# COMMENTARY

# The Problem With Overreliance on the PCL–5 as a Measure of PTSD Diagnostic Status

Michelle J. Bovin and Brian P. Marx
National Center for PTSD, VA Boston Healthcare System, Boston, Massachusetts, United States
Department of Psychiatry, Boston University School of Medicine

In their systematic review, Forkus et al. (2022) synthesized results from the research on the psychometric properties of the posttraumatic stress disorder (PTSD) Checklist for the *Diagnostic and Statistical Manual of Mental Disorders, 5th edition* (*DSM–5*; PCL–5; Weathers, Litz, et al., 2013). The PCL–5 is one of the most widely used self-rating scales for determining provisional PTSD diagnostic status, quantifying PTSD symptom severity, and monitoring symptom change across time and in response to treatment (Forkus et al., 2022). In their review—which examined findings from 64 studies on the reliability, validity, factor structure, optimal cutoff scores, and sensitivity to change—PCL–5 scores demonstrated high internal consistency, test-retest reliability, construct validity, and sensitivity to change. Further, the authors found evidence that the recommended cutoff for a PCL–5 total score in providing a provisional PTSD diagnosis is between 31 and 33. Taken together, these results present a compelling case that the PCL–5 is psychometrically robust and appropriate for use with a variety of samples and in various settings.

Forkus et al.'s (2022) review affords an opportunity for us to consider the parameters within which the PCL–5 should and should not be used. The PCL–5 was designed primarily to capture information from respondents about the extent to which they may be experiencing distress related to PTSD symptoms. Consequently, a PCL–5 total score is an indicator of respondent-reported PTSD symptom severity. Forkus et al. noted that PCL–5 total scores were consistently associated with total scores of *DSM*-correspondent PTSD diagnostic interviews, such as the Clinician Administered PTSD Scale for *DSM–5* (CAPS-5; Weathers, Blake, et al., 2013). However, the association between the PCL–5 and diagnostic interviews is never perfect, leading some to caution against using it as a stand-alone diagnostic tool (McDonald & Calhoun, 2010). Despite this, due to time constraints,

concerns about respondent burden, and/or lack of access to or adequate training on the CAPS-5 or similar instrumentation, it is commonplace for PCL–5 scores to be used as a substitute for results from a more thorough assessment of PTSD symptoms. In so doing, the field has reified PCL–5 cutoff scores in a way that was never intended. While understandable, this practice is problematic, with important implications for everything from estimating the population prevalence of PTSD, to determining the extent to which PTSD may be associated with various biomarkers and other risk factors for the disorder, to understanding the association between PTSD and psychiatric and medical comorbidities and suicide risk, to evaluating the effectiveness of evidence-based treatments, as well as introducing additional noise into an already complex diagnostic profile. In this commentary, we discuss key limitations to conceptualizing PCL–5 scores as equivalent to PTSD status determined by a structured or semi-structured diagnostic interview, review why this practice is so compelling despite obvious threats to validity, and conclude with suggestions for moving away from this reification.

## Problems With Conflating the PCL–5 With a PTSD Diagnostic Interview

### Respondent Interpretation

Forkus and colleagues' (2022) review suggest that, overall, PCL–5 scores generally performed well in terms of all important psychometric indices. However, findings from a minority of studies suggested that the PCL–5 performed poorly, as evidenced by low internal consistency and temporal stability, as well as lower-than-expected correlations between PCL–5 scores and scores of measures ostensibly assessing related constructs (e.g., trauma exposure) and higher-than-expected correlations with scores of measures assessing theoretically dissimilar constructs (e.g., psychosis). The authors proposed several reasonable hypotheses to explain these discrepancies, including sample-specific characteristics, qualities of selected measures, and high levels of comorbidity.

One possibility not mentioned by Forkus et al. (2022) that may explain both between- and within-sample variability merits consideration, namely patient interpretation of the PCL–5 items. Despite the strengths of the PCL–5 and the correspondence between its items and the 20 *DSM–5* PTSD symptoms, it is still an

Brian P. Marx https://orcid.org/0000-0001-8988-1796

Michelle J. Bovin served as lead for writing–original draft and contributed equally to conceptualization and writing–review and editing. Brian P. Marx contributed equally to conceptualization and writing–review and editing and served in a supporting role for writing–original draft.

Correspondence concerning this article should be addressed to Brian P. Marx, National Center for PTSD, VA Boston Healthcare System, 150 South Huntington Avenue, 116B-4, Boston, MA 02130, United States. Email: brian.marx@va.gov

instrument on which respondents rate themselves. As such, the PCL–5 is susceptible to bias and misinterpretation that, unlike a diagnostic interview, cannot be addressed and corrected for by a trained assessor in real time. Consistent with this concern, Kramer et al. (2022) recently conducted a study to better understand common discrepancies between results obtained from the PCL–5 and results obtained from the CAPS-5. They administered both instruments to participants, identified item-level discrepancies, and then conducted both quantitative and qualitative analyses to gain insight into the reasons for these discrepancies. Results highlighted that, although PCL–5 and CAPS-5 scores and diagnostic results may be strongly associated, they are not interchangeable, with the PCL–5 yielding higher PTSD prevalence estimates and higher severity scores than the CAPS-5. Further, elevated PCL–5 scores were more strongly associated with neuroticism, response bias, verbal intelligence estimates, self-rating errors, lack of opportunity to clarify symptoms, and difficulty differentiating distress associated with PTSD symptoms from general distress.

Kramer et al.'s (2022) work underscores the problem with equating results obtained from the PCL–5 with those obtained from the CAPS-5 and other diagnostic interviews. Self-rating measures inherently present a greater risk for misinterpretation, misunderstanding, and other forms of bias, reducing concordance with the respondent's true score. The implications of this are that even if an individual meets criteria for PTSD according to the DSM–5 algorithm based on information collected via the PCL–5 (i.e., at least one Criterion B symptom, at least one Criterion C symptom, at least two Criterion D symptoms, and at least two Criterion E symptoms are scored by the patient as "2" [moderately] or greater), that individual will not necessarily meet criteria for PTSD on a diagnostic interview.

## The Problem With Cutoff Scores

Clearly, patient misinterpretation, error, and current distress levels importantly influence PCL–5 scores, reducing validity. However, even if we could ensure that patients perfectly understood each PCL–5 item and responded to each item by carefully separating PTSD-related distress from the impact of current stressors, the use of the PCL–5 to represent PTSD diagnostic status would remain problematic. This problem is inherent in the use of a single cutoff score to characterize diagnostic status.

In settings or conditions under which the use of a diagnostic interview to determine diagnostic status (e.g., the CAPS-5) is prohibitively long or where trained assessors are unavailable, the PCL–5 is often used as a proxy means of determining PTSD diagnostic status. In these cases, clinicians or researchers typically employ a predetermined severity score as a cutoff, with individuals scoring above the cutoff score labeled as having "probable" PTSD. The ideal cutoff score, which is determined by balancing levels of sensitivity (i.e., the probability that individuals with a positive diagnosis will have a positive test) and specificity (i.e., the probability that individuals with a negative diagnosis have a negative test) against a well-accepted criterion measure (typically the CAPS-5 or another structured diagnostic interview), varies as a product of the base rate of the disorder in the sample and can be set to vary in stringency based on the goal of the assessment (e.g., to reduce the chances of missing anyone with PTSD; to minimize diagnostic errors; to confirm a diagnosis).

Forkus et al. (2022) reported that the most frequently identified cutoff score using diagnostic utility analyses was a PCL–5 score between 31-33, regardless of sample type. Scores in this range tend to best balance sensitivity and specificity. Although this is the most frequently identified range of optimal cutoff scores, reification of this range is problematic; just because this was the most frequently observed optimal cutoff score range does not mean that it is appropriately suited for all trauma exposed individuals or settings. Indeed, across the studies in Forkus et al.'s review that examined cutoff scores, the optimal cutoff score ranged from 23-49. As Forkus et al. noted, "Consistent with findings for the PCL for DSM–IV (McDonald & Calhoun, 2010), our review results suggest that a universal cutoff score applicable across diverse samples and settings does not exist for the PCL–5" (p. 20; emphasis added).

Emerging evidence suggests that the validity of PCL–5 cutoff scores may systematically vary by group membership. Of the 21 studies in Forkus et al. (2022) review that established an optimal cutoff score using diagnostic utility analyses, only one study considered differential cutoff scores as a product of patient-level factors. In this study, Geier et al. (2019) examined the cutoff score for individuals seen in an emergency department whose index event was a physical injury. Whereas the optimal cutoff score for the overall sample (N = 251) was a PCL–5 severity score of 30, the optimal cutoff score for participant with an intentional injury (n = 72) was 34 and the optimal cutoff score for participants with a nonintentional injury (n = 179) was 22. These findings not only highlight how substantially individual factors can influence the optimal cutoff score, but how much within group variance exists even within a single group (i.e., those exposed to physical injury).

The frequency of 31-33 being identified as the optimal cutoff score on the PCL–5 also varies importantly as a product of the criterion measure used. Whereas 80% (8 of 10) of the analyses that used the CAPS-5 as the criterion identified a cutoff score ranging from 30–34 as optimal, only 50% (6 of 12) of the analyses using another version of the PCL as the criterion (e.g., the PCL for DSM–IV) identified an optimal cutoff score in this range, and 0% (0 of 8) of the analyses using a criterion other than the CAPS or PCL identified a cutoff score in this range.

The variation of optimal cutoff scores as a product of PTSD population base rates, patient-level factors, and the choice of criterion measure underlines the importance of not assuming that any given cutoff score is ideal for every group of trauma-exposed individuals. This challenge highlights a real tension between research findings and clinical practice. In busy clinical settings, where conducting a thorough (semi)structured diagnostic interview (e.g., a CAPS-5) for every patient is not possible, other means of ensuring PTSD status is necessary prior to providing care. However, use of the PCL–5 in this way guarantees that whereas some patients with PTSD are not identified, other patients who do not have PTSD are referred for care they may not need.

To address this tension in the field, other researchers have worked to establish cutoff scores that vary by patient-level factors. For example, due to research demonstrating systematic differences across sex, the abbreviated Alcohol Use Disorder Identification Test for Consumption (AUDIT-C) uses different cutoff scores for men and women in response to research indicating that the higher cutoff score consistently yielded lower sensitivities and higher specificities for women than for men (Reinert & Allen, 2007).

However, even if we could identify and implement the use of multiple cutoff scores as a product of group membership, these efforts would not fully correct for within group variability. Regardless of how we slice the PCL–5, we will never be able to have it perfectly reflect the results of a PTSD diagnostic interview. This is unsurprising, considering that even if the PCL–5 and other self-report measures were not susceptible to bias and that a single cutoff score could be applied efficiently to all trauma-exposed groups, the PCL–5 also does not assess symptom duration or clinically significant distress or functional impairment resulting from these symptoms, which are key factors in determining diagnostic status. Further, depending upon the version used, the assessment of Criterion A, as well as linkage between trauma exposure and PTSD symptoms, may be weak or even nonexistent.

## The Medical Model

Knowing the pitfalls of using PCL–5 scores as a proxy for diagnostic status, why does this practice remain rampant in the field? To answer this question, we need to look no further than the context within which these probable PTSD diagnoses are produced—the medical model of mental disorders. This medical, or disease oriented, approach has long dominated the context in which mental disorders are identified and treated. This model is predicated on the assumption that psychopathology is the result of one's biology (e.g., Benight, 2012). As such, it presents an all-or-none proposition; either the individual has the disorder, or the individual does not. In the case of PTSD, it proposes the existence of a categorically discrete syndrome that can be qualitatively distinguished from normal stress states (e.g., Ruscio et al., 2002).

The medical model can be a useful guide for diagnosis, prognosis, and research. However, from prior research, we know that PTSD does not tend to behave in an all or none fashion (Ruscio et al., 2002). Thus, exclusive reliance on the medical model may lead to an incomplete understanding of an individual's response to trauma, and, frequently, to incomplete or ineffective treatment interventions. Importantly, the medical model is inherently incompatible with the way the PCL–5 assesses PTSD symptomatology, as the latter provides a continuous total severity score which indicates a range of PTSD symptom expression.

Despite the limitations of the medical model, particularly in the case of PTSD, it continues to be the dominant paradigm through which mental illness is conceptualized and treated. Within this context, patients must be quickly categorized by clinicians for treatment and billing purposes. Insurance companies are not interested in subjective levels of patient distress, but rather whether a patient meets criteria for a diagnosis or not. Busy clinicians, working to balance patient need with system requirements, understandably look for quick ways to determine treatment plans for patients. In these settings, where diagnostic interviews are typically not feasible, PCL–5 cutoff scores may be treated as the *de facto* diagnostic standard.

In the trickle-down effect that this paradigm produces, we acknowledge our own role. To give the field what it has wanted, we have conducted diagnostic utility analyses, providing cutoff scores for measures like the PCL–5. Although we regularly provide caveats about the use of such cutoff scores, cautioning clinicians and researchers alike to view self-rating measures of PTSD like the PCL–5 as one tool among many and to interpret cutoff scores as invitations for additional assessment, it does not escape us that the cutoff scores we provide are frequently adopted in clinical care as infallible diagnostic metrics.

The attraction of this reification is the ability to quickly dichotomize patients based on need, so that they can be efficiently directed to appropriate care. Ironically, in practice, this reductionist approach often has the opposite effect. By ignoring individual differences for the purposes of quickly cataloging the likely presence or absence of a disease, individuals with PTSD can fall through the cracks and not be referred for treatment. In addition, individuals without PTSD may be referred for unnecessary care, further increasing the caseload of already busy clinicians. This pattern risks limiting clinician time even more, which increases the need for an efficient manner for categorizing patients. In this way, the use of the PCL–5 as a quick method for diagnosing patients becomes that much more attractive, moving us further away from the ability to provide quality care.

## The Path Forward

Clearly, the overreliance on the PCL–5 as a measure of PTSD diagnostic status is problematic. As a self-rating measure, the PCL–5 is subject to patient misinterpretation and bias. The complexity of the symptoms of PTSD makes any assumption that respondents understand each PCL–5 item in the way trained clinicians do doubtful at best. Further, dichotomizing PCL–5 severity scores to assign a probable diagnosis introduces additional error. Despite these challenges, the healthcare system in which mental health is assessed and treated typically necessitates methods for quickly classifying individuals as either having or not having PTSD. Is there a path forward that does not reinforce this practice?

In a perfect world, we would have the time and expertise to move away from simple diagnostic categorization. We would shift the paradigm to celebrate the exploration of individual differences in the context of trauma exposure, relying on dimensional, rather than categorical, conceptualizations. The PCL–5 could then be used exclusively as a guide to patient-reported symptom severity, rather than as a proxy for diagnosis.

Even within the current mental healthcare paradigm, the use of a multimethod assessment approach that incorporates diagnostic interviews, self-rating instruments, and other relevant data (e.g., psychophysiological data, other measures of distress, informant reports) would be ideal. We have previously highlighted the importance of using multiple methods of assessment to adequately capture data from the three response systems (self-reported emotional experience, expressive behavior, and objective physiological indicators of distress or arousal) to fully capture each patient's experience before, during, and after treatment for PTSD (Bovin et al., 2015).

Of course, we do not live in a perfect world. Diagnoses are still needed for billing purposes, and overwhelmed clinicians often do not have adequate time or resources to administer comprehensive assessment batteries. Despite this, we can still make changes to our assessment practices that will avoid overreliance on PCL–5 scores as a proxy for PTSD diagnosis. First and foremost, the field needs to move away from reifying cutoff scores, and instead recognize a cutoff score as one (fallible) piece of information about a patient's PTSD diagnostic status and treatment needs. Clinical judgment, the self-reported distress of the patient, and attention to

other factors that may influence reporting style must be considered in concert with the PCL–5 score prior to giving a "probable PTSD" label. A score on the PCL–5 that is above a given cutoff score should be equated with the need for additional assessment, rather than as a PTSD oracle.

Second, every effort should be made to ensure that patients completing the PCL–5 understand both the instructions and content. Kramer et al. (2022) end their paper with several suggestions for improvement in this domain, including providing additional instruction to patients prior to their answering any of the PCL–5 items, reminding patients to carefully read each question before responding and to consult the clinician or assessor with any questions, providing cues to the trauma-relatedness and time frame of the test items, and discussing the purpose of the PCL–5 with patients in detail prior to administration. These suggestions and others (e.g., examining and potentially adjusting the reading level for each PCL–5 item) may demonstrate empirical support for improving patient understanding and therefore increasing the ability of the PCL–5 to more closely reflect information gleaned from a diagnostic interview.

Finally, clinicians and researchers alike must remember what the PCL–5 is: a measure of self-reported PTSD symptom severity. In this way, it is a tool our patients use to communicate with us about how they are perceiving their own distress. This is key information that can be used to inform patient care and is at its most valuable when it is considered continuously rather than categorically.

The strong psychometric properties of the PCL–5 at a group level make the desire to use it as a proxy of PTSD diagnostic status compelling, particularly within the context of any healthcare system in which the medical model of psychopathology is the primary paradigm. We do not harbor any delusions that clinicians and researchers will immediately stop using the PCL–5 as a proxy based on the considerations raised in this commentary. Nor do we wish to discourage people from using the PCL–5 at all. Instead, what we do hope is that the points made here will better inform members of the field about how PCL–5 scores can be misused or misinterpreted. We hope that it will encourage clinicians to ask additional questions prior to assigning a PTSD diagnosis, and that it will encourage researchers to investigate other avenues for increasing the validity of the PCL–5. Such efforts have the potential to make the use of the PCL–5 suited to its capabilities, which in turn will aid us in achieving our ultimate goal of providing our patients with treatment plans best suited to their needs.

## References

Benight, C. C. (2012). Understanding human adaptation to traumatic stress exposure: Beyond the medical model. *Psychological Trauma: Theory, Research, Practice, and Policy*, *4*(1), 1–8. https://doi.org/10.1037/a0026245

Bovin, M. J., Marx, B. P., & Schnurr, P. P. (2015). Evolving DSM diagnostic criteria for PTSD: Relevance for assessment and treatment. *Current Treatment Options in Psychiatry*, *2*(1), 86–98. https://doi.org/10.1007/s40501-015-0032-y

Forkus, S. R., Raudales, A., Rafiuddin, H., Weiss, N. H., Messman, B. A., & Contractor, A. A. (2022). The PTSD checklist for *DSM–5*: A systematic review of existing psychometric evidence. *Clinical Psychology: Science and Practice*, *30*(1), 110–121.

Geier, T. J., Hunt, J. C., Nelson, L. D., Brasel, K. J., & deRoon-Cassini, T. A. (2019). Detecting PTSD in a traumatically injured population: The diagnostic utility of the PTSD Checklist for *DSM–5*. *Depression and Anxiety*, *36*(2), 170–178. https://doi.org/10.1002/da.22873

Kramer, L. B., Whiteman, S. E., Petri, J. M., Spitzer, E. G., & Weathers, F. W. (2022). Self-rated versus clinician-rated assessment of posttraumatic stress disorder: An evaluation of discrepancies between the PTSD Checklist for *DSM–5* (PCL–5) and the Clinician-Administered PTSD Scale for *DSM–5* (CAPS-5). *Assessment*. Advance online publication. https://doi.org/10.1177/10731911221113571

McDonald, S. D., & Calhoun, P. S. (2010). The diagnostic accuracy of the PTSD checklist: A critical review. *Clinical Psychology Review*, *30*(8), 976–987. https://doi.org/10.1016/j.cpr.2010.06.012

Reinert, D. F., & Allen, J. P. (2007). The alcohol use disorders identification test: An update of research findings. *Alcoholism, Clinical and Experimental Research*, *31*(2), 185–199. https://doi.org/10.1111/j.1530-0277.2006.00295.x

Ruscio, A. M., Ruscio, J., & Keane, T. M. (2002). The latent structure of posttraumatic stress disorder: A taxometric investigation of reactions to extreme stress. *Journal of Abnormal Psychology*, *111*(2), 290–301. https://doi.org/10.1037/0021-843X.111.2.290

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*. U.S. Veterans Affairs National Center for Posttraumatic Stress Disorder.

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). https://www.ptsd.va.gov/professional/assessment/index.asp