Received: 14 March 2024 | Revised: 6 July 2024 | Accepted: 7 July 2024

DOI: 10.1002/jts.23093

RESEARCH ARTICLE

ISTSS International Society for Traumatic Stress Studies WILEY

The revised Clinician-Administered PTSD scale for *DSM-5* (CAPS-5-R): Initial psychometric evaluation in a trauma-exposed community sample

Brianna N. Jackson[1] | **Frank W. Weathers**[1,2,3,4] | **Stephanie M. Jeffirs**[1] | **Thomas J. Preston**[1] | **Cassidy M. Brydon**[1]

[1]Department of Psychological Sciences, Auburn University, Auburn, Alabama, USA

[2]Behavioral Science Division, National Center for PTSD, Boston, Massachusetts, USA

[3]VA Boston Healthcare System, Boston, Massachusetts, USA

[4]Boston University Chobanian & Avedisian School of Medicine, Boston, Massachusetts, USA

Correspondence
Frank W. Weathers, Department of Psychological Sciences, Auburn University, Auburn, Alabama, United States.
Email: frank.weathers@va.gov

Funding information
Psychological Assessment Resources (PAR)

Abstract

The Clinician-Administered PTSD Scale for *DSM-5* (CAPS-5) is a widely used, well-validated structured interview for posttraumatic stress disorder (PTSD). It was recently revised to improve various aspects of administration and scoring. We conducted a psychometric evaluation of the revised version, known as the CAPS-5-R. Participants were 73 community residents with mixed trauma exposure (e.g., sexual assault, physical assault, transportation accident, the unnatural death of a loved one). CAPS-5-R PTSD diagnosis demonstrated good test–retest reliability, κs = .73–.79; excellent interrater reliability, κs = .86–.93; and good-to-excellent alternate forms reliability with the CAPS-5, κs = .79–.93. In addition, the CAPS-5-R total PTSD severity score demonstrated excellent test–retest reliability, intraclass correlation coefficient (ICC) = .86; interrater reliability, ICC = .98; and alternate forms reliability with the CAPS-5, $r = .95$. Further, the CAPS-5-R demonstrated good convergent validity with other measures of PTSD and good discriminant validity with measures of other constructs (e.g., depression, anxiety, alcohol problems, somatic concerns, mania). Given its strong psychometric performance in this study, as well as its improvements in administration and scoring, the CAPS-5-R appears to be a valuable update of the current CAPS-5.

The Clinician-Administered PTSD Scale (CAPS) is a widely used, well-validated structured interview for posttraumatic stress disorder (PTSD; Weathers et al., 2001, 2018; Weathers, Blake, et al., 2013a). Developed in 1990 and originally based on PTSD criteria in the *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., text rev.; *DSM-III-R*; American Psychiatric Association [APA], 1987), the CAPS was subsequently revised for the *DSM-IV* (APA, 1994) and *DSM-5* (APA, 2013). All versions of the CAPS share several distinctive features, including separate ratings for the frequency and intensity of symptoms; behaviorally anchored prompts and rating scales; and dichotomous (i.e., present/absent) and dimensional scores for items, symptom clusters, and the full syndrome.

The CAPS-5

Revisions for the CAPS-5, the current *DSM-5* version, included (a) adding and updating items to parallel *DSM-5* changes to the PTSD criteria, (b) streamlining administration by reorganizing prompts into a top-to-bottom format, and (c) simplifying scoring by creating a system for converting the frequency and intensity ratings for

individual items into a single 0–4 severity score. A symptom is considered present if the item severity score is 2 (*moderate/threshold*) or higher (SEV2 rule), and a PTSD diagnosis is given if the *DSM-5* criteria are met (i.e., the presence of at least one intrusion [i.e., *DSM-IV* reexperiencing] symptom, one avoidance symptom, two negative alterations in cognition and mood [NACM] symptoms, and two alterations in arousal and reactivity [AAR] symptoms). As were previous versions of the CAPS, the CAPS-5 is psychometrically sound. In a dedicated psychometric study (Weathers et al., 2018), with Cohen's kappa (κ) for diagnosis and intraclass correlation coefficients (ICCs) and Pearson correlations (r) for continuous variables, both the CAPS-5 PTSD diagnosis and total severity score demonstrated strong test–retest reliability ($\kappa = .83$, ICC = .78), interrater reliability (κs = .78–1.00, ICC = .91), and correspondence with the CAPS-IV ($\kappa = .84$, $r = .83$). In addition, the total severity score demonstrated good convergent validity with the PTSD Checklist for *DSM-5* (PCL-5 [Weathers, Litz, et al., 2013]; $r = .66$) and good discriminant validity with measures of anxiety, depression, somatization, psychopathy, alcohol abuse, and functional impairment (rs = .02–.54). In other studies, the CAPS-5 has demonstrated high interrater reliability (ICC = .97 [Schnurr et al., 2022]; $\kappa = .90$, $r = .98$ [Resick et al., 2023]), a conceptually meaningful factor structure (Lee et al., 2019), and sensitivity to clinical change (Lee et al., 2022; Schnurr et al., 2022).

Changes for the revised CAPS-5 (CAPS-5-R)

However, despite its strong psychometric performance, it became evident through feedback from a wide variety of raters and the observation of hundreds of interviews over the past decade that the CAPS-5 could benefit from additional improvements. Accordingly, it was recently revised to incorporate many of the suggested changes and, thereby, create an improved version that would supersede the current one. The goals for the revised CAPS-5 (CAPS-5-R; Weathers et al., 2022), were to (a) increase the standardization of administration and scoring, (b) improve the quality of the information obtained, (c) expand the rating scale to capture more variability in severity scores, (d) make it easier to learn, and (e) maintain backward compatibility with the CAPS-5.

The following sections outline how these goals were accomplished. To ensure a high level of content validity for the CAPS-5-R, all proposed changes were discussed among numerous PTSD experts until agreement was reached on the final form of the interview. In addition to the CAPS-5-R authors, other primary contributors to the revision process included Michelle Bovin, Daniel Lee, and Sarah Kleiman. To illustrate the various changes, the CAPS-5 and CAPS-5-R versions of Item D2 are reproduced in Supplementary Figures S1 and S2.

Modification and addition of prompts

First, some existing prompts were modified, and new prompts were added to clarify inquiries and elicit more detailed information. Although the CAPS-5 is fully scripted, the standard prompts do not always elicit sufficient information, in which case raters are instructed to improvise additional prompts. Going off script is necessary at times, but it may introduce unwanted variability in the inquiry, which could reduce both reliability (e.g., an improvised prompt may not be used again in a subsequent interview) and validity (e.g., an improvised prompt may not accurately reflect the conceptual basis of a given symptom). Therefore, to minimize improvisation, an effort was made to include standard prompts covering all key aspects of symptoms and all points in the interview that commonly require clarification. Prompts were carefully reviewed and tested in role plays and actual interviews with trauma survivors to ensure they were conceptually accurate, easily understood, and effective at eliciting adequate responses. For example, as shown in Supplementary Figures S1 and S2, Item D2 was revised by (a) moving examples of qualifying beliefs to the initial prompt to increase comprehension; (b) adding the prompt, "Any other strong negative beliefs?" along with the instruction to follow up on beliefs respondents may have overlooked; (c) adding the phrase "like seeing another side to an argument" to increase comprehension and elicit and normalize ambivalence; and (d) adding the prompt, "How hard is that for you to do?" to encourage respondents to elaborate on what is often a "yes" or "no" response to the previous prompt.

Expansion of the scoring guidelines

Second, the scoring guidelines for each item were substantially expanded. On the CAPS-5, the scoring guidelines consisted only of behavioral referents to anchor the severity score ratings of 2 (*moderate/threshold*) and 3 (*severe/markedly elevated*). These were retained for the CAPS-5-R but were supplemented with much more detailed information, including conceptual clarifications (e.g., for Item D2, "Focus on beliefs, and do not count feelings, behaviors, facts, or abstract concepts") and specific rating instructions (e.g., for Item D2, "Base the intensity rating on both the degree of distortion and the degree of conviction"). Thus, the new scoring guidelines provide the

rater with what is essentially a brief user guide for each symptom.

Expansion of the severity rating scale

Third, the rating scale for item severity was expanded from 0–4 on the CAPS-5 to 0–10 on the CAPS-5-R, resulting in an increase in the possible range for total symptom severity from 0–80 to 0–200. This created a more granular rating scale to detect smaller differences in symptom severity and address a possible restriction of range in CAPS-5 scores. Evidence from clinical trials indicates that most nonzero CAPS-5 item severity scores are either 2 (*moderate/threshold*) or 3 (*severe/markedly elevated*; Parmenter et al., 2024; Resick et al., 2023). Scores of 1 (*mild/subthreshold*) and 4 (*extreme/incapacitating*) are used much less frequently. This pattern of CAPS-5 item severity scores (i.e., a preponderance of items scored as moderate or severe, with relatively few scored as mild or extreme) may accurately reflect the actual distribution of symptom severity in PTSD clinical samples. However, at least to some extent, it may indicate a restriction of range, which is of particular concern at the high end of the scale for a measure of psychopathology given the need to represent the full range of symptom severity and accurately differentiate individuals who all fall above a clinically significant threshold but vary in symptom severity. It is not possible to distinguish between these two possibilities using the CAPS-5 because there is no other gold-standard measure to use as a comparison. However, this question can be investigated with the expanded item severity scale on the CAPS-5-R.

The new 0–10 severity scale was created by increasing both the number of rating options for intensity and the number of thresholds for frequency. The intensity rating options on the CAPS-5 were *minimal*, *clearly present*, *pronounced*, and *extreme*. For the CAPS-5-R, intermediate ratings were added between *clearly present* and *pronounced* and between *pronounced* and *extreme*. For items for which frequency reflects the number of times the respondent has experienced the symptom, the thresholds on the CAPS-5 were *once a month*, *twice a month*, *twice a week*, and *daily or almost every day*. For the CAPS-5-R, these were increased to *once a month*, *two or three times a month*, *once a week*, *two or three times a week*, *four or five times a week*, and *six or seven times a week*. Items for which frequency is rated as a percentage of time were revised similarly. An expanded severity scoring grid for the CAPS-5-R was created by crossing intensity anchors and frequency thresholds, and then assigning severity scores to each combination of intensity and frequency. This new scoring system is more granular in that it has smaller increments between ratings. It also addresses a possible restriction of range by extending the high end of the scale (i.e., from ratings of 3 and 4 on the CAPS-5 to ratings of 6, 7, 8, 9, and 10 on the CAPS-5-R). As shown in Supplementary Figure S2, the severity scoring grid is included below each symptom rather than being presented in a separate document as for the CAPS-5.

Importantly, backward compatibility with the CAPS-5 was maintained in that CAPS-5 intensity anchors and frequency thresholds are nested within the CAPS-5-R severity scoring grid (see Supplementary Figure S3). Therefore, both CAPS-5-R 0–10 and CAPS-5 0–4 severity scores can be determined from the same inquiry by using the corresponding scoring grid for each version of the CAPS. On the CAPS-5-R, a symptom is considered present if the item severity score is 3 or higher (i.e., SEV3 rule), and, as with the CAPS-5, a PTSD diagnosis is given if the *DSM-5* criteria are met.

Addition of a frequency response card

Fourth, a frequency response card was developed to reduce respondent burden and promote more accurate estimates of symptom frequency. The card contains four rating scales corresponding to each of the four types of frequency responses required on the CAPS-5-R. Each scale provides a range of fixed options intended to help respondents more quickly identify an appropriate frequency rating for a given symptom (see Supplementary Figure S4). Card use is also intended to mitigate the potential problems of satisficing (i.e., giving an easy response rather than a more accurate but more effortful response) and perseverating (i.e., giving the same response repeatedly without plausible variation across symptoms).

Improved formatting

Last, the format for the CAPS-5-R was streamlined to make it easier to follow the sequence of prompts, distinguish between required and conditional prompts, and encourage raters to ask just one prompt at a time. Notably, to simplify inquiry, Item 10, which assesses distorted blame of oneself or others, was split into two items that separately assess self-blame and other-blame. The higher of the two ratings is subsequently used for Item 10 when determining the respondent's diagnostic status and calculating a total severity score.

With all these changes, from a content validity perspective, the CAPS-5-R represents a significant improvement over the CAPS-5. However, the psychometric performance of the CAPS-5-R has yet to be investigated. Accordingly, our aim in the present study was to conduct a psycho-

metric evaluation of the CAPS-5-R in a community sample of trauma survivors. Specifically, we examined the CAPS-5-R's internal consistency, test–retest and interrater reliability, convergent and discriminant validity, and backward compatibility with the CAPS-5. We hypothesized that the CAPS-5-R would demonstrate levels of reliability and validity at least as high as previously reported for the CAPS-5 and would be strongly correlated with the CAPS-5.

METHOD

Participants and procedure

We recruited 76 adult community members from the southeastern United States using flyers and social media advertisements. The Auburn University Institutional Review Board reviewed and approved the following procedures. After completing an online screener, individuals were invited to participate in the study if they (a) reported experiencing a *DSM-5-TR* Criterion A traumatic event, as assessed using the Life Events Checklist–Extended Version (LEC-5–Extended; Weathers, Blake, et al., 2013b); (b) endorsed at least one symptom of PTSD on the Primary Care PTSD Screen for *DSM-5* (PC-PTSD-5; Prins et al., 2016); (c) allowed their interview sessions to be video-recorded; and (d) demonstrated proficiency in English. Participants who elected to complete the interviews remotely were additionally required to have access to the internet, a webcam, and a confidential space and to provide emergency contact information. Exclusion criteria were having an invalid Personality Assessment Inventory (PAI) profile using Morey's (1996) cutoffs (i.e., infrequency [INF] *T* score greater than 75, inconsistency [ICN] *T* score greater than 73, and positive impression management [PIM] *T* score greater than 68), as this suggested that participants were not adequately attending to questionnaire items.

A total of 186 individuals met the initial eligibility criteria based on their screener responses, and 118 participants provided written consent to engage in study procedures. After providing informed consent, participants were asked to complete an online battery of self-report questionnaires, followed by two interviews conducted approximately 1 week apart. Participants were randomly assigned to either a test–retest condition wherein the CAPS-5-R was administered twice or an alternate forms condition wherein the CAPS-5 and CAPS-5-R were administered in counterbalanced order. Of the 90 initial participants who completed the online questionnaire, 76 completed the first interview session. Participants completed interviews either remotely ($n = 57$) or in person ($n = 18$); one participant completed one session remotely and one in person. When the CAPS-5-R was administered remotely, the screen-share function of our video conferencing software was used to intermittently display the frequency rating card to the participant. Three participants were excluded from the final analyses based on invalid PAI profiles, resulting in a final sample size of 73 participants at Time 1. At Time 2, 58 participants completed an interview (test–retest condition: $n = 29$, alternate forms condition: $n = 29$). Dropout at Time 2 was not significantly associated with PTSD symptom severity or PTSD diagnosis on either the CAPS-5 or the CAPS-5-R nor was it associated with sociodemographic characteristics such as age, sex, or race.

Raters were five clinical psychology doctoral students. Prior to joining the current study, they were trained in diagnostic interviewing as part of their required coursework and clinical caseload. They were also trained by the second author on the CAPS-5 and CAPS-5-R and participated in ongoing scoring calibration meetings. Interview assignments were based on raters' availability. To estimate interrater reliability, four of the five raters independently scored six to eight video-recorded interviews each (i.e., 29 total); one rater was not able to contribute to this analysis.

For a detailed description of the final sample, see Supplementary Table S1. Participants were primarily female (69.9%), and the mean participant age was 42.6 years (range: 19–71, $SD = 13.1$). Most participants identified as White (83.6%), followed by Black (11%), American Indian/Alaskan Native (1.4%), and "other race" (4.1%). Based on the LEC-5 and a written narrative of the index trauma, the most prevalent index traumatic events were sexual assault (24.7%), physical assault (21.9%), the sudden unnatural death of a loved one (20.5%), and transportation accident (9.6%). Based on Time 1 CAPS-5 and CAPS-5-R interviews, 16.4% of the sample was asymptomatic or had few PTSD symptoms, 31.5% reported mild symptom levels, 28.8% reported moderate symptom levels, 16.4% reported severe symptom levels, and 6.8% reported extreme symptom levels. At Time 1, 49.3% of the sample met *DSM-5* PTSD diagnostic criteria.

Measures

See Table 1 and Supplementary Table S2 for descriptive statistics and internal consistency estimates for all measures used in the study. In addition to the CAPS-5 and the CAPS-5-R, the following measures were administered.

Psychiatric disorders

The Mini-International Neuropsychiatric Interview version 7.0.2 (MINI-7; Sheehan et al., 1998) is a structured

TABLE 1 Descriptive statistics for the Clinician-Administered PTSD Scale for *DSM-5*–Revised total, cluster, and item scores at Time 1

Scale and item	*M*	*SD*	Observed range	Possible range	Cronbach's α	Ω	CEP (%)[a]
Total	53.3	30.1	7–124	0–200	.89	.90	–
INT	12.7	7.9	0–30	0–50	.75	.76	–
B1	3.7	2.2	0–9	0–10	–	–	72.7
B2	2.4	2.8	0–9	0–10	–	–	41.8
B3	0.7	1.6	0–7	0–10	–	–	16.4
B4	3.2	2.1	0–10	0–10	–	–	69.1
B5	2.7	2.4	0–8	0–10	–	–	54.5
AV	6.5	4.8	0–18	0–20	.61	.61	–
C1	3.5	3.0	0–9	0–10	–	–	56.4
C2	3.1	2.7	0–9	0–10	–	–	56.4
NACM	18.8	12.9	0–48	0–70	.77	.80	–
D1	1.0	2.3	0–10	0–10	–	–	18.2
D2	3.3	3.3	0–10	0–10	–	–	52.7
D3	3.3	2.8	0–9	0–10	–	–	61.8
D4	3.3	2.7	0–9	0–10	–	–	61.8
D5	1.7	2.5	0–8	0–10	–	–	34.5
D6	3.3	3.3	0–9	0–10	–	–	50.9
D7	2.9	2.8	0–9	0–10	–	–	56.4
AAR	15.2	10.4	0–37	0–60	.73	.79	–
E1	1.7	1.9	0–6	0–10	–	–	34.5
E2	0.2	0.9	0–4	0–10	–	–	5.5
E3	4.4	3.3	0–10	0–10	–	–	70.9
E4	1.9	2.1	0–6	0–10	–	–	43.6
E5	2.6	3.0	0–8	0–10	–	–	45.5
E6	4.4	3.7	0–10	0–10	–	–	60.0
Depersonalization	0.4	1.0	0–4	0–10	–	–	9.1
Derealization	0.5	1.2	0–6	0–10	–	–	10.9

Note. $N = 55$. PTSD = posttraumatic stress disorder; *DSM-5* = *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.); INT = intrusions; AV = avoidance; NACM = negative alterations in cognition and mood; AAR = alterations in arousal and reactivity.

[a]Clinical elevation prevalence (i.e., the percentage of respondents with item severity scores ≥ 3).

clinical interview used to assess *DSM-5* psychiatric disorders. Questions are rated as "yes" or "no" to indicate the presence or absence of current and past symptoms of most major psychiatric disorders. In the current study, all MINI modules were administered except for the Suicidality, Suicidal Behavior Disorder, PTSD, Antisocial Personality Disorder, and Borderline Personality Disorder modules. The MINI was only used to describe the prevalence of comorbidities in the sample. In its initial validation studies, the MINI demonstrated excellent interrater reliability ($\kappa > .75$), good test–retest reliability, and good sensitivity and specificity (Sheehan et al., 1998).

PTSD symptom severity

PTSD Checklist for *DSM-5* (PCL-5; Weathers, Litz, et al., 2013) is a self-report, 20-item, *DSM-5*-correspondent questionnaire used to measure PTSD symptom severity. Respondents are asked to indicate how much they are bothered by each PTSD symptom using a 5-point Likert scale ranging from 0 (*not at all*) to 4 (*extremely*), with higher scores indicating higher levels of symptom severity. The PCL-5 has been found to have good test–retest reliability ($rs = .82–.86$), high internal consistency for the full scale (Cronbach's $\alpha = .94$), and good diagnostic utility (Bovin et al., 2016). PCL-5 scores have also demonstrated convergent and discriminant validity with relevant external correlates (Blevins et al., 2015). In the current sample, internal consistency estimates were high, Cronbach's $\alpha = .94$, McDonald's $\Omega = .94$.

Dissociative symptoms

The Multiscale Dissociation Inventory (MDI; Briere, 2002) is a 30-item questionnaire designed to assess a wide range of dissociative symptoms. The frequency of each symptom

is rated on a 5-point Likert scale ranging from 1 (*never*) to 5 (*very often*), with higher scores indicating higher levels of symptom severity. MDI scales have been found to have high internal consistency and convergent and discriminant validity with relevant external correlates (Jeffirs et al., 2023). In the current sample, internal consistency estimates were high, Cronbach's $\alpha = .96$, McDonald's $\Omega = .97$.

Psychopathology and personality traits

The Personality Assessment Inventory (PAI; Morey, 2007) is a 344-item multiscale questionnaire that is used to assess a broad range of psychopathology and personality traits. The PAI includes 11 subscales that assess clinical concerns (i.e., Anxiety, Depression, Anxiety-Related disorders [e.g., obsessive–compulsive disorder], Mania, Paranoia, Schizophrenia, Borderline Features, Antisocial Features, Somatic Concerns, Alcohol Problems, and Drug Problems), as well as other scales that assess treatment considerations, interpersonal functioning, and response validity. Items are rated on a 4-point Likert scale with response options of F (*false, not at all true*), ST (*slightly true*), MT (*mainly true*), and VT (*very true*). PAI scale scores have been found to have high internal consistency (Cronbach's αs = .81–.86), excellent test–retest reliability (rs = .79–.92), and convergent and discriminant validity (Morey, 2007). The PAI Traumatic Stress subscale (PAI ARD-T) has demonstrated excellent convergent validity with other self-report measures of PTSD, such as the PCL-5 ($r = .74$; Blevins et al., 2015). In the current sample, internal consistency estimates of the PAI scale scores were high, Cronbach's αs = .81–.93, McDonald's Ωs = .81–.93).

Trauma exposure and trauma-related symptoms

Recently updated from the original Detailed Assessment of Posttraumatic Stress (DAPS), the DAPS-2 (Briere, 2001) is a 119-item multiscale questionnaire used to assess trauma exposure and related symptoms. In addition to the trauma exposure assessment, the DAPS-2 contains a total PTSD scale, four subscales corresponding to *DSM-5* PTSD criteria, and additional subscales for assessment of related concerns (e.g., dissociation). Participants indicate the frequency or intensity of traumatic stress reactions on a 5-point Likert scale ranging from 1 (*not at all/never*) to 5 (*very much/4 or more times a week*). The current study only utilized the Posttraumatic Stress Total scale (PTS-T) as a measure of convergent validity. The PTS-T scale has demonstrated high internal consistency (Cronbach's $\alpha = .99$), excellent convergent validity with the PCL-5 ($r = .91$), and excellent discriminant validity (Petri et al., 2020). In the current sample, internal consistency estimates for the PTS-T scale were high, Cronbach's $\alpha = .97$, McDonald's $\Omega = .97$.

Data analysis

To evaluate the internal consistency of CAPS-5-R cluster and total scores, we calculated Cronbach's alpha and McDonald's omega coefficients and examined item-total and interitem correlations for all CAPS-5-R Time 1 administrations, which included interviews from single (i.e., participants who dropped out at Time 2), test–retest, and alternate forms administrations ($n = 55$). To evaluate CAPS-5-R test–retest and interrater reliability, we calculated kappa coefficients for PTSD diagnosis using the SEV3 scoring rule and ICCs for dimensional severity scores for all test–retest administrations ($n = 29$). Regarding test–retest reliability, as a different rater administered the second interview, reliability estimates derived from this method are more precisely referred to as "coefficients of stability and interrater equivalence" because they model two sources of measurement error, namely testing occasion and rater (Crocker & Algina, 1986). The specific ICC was a one-way random effects model for a single measurement, which estimates absolute agreement between independent measurements (McGraw & Wong, 1996).

To evaluate the concordance between the CAPS-5-R scored 0–10 with the CAPS-5 scored 0–4, we calculated Cohen's kappa for PTSD diagnosis, as well as Pearson correlations for severity scores for all alternate forms administrations ($n = 29$). Given that the CAPS-5-R and CAPS-5 are distinct but highly similar, we conceptualized this concordance as alternate forms reliability. Further, to evaluate the performance of CAPS-5 0–4 scores derived from the CAPS-5-R, we recoded CAPS-5-R 0–10 scores into 0–4 scores and calculated ICCs using these recoded CAPS-5-R scores for test–retest, interrater, and alternate forms reliability.

To examine the convergent and discriminant validity of the CAPS-5-R with measures of relevant constructs, we calculated zero-order correlations using all CAPS-5-R Time 1 administrations ($n = 55$). To compare the CAPS-5-R more directly with the CAPS-5, we also calculated zero-order correlations using all administrations in the alternate forms condition ($n = 29$). Following Westen and Rosenthal (2003), we predicted a pattern of correlations based on existing validity evidence. Effect size statistics generated from this approach summarize the fit between predicted and observed patterns of correlations. The first statistic, $r_{\text{alerting-CV}}$, indexes the level of agreement between the predicted and observed correlations. The second statistic, $r_{\text{contrast-CV}}$, additionally considers sample size, median intercorrelation among criterion variables, and the

magnitudes of correlations between the measure being validated and criterion variables. Due to missing questionnaire data, we calculated the unweighted harmonic mean of *n* for each set of criterion measures (see, for example, Poythress et al., 2010). All analyses were conducted in SPSS (Version 29).

RESULTS

Descriptive statistics and internal consistency

CAPS-5-R descriptive statistics are presented in Table 1. Cronbach's alpha, $\alpha = .89$, and omega, $\omega = .90$, values were both high for the CAPS-5-R full scale. Internal consistency was more variable for the symptom cluster, Cronbach's αs = .61–.77, ωs = .61–.80. In line with previous CAPS-5 findings, the internal consistency for the CAPS-5-R avoidance cluster was lower than that of the other three clusters.

Corrected item-total correlations for the full scale ranged from .04 to .71. Item D1 (amnesia), had a very low item-total correlation of .04. After removing D1, the item-total correlation of the remaining 19 items ranged from .23 to .71. Corrected item-total correlations were also calculated for each of the four symptom clusters. For the intrusions cluster, corrected item-total correlations ranged from .37 to .63. Both items within the avoidance cluster had a corrected item-total correlation of .44. For the NACM cluster, corrected item-total correlations ranged from .15 to .68; however, after removing Item D1, item-total correlations ranged from .34 to .68. Lastly, for the AAR cluster, corrected item-total correlations ranged from .20 to .65.

Interitem correlations for the full scale ranged from −.16 to .69, with a mean of .28. Items D1 (amnesia) and E2 (recklessness) had very low interitem correlations, ranging from −.16 to .20 for D1 and from −.11 to .38 for E2. When these items were removed, the remaining 18 symptoms had a mean interitem correlation of .33. We also calculated interitem correlations for each of the four symptom clusters. Interitem correlations for the intrusions cluster ranged from .19 to .51, with a mean of .38. The interitem correlation for the avoidance cluster was .44. For the NACM cluster, interitem correlations ranged from −.02 to .67, with a mean of .32; however, when Item D1 was removed, interitem correlations ranged from .13 to .67, with a mean of .40. For the AAR cluster, interitem correlations ranged from .07 to .69, with a mean of .29; however, when Item E2 was removed, interitem correlations ranged from .09 to .69, with a mean of .37.

Test–retest reliability

A paired samples *t* test was conducted to determine whether the difference between CAPS-5-R Time 1 total score ($M = 53.3$, $SD = 30.1$) and Time 2 total score ($M = 55.1$, $SD = 27.8$) significantly differed from zero. The results were not significant, $t(28) = -0.73$, $p = .472$, indicating total scores did not significantly differ between time points. Test–retest reliability for the CAPS-5-R PTSD diagnosis was substantial, $\kappa = .73$, 95% CI [.48, .97], resulting in 25 of 29 correct classifications. Three participants obtained a PTSD diagnosis at Time 1 but not Time 2, and one participant obtained a PTSD diagnosis at Time 2 but not Time 1. Following Weathers et al. (2018), we added a requirement of a minimum total severity score to improve diagnostic correspondence between the CAPS-5-R and CAPS-5, as discussed later. Using signal detection analysis to optimize the quality of efficiency (κ [.5]; Kraemer, 1992), we found the highest correspondence by requiring minimum total severity scores of 54 for the CAPS-5-R (SEV3/54 scoring rule) and 26 for the CAPS-5 (SEV2/26 scoring rule). Applying the SEV3/54 rule improved test–retest reliability for the CAPS-5-R, $\kappa = .79$, 95% CI [.56, 1.00], resulting in 26 of 29 correct classifications.

Test–retest, interrater, and alternate forms reliability results for dimensional scores are presented in Table 2. Test–retest reliability was very high for the total severity score, ICC = .86, and moderate to high for the four symptom clusters, intrusions: ICC = .81, avoidance: ICC = .66, NACM: ICC = .79, AAR: ICC = .82. Test–retest reliability varied at the item level. Most items demonstrated moderate-to-high test–retest reliability (i.e., ICC values of .50 or higher); however, Items B4, B5, D4, and E2 demonstrated lower reliability across time points.

Interrater reliability

Interrater reliability for the CAPS-5-R PTSD diagnosis was very high, $\kappa = .86$, 95% CI [.68, 1.00], with 27 of 29 correct classifications. This increased when the SEV3/54 scoring rule was implemented, $\kappa = .93$, 95% CI [.80, 1.00], with 28 of 29 correct classifications. Interrater reliability was also very high for total severity score, ICC = .98, and for the four symptom clusters, intrusions: ICC = .96, avoidance: ICC = .92, NACM: ICC = .97, AAR: ICC = .98. At the item level, interrater reliability was good to excellent (i.e., ICC values of .75 or higher); however, derealization demonstrated only moderate interrater reliability, ICC = .63. Interrater reliability was not evaluated for the CAPS-5.

TABLE 2 Reliability coefficients for Clinician-Administered PTSD Scale for *DSM-5*–Revised (CAPS-5-R) total, cluster, and item scores[a]

	Test–retest[b]				Interrater[c]				Alternate forms[d]			
	CAPS-5-R: 0–10 scoring		CAPS-5-R: 0–4 scoring		CAPS-5-R: 0–10 scoring		CAPS-5-R: 0–4 scoring		CAPS-5-R: 0–10 scoring		CAPS-5-R: 0–4 scoring	
Scale/item	ICC	95% CI	ICC	95% CI	ICC	95% CI	ICC	95% CI	*r*	95% CI	ICC	95% CI
Total	.86	[.73, .93]	.87	[.74, .94]	.98	[.96, .99]	.98	[.97, .99]	.95	[.89, .98]	.94	[.88, .97]
INT	.81	[.64, .91]	.74	[.53, .87]	.96	[.92, .98]	.97	[.93, .98]	.82	[.65, .91]	.83	[.66, .91]
B1	.63	[.36, 81]	.49	[.16, .72]	.85	[.71, .93]	.81	[.64, .91]	.62	[.33, .80]	.64	[.36, .81]
B2	.76	[.56, .88]	.70	[.45, .84]	.99	[.97, .99]	.97	[.94, .99]	.81	[.62, .90]	.78	[.58, .89]
B3	.56	[.26, .77]	.66	[.40, .82]	.83	[.68, .92]	.96	[.93, .98]	.68	[.42, .84]	.62	[.34, .80]
B4	.43	[.08, .68]	.12	[−.25, .46]	.90	[.81, .95]	.88	[.76, .94]	.59	[.29, .79]	.57	[.27, .77]
B5	.49	[.16, .72]	.57	[.26, .77]	.95	[.91, .98]	.96	[.91, .98]	.65	[.37, .82]	.63	[.35, .81]
AV	.66	[.40, .83]	.73	[.50, .86]	.92	[.85, .96]	.93	[.85, .96]	.84	[.68, .92]	.88	[.76, .94]
C1	.59	[.29, .78]	.47	[.14, .71]	.90	[.80, .95]	.89	[.79, .95]	.69	[.44, .85]	.71	[.48, .85]
C2	.50	[.17, .73]	.66	[.39, .82]	.95	[.90, .98]	.92	[.83, .96]	.81	[.64, .91]	.86	[.72, .93]
NACM	.79	[.61, .90]	.79	[.60, .89]	.97	[.93, .98]	.96	[.93, .98]	.88	[.75, .94]	.85	[.70, .92]
D1	.71	[.47, .85]	.80	[.62, .90]	.84	[.69, .92]	.90	[.79, .95]	.62	[.33, .81]	.65	[.38, .82]
D2	.64	[.37, .81]	.64	[.37, .81]	.96	[.91, .98]	.97	[.94, .99]	.80	[.61, .90]	.74	[.52, .87]
D3	.57	[.30, .79]	.61	[.32, .79]	.91	[.83, .96]	.92	[.85, .96]	.50	[.17, .73]	.56	[.26, .77]
D4	.30	[−.06, .60]	.24	[−.13, .55]	.89	[.78, .95]	.87	[.75, .94]	.58	[.27, .78]	.52	[.19, .74]
D5	.58	[.29, .78]	.64	[.37, .81]	.93	[.86, .97]	.85	[.70, .92]	.76	[.55, .88]	.68	[.42, .83]
D6	.58	[.28, .78]	.43	[.09, .69]	.96	[.92, .98]	.96	[.92, .98]	.84	[.68, .92]	.77	[.57, .89]
D7	.66	[.40, .83]	.70	[.45, .85]	.92	[.83, .96]	.93	[.85, .96]	.71	[.46, .85]	.70	[.46, .85]
AAR	.82	[.65, .91]	.77	[.57, .89]	.98	[.96, .99]	.98	[.95, .99]	.84	[.69, .92]	.81	[.64, .91]
E1	.62	[.34, .80]	.59	[.30, .78]	.87	[.74, .94]	.92	[.83, .96]	.42	[.06, .68]	.39	[.03, .65]
E2	.41	[.06, .67]	.44	[.10, .69]	1.00	[1.00, 1.00]	1.00	[1.00, 1.00]	1.00	[1.00, 1.00]	1.00	[1.00, 1.00]
E3	.65	[.38, .82]	.65	[.38, .82]	.92	[.85, .96]	.93	[.86, .97]	.76	[.54, .88]	.74	[.52, .87]
E4	.64	[.37, .81]	.61	[.32, .79]	.87	[.74, .93]	.93	[.85, .97]	.73	[.49, .86]	.64	[.37, .81]
E5	.83	[.67, .92]	.79	[.60, .90]	.96	[.92, .98]	.96	[.92, .98]	.91	[.81, .96]	.89	[.78, .95]
E6	.76	[.56, .88]	.71	[.47, .85]	.96	[.93, .98]	.96	[.91, .98]	.44	[.08, .69]	.43	[.08, .68]
Depers.	.63	[.35, .81]	.50	[.17, .73]	.96	[.91, .98]	1.00	[1.00, 1.00]	.82	[.64, .91]	.82	[.65, .91]
Dereal.	.55	[.24, .76]	.32	[−.05, .60]	.63	[.35, .80]	.72	[.48, .86]	.83	[.66, .92]	.82	[.65, .91]

Note: Participants in the test–retest condition ($n = 29$) were used to calculate intraclass correlation coefficients (ICCs) for test–retest and interrater reliability; participants in the alternate forms condition ($n = 29$) were used to calculate ICCs for alternate forms reliability. PTSD = posttraumatic stress disorder; *DSM-5* = *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.); INT = intrusions; AV = avoidance; NACM = negative alterations in cognition and mood; AAR = alterations in arousal and reactivity; Depers. = depersonalization; Dereal = derealization.

[a]CAPS-5-R 0–10 scoring represents CAPS-5-R original scoring, CAPS-5-R 0–4 scoring represents CAPS-5-R scored using CAPS-5 scoring.

[b]Test–retest: CAPS-5-R was administered twice by different raters.

[c]Interrater: Two raters rated the same CAPS-5-R interview.

[d]Alternate forms: CAPS-5-R and CAPS-5 were administered by different raters.

Alternate forms reliability

The CAPS-5-R demonstrated excellent concordance with the CAPS-5. With respect to PTSD diagnostic status, we observed a substantial association between the CAPS-5-R SEV3 and CAPS-5 SEV2 scoring rules, $\kappa = .79$, 95% CI [.57, 1.00]. This resulted in 26 out of 29 correct classifications, with two participants with a PTSD diagnosis on the CAPS-5 but not the CAPS-5-R and one participant with a PTSD diagnosis on the CAPS-5-R but not the CAPS-5. Concordance increased when using the CAPS-5-R SEV3/54 and CAPS-5 SEV2/26 scoring rules, $\kappa = .93$, 95% CI [.80, 1.00]. This resulted in 28 out of 29 correct classifications. At the dimensional level, CAPS-5-R and CAPS-5 scores were strongly correlated, full scale: $r = .95$, intrusions: $r = .82$, avoidance: $r = .84$, NACM: $r = .88$, AAR: $r = .84$. We used Pearson correlations for these analyses given that ratings were not on the same scale (i.e., 0–10 for CAPS-5-R vs. 0–4 for CAPS-5). Most items revealed large associations (i.e., r values of .50 or higher) between the CAPS-5-R and the CAPS-5. However, Items E1 and E6 revealed comparatively weaker associations. We found a very similar pattern of associations when CAPS-5-R scores were recoded into 0–4 and evaluated against CAPS-5 scores using ICCs.

We also examined the possibility that raters may have inadvertently used novel CAPS-5-R prompts while administering the CAPS-5, thereby inflating the degree of concordance between the two versions. The first and second authors reviewed eight randomly selected CAPS-5 interviews and found no instances of novel CAPS-5-R prompts occurring in a CAPS-5 interview, suggesting that raters maintained strict fidelity to the CAPS-5 script.

Convergent and discriminant validity

The CAPS-5-R was correlated as expected with measures of PTSD and constructs theoretically related to PTSD (see Tables 3 and 4). Regarding convergent validity, the strongest associations were observed between the CAPS-5-R and other measures of PTSD (i.e., the PCL-5, DAPS-2 PTS-T scale, and PAI ARD-T scale). Regarding discriminant validity, the CAPS-5-R was modestly correlated with measures of theoretically related constructs (e.g., dissociation, depression, anxiety, borderline features) and least strongly correlated with measures of unrelated constructs (e.g., alcohol use, drug use, antisocial features). We observed a few discriminant correlations that were higher than expected (i.e., schizophrenia, paranoia, and mania). Construct validity effect size correlations between the CAPS-5-R and criterion measures were high, Time 1 interviews only ($n = 55$): $r_{\text{contrast-CV}} = 0.81$; $r_{\text{alerting-CV}} = 0.88$; alternate forms interviews: ($n = 29$): $r_{\text{contrast-CV}} = 0.98$; $r_{\text{alerting-CV}} = 0.90$.

Rater feedback and administration time

At the conclusion of the study, we anonymously surveyed raters to evaluate their receptiveness to CAPS-5-R changes as well as their perceptions of the CAPS-5-R's improvement in various domains (e.g., administration, ease of scoring). Overall, raters provided an overwhelmingly positive review of the revised instrument, particularly regarding unique revisions, such as the frequency rating card and scoring guidelines, and their impact on participant and rater burden. This survey is, of course, limited in that raters may have felt compelled to endorse the CAPS-5-R due to their knowledge of study aims; however, it does provide at least some evidence that the CAPS-5-R revisions are well-received by raters. Although we are only able to provide a narrative summary of these results given the scope of this paper, interested readers may contact the corresponding author for further details regarding the survey. In addition, given that the CAPS-5-R has additional prompts and rates self-blame and other-blame as separate items, we found that it took slightly longer to administer (CAPS-5-R: 49.1 min vs. CAPS-5: 43.3 min). However, this difference was not statistically significant, suggesting that the revisions do not significantly lengthen administration time.

DISCUSSION

This study was the first psychometric evaluation of the CAPS-5-R. Utilizing a trauma-exposed community sample, we found that the CAPS-5-R is highly reliable and valid for making PTSD diagnoses and generating PTSD symptom severity scores. First, CAPS-5-R total and cluster scores demonstrated high internal consistency. As in studies of the CAPS-5, internal consistency for the avoidance cluster was lower than that for the other three symptom clusters, likely due to the avoidance cluster consisting of only two items (Weathers et al., 2018). Mean interitem correlations for the total scale and the four clusters were within the recommended range proposed by Clark and Watson (1995). However, Items D1 (amnesia) and E2 (recklessness) had very low interitem correlations, likely due to their infrequent endorsement.

Second, CAPS-5-R test–retest reliability was substantial for both diagnosis and severity scores. At the item level, CAPS-5-R severity scores generally revealed adequate test–retest reliability. However, several items revealed comparatively lower reliability across time points, including Items

TABLE 3 Predicted and observed correlations between the Clinician-Administered PTSD Scale for *DSM-5*–Revised (CAPS-5-R), CAPS-5, and criterion measures

	Observed correlations					
	Time 1 only[a]		Alternate forms[b]			
Variable	CAPS-5-R: 0–10 scoring	CAPS-5-R: 0–4 scoring	CAPS-5-R: 0–10 scoring	CAPS-5-R: 0–4 scoring	CAPS-5	Predicted correlations
PCL-5	.72	.67	.80	.76	.81	.70
DAPS-2-PTS-T	.69	.66	.84	.80	.85	.70
PAI- ARD-T	.72	.70	.82	.80	.86	.70
MDI	.56	.55	.70	.67	.74	.60
PAI-DEP	.58	.56	.70	.70	.75	.60
PAI-ANX	.46	.43	.70	.67	.68	.50
PAI-BOR	.49	.48	.55	.56	.61	.50
PAI-SCZ	.55	.53	.76	.75	.79	.40
PAI-PAR	.63	.61	.59	.61	.71	.40
PAI-SOM	.33*	.32*	.46*	.45*	.46*	.40
PAI-ALC	.14	.13	.13	.08	.21	.20
PAI-DRG	.24	.23	−.18	−.28	−.09	.20
PAI-ANT	.25	.25	.12	.05	.22	.10
PAI-MAN	.33*	.34*	.24	.27	.31	.10

Note: Lambda values are available upon request. PTSD = posttraumatic stress disorder; *DSM-5* = *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.); PCL-5 = PTSD Checklist for *DSM-5*; DAPS-2-PTS-T = Detailed Assessment of Posttraumatic Stress (2nd ed.)–Posttraumatic Stress Total Scale; PAI = Personality Assessment Inventory; PAI-ARD-T = PAI Anxiety-Related Disorders Traumatic Stress subscale; MDI = Multiscale Dissociation Inventory; PAI-DEP = PAI Depression subscale; PAI-ANX = PAI Anxiety subscale; PAI-BOR = PAI Borderline subscale; PAI-SCZ = PAI Schizophrenia subscale; PAI-PAR = PAI Paranoia subscale; PAI-SOM = PAI Somatization subscale; PAI-ALC = PAI Alcohol Problems subscale; PAI-DRG = PAI Drug Problems subscale; PAI-ANT = PAI Antisocial subscale; PAI-MAN = PAI Mania subscale.

[a]N = 49–55, which includes only participants given a CAPS-5-R at the first interview session

[b]N = 27–29, which includes only participants in the alternate forms condition.

*$p < .05$; **$p < .01$.

B4 (cued distress), B5 (cued physiological reactions), D4 (persistent negative emotional state), and E2 (recklessness). The relative instability of these items may be due to several factors, including a focus on more transient experiences that may differ from week to week, discrepancies in participant responses across time points, or rater error. However, differences due to rater error are likely to be minimal given that interrater reliability was very high for these items.

Third, the interrater reliability of the CAPS-5-R was almost perfect for both diagnosis and severity scores, demonstrating remarkable consistency between raters when scoring the same interview. The revisions for the CAPS-5-R, such as increased variability in item scores, clarification of item prompts, and expanded scoring guidelines, likely contributed to the very high level of agreement between raters. Interrater reliability at the item level was generally excellent, with only one item (derealization) demonstrating moderate consistency between raters.

Fourth, the CAPS-5-R was very strongly associated with the CAPS-5 for both diagnosis and severity scores. Diagnostic concordance between the CAPS-5-R and the CAPS-5 was substantial using basic scoring rules (i.e., SEV3 for the CAPS-5-R and SEV2 for the CAPS-5) and almost perfect when a minimum total severity score was added as a requirement (54 for the CAPS-5-R and 26 for the CAPS-5). Regarding severity scores, CAPS-5-R standard scores (0–10) performed comparably to CAPS-5-R scores recoded to CAPS-5 scores (0–4), with both scoring methods demonstrating excellent internal consistency, test–retest and interrater reliability, and convergent and discriminant validity. Taken together, these findings demonstrate that the CAPS-5-R corresponds very closely to its predecessor, indicating its backward compatibility with and suitability as a replacement for the CAPS-5.

Regarding convergent and discriminant validity, the CAPS-5-R was strongly correlated with other measures of PTSD, moderately correlated with measures of related constructs (i.e., dissociation, depression, anxiety, borderline personality features, schizophrenia, paranoia, and somatization), and weakly and nonsignificantly correlated with measures of unrelated constructs (i.e., alcohol problems, drug problems, antisocial personality features). Although associations between the CAPS-5-R and some constructs

TABLE 4 Effect size statistics for the Clinician-Administered PTSD Scale for *DSM-5*–Revised (CAPS-5-R) and CAPS-5

	Time 1 only[a]		Alternate forms[b]		
Quantity	CAPS-5-R 0–10 scoring	CAPS-5-R 0–4 scoring	CAPS-5-R 0–10 scoring	CAPS-5-R 0–4 scoring	CAPS-5
$r_{\text{alerting-CV}}$	.88	.86	.90	.88	.89
$r_{\text{contrast-CV}}$	.81	.77	.98	.97	.97
95% CI	[.69, .89]	[.63, .86]	[.95, .99]	[.94, .99]	[.94, .99]
Z_{contrast}	7.20	6.61	8.87	8.61	8.46
t_{contrast}	9.62	8.42	23.29	21.25	20.14
p_{contrast}	< .001	< .001	< .001	< .001	< .001

Note: PTSD = posttraumatic stress disorder; *DSM-5* = *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.); CV = construct validity.

[a]N = 49–55, which includes only participants who completed the CAPS-5-R at the first interview session; however, due to missing questionnaire data, the harmonic mean used for computation was n = 52.

[b]N = 27–29, which includes only participants in the alternate forms condition; however, due to missing questionnaire data, the harmonic mean used for computation was n = 28.

[c]$r_{\text{alerting-CV}}$ indexes the level of agreement between the predicted and observed correlations; $r_{\text{contrast-CV}}$ additionally considers sample size, median intercorrelation among criterion variables, and the magnitudes of correlations between the measure being validated and criterion variables.

were stronger than expected (i.e., schizophrenia, mania, and paranoia), this is likely due to item content that overlaps with aspects of PTSD (i.e., social detachment, hypervigilance, irritability). Nonetheless, large construct validity effect sizes were found for the CAPS-5-R, indicating a close fit between predicted and observed patterns of association with a variety of external correlates.

Our results should be evaluated in light of several limitations. First, the sample size was modest, which may have resulted in insufficient power for some analyses, including evaluations of differences between Time 1 and Time 2 mean CAPS-5-R scores and differences between kappa values based on different diagnostic scoring rules. Larger samples would increase confidence in the stability and replicability of the results. Second, we recruited and studied a trauma-exposed community sample, so it is not clear how well the results generalize to other trauma populations, such as veterans or treatment-seeking trauma survivors. Third, our sample had limited racial/ethnic diversity, which precluded further investigation of how these demographic factors may have influenced the study results. Fourth, interview modality (i.e., in-person vs. remote) was self-selected and not randomized, which may have influenced study results; however, post hoc analyses revealed that PTSD severity was not significantly different between remote and in-person participants. Fifth, although we found no evidence for contamination effects during CAPS-5 administrations, it is likely that CAPS-5 raters retained knowledge of the scoring guidelines that are extensively detailed in the CAPS-5-R, which may have influenced CAPS-5 ratings and resulted in a high concordance rate between both measures. Lastly, although not necessarily a limitation per se, it should be noted that raters were well-qualified, carefully trained, and supervised by the second author. Less qualified raters or raters with less training may not generate the same high levels of reliability and validity.

In conclusion, this study provides promising evidence that the CAPS-5-R is a psychometrically sound measure of *DSM-5-TR* PTSD and a valuable update of the current CAPS-5. The measure demonstrated reliability and validity at levels comparable to or even exceeding those reported in published evidence for the CAPS-5, and it demonstrated a high degree of backward compatibility through its strong association with the CAPS-5. The CAPS-5-R's simplified formatting streamlines administration; its revised prompts facilitate accurate inquiry; its new intermediate intensity anchors allow more nuanced ratings; and its expanded scoring grid increases the range and variability of item, symptom cluster, and total severity scores. Indeed, when comparing the distribution of CAPS-5-R (0–10) scores to CAPS-5 (0–4) scores, it is clear that the expanded scoring grid extends the higher end of the scale and addresses the restriction of range (see Supplementary Figure S3).

Further, despite all its revisions, the CAPS-5-R remains highly similar to the CAPS-5. Experienced CAPS-5 raters can easily adapt to the new formatting, prompts, and scoring. This is due, in large part, to the fact that the changes were motivated by rater feedback and provides useful solutions to some limitations of the CAPS-5. More importantly, raters who are new to the CAPS will be able to learn the CAPS-5-R more quickly and to a much higher level of quality due to the more specific and comprehensive prompts, specific scoring guidelines, and streamlined format. These conclusions are supported, in part, by the responses to the rater feedback survey as well as informal observations of the second author, who has now trained dozens of raters on the CAPS-5-R, ranging from those who are new

to the CAPS to highly experienced CAPS-5 raters. Raters have been found to master standard administration more quickly on the CAPS-5-R, which allows training to focus on the crucial tasks of scoring calibration and developing a conceptual understanding of specific PTSD symptoms.

Future directions include replicating these findings in larger samples in other trauma populations and evaluating the CAPS-5-R's sensitivity to clinical change in clinical trials. At this point, researchers and clinicians are encouraged to switch to the CAPS-5-R when feasible. Experienced CAPS raters will appreciate the various improvements and adjust quickly to the new scoring grid, and new CAPS raters will be able to learn standard administration and scoring more efficiently. Further, the backward compatibility permits straightforward integration of new CAPS-5-R data with existing CAPS-5 datasets. This will ensure continuity in the transition to the CAPS-5-R, the latest iteration in the long-standing tradition of the CAPS.

OPEN PRACTICES STATEMENT

Per Auburn University Institutional Review Board regulatory approval, the data from this study are not publicly available. However, interested researchers may contact the corresponding author at frank.weathers@va.gov and can apply for institutional approval to use the data.

AUTHOR NOTE

Brianna N. Jackson, Frank W. Weathers, and Thomas J. Preston receive consultant fees for providing Clinician-Administered PTSD Scale for *DSM-5* (CAPS-5) training and interviewing services to multiple research projects. Funding for participant compensation was provided by Psychological Assessment Resources (PAR).

ORCID

Brianna N. Jackson https://orcid.org/0000-0002-9803-676X

REFERENCES

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., text rev.).

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). https://doi.org/10.1176/appi.books.9780890425787

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*(6), 489–498. https://doi.org/10.1002/jts.22059

Bovin, M. J., Marx, B. P., Weathers, F. W., Gallagher, M. W., Rodriguez, P., Schnurr, P. P., & Keane, T. M. (2016). Psychometric properties of the PTSD Checklist for *Diagnostic and Statistical Manual of Mental Disorders*–Fifth Edition (PCL-5) in veterans. *Psychological Assessment*, *28*(11), 1379–1391. https://doi.org/10.1037/pas0000254

Briere, J. (2001). *Detailed Assessment of Posttraumatic Stress*. Psychological Assessment Resources, Inc.

Briere, J. (2002). *Multiscale Dissociation Inventory: Professional manual*. Psychological Assessment Resources.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart, & Winston.

Jeffirs, S. M., Petri, J. M., Camden, A. A., Jackson, B. N., & Weathers, F. W. (2023). Psychometric evaluation of the Multiscale Dissociation Inventory (MDI): A replication and extension in two trauma-exposed samples. *European Journal of Trauma & Dissociation*, *7*(2), Article 100324. https://doi.org/10.1016/j.ejtd.2023.100324

Kraemer, H. C. (1992). *Evaluating medical tests: Objective and quantitative guidelines*. SAGE Publications.

Lee, D. J., Bovin, M. J., Weathers, F. W., Palmieri, P. A., Schnurr, P. P., Sloan, D. M., Keane, T. M., & Marx, B. P. (2019). Latent factor structure of *DSM–5* posttraumatic stress disorder: Evaluation of method variance and construct validity of novel symptom clusters. *Psychological Assessment*, *31*(1), 46–58. https://doi.org/10.1037/pas0000642

Lee, D. J., Weathers, F. W., Thompson-Hollands, J., Sloan, D. M., & Marx, B. P. (2022). Concordance in PTSD symptom change between *DSM-5* versions of the Clinician-Administered PTSD Scale (CAPS-5) and PTSD Checklist (PCL-5). *Psychological Assessment*, *34*(6), 604–609. https://doi.org/10.1037/pas0001130

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46. http://doi.org/10.1037/1082-989X.1.1.30

Morey, L. C. (1996). *An interpretive guide to the Personality Assessment Inventory (PAI)*. Psychological Assessment Resources.

Morey, L. C. (2007). *The Personality Assessment Inventory professional manual* (2nd ed.). Psychological Assessment Resources.

Parmenter, M. E., Lederman, S., Weathers, F. W., Davis, L. L., Vaughn, B., Engels, J., & Sullivan, G. M. (2024). A Phase 3, randomized, placebo-controlled, trial to evaluate the efficacy and safety of bedtime sublingual cyclobenzaprine (TNX-102 SL) in military-related posttraumatic stress disorder. *Psychiatry Research*, *334*, Article 115764. https://doi.org/10.1016/j.psychres.2024.115764

Petri, J. M., Weathers, F. W., Witte, T. K., & Silverstein, M. W. (2020). The Detailed Assessment of Posttraumatic Stress–Second Edition (DAPS-2): Initial psychometric evaluation in an MTurk-recruited, trauma-exposed community sample. *Assessment*, *27*(6), 1116–1127. https://doi.org/10.1177/1073191119880963

Poythress, N. G., Edens, J. F., Skeem, J. L., Lilienfeld, S. O., Douglas, K. S., Frick, P. J., Patrick, C. J., Epstein, M., & Wang, T. (2010).

Identifying subtypes among offenders with antisocial personality disorder: A cluster-analytic study. *Journal of Abnormal Psychology*, *119*(2), 389–400. https://doi.org/10.1037/a0018611

Prins, A., Bovin, M. J., Smolenski, D. J., Marx, B. P., Kimerling, R., Jenkins-Guarnieri, M. A., Kaloupek, D. G., Scnurr, P. P., Kaiser, A. P., Leyva, Y. E., & Tiet, Q. Q. (2016). The Primary Care PTSD Screen for *DSM-5* (PC-PTSD-5): Development and evaluation within a veteran primary care sample. *Journal of General Internal Medicine*, *31*(10), 1206–1211. https://doi.org/10.1007/s11606-016-3703-5

Resick, P. A., Straud, C. L., Wachen, J. S., LoSavio, S. T., Peterson, A. L., McGeary, D. D., Young-McCaughan, S., Taylor, D. J., Mintz, J., & for the STRONG STAR Consortium and the Consortium to Alleviate PTSD. (2023). A comparison of the CAPS-5 and PCL-5 to assess PTSD in military and veteran treatment-seeking samples. *European Journal of Psychotraumatology*, *14*(2), Article 2222608. https://doi.org/10.1080/20008066.2023.2222608

Schnurr, P. P., Chard, K. M., Ruzek, J. I., Chow, B. K., Resick, P. A., Foa, E. B., Marx, B. P., Friedman, M. J., Bovin, M. J., Caudle, K. L., Castillo, D., Curry, K. T., Hollifield, M., Huang, G. D., Chee, C. L., Astin, M. C., Dickstein, B., Renner, K., Clancy, C. P., … Shih, M. C. (2022). Comparison of prolonged exposure vs cognitive processing therapy for treatment of posttraumatic stress disorder among U.S. veterans: A randomized clinical trial. *JAMA Network Open*, *5*(1), Article e2136921. https://doi.org/10.1001/jamanetworkopen.2021.36921

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for *DSM-IV* and *ICD-10*. *Journal of Clinical Psychiatry*, *59*(Suppl 20), 22–33.

Weathers, F. W., Blake, D. D., Schnurr, P. P., Kaloupek, D. G., Marx, B. P., & Keane, T. M. (2013a). *The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5)*. https://www.ptsd.va.gov/professional/assessment/adult-int/caps.asp

Weathers, F. W., Blake, D. D., Schnurr, P. P., Kaloupek, D. G., Marx, B. P., & Keane, T. M. (2013b). *The Life Events Checklist For DSM-5 (LEC-5)*. https://www.ptsd.va.gov/professional/assessment/te-measures/life_events_checklist.asp

Weathers, F. W., Bovin, M. J., Lee, D. J., Sloan, D. M., Schnurr, P. P., Kaloupek, D. G., Keane, T. M., & Marx, B. P. (2018). The Clinician-Administered PTSD Scale for *DSM–5* (CAPS-5): Development and initial psychometric evaluation in military veterans. *Psychological Assessment*, *30*(3), 383–395. https://doi.org/10.1037/pas0000486

Weathers, F. W., Byrne, C. A., Marx, B. P., Schnurr, P. P., & Keane, T. M. (2022). *The Clinician-Administered PTSD Scale for DSM-5, Revised (CAPS-5-R)–Past Month–Initial Interview Version*.

Weathers, F. W., Keane, T. M., & Davidson, J. R. T. (2001). Clinician-Administered PTSD scale: A review of the first ten years of research. *Depression and Anxiety*, *13*(3), 132–156. https://doi.org/10.1002/da.1029

Weathers, F. W., Litz, B. T., Keane, T. M., Palmier, P. A., Marx, B. P., & Schnurr, P. P. (2013). *The PTSD Checklist for DSM-5 (PCL-5)*. https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp

Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology*, *84*(3), 608–618. https://doi.org/10.1037/0022-3514.84

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Jackson, B. N., Weathers, F. W., Jeffirs, S. M., Preston, T. J., & Brydon, C. M. (2024). The revised clinician-administered PTSD scale for *DSM-5* (CAPS-5-R): Initial psychometric evaluation in a trauma-exposed community sample. *Journal of Traumatic Stress*, 1–13. https://doi.org/10.1002/jts.23093