they are engaging in avoidance behavior and are less likely to report it on self-report measures. Additionally, covert symptoms, including cued distress, cued physical reactions, loss of interest, and blame, were also found to be different across the different assessment modalities.

Fourth, discrepancies between PCL-5 and CAPS-5 dichotomous scores were examined based on misses (absent on the PCL-5 but present on the CAPS-5) and false alarms (present on the PCL-5 but absent on the CAPS-5). The PCL-5 was found to be less efficient in determining the presence of individual PTSD symptoms and PTSD diagnosis. The properties of individual items are important to make diagnostic judgments about the possible presence of PTSD. However, considerable variation was apparent at this level of diagnostic accuracy. This finding is

similar to previous literature examining diagnostic utility of individual items (e.g., Blanchard, 1996; Forbes, et al., 2001).

Negative beliefs, negative emotions, detachment or estrangement, and sleep problems were found to be the highest prevalence of misses. Symptoms of flashbacks, avoidance of external reminders, blame, hypervigilance, and startle were found to be the highest prevalence of false alarms. Items with the lowest kappas (i.e., indicating chance-corrected overall agreement between the PCL-5 and CAPS-5), and thus lowest agreement, were flashbacks, avoidance of internal reminders, negative emotions, and startle. Items with the highest kappas, and thus highest agreement, include nightmares, blame, inability to experience positive emotions, and reckless behavior. This item information is valuable for potential diagnostic purposes. While the PCL-5 appears to be useful in diagnosis based on total score, it does not seem to accurately reflect individual symptom patterns. Overall, through the four different methods of examining discrepancies, it was demonstrated at each level that individual symptoms, clusters, and diagnostic discrepancies exist.

The reasons for discrepancies have remained a matter of speculation. As noted earlier, various theories have been proposed to explain why discordance occurs. However, no study has empirically examined the quantitative and qualitative reasons behind discrepancies. Specifically, the second aim of the study examined six different potential objective predictors of discrepancy including response validity indicators of negative impression management, inconsistency, infrequency; personality traits of conscientiousness and neuroticism; and verbal IQ. The predictors chosen are exploratory variables; and therefore, no prior hypotheses were formed for their relationship with individual PTSD symptoms.

Utilizing multiple regression analyses, it was indicated that neuroticism was the most predictive of discrepancy. Neuroticism has been defined as commonly experiencing unpleasant and negative emotions, such as anxiousness, pessimism, sadness, and insecurity (Raynor & Levine, 2009). Therefore, highly neurotic individuals have been found to assess even banal, everyday situations as threatening (Ebstrup, Eplov, Pisinger, & Jørgensen, 2011). Neuroticism was found to predict the discrepant symptoms of feeling irritable and acting out and of concentration, which are both symptoms that are comprised in the arousal and reactivity cluster. The variable was also predictive of total score of PTSD symptoms. It is suggested that symptoms that have more of a negative connotation (i.e., feeling irritable and acting out) or symptoms that are connected to more everyday situations (i.e., concentration) are more likely to be viewed as more frequent and severe in a highly neurotic individual.

Additionally, response bias, which includes "faking good," "faking bad," or random responding, could impact assessment of symptoms (Weathers & Keane, 1999). Therefore, it was found that response validity indicators, including negative impression management and infrequency, influenced discrepancies between the PCL-5 and CAPS-5. The symptoms of PTSD that were significantly associated with these response validity indicators were cued physical reactions, irritability or aggressive behavior, and hypervigilance. It is suggested that symptoms that are both conceptually difficult to understand and have a more external manifestation could be more likely influenced by response bias.

The association between verbal IQ and discrepancies was also examined. It was hypothesized that lower verbal IQ would impact comprehension of the more abstract symptoms. Lower IQ was associated with cued physical reactions, while higher IQ was associated with

concentration. Therefore, this hypothesis was found to be partially supported, as the symptom of cued physical reactions is considered conceptually difficult for participants to understand.

Last, the objective predictors were combined in logistic regression to examine their combined ability to predict dichotomized discrepant items and diagnostic score (i.e., combinations of PCL-5 present/CAPS-5 absent and PCL-5 absent/CAPS-5 present). Analyses revealed that none of the predictors significantly predicted dichotomized discrepant items and diagnostic discrepancy score.

In conclusion, some of the objective predictors were significantly related to discrepancy scores. The objective predictors chosen were exploratory variables, as this is the first known study to examine quantitative sources of discrepancies. It is recommended that these analyses are replicated in future studies.

The third aim of the study was to identify subjective predictors of discrepancies through qualitative analysis of participant feedback. Participants provided reasons that their responses were discrepant across the different assessment methods. Twenty-one different themes were generated based on these responses. The top five themes were: time-frame reminders, comprehension of symptoms, trauma-related attribution error, increased awareness, and general errors. Surprisingly, themes of opportunity for clarification and elaboration was not among the most frequent reasons behind discrepancies. Structured interviews have been generally considered the diagnostic standard because of the opportunity to clarify questions and to ask for examples (Palmieri et al., 2007; Weathers, Keane, Davidson, 2001). Although it seems that the ability for clarification and elaboration is important, it may not be as pivotal as thought for an accurate assessment.

One of the most significant reasons for discrepancy was participants answering based on

general distress instead of reporting symptoms from the past month. Marmar et al. (2015) suggested that self-reports could be more influenced by general distress compared to clinical interviews. It appears that this theme became more frequent for the PTSD symptoms in the latter half of the assessment. Since the CAPS-5 inquires about symptoms from the past-month with each question, it could be beneficial that the PCL-5 includes a repetition of the instructions around the latter half to ensure that the participants are reminded of the time-frame.

Additionally, participants struggled to comprehend some of the symptoms. Specifically, they struggled to differentiate between intrusive memories and flashbacks of the traumatic memory. While self-reporting, participants may lack the ability to distinguish between vivid recollections and dissociative states where they feel as if the event was actually happening again (Marmar et al., 2015; Moshier et al., 2019). Psychopathology can impact an individual's ability to reflect on their symptomatology (e.g., Enns, Larsen, & Cox, 2000). PTSD symptoms are also conceptually complex, which makes them difficult for respondents to comprehend the meaning of the items. Furthermore, Schwarz (1999) emphasizes the importance that the respondent's understanding of the questions matches what the researcher had in mind. Therefore, due to the complexity of some of the symptoms of PTSD, it is recommended for revision to the wording of some of the more conceptually complex symptoms (i.e., symptoms in the reexperiencing cluster).

Participants also struggled to recognize how the traumatic event had impacted their daily functioning. Some PTSD symptoms can be connected to everyday struggles (i.e., exhibiting caution in dangerous situations and concentration and sleeping difficulties during stressful times). When assessing PTSD symptomatology, it is important to consider more than subjective distress. Assessment of symptoms should also be focused on the nature and pervasiveness of behaviors (Moshier et al., 2019). Individuals reported that some of their responses were based on

how they had always acted in certain situations. Helping participants distinguish between internal characteristics and the external contribution of the traumatic event could lead to a more accurate assessment of a PTSD diagnosis.

Another pervasive theme was the ability to have more time to reflect on symptoms. Enns, Larson, and Cox (2000) suggested that individuals can struggle to think through changes in their symptomatology. Structured interviews allow for a more detailed exploration of symptomatology, as respondents are able to discuss their experiences with a trained clinician. Therefore, they provide the opportunity for individuals to process their thoughts, especially regarding more emotionally charged symptoms (i.e., memories, avoidance of internal and external reminders, and blame).

Last, participants reported making several general errors throughout responding to both the PCL-5 and the CAPS-5. For example, they reported forgetting relevant experiences, selecting the wrong answer, or not reading or hearing the entire question. Specifically, when filling out the PCL-5, they reported considering frequency of their symptoms. Having participants report on their symptoms assumes that respondents have sufficient memory and insight to provide accurate reports and are motivated to respond accurately. However, participants reported frequently that their responses were unknowingly inaccurate or dishonest.

The present study did have several limitations. First, participants were a convenience sample of college students, and thus the generalizability of the results to clinical samples is unknown. However, all participants were trauma-exposed and reported PCL-5 symptom scores in the moderate to severe range. Additionally, generalizability of the current results is also limited with respect to demographic variables such as age, race, ethnicity, and socioeconomic status, and future research using samples with greater diversity is warranted. Furthermore, given

the cross-sectional design, longitudinal research is needed to build on these initial findings. It would be important to see if the findings replicate in a treatment outcome research study. Further limiting these findings, administration of the CAPS-5 and the PCL-5 was not counterbalanced, potentially leading to order effects. Future research should examine if assessment order has an effect on quantitative and qualitative discrepancies between the CAPS-5 and the PCL-5.

The present study, along with previous research, provides evidence that the PCL-5 and the CAPS-5 are strongly associated, but are not interchangeable. It is recommended that researchers and clinicians are aware of both the concordance and the discordance of the PCL-5 and the CAPS-5 to further understand the utility of the measures. The PCL-5 is useful as a screening tool and to identify individuals who are symptomatic to warrant further assessment. It is also comparable to the CAPS-5 as a dimensional measure of the construct of PTSD (Lee et al., 2019). Finally, the PCL-5 is an efficient, inexpensive method for assessing PTSD.

However, the PCL-5 yields higher prevalence estimates (Marmar et al., 2015) and higher ratings of symptom severity compared to the CAPS-5. Furthermore, it is unclear which measure is more sensitive to clinical change because of contradictory findings from Forbes et al. (2001) and Monson et al. (2008). However, it is clear that they are differentially sensitive to change. Given some of the limitations of the PCL-5 in representing the complexity of the PTSD symptom criteria, the CAPS-5 should be favored for establishing diagnostic status because it allows individual to think through their symptoms, evaluate respondents' self-reported examples for conceptual fit to *DSM-5* criteria, and use clinical judgment to make final ratings.

In conclusion, this study provides evidence for moderate concordance between the PCL-5 and the CAPS-5. However, a considerable number of discrepancies was still evident across individual symptoms, cluster scores, and total score. Additionally, the study provides

quantitative and qualitative predictors of sources of discordance between the PCL-5 and the CAPS-5. Therefore, the present study provides insight to several way that can improve the measures' diagnostic and clinical utility.

References

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

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

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

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. doi: 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. doi: 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. doi: 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. doi: 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. doi: 10.1037/pas0000504

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

Brewin, C. R. (2005). Systematic review of screening instruments for adults at risk of PTSD. *Journal of Traumatic Stress, 18,* 53−62. doi: 10.1002/jts.20007

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

Davidson, J.R.T., Book, S.W., Colket, J.T., Tupler, L.A., Roth, S.H., David, D., Hertzberg,… & Feldman, M.E. (1997). Assessment of a new self-rating scale for posttraumatic stress disorder. *Psychological Medicine, 27*, 153–160. doi: 10.1017/s0033291796004229

Davidson, J., Smith, R., & Kudler, H. (1989). Validity and reliability of the DSM-III criteria for posttraumatic stress disorder: Experience with a structured interview. *Journal of Nervous and Mental Disease*, *177,* 336-341. doi: 10.1097/00005053-198906000-00003

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. doi: 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. doi: 10.1002/jts.20062

Enns, M. W., Larsen, D. K., & Cox, B. J. (2000). Discrepancies between self and observer

ratings of depression: The relationship to demographic, clinical and personality variables. *Journal of Affective Disorders, 60,* 33–41. doi: 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. doi:10.7205/milmed.172.4.359

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

Foa, E.B., Riggs, D.S., Dancu, C.V., & Rothbaum, B.O. (1993). Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *Journal of Traumatic Stress, 6,* 459–473. doi:10.1007/bf00974317

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

Frisch, M. B. (1998). Social desirability responding in the assessment of social skill and anxiety. *Psychological Reports, 63,* 763–766. doi:10.2466/pr0.1988.63.3.763

Frueh, B. C., Elhai, J. D., Gold, P. B., Monnier, J., Magruder, K. M., Keane, T. M., & Arana, G. W. (2003). Disability compensation seeking among veterans evaluated for posttraumatic stress disorder. *Psychiatric Services*, *54*, 84-91. doi: 10.1176/appi.ps.54.1.84

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

Hickling, E. J., Blanchard, E. B., Mundy, E., & Galovski, T. E. (2002). Detection of malingered

MVA related posttraumatic stress disorder: An investigation of the ability to detect professional actors by experienced clinicians, psychological tests and psychophysiological assessment. *Journal of Forensic Psychology Practice*, *2*, 33-53. doi: 10.1300/J158v02n01_02

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. *The New England Journal of Medicine, 351*, 13−22. doi:10.1056/nejmoa040603

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. doi: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. doi:10.1037/1040-3590.20.1.81

Horowitz, M.J., Wilner, N., & Alvarez, W. (1979). Impact of Event Scale: A measure of subjective stress. *Psychosomatic Medicine, 41,* 209–218. doi:10.1097/00006842-197905000-00004

Hunt, C., & Andrews, G. (1992). Measuring personality disorder: The use of self-report questionnaires. *Journal of Personality Disorders*, *6*, 125-133. doi: 10.1521/pedi.1992.6.2.125

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). New York: Guilford Press.

Keen, S. M., Kutter, C. J., Niles, B. L., & Krinsley, K. E. (2008). Psychometric properties of PTSD checklist in sample of male veterans. *Journal of Rehabilitation Research and Development, 45,* 465−474. doi:10.1682/jrrd.2007.09.0138

Kraemer, H. C. (1992). *Evaluating medical tests: Objective and quantitative guidelines.* Newbury Park, CA: Sage.

Lee, D. J., Bovin, M. J., Weathers, F. W., Palmieri, P. A., Schnurr, P. P., Sloan, D. 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. doi: 10.1037/pas0000642

Marmar, C. R., Schlenger, W., Henn-Haase, C., Qian, M., Purchia, E., Li, M., . . . 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. doi:10.1001/jamapsychiatry.2015.0803

McDevitt-Murphy, M. E., Weathers, F. W., Flood, A. M., Eakin, D. E., & Benson, T. A. (2007). The utility of the PAI and the MMPI-2 for discriminating PTSD, depression, and social phobia in trauma-exposed college students. *Assessment, 14,* 181–195. doi:10.1177/1073191106295914

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

McNally, R. J., & Frueh, B. C. (2013). Why are Iraq and Afghanistan War veterans seeking PTSD disability compensation at unprecedented rates?. *Journal of Anxiety Disorders*, *27*, 520-526. doi: 10.1016/j.janxdis.2013.07.002

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. doi:10.1037/e558112009-009

Moshier, S. J., Bovin, M. J., Gay, N. G., Wisco, B. E., Mitchell, K. S., Lee, D. J., ... & 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. doi: 10.1037/abn0000368

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

Morey, L.C. (2007). The Personality Assessment Inventory Professional Manual. Odessa, FL: Psychological Assessment Resources.

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. doi:10.1037/0021-843x.116.2.329

Parker-Guilbert, K. S., Leifker, F. R., Sippel, L. M., & Marshall, A. D. (2014). The differential diagnostic accuracy of the PTSD Checklist among men versus women in a community sample. *Psychiatry Research*, *220*, 679-686. doi:10.1016/j.psychres.2014.08.001

Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality* (pp. 224–239). London, England: Guilford.

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. doi: 10.3200/JACH.58.1.73-82

Reardon, A. F., Brief, D., Miller, M., & Keane, T. (2014). Assessment of PSTD and its comorbidities in adults. In M.J. Friedman, T.M.Keane, & P.A. Resick (Eds.), *Handbook of PTSD. Science and Practice* (pp. 269 –290). New York: Guilford Press.

Resnick, H. S., Kilpatrick, D. G., & Lipovsky, J. A. (1991). Assessment of rape-related posttraumatic stress disorder: Stressor and symptom dimensions. *Psychological Assessment, 3*, 561-572. doi:10.1037/1040-3590.3.4.561

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54,* 93–105. doi: 10.1037/0003-066X.54.2.93

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

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.

Van Etten, M. L., & Taylor, S. (1998). Comparative efficacy of treatments for post-traumatic stress disorder: A meta-analysis. *Clinical Psychology & Psychotherapy, 5,* 126−144.

Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96,* 465-490. doi: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. doi: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)*. Boston, MA: 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. (2017). The Clinician-Administered PTSD Scale for DSM–5 (CAPS-5): Development and initial psychometric evaluation in military veterans. *Psychological Assessment, 30,* 383-395. doi: 10.1037/pas0000486

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). Boston: Allyn & Bacon.

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. doi: https://doi.org/10.1002/da.1029

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., Keane, T. M. (1993). The PTSD checklist (PCL): reliability, validity, and diagnostic utility. Paper presented at the Ninth Annual Conference of the ISTSS, San Antonio.

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). Boston, MA: 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. doi: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. doi:10.1037/1040-3590.11.2.124

Weiss, D.S., Marmar, C.R. (1997). The Impact of Event Scale – Revised. In J.P. Wilson & T.M.

Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 399–411). New York: Guilford Press.

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. doi: 10.1002/da.20837

Wortmann, J. H., Jordan, A. H., Weathers, F. W., Resick, P. A., Dondanville, K. A., Hall-Clark, B., ... Litz, B. T. (2016). Psychometrics analysis of the PTSD Checklist-5 (PCL-5) among treatment-seeking military members. *Psychological Assessment*, *28*, 1392–1403. doi:10.1037/pas0000260

Table 1

Correlations and Mean Differences Between PCL-5 and CAPS-5 Intensity and Severity Scores

	PCL-5		CAPS-5 Intensity		PCL-5 vs. CAPS-5 Intensity		CAPS-5 Severity		PCL-5 vs. CAPS-5 Severity	
Item/cluster	*M*	*SD*	*M*	*SD*	*r*	*T*	*M*	*SD*	*r*	*t*
1. B1	1.78	1.09	1.53	1.02	.55**	1.93	1.45	.98	.60**	2.77*
2. B2	1.27	1.27	1.20	1.23	.74**	.57	1.08	1.12	.77**	1.70
3. B3	1.23	1.09	.62	.94	.43**	4.38**	.48	.75	.46**	5.80**
4. B4	2.58	1.15	1.75	1.00	.50**	5.91**	1.57	.85	.56**	8.01**
5. B5	2.22	1.21	1.68	1.00	.49**	3.67*	1.60	.96	.57**	4.65**
6. C1	2.80	1.18	1.92	1.01	.40**	5.67**	1.82	1.00	.34**	6.07**
7. C2	2.48	1.32	1.27	1.19	.44**	7.09**	1.18	1.13	.47**	7.95**
8. D1	1.63	1.54	.90	1.30	.61**	4.46**	.88	1.28	.62**	4.67**
9. D2	1.70	1.32	1.72	1.33	.54**	-.10	1.65	1.29	.52**	.30
10. D3	2.13	1.48	1.43	1.33	.51**	3.89**	1.30	1.24	.55**	4.95**
11. D4	2.00	1.25	1.73	.97	.48**	1.77	1.57	.91	.54**	3.11*
12. D5	1.13	1.23	.87	1.26	.47**	1.62	.73	1.12	.51**	2.65*
13. D6	1.22	1.11	1.27	1.16	.47**	-.33	1.22	1.12	.51**	.00
14. D7	1.13	1.19	.92	1.06	.62**	1.69	.88	1.04	.60**	1.93
15. E1	.83	1.06	.77	1.06	.64**	.57	.72	1.01	.60**	.98
16. E2	.52	1.05	.27	.78	.76**	2.85*	.23	.67	.81**	3.43*
17. E3	2.30	1.25	1.65	1.12	.43**	3.95**	1.53	1.08	.51**	5.08**
18. E4	1.72	1.28	.90	1.12	.41**	4.84**	.83	1.03	.47**	5.66**
19. E5	1.72	1.28	1.07	1.15	.38**	3.73**	1.05	1.14	.38**	3.83**
20. E6	1.65	1.40	1.47	1.28	.58**	1.16	1.35	1.19	.59**	1.97
Cluster B	9.08	4.58	6.78	3.65	.67**	5.17**	6.18	3.28	.76**	7.53**
Cluster C	5.28	2.20	3.18	1.75	.52**	8.21**	3.00	1.67	.51**	9.00**
Cluster D	10.95	6.42	8.83	5.75	.63**	3.12*	8.23	5.43	.65**	4.17**
Cluster E	8.73	4.64	6.12	3.82	.57**	5.07**	5.72	3.62	.60**	6.11**
Total Score	34.05	14.34	24.92	12.22	.63**	6.10**	23.13	11.36	.69**	8.05**

Note. $N = 60$. PCL-5 =PTSD Checklist for *DSM-5*; CAPS-5 = Clinician-Administered PTSD Scale for *DSM-5;* For the purpose of the analyses, CAPS-5 Intensity anchors were coded the following values: 0 = *absent*, 1= *minimal*, 2 = *clearly present*, 3 = *pronounced*, and 4 = *extreme*; *$p < .05$; **$p < .001$.

Table 2

Frequency of Raw Discrepancies between PCL-5 Item Scores and CAPS-5 Intensity Scores

	Raw Discrepancy								
	-4	-3	-2	-1	0	+1	+2	+3	+4
Item	*n*(%)	*n*(%)	*n*(%)	*n*(%)	*n*(%)	*n*(%)	*n*(%)	*n*(%)	*n*(%)
1. B1	0(0)	0(0)	2(3.3)	11(18.3)	24(40.0)	16(26.7)	7(11.7)	0(0)	0(0)
2. B2	0(0)	0(0)	3(5.0)	9(15.0)	32(53.3)	14(23.3)	1(1.7)	1(1.7)	0(0)
3. B3	0(0)	1(1.7)	1(1.7)	3(5.0)	24(40.0)	19(31.7)	10(16.7)	2(3.3)	0(0)
4. B4	0(0)	0(0)	0(0)	5(8.3)	20(33.3)	20(33.3)	11(18.3)	3(5.0)	1(1.7)
5. B5	0(0)	0(0)	3(5.0)	7(11.7)	16(26.7)	25(41.7)	8(13.3)	0(0)	1(1.7)
6. C1	0(0)	0(0)	2(3.3)	3(5.0)	19(31.7)	18(30.0)	13(21.7)	4(6.7)	1(1.7)
7. C2	0(0)	0(0)	1(1.7)	1(1.7)	20(33.3)	15(25.0)	11(18.3)	9(5.0)	3(5.0)
8. D1	0(0)	0(0)	1(1.7)	4(6.7)	27(45.0)	16(26.7)	5(8.3)	4(6.7)	3(5.0)
9. D2	0(0)	1(1.7)	9(15.0)	6(10.0)	26(43.3)	11(18.3)	6(10.0)	1(1.7)	0(0)
10. D3	0(0)	2(3.3)	3(5.0)	2(3.3)	18(30.0)	20(33.3)	10(16.7)	4(6.7)	1(1.7)
11. D4	0(0)	1(1.7)	3(5.0)	11(18.3)	17(28.3)	20(33.3)	8(13.3)	0(0)	0(0)
12. D5	0(0)	1(1.7)	5(8.3)	4(6.7)	29(48.3)	13(21.7)	5(8.3)	2(3.3)	1(1.7)
13. D6	0(0)	1(1.7)	5(8.3)	13(21.7)	25(41.7)	10(16.7)	5(8.3)	1(1.7)	0(0)
14. D7	0(0)	0(0)	2(3.3)	7(11.7)	34(56.7)	13(21.7)	2(3.3)	1(1.7)	1(1.7)
15. E1	0(0)	0(0)	5(8.3)	4(6.7)	36(60.0)	12(20.0)	3(5.0)	0(0)	0(0)
16. E2	0(0)	0(0)	1(1.7)	1(1.7)	44(73.3)	10(16.7)	4(6.7)	0(0)	0(0)
17. E3	0(0)	0(0)	2(3.3)	8(13.3)	19(31.7)	16(26.7)	12(20.0)	1(1.7)	2(3.3)
18. E4	0(0)	1(1.7)	1(1.7)	3(5.0)	21(35.0)	19(31.7)	10(16.7)	2(3.3)	3(5.0)
19. E5	0(0)	0(0)	3(5.0)	7(11.7)	20(33.3)	16(26.7)	7(11.7)	6(10.0)	1(1.7)
20. E6	0(0)	2(3.3)	4(6.7)	5(8.3)	27(45.0)	16(26.7)	5(8.3)	0(0)	1(1.7)

Note. $N = 60$. PCL-5 = PTSD Checklist for *DSM-5*; CAPS-5 = Clinician-Administered PTSD Scale for *DSM-5*; Discrepancy scores for a given item were calculated by subtracting the CAPS-5 Intensity score from the PCL-5 item score.

Table 3

Frequency of Raw Discrepancies between PCL-5 Item Scores and CAPS-5 Severity Scores

	Raw Discrepancy								
	-4	-3	-2	-1	0	+1	+2	+3	+4
Item	*n*(%)	*n*(%)	*n*(%)	*n*(%)	*n*(%)	*n*(%)	*n*(%)	*n*(%)	*n*(%)
1. B1	0(0)	0(0)	1(1.7)	9(15.0)	26(43.3)	17(28.3)	7(11.7)	0(0)	0(0)
2. B2	0(0)	0(0)	1(1.7)	8(13.3)	34(56.7)	14(23.3)	2(3.3)	1(1.7)	0(0)
3. B3	0(0)	0(0)	2(3.3)	1(1.7)	22(36.7)	22(36.7)	11(18.3)	2(3.3)	0(0)
4. B4	0(0)	0(0)	0(0)	2(3.3)	16(26.7)	26(43.3)	12(20.0)	3(5.0)	1(1.7)
5. B5	0(0)	0(0)	1(1.7)	7(11.7)	17(28.3)	26(43.3)	8(13.3)	0(0)	1(1.7)
6. C1	0(0)	0(0)	2(3.3)	3(5.0)	17(28.3)	18(30.0)	13(21.7)	6(10.0)	1(1.7)
7. C2	0(0)	0(0)	0(0)	1(1.7)	19(31.7)	17(28.3)	10(16.7)	10(16.7)	3(5.0)
8. D1	0(0)	0(0)	0(0)	5(8.3)	27(45.0)	16(26.7)	5(8.3)	4(6.7)	3(5.0)
9. D2	0(0)	1(1.7)	9(15.0)	5(8.3)	24(40.0)	14(23.3)	6(10.0)	1(1.7)	0(0)
10. D3	0(0)	1(1.7)	2(3.3)	3(5.0)	18(30.0)	18(30.0)	13(21.7)	4(6.7)	1(1.7)
11. D4	0(0)	0(0)	3(5.0)	8(13.3)	19(31.7)	20(33.3)	10(16.7)	0(0)	0(0)
12. D5	0(0)	0(0)	3(5.0)	7(11.7)	25(41.7)	17(28.3)	5(8.3)	2(3.3)	1(1.7)
13. D6	0(0)	0(0)	5(8.3)	13(21.7)	26(43.3)	10(16.7)	5(8.3)	1(1.7)	0(0)
14. D7	0(0)	0(0)	2(3.3)	6(10.0)	35(58.3)	12(20.0)	3(5.0)	1(1.7)	1(1.7)
15. E1	0(0)	0(0)	5(8.3)	4(6.7)	33(55.0)	15(25.0)	3(5.0)	0(0)	0(0)
16. E2	0(0)	0(0)	0(0)	1(1.7)	43(71.7)	11(18.3)	4(6.7)	0(0)	0(0)
17. E3	0(0)	0(0)	0(0)	7(11.7)	21(35.0)	16(26.7)	13(21.7)	1(1.7)	2(3.3)
18. E4	0(0)	0(0)	1(1.7)	3(5.0)	21(35.0)	20(33.3)	10(16.7)	2(3.3)	3(5.0)
19. E5	0(0)	0(0)	3(5.0)	7(11.7)	19(31.7)	17(28.3)	7(11.7)	6(10.0)	1(1.7)
20. E6	0(0)	1(1.7)	3(5.0)	6(10.0)	28(46.7)	13(21.7)	8(13.3)	0(0)	1(1.7)

Note. $N = 60$. PCL-5 = PTSD Checklist for *DSM-5*; CAPS-5 = Clinician-Administered PTSD Scale for *DSM-5*; Discrepancy scores for a given item were calculated by subtracting the CAPS-5 Severity score from the PCL-5 item score.

Table 4

Frequency of Any Discrepancy and Mean Discrepancy between PCL-5 and CAPS-5 Intensity and Severity Scores

	CAPS-5 Intensity			CAPS-5 Severity		
Item/cluster	Freq *n* (%)	Mean	SD	Freq *n*(%)	Mean	SD
1. B1	36(60.0)	.75	.70	34(56.7)	.70	.70
2. B2	28(46.7)	.57	.70	26(43.3)	.52	.68
3. B3	36(60.0)	.89	.88	38(63.3)	.92	.85
4. B4	40(66.7)	1.00	.94	44(73.3)	1.08	.91
5. B5	44(73.3)	.97	.78	43(71.7)	.92	.77
6. C1	41(68.3)	1.12	.99	43(71.7)	1.22	1.03
7. C2	40(66.7)	1.32	1.23	41(68.3)	1.33	1.23
8. D1	31(55.0)	.93	1.13	33(55.0)	.92	1.12
9. D2	34(56.7)	.88	.90	36(60.0)	.92	.89
10. D3	42(70.0)	1.17	1.03	42(70.0)	1.17	1.01
11. D4	43(71.7)	.93	.73	41(68.3)	.90	.73
12. D5	31(51.7)	.83	.99	35(58.3)	.83	.91
13. D6	35(58.3)	.82	.83	34(56.7)	.77	.79
14. D7	26(43.3)	.58	.83	25(41.7)	.58	.85
15. E1	24(40.0)	.53	.72	27(45.0)	.58	.72
16. E2	16(26.7)	.35	.63	17(28.3)	.35	.61
17. E3	41(68.3)	1.05	.96	39(65.0)	1.00	.97
18. E4	39(65.0)	1.08	1.09	39(65.0)	1.05	1.06
19. E5	40(66.7)	1.08	1.03	41(68.3)	1.10	1.02
20. E6	33(55.0)	.82	.93	32(53.3)	.80	.92
Cluster B	58(96.7)	4.17	2.18	58(96.7)	4.13	2.24
Cluster C	53(88.3)	2.43	1.65	55(91.7)	2.50	1.70
Cluster D	100(100.0)	6.15	3.25	58(96.7)	6.08	3.28
Cluster E	58(96.7)	4.92	2.67	59(98.3)	4.88	2.57
Total Score	100(100.0)	17.67	6.93	100(100.0)	17.65	6.99

Note. N= 60. PCL-5=PTSD Checklist for *DSM-5;* CAPS-5 = Clinician Administered PTSD Scale for *DSM-5;* Freq= Frequency; For the purpose of the analyses, CAPS-5 Intensity anchors were coded the following values: 0 = *absent*, 1= *minimal*, 2 = *clearly present*, 3 = *pronounced*, and 4 = *extreme*

Table 5

Prevalence of PTSD Symptoms and Diagnosis Based on PCL-5 and CAPS-5

Item	PCL-5 prevalence *n*(%)	CAPS-5 prevalence *n*(%)	Hits *n*(%)	Misses *n*(%)	False Alarms *n*(%)	Correct Rejections *n*(%)	Efficiency *n*(%)	κ (.5)
1. B1	35(58.3)	35(58.3)	27(45.0)	8(13.3)	8(13.3)	17(28.3)	44(73.3)	.45
2. B2	23(38.3)	26(43.3)	18(30.0)	8(13.3)	5(8.3)	29(48.3)	47(78.3)	.56
3. B3	25(41.7)	9(15.0)	7(11.7)	2(3.3)	18(30.0)	33(55.0)	40(66.7)	.25
4. B4	49(81.7)	38(63.3)	37(61.7)	1(1.7)	12(20.0)	10(16.7)	47(78.3)	.45
5. B5	40(66.7)	40(66.7)	33(55.0)	7(11.7)	7(11.7)	13(21.7)	46(76.7)	.48
6. C1	51(85.0)	41(68.3)	38(63.3)	3(5.0)	13(21.7)	6(10.0)	44(73.3)	.25
7. C2	47(78.3)	28(46.7)	27(45.0)	1(1.7)	20(33.3)	12(20.0)	39(65.0)	.37
8. D1	27(45.0)	21(35.0)	16(26.7)	5(8.3)	11(18.3)	28(46.7)	44(73.3)	.45
9. D2	31(51.7)	36(60.0)	24(40.0)	12(20.0)	7(11.7)	17(28.3)	41(68.3)	.36
10. D3	40(66.7)	28(46.7)	23(38.3)	5(8.3)	17(28.3)	15(25.0)	38(63.3)	.57
11. D4	39(65.0)	37(61.7)	28(46.7)	9(15.0)	11(18.3)	12(20.0)	40(66.7)	.28
12. D5	19(31.7)	14(23.3)	9(15.0)	5(8.3)	10(16.7)	36(60.0)	45(75.0)	.38
13. D6	24(40.0)	26(43.3)	15(25.0)	11(18.3)	9(15.0)	25(41.7)	40(66.7)	.32
14. D7	21(35.0)	21(35.0)	16(26.7)	5(8.3)	5(8.3)	34(56.7)	50(83.3)	.63
15. E1	15(25.0)	15(25.0)	9(15.0)	6(10.0)	6(10.0)	39(65.0)	48(80.0)	.47
16. E2	6(10.0)	4(6.7)	3(5.0)	1(1.7)	3(5.0)	53(88.3)	56(93.3)	.64
17. E3	43(71.7)	36(60.0)	32(53.3)	4(6.7)	11(18.3)	13(21.7)	45(75.0)	.45
18. E4	33(55.0)	21(35.0)	18(30.0)	3(5.0)	15(25.0)	24(40.0)	42(70.0)	.39
19. E5	32(53.3)	24(40.0)	17(28.3)	7(11.7)	15(25.0)	21(35.0)	38(63.3)	.24
20. E6	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. *N*= 60. 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 6
Descriptive Statistics of Objective Predictors

Variable	*M*	*SD*	Variance	Skew	Kurtosis	Range
PAI NIM	58.81	14.07	197.982	1.46	2.55	44-110
PAI INC	52.78	7.42	55.12	.79	1.38	37-76
PAI INF	53.42	9.38	87.94	.88	.61	40-82
BFI CON	3.78	.67	.45	-.25	-.63	2-5
BFI NEURO	3.54	.83	.69	-.18	-.87	1-5
Verbal IQ	30.05	3.52	12.39	-.36	-.76	21-39

Note. N = 60. PAI = Personality Assessment Inventory; NIM= Negative Impression Management; INC= Inconsistency; INF= Infrequency; BFI= Big Five Inventory; CON= Conscientiousness; NEURO= Neuroticism; Verbal IQ= Shipley-2.

Table 7
Zero-Order Correlations for PCL-5 and CAPS-5 with Objective Predictor Variables

Item	PAI NIM	PAI INC	PAI INF	BFI CON	BFI NEURO	Verbal IQ
1 (B1)	-.03	-.10	.07	.11	-.03	.12
2 (B2)	.14	.01	-.06	.05	.20	-.13
3 (B3)	-.13	-.08	.07	-.10	-.07	-.07
4 (B4)	.01	.07	-.12	-.16	.11	-.05
5 (B5)	.42**	-.06	-.02	-.16	.04	-.25
6 (C1)	-.15	.20	-.10	-.05	.17	.06
7 (C2)	.17	-.03	-.06	-.15	.14	.11
8 (D1)	.08	.13	-.17	-.14	.18	-.03
9 (D2)	.02	.19	.11	.14	.00	.09
10 (D3)	-.01	.13	.09	-.20	.11	-.03
11 (D4)	.03	.07	.25	-.07	.07	.14
12 (D5)	.03	.11	-.09	.15	.20	.10
13 (D6)	-.05	.14	-.06	.22	.21	-.03
14 (D7)	.27*	.02	.19	.04	.07	-.14
15 (E1)	-.02	.16	.35**	.11	.24	-.19
16 (E2)	.36*	.11	-.08	-.15	.26*	-.10
17 (E3)	.34*	.10	.11	-.29*	.19	-.08
18 (E4)	.19	-.15	-.14	.03	.21	-.01
19 (E5)	.10	.14	.13	-.29*	.37**	.34**
20 (E6)	.08	.21	.08	-.32*	.27*	.10
Cluster B	.13	-.06	-.03	-.07	.08	-.13
Cluster C	.03	.10	-.10	-.14	.20	.12
Cluster D	.11	.22	.07	.02	.24	.02
Cluster E	.36**	.18	.14	.33*	.52**	.07
Total Score	.23	.17	.05	-.17	.38*	.02

Note. N= 60. PAI = Personality Assessment Inventory; NIM= Negative Impression Management; INC= Inconsistency; INF= Infrequency; BFI= Big Five Inventory; CON= Conscientiousness; NEURO= Neuroticism; Verbal IQ= Shipley-2. *$p < .05$; **$p < .01$.

Table 8

Linear Multiple Regression for Predicting Discrepancy Scores for Individual Items, Cluster, and Total PTSD

Variables	R^2	F	*P*	*Standardized β*	*t*	*p*
Cued Distress	.26	2.90	.02*			
PAI NIM				.46	3.35	.001**
PAI INC				-.04	-.27	.79
PAI INF				-.05	-.37	.71
BFI CON				-.13	-.95	.35
BFI NEURO				-.21	-1.44	.16
Verbal IQ				-.27	-2.08	.04*
Irritability or Aggressive Behavior	.31	3.67	.004*			
PAI NIM				-.10	-.72	.47
PAI INC				-.02	-.16	.87
PAI INF				.30	2.34	.02*
BFI CON				.25	1.95	.06
BFI NEURO				.50	3.58	.001**
Verbal IQ				-.14	-1.10	.28
Hypervigilance	.13	2.45	.04*			
PAI NIM				.35	2.48	.02*
PAI INC				.07	.50	.62
PAI INF				.05	.37	.71
BFI CON				-.26	-1.95	.06
BFI NEURO				-.04	-.26	.80
Verbal IQ				-.08	-.64	.52
Concentration	.33	4.18	.001**			
PAI NIM				-.09	-.66	.51
PAI INC				.11	.85	.40
PAI INF				.17	1.40	.17
BFI CON				-.08	-.61	.55
BFI NEURO				.35	2.62	.01*
Verbal IQ				.40	3.31	.001**

Arousal and Reactivity Cluster	.45	6.92	.001**			
PAI NIM				.21	1.73	.09
PAI INC				.11	.92	.36
PAI INF				.08	.69	.50
BFI CON				-.14	-1.24	.22
BFI NEURO				.46	3.73	.001*
Verbal IQ				.09	.78	.44
Total Score	.24	2.61	.03*			
PAI NIM				.15	1.04	.31
PAI INC				-.14	.10	.32
PAI INF				.00	.00	1.0
BFI CON				-.06	-.43	.67
BFI NEURO				.35	2.41	.02*
Verbal IQ				.04	.32	.75

Note. N=60. PAI = Personality Assessment Inventory; NIM= Negative Impression Management; INC= Inconsistency; INF= Infrequency; BFI= Big Five Inventory; CON= Conscientiousness; NEURO= Neuroticism; Verbal IQ= Shipley-2. *$p < .05$; **$p < .001$.

Table 9

Frequencies of the Themes of Discrepancies between the PCL-5 and the CAPS-5: Reexperiencing and Avoidance Clusters

Themes	B1 #(%)	B2 #(%)	B3 #(%)	B4 #(%)	B5 #(%)	C1 #(%)	C2 #(%)
#	61	25	49	54	54	48	56
1. Item Order	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	3(5.4)
2. Response Options	5(8.2)	2(8.0)	1(2.0)	4(7.4)	2(3.7)	2(4.2)	1(1.8)
3. Time- Frame	4(6.6)	1(4.0)	4(8.2)	4(7.4)	2(3.7)	5(10.4)	5(8.9)
4. Trauma-Related	1(1.6)	0(0)	1(2.0)	1(1.9)	1(1.9)	0(0)	1(1.8)
5. Clarification	5(8.2)	1(4.0)	3(6.1)	0(0)	1(1.9)	0(0)	2(3.6)
6. Elaboration	7(11.5)	1(4.0)	4(8.2)	5(9.3)	1(1.9)	2(4.2)	1(1.8)
7. Length and Complex	5(8.2)	0(0)	0(0)	0(0)	0(0)	0(0)	1(1.8)
8. Clinical Judgment	1(1.6)	0(0)	1(2.0)	0(0)	0(0)	0(0)	1(1.8)
9. Oral vs. Written	2(3.3)	1(4.0)	0(0)	0(0)	1(1.9)	2(4.2)	1(1.8)
10. Comp. of Symptoms	9(14.8)	5(20.0)	14(28.6)	7(13.0)	9(16.7)	5(10.4)	4(7.2)
11. Item Content	1(1.6)	0(0)	0(0)	2(3.7)	1(1.9)	2(4.2)	3(5.4)
12. Avoid. of Extreme	0(0)	0(0)	1(2.0)	0(0)	0(0)	0(0)	0(0)
13. Minimization of Symptoms	1(1.6)	3(12.0)	2(4.1)	5(9.3)	6(11.1)	6(12.5)	5(8.9)
14. Avoid. of Questions	0(0)	0(0)	1(2.0)	3(5.6)	0(0)	1(2.1)	1(1.8)
15. Trauma-Related Error	0(0)	0(0)	3(6.1)	3(5.6)	4(7.4)	3(6.3)	3(5.6)
16. Deliberate Calibration	0(0)	1(4.0)	2(4.1)	4(7.4)	4(7.4)	0(0)	1(1.8)
17. Perception vs. Reality	0(0)	0(0)	0(0)	0(0)	3(5.6)	0(0)	5(8.9)
18. Increased Awareness	13(21.3)	4(16.0)	4(8.2)	6(11.1)	6(11.1)	8(16.7)	8(14.8)
19. Impression Management	0(0)	1(4.0)	1(2.0)	1(1.9)	0(0)	0(0)	0(0)
20. General Errors	6(9.8)	5(20.0)	7(14.3)	8(14.8)	10(18.5)	11(22.9)	9(16.7)
21. Clinical Error	1(1.6)	0(0)	0(0)	1(1.9)	3(5.6)	1(2.1)	1(1.8)

Note. Theme 1= Item Order; Theme 2= Response Options; Theme 3= Time-Frame Reminders; Theme 4= Trauma-Related Reminders; Theme 5= Opportunity for Clarification; Theme 6= Opportunity for Elaboration; Theme 7= Question Length and Complexity; Theme 8= Clinical Judgment; Theme 9= Oral vs. Written Comprehension; Theme 10= Comprehension of Symptoms; Theme 11= Item Content; Theme 12= Avoidance of Extreme Responses; Theme 13= Minimization of Symptomatology; Theme 14= Avoidance of Specific Questions; Theme 15= Trauma-Related Attribution Error; Theme 16= Deliberate Calibration; Theme 17= Perception vs. Reality; Theme 18= Increased Awareness; Theme 19= Impression Management; Theme 20= General Errors; Theme 21= Clinical Judgment Contributed to Error

Table 10

Frequencies of the Themes of Discrepancies between the PCL-5 and the CAPS-5: Negative Alterations in Cognition and Mood Cluster

Themes	D1 #(%)	D2 #(%)	D3 #(%)	D4 #(%)	D5 #(%)	D6 #(%)	D7 #(%)
#	35	45	61	63	35	43	34
1. Item Order	0(0)	0(0)	0(0)	2 (3.2)	0(0)	0(0)	1(2.9)
2. Response Options	0(0)	3(6.7)	1(1.6)	1(1.6)	0(0)	3(7.0)	1(2.9)
3. Time- Frame	3(8.6)	3(6.7)	4(6.6)	4(6.3)	5(14.3)	9(20.9)	5(14.7)
4. Trauma-Related	0(0)	0(0)	1(1.6)	5(7.9)	2(5.7)	4(9.3)	3(8.8)
5. Clarification	2(5.7)	0(0)	1(1.6)	4(6.3)	0(0)	4(9.3)	1(2.9)
6. Elaboration	2(5.7)	2(4.4)	1(1.6)	3(4.8)	3(8.6)	0(0)	0(0)
7. Length and Complex	1(2.9)	5(11.1)	4(6.6)	7(11.1)	0(0)	0(0)	0(0)
8. Clinical Judgment	1(2.9)	0(0)	0(0)	1(1.6)	1(2.9)	1(2.3)	2(5.9)
9. Oral vs. Written	0(0)	1(2.2)	2(3.3)	0(0)	0(0)	1(2.3)	0(0)
10. Comp. of Symptoms	5(14.3)	7(15.6)	4(6.6)	7(11.1)	6(17.1)	3(7.0)	4(11.8)
11. Item Content	5(14.3)	1(2.2)	14(23.0)	0(0)	0(0)	1(2.3)	1(2.9)
12. Avoid. of Extreme	0(0)	0(0)	0(0)	0(0)	1(2.9)	0(0)	2(5.9)
13. Minimization of Symptoms	3(8.6)	4(8.9)	4(6.6)	5(7.9)	4(11.4)	3(7.0)	1(2.9)
14. Avoid. of Questions	1(2.9)	1(2.2)	0(0)	1(1.6)	1(2.9)	1(2.3)	0(0)
15. Trauma-Related Error	0(0)	3(6.7)	0(0)	3(4.8)	4(11.4)	3(7.0)	3(8.8)
16. Deliberate Calibration	0(0)	0(0)	0(0)	3(4.8)	1(2.9)	0(0)	0(0)
17. Perception vs. Reality	0(0)	2(4.4)	2(3.3)	5(7.9)	1(2.9)	0(0)	1(2.9)
18. Increased Awareness	4(11.4)	6(13.3)	10(16.4)	7(11.1)	3(8.6)	7(16.3)	3(8.8)
19. Impression Management	1(2.9)	2(4.4)	6(9.8)	2(3.2)	0(0)	0(0)	3(8.8)
20. General Errors	5(14.3)	3(6.7)	7(11.4)	2(3.2)	3(8.6)	2(4.7)	3(8.8)
21. Clinical Error	2(5.7)	2(4.4)	0(0)	1(1.6)	0(0)	1(2.9)	0(0)

Note. Theme 1= Item Order; Theme 2= Response Options; Theme 3= Time-Frame Reminders; Theme 4= Trauma-Related Reminders; Theme 5= Opportunity for Clarification; Theme 6= Opportunity for Elaboration; Theme 7= Question Length and Complexity; Theme 8= Clinical Judgment; Theme 9= Oral vs. Written Comprehension; Theme 10= Comprehension of Symptoms; Theme 11= Item Content; Theme 12= Avoidance of Extreme Responses; Theme 13= Minimization of Symptomatology; Theme 14= Avoidance of Specific Questions; Theme 15= Trauma-Related Attribution Error; Theme 16= Deliberate Calibration; Theme 17= Perception vs. Reality; Theme 18= Increased Awareness; Theme 19= Impression Management; Theme 20= General Errors; Theme 21= Clinical Judgment Contributed to Error

Table 11

Frequencies of the Themes of Discrepancies between the PCL-5 and the CAPS-5: Arousal and Reactivity Cluster

Themes	E1 #(%)	E2 #(%)	E3 #(%)	E4 #(%)	E5 #(%)	E6 #(%)
#	29	21	36	44	48	37
1. Item Order	0(0)	0(0)	0(0)	0(0)	2 (4.2)	0(0)
2. Response Options	1(3.4)	0(0)	2(5.6)	0(0)	0(0)	1(2.7)
3. Time- Frame	2(6.9)	2(9.5)	1(2.8)	6(13.6)	6(12.5)	3(8.1)
4. Trauma-Related	3(10.3)	0(0)	0(0)	6(13.6)	8(16.7)	9(24.3)
5. Clarification	0(0)	1(4.8)	1(2.8)	2(4.5)	0(0)	0(0)
6. Elaboration	3(10.3)	2(9.5)	1(2.8)	3(6.8)	1(2.1)	0(0)
7. Length and Complex	2(6.9)	0(0)	0(0)	0(0)	0(0)	1(2.7)
8. Clinical Judgment	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
9. Oral vs. Written	0(0)	0(0)	0(0)	0(0)	4(4.8)	0(0)
10. Comp. of Symptoms	5(17.2)	4(19.0)	3(8.3)	5(11.4)	2(4.2)	4(10.8)
11. Item Content	5(17.2)	0(0)	4(11.1)	1(2.3)	0(0)	1(2.7)
12. Avoid. of Extreme	1(3.4)	2(9.5)	0(0)	0(0)	0(0)	1(2.7)
13. Minimization of Symptoms	2(6.9)	0(0)	2(5.6)	0(0)	2(4.2)	2(5.4)
14. Avoid. of Questions	0(0)	0(0)	0(0)	2(4.5)	1(2.1)	0(0)
15. Trauma-Related Error	0(0)	2(9.5)	8(22.2)	5(11.4)	6(12.5)	10(27.0)
16. Deliberate Calibration	0(0)	0(0)	0(0)	1(2.3)	1(2.1)	0(0)
17. Perception vs. Reality	0(0)	0(0)	1(2.8)	1(2.3)	0(0)	0(0)
18. Increased Awareness	3(10.3)	2(9.5)	7(19.4)	3(6.8)	8(16.7)	2(5.4)
19. Impression Management	0(0)	1(4.8)	0(0)	0(0)	0(0)	0(0)
20. General Errors	2(6.9)	4(19.0)	3(8.3)	8(18.2)	7(14.6)	3(8.1)
21. Clinical Error	0(0)	1(4.8)	3(8.3)	1(2.3)	0(0)	0(0)

Note. Theme 1= Item Order; Theme 2= Response Options; Theme 3= Time-Frame Reminders; Theme 4= Trauma-Related Reminders; Theme 5= Opportunity for Clarification; Theme 6= Opportunity for Elaboration; Theme 7= Question Length and Complexity; Theme 8= Clinical Judgment; Theme 9= Oral vs. Written Comprehension; Theme 10= Comprehension of Symptoms; Theme 11= Item Content; Theme 12= Avoidance of Extreme Responses; Theme 13= Minimization of Symptomatology; Theme 14= Avoidance of Specific Questions; Theme 15= Trauma-Related Attribution Error; Theme 16= Deliberate Calibration; Theme 17= Perception vs. Reality; Theme 18= Increased Awareness; Theme 19= Impression Management; Theme 20= General Errors; Theme 21= Clinical Judgment Contributed to Error

Appendix A

Table A1

Item-Level Descriptive Statistics for the PCL-5

PCL-5 item	Item description	*M*	*SD*	Variance	Skew	Kurtosis	Range
1 (B1)	Memories	1.78	1.09	1.19	.21	-.53	0-4
2 (B2)	Dreams	1.27	1.27	1.62	.60	-.88	0-4
3 (B3)	Flashbacks	1.23	1.09	1.20	.40	-.82	0-4
4 (B4)	Cued Distress	2.58	1.15	1.33	-.48	-.55	0-4
5 (B5)	Cued physical reactions	2.22	1.21	1.46	-.08	-1.06	0-4
6 (C1)	Avoiding internal reminders	2.80	1.18	1.38	-.76	-.26	0-4
7 (C2)	Avoiding external reminders	2.48	1.32	1.75	-.65	-.63	0-4
8 (D1)	Amnesia	1.63	1.54	2.37	.42	-1.34	0-4
9 (D2)	Negative beliefs	1.70	1.32	1.74	.21	-1.14	0-4
10 (D3)	Blame	2.13	1.48	2.19	-.27	-1.33	0-4
11 (D4)	Negative feelings	2.00	1.25	1.56	-.27	-1.06	0-4
12 (D5)	Loss of interest	1.13	1.23	1.51	.93	-.05	0-4
13 (D6)	Detachment or estrangement	1.22	1.11	1.22	.56	-.39	0-4
14 (D7)	Numbing	1.13	1.19	1.41	.81	-.21	0-4
15 (E1)	Irritability or aggressive behavior	.83	1.06	1.12	.96	-.42	0-3
16 (E2)	Reckless behavior	.52	1.05	1.10	2.37	4.98	0-4
17 (E3)	Hypervigilance	2.30	1.25	1.16	-.17	-.98	0-4
18 (E4)	Startle	1.72	1.28	1.16	.20	-.97	0-4
19 (E5)	Concentration	1.72	1.28	1.16	.15	-1.12	0-4
20 (E6)	Sleep	1.65	1.40	1.97	.28	-1.28	0-4
Cluster B	Reexperiencing	9.08	4.58	20.93	.38	-.35	0-20
Cluster C	Avoidance	5.28	2.20	4.85	-.64	-.53	0-8
Cluster D	NACM	10.95	6.42	41.20	.09	-.86	0-25
Cluster E	Hyperarousal	8.73	4.64	21.49	.21	-.44	0-21
Total score		34.05	14.34	205.85	.12	-.49	6-61

Note. *N*= 60. PCL-5 =PTSD Checklist for *DSM-5*; NACM = Negative Alterations in Cognitions and Mood.

Table A2
Item-Level Descriptive Statistics for the CAPS-5

CAPS-5 item	Item description	*M*	*SD*	Variance	Skew	Kurtosis	Range
1 (B1)	Memories	1.45	.98	.96	-.13	-.52	0-4
2 (B2)	Dreams	1.08	1.12	1.26	.35	-1.44	0-3
3 (B3)	Flashbacks	.48	.75	.56	1.19	-.12	0-2
4 (B4)	Cued Distress	1.57	.85	.72	-.56	-.38	0-3
5 (B5)	Cued physical reactions	1.60	.96	.92	-.53	-.70	0-3
6 (C1)	Avoiding internal reminders	1.82	1.00	1.00	-.35	-.49	0-4
7 (C2)	Avoiding external reminders	1.18	1.13	1.27	.21	-1.47	0-3
8 (D1)	Amnesia	.88	1.28	1.63	.98	-.54	0-4
9 (D2)	Negative beliefs	1.65	1.29	1.66	-.20	-1.52	0-4
10 (D3)	Blame	1.30	1.24	1.54	.28	-1.36	0-4
11 (D4)	Negative feelings	1.57	.91	.83	-.42	-.62	0-3
12 (D5)	Loss of interest	.73	1.12	1.25	1.31	-.46	0-4
13 (D6)	Detachment or estrangement	1.22	1.12	1.26	.38	-.95	0-4
14 (D7)	Numbing	.88	1.04	1.09	.61	-1.16	0-3
15 (E1)	Irritability or aggressive behavior	.72	1.01	1.02	1.22	.68	0-4
16 (E2)	Reckless behavior	.23	.67	.45	3.13	9.47	0-3
17 (E3)	Hypervigilance	1.53	1.08	1.17	-.13	-.91	0-4
18 (E4)	Startle	.83	1.03	1.06	.64	-1.21	0-3
19 (E5)	Concentration	1.05	1.14	1.30	.47	-1.35	0-3
20 (E6)	Sleep	1.35	1.19	1.42	.09	-1.32	0-4
Cluster B	Reexperiencing	6.18	3.28	10.76	-.03	-.16	0-14
Cluster C	Avoidance	3.00	1.67	2.78	-.25	-.87	0-6
Cluster D	NACM	8.23	5.43	29.50	.36	-.63	0-20
Cluster E	Hyperarousal	5.72	3.62	13.12	.56	-.10	0-14
Total score		23.13	11.36	128.97	.16	-.24	1-49

Note. $N = 60$. CAPS-5 = Clinician-Administered Posttraumatic Stress Disorder Scale for *DSM-5*; NACM = Negative Alterations in Cognitions and Mood.

Appendix B

Scoring Guidelines

I. Structural

1. Item Order: Scored if participant stated that the placement of the items within the measure impacted their answer.

> **Example:**
> The question was one of the first symptoms. I hadn't had time to think about my reactions to the traumatic event. Later questions were easier to answer because by answering earlier questions, I was able to work through my reactions and better able to answer later questions.
>
> **KEYWORDS:** Earlier/later items

2. Response Options: Scored if participant reported difficulty understanding the rating scale or trouble choosing a rating.

> **Example:**
> I couldn't decide what to select on the scale. I wish there were different options to choose from.
>
> **KEYWORD:** Rating scale

3. Time-frame reminders: Scored if participant explained that they were not answering based on the past month or that they wanted reminders of the specific timeframe.

> **Example:**
> When I answered the question, I was thinking about my overall distress. I did not remember to only answer based on the past month.
>
> **KEYWORDS:** Past month

4. Trauma-related reminders: Scored if participant stated that their answer was not connected to the trauma itself or that they wanted a reminder to answer the question based on the specific traumatic event.

> **Example:**
> It was hard for me to relate that question back to the trauma. My answer was a general one because I forgot to relate it back to how the trauma impacted my distress.

KEYWORD: Trauma or Event, "in general," "wasn't thinking about event"

5. Opportunity for Clarification: Scored if participant labeled the ability to ask questions, ability to ask for the question to be reworded, or the ability to hear additional prompts to clarify confusing words as a reason for the discrepancy.

Example:
I misinterpreted the question. The wording of the question was confusing. It helped to hear it out loud and to have additional questions asked during the interview.

KEYWORD: Ask, clarify

6. Opportunity for Elaboration: Scored if participant described that the assessment method allowed for more detailed exploration of symptoms and allowed them to explain their answer with further information and detail.

Example:
I was able to explain more in greater detail about what was going on. It wasn't just circling a number, rather it was talking it through with a person. I was able to provide a fuller and more accurate description of my symptoms because I was able to provide more information to the question.

KEYWORD: Explain

7. Question Length and Complexity: Scored if the participant stated that the length of the question impacted their ability to understand the question. Additionally, scored if they discussed how having a list of symptoms impacted their answer.

Example:
It was too long of a question and I didn't see the last part of it. I wish each part of the list was a separate question.

KEYWORD: Too long OR List

8. Clinical Judgment: Scored if the participant stated that having the clinician make the final rating decision is the reason for discrepancy.

Example:
As the researcher, you are able to compile all the information and use your judgment to choose a rating.

9. Oral vs. Written Comprehension: Scored if participant indicated that they answered differently when reading the questions compared to when hearing the questions out loud.

Participant must mention specifically that reading the questions was different from hearing them out loud (i.e., have to state a comparison)

Example:
I answered differently because it was more difficult for me to read the questions on paper. When I heard the same question, it was easier to understand and to respond more accurately compared to seeing the question.

KEYWORDS: On paper versus aloud

II. Content

10. Comprehension of Symptoms: Scored if participant reported that they misunderstood or misinterpreted the question

Example:
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.

KEYWORD: Misunderstood, confused

11. Item Content: Scored if the participant labeled specific words that they did not understand or the differences in wording of the assessment measures as causing the discrepancy. Also scored if the discrepancy is due to a predetermined difference between the two measures. For example, for symptoms 8 and 10 where the PCL does not include the clarifying statement that intoxication or blaming the perpetrator are absent symptoms.

Example:
The question is written differently on the self-report compared to the interview measure. The self-report talks about avoiding "conversations" while the interview does not mention "conversations."

KEYWORD: Specific word

III. Intrapersonal

12. Avoidance of Extreme Responses: Scored if participant specifically said that they did not want to select a 0 or a 4 rating on the assessment measure.

Example:
I didn't want to select any extreme responses. Putting a 0 means that I don't have distress and that is not the case.

KEYWORD: 0 OR 4 rating

13. Minimization of Symptoms: Scored if participant explained that they provided lower ratings that did not match their accurate distress level.

> **Example:** I was unwilling to admit that it is a problem. I was in denial that I avoid certain places, so I marked a lower rating because it was too difficult to admit that I am in distress.
>
> **KEYWORD:** Lower symptoms, denial, "not a problem"

14. Avoidance of Specific Questions: Scored if participant specifically labeled certain types of questions or specific questions that they were reluctant to answer.

> **Example:** I completely avoided the question about emotional reactions. It was difficult for me to see/hear that question, so I gave a low rating, and moved on immediately.
>
> **KEYWORDS:** avoid, emotional

15. Trauma-related Attribution Error: Scored if participant interpreted their symptoms as an internal characteristic instead of recognizing the contribution of the traumatic event.

> **Example:** I have always had problems with sleeping so I didn't realize the trauma had progressively changed my sleeping habits.
>
> **KEYWORD:** Always, "always been this way"

16. Deliberate Calibration: Scored if participant reported that they intentionally chose a rating (not necessarily an extreme response) even though that rating did not accurately reflect their current distress level.

> **Example:** On the self-report, if I was between two ratings, I always chose the more severe rating to make it clear I am suffering**.**

17. Perception vs. Reality: Scored if participant explained that they answered differently on one of the measures compared to the other because they have different views of themselves.

> **Example:** On the self-report, my answer was what I wanted to be doing. Like I wanted to not be avoiding places and I wanted to not still have memories of it. But when answering the interview, reality hit and I answered what I am actually doing and how I am actually handling my distress.

18. Increased Awareness: Scored if participant discussed the overall internal process that occurred while answering the assessment measure.

> **Example:** Having the time to think and talk about my symptoms helped me understand what I am actually going through. Thinking it through helped me to process and organize my thoughts and helped me remember what I have been struggling with this past month.
>
> **KEYWORD:** time, think through

IV. Interpersonal

19. Impression Management: Scored if participant reported that they answered in a particular way because they were trying to influence how the clinician perceived them. Additionally, scored if participant discussed the influence of the connection/relationship with the clinician.

> **Example:** I tried to act a certain way on the self-report. I wanted you to think of me as a positive person.

V. General Errors

20. Scored if 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.

> **Example:** I don't remember marking that answer. I must not have read the question and I marked incorrectly.
>
> **KEYWORD:** mismarked, don't know, forgot

VI. Clinical Judgment Contributed to Error

21. Scored if the clinician looks back on the interview measure and realizes that she made a mistake in marking a response.