

# Change in Posttraumatic Stress Disorder Symptoms: Do Clinicians and Patients Agree?

Candice M. Monson and Jaimie L. Gradus  
National Center for PTSD and Boston University School of  
Medicine

Yinong Young-Xu  
National Center for PTSD

Paula P. Schnurr  
National Center for PTSD and Dartmouth Medical School

Jennifer L. Price  
Georgetown College

Jeremiah A. Schumm  
Harvard Medical School and Veterans Affairs Boston Healthcare System

This study assessed the longitudinal association between clinician and patient ratings of posttraumatic stress disorder (PTSD) symptoms over the course of 2 different randomized clinical trials of veterans with chronic PTSD. One trial, the Department of Veterans Affairs Cooperative Study 420 (CSP 420;  $N = 360$ ) compared trauma-focused and present-centered group therapies, and the 2nd trial compared cognitive processing theory and a waitlist control condition ( $N = 60$ ). Linear mixed effects modeling revealed significant associations between clinician ratings (Clinician-Administered PTSD Scale; CAPS; D. D. Blake et al., 1990) and patient ratings (Posttraumatic Stress Disorder Checklist; PCL; F. W. Weathers, B. T. Litz, J. A. Herman, J. A. Huska, & T. M. Keane, 1993) in total and symptom clusters of PTSD. Contrary to hypothesis, the amount of change on the CAPS ranged from .75 to .82 standard deviations for every 1 standard deviation change on the PCL. The CAPS and PCL were more closely associated in the trauma-focused vs. present-centered treatment condition in CSP 420, and especially regarding hyperarousal symptoms. When comparing categorization of clinically significant change on the CAPS and PCL, the authors found no differences in the percentages of agreement between clinicians and patients in improvement and exacerbation. The value of multimodal assessment of PTSD treatment outcomes is discussed.

**Keywords:** PTSD, assessment, treatment outcomes, concordance, psychometrics

Most randomized controlled trials of medication or psychotherapy for posttraumatic stress disorder (PTSD) utilize clinician ratings of symptomatology as the primary measure in assessing

outcomes. Although patient self-ratings of symptoms are often simultaneously assessed, there has been less emphasis placed on these outcomes and surprisingly little research that has evaluated the agreement between clinicians and patients about the direction and degree of changes in symptoms following treatment. Meanwhile, there are several contexts in which patient-rated measures of PTSD might be used alone if researchers can determine that they adequately characterize changing symptomatology over time, including repeated assessments during the course of treatment, busy clinical practices, and pilot studies with limited resources. Thus, we examined how changes in patient-rated symptoms of PTSD correspond with changes in clinician-rated symptoms over the course of treatment and time in two randomized clinical trials of veterans with chronic PTSD using data analytic methods that capitalize on the longitudinal nature of the outcome data.

We are aware of only one study to date that has sought to evaluate the correspondence between clinicians' ratings and patients' ratings of PTSD over the course of treatment and follow-up (Forbes, Creamer, & Biddle, 2001). This study utilized the Clinician-Administered PTSD Scale (CAPS; Blake et al., 1990), a well-established clinician interview for evaluating PTSD symptom

---

Candice M. Monson and Jaimie L. Gradus, Women's Health Sciences Division, National Center for PTSD, Boston, Massachusetts, and Department of Psychiatry, Boston University School of Medicine; Yinong Young-Xu, Executive Division, National Center for PTSD, White River Junction, Vermont; Paula P. Schnurr, Executive Division, National Center for PTSD; and Department of Psychiatry, Dartmouth Medical School; Jennifer L. Price, Department of Psychology, Georgetown College; Jeremiah A. Schumm, Department of Psychiatry, Harvard Medical School; and Mental Health Service Line, Veterans Affairs Boston Healthcare System, Boston, Massachusetts.

This research was supported by a Department of Veterans Affairs Cooperative Studies Program Clinical Research Career Development Award to Candice M. Monson. The treatment trials were also supported by the Department of Veterans Affairs Cooperative Studies Program.

Correspondence concerning this article should be addressed to Candice M. Monson, Women's Health Sciences Division, National Center for PTSD, Veterans Affairs Boston Healthcare System, 150 South Huntington Avenue (116B-3), Boston, MA 02130. E-mail: Candice.Monson@va.gov

severity and diagnosis, as well as the PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993). The PCL is a frequently used self-report measure of PTSD symptoms that has good psychometric properties (Blanchard, Jones-Alexander, Buckley, & Forneris, 1996). This sample included 97 Vietnam veterans with combat-related PTSD who participated in a multifaceted PTSD treatment program. To assess the association between the CAPS and the PCL, the authors calculated cross-sectional correlations prior to treatment and again at the 9-month follow-up. The correlations at the respective time points varied from  $r = .30$  at pretreatment to  $r = .62$  at the 9-month follow-up. They also found that the effect sizes for change from pretreatment to the 9-month follow-up were greater on the CAPS compared with the PCL ( $d = 0.84$  vs.  $d = 0.59$ , respectively).

Although Forbes et al. (2001) did not use some of the methods typically found in efficacy studies (i.e., independent reliability monitoring, assessor blinding to patient treatment status and time period of assessment) and analyzed the data cross-sectionally, their study represents an important step in the evaluation of the correspondence in patient and clinician ratings of PTSD during the course of treatment. It might be expected that clinician- and patient-rated PTSD symptoms would be highly correlated, given that both the CAPS and the PCL are based on patient self-report and both include the rating of the 17 symptoms of PTSD found in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; *DSM-IV-TR*; American Psychiatric Association, 2000). Consistent with this notion, psychometric research has revealed high cross-sectional correlations between the CAPS and the PCL across a range of traumatized samples (i.e.,  $r_s \geq .90$ ; e.g., Blanchard et al., 1996; Forbes et al., 2001). Although the CAPS and the PCL both involve the rating of PTSD symptoms listed in the *DSM-IV-TR*, the orientation to answering the questions differs on each measure. The CAPS asks clinicians to inquire separately about the frequency and intensity of PTSD symptoms, and the PCL asks patients to rate the degree to which they are "bothered" by each of the 17 symptoms. Also, when administering the CAPS, clinicians must make the determination that a given emotional numbing or hyperarousal symptom is not better accounted for by another mental health condition. Patients do not make this differential diagnostic determination on the PCL. Thus, discrepancies in the results of these two measures might be found.

Forbes et al.'s study provides evidence for differences between CAPS and PCL measurements through findings of lower and variable cross-sectional correlations, as well as of an effect size advantage in symptom change on the CAPS versus the PCL. The effect size advantage on the CAPS, relative to the PCL, has been found in a few other studies with veterans and nonveterans (Blanchard et al., 2003; Blanchard, Hickling, & Forneris, 1997; Monson, Schnurr, Stevens, & Guthrie, 2004). However, in a recent large study of women veterans with PTSD, Schnurr et al. (2007) found an effect size advantage on the PCL compared with the CAPS. Researchers conducting meta-analyses of studies that utilized a range of clinician and patient assessment measures have indicated greater effect size changes from clinician than from patient ratings (Van Etten & Taylor, 1998).

To further assess the correspondence between clinician and patient ratings of PTSD symptoms over the course of psychological treatment and time, we used data from two different controlled

trials. The first is the Department of Veterans Affairs Cooperative Study 420 (CSP 420; Schnurr et al., 2003), a multisite randomized clinical trial comparing trauma-focused group therapy (TFGT) with present-centered group therapy (PCGT) in a sample of 360 Vietnam veterans with chronic PTSD. The other trial included in this investigation compared cognitive processing therapy (CPT) with a waitlist condition in a sample of 60 veterans with chronic PTSD (Monson et al., 2006). We chose to examine these two studies because both employed the CAPS and the PCL. However, the two studies had different treatment designs, sample sizes, and outcomes. Specifically, the CSP 420 was a large trial that compared two active treatments, and researchers found no differences between the treatments on any of the outcome measures in intention-to-treat analyses. Moreover, there were only modest improvements in PTSD in both conditions. In contrast, the CPT trial was a smaller waitlist-controlled trial that found significant improvements in PTSD outcomes in the active treatment condition versus the waitlist condition, and the improvements were closer to those found in civilian samples (Bradley, Greene, Russ, Dutra, & Westen, 2005). The use of a waitlist in this trial also allowed us to examine correspondence between the CAPS and the PCL over time only. Through the use of these disparate studies, we could evaluate possible variation in the longitudinal association between the PCL and the CAPS.

To take full advantage of the longitudinal nature of the two data sets, we used linear mixed effects modeling for repeated measures, a data analytic method that evaluates the association between two continuous variables assessed on multiple occasions. These analyses allowed us to evaluate how well CAPS and PCL ratings corresponded with each other over the course of time in each of the studies. This analytic method also measures individual trajectories in symptom change over time instead of overall group averages at discrete time points. It is also able to accommodate missing data.

Based on prior research, our primary hypothesis was that there would be longitudinal associations between total CAPS and total PCL scores, as well as their cluster subtotals, across the course of treatment and in follow-up assessments in the two studies. However, consistent with several prior studies using the CAPS and the PCL, and prior meta-analyses of clinician- and patient-rated PTSD outcomes more generally, we hypothesized that the CAPS would reveal greater change in symptomatology relative to the PCL. The second hypothesis was that the longitudinal associations between the CAPS and the PCL would be stronger in the TFGT and the CPT conditions than in the PCGT and the waitlist control conditions, given the symptom-focused nature of the trauma-focused treatments. We also evaluated the correspondence between categorizations of clinically significant change on the CAPS and the PCL. We hypothesized that the clinicians and patients would be less likely to agree on clinically significant improvements than on no change and clinically significant exacerbations given the above cited literature.

## Method

### *Participants and Procedure*

**CSP 420.** Three hundred sixty male Vietnam veteran outpatients diagnosed with combat-related PTSD were recruited from 10

Department of Veterans Affairs Medical Centers across the nation. Participants were about 50 years old, mostly White, and most had a post-high school education. Approximately half were married and approximately half were unemployed. Slightly over half of the sample received disability payments for their service-connected PTSD. Participants' combat traumas had occurred approximately 25 years prior to treatment.

Participants were randomly assigned to either TFGT or PCGT. TFGT included psychoeducation about PTSD, identification of coping resources for managing symptoms, exposure and cognitive restructuring related to members' identified war zone scenes, and relapse prevention to help members maintain gains made during the course of treatment. PCGT included psychoeducation about PTSD and its impact on relationships and problem-solving, as well as a focus on processing current experiences within the sessions. Both treatment groups consisted of 30 weekly sessions followed by 5 monthly booster sessions. Fidelity to the treatments, as assessed by independent evaluators, was good. Average adherence ratings for the two forms of treatment were 0 for TFGT and  $-0.07$  for PCGT ( $0 = \text{just right}$ ), and the average competence ratings for the two forms of treatment were 2.24 for TFGT and 1.98 for PCGT ( $0 = \text{poor}$ ,  $4 = \text{highly competent}$ ). All participants were assessed with the CAPS and the PCL at pretreatment, following treatment completion (7 months), and following completion of booster sessions (12 months). A portion of the participants were also assessed at 18 ( $n = 201$ ) and 24 ( $n = 102$ ) months postrandomization; these data were included in the linear mixed effects modeling.

**CPT.** Sixty veterans (54 men, 6 women) with military-related PTSD were randomized into the trial. Most participants served in Vietnam (78%), were White, were in their mid-50s, and had a post-high school education. Nearly two thirds were married and nearly two thirds were unemployed. Slightly over half of the sample received disability payments for their service-connected PTSD. Approximately 80% of the participants' index traumas were combat-related; the remaining participants identified non-combat sexual (17%) and physical (3%) assaults as their index traumas. All index traumas had occurred at least 15 years prior to treatment.

Participants were assigned at random to receive CPT or to be assigned to a waiting list—a delayed treatment condition. CPT is a manualized, 12-session, cognitive-behavioral therapy that has a primary focus on cognitive interventions for PTSD (Resick, Monson, & Chard, 2007). Treatment fidelity, as rated by an independent evaluator, was good (i.e., 93% adherence to the specific elements of the therapy, and 100% adherence to the nonspecific but essential elements of the therapy). All participants were assessed with the CAPS and the PCL at baseline, midtreatment or after 3 weeks of waiting, postrandomization or after 6 weeks of waiting, and 1-month follow-up or after 10 weeks of waiting, depending on study condition.

### Measures

**CAPS** (Blake et al., 1990). The CAPS was used to assess current clinician-rated PTSD symptom severity. The CAPS is a widely used, semistructured clinical interview that measures the frequency and intensity of PTSD symptoms as outlined in the *DSM-IV-TR*. The frequency and intensity ratings for each of the

17 items on the CAPS are summed to create a total score, with possible scores ranging from 0 to 136. Cluster subscale scores are created by summing the frequency and intensity of the items related to the respective PTSD cluster. The CAPS has excellent psychometric properties (see Weathers, Keane, & Davidson, 2001, for review). In both trials, master's- and doctoral-level independent clinician interviewers, who were blind to the condition, administered the CAPS. There were at least 10 different assessors in each of the trials, and the assessors did not necessarily follow the same patient over the course of the assessment points of the respective trials. At the initial assessment, the time frame of reference for the symptomatology was the prior month in both trials. In the CSP 420, the time of reference for subsequent administrations was 1 month, and in the CPT the time frame of reference for subsequent administrations was 1 week. These varying time frames are allowable in the CAPS administration. The interrater reliability of the CAPS interviews was determined by independent experts who rated a proportion of the interviews conducted in the studies. The intraclass correlations between the assessors and reliability monitors for total CAPS were excellent in both trials (correlations ranged from .85 to .87).

**PCL** (Weathers et al., 1993). On the PCL, participants rate the degree to which they are bothered by each of the symptoms of PTSD from the *DSM-IV-TR* on a Likert-type scale ( $1 = \text{not at all}$ ,  $5 = \text{extremely}$ ), and items are summed to create a total PCL score ranging from 17 to 85. Cluster subscale scores are created by summing the items related to the respective PTSD cluster. The PCL has good psychometric properties across various trauma populations (e.g., Blanchard et al., 1996; Forbes et al., 2001). Consistent with the CAPS, the time of reference for reporting was the prior month in both trials at initial assessment. In the CSP 420, the time of reference for subsequent administrations was 1 month, and in the CPT trial, the time of reference for subsequent administrations was 1 week. At baseline, the alpha coefficients for the PCL in the CSP 420 and the CPT trials were .90 and .81, respectively.

### Data Analysis

The primary hypothesis of this study was that patient-rated PTSD symptoms would correspond with clinician-rated PTSD symptoms on a longitudinal, or within-subject, basis when repeatedly assessed over the course of time. Thus, we used linear mixed effects modeling for longitudinal analysis (SAS PROC MIXED, Version 8). Like other regression techniques that include random effect estimation, these analyses account for individual participants' trajectories of symptomatology and control for correlations between repeated assessments among individual participants. As a result, they provide accurate estimations of regression coefficients that represent the associated unit change between the dependent and independent variables. We controlled for the cross-sectional (between-subject) correlation in CAPS and PCL scores by including baseline scores in the model in order to accurately estimate their longitudinal (within-subject) association. We also allowed the intercept and slope for change to vary randomly among participants and used an unstructured covariance structure in modeling

the data.<sup>1</sup> Separate analyses were conducted on each trial to evaluate the consistency of the findings across the two trials. We present both the unstandardized and standardized coefficients from these analyses. The unstandardized coefficients allow for prediction of change on the CAPS based on changes on the PCL. The standardized coefficients take into account the different scaling on the CAPS and on the PCL, and thereby allow a direct comparison of the amount of change on the PCL relative to the amount of change on the CAPS over time. We also calculated  $R^2$  as an effect size measure, which represents the estimated proportion of variance in the CAPS scores accounted for by the PCL scores (Hardin & Hilbe, 2007). Means and standard deviations for available participants at each time in each of the trials were also calculated and are presented in Table 1.

To test the hypothesis that there would be differences in the longitudinal association between the CAPS and the PCL by study condition we subsequently included an interaction term involving treatment type in the models. In the interpretation of the interaction terms, we treated the control conditions in each trial (i.e., PCGT and waitlist) as the reference groups. These were coded as 0. The trauma-focused interventions (i.e., TFGT and CPT) were coded as 1. Thus, the coefficients presented in Table 2 represent the significant difference in longitudinal association between the CAPS and PCL over time in the TFGT and CPT conditions compared with their respective control conditions. The full mixed effects model used to test the first two hypotheses was

$$E(Y_{ij}|b_i) = \beta_1 + \beta_2\text{PCL}_{ij} + \beta_3\text{PCL}_{i1} + \beta_4\text{Treatment}_i \\ + \beta_5\text{PCL}_{ij} \times \text{Treatment}_i + b_{1i} + b_{2i}\text{PCL}_{ij},$$

where  $Y_{ij}$  is the CAPS score for the  $j$ th measurement on the  $i$ th individual, and  $\text{PCL}_{ij}$  denotes the PCL measurement time for the  $j$ th measurement on the  $i$ th individual, whereas  $\text{Treatment}_i = 1$  if the  $i$ th individual was assigned to the experimental treatment, and  $\text{Treatment} = 0$  otherwise. In the above model, the CAPS and the PCL were treated as continuous variables.

To further assess the degree to which changes on these measures correspond at the individual participant level, we assessed the proportion of concordance between the two measures using clinically significant change criteria. First, we calculated clinically significant change criteria on the CAPS and the PCL (Jacobson & Truax, 1991), which are 20 and 10 points, respectively. We then generated difference scores on the PCL and on the CAPS between all possible pairs of study measurement time points in each study (i.e., baseline to midtreatment, midtreatment to postrandomization, baseline to postrandomization). Finally, we used McNemar's (1947) test of paired proportions to compare the proportion of agreement in clinically significant change criteria on the CAPS and the PCL across the possible pairs of study measurement. This was used to examine the hypothesis that clinicians and patients would agree less on clinically significant improvement than on no change and clinically significant exacerbation.

## Results

As expected, there were longitudinal associations between total CAPS and total PCL scores across the different conditions in both studies (see Table 2). The unstandardized coefficients for the two

studies indicate that for every 1-point change in total PCL, there was a corresponding 1.28- to 1.50-point change in the same direction in total CAPS scores. The standardized coefficients indicate that for every standard deviation change in total PCL scores, there was a 0.75- to 0.82- $SD$  change in total CAPS scores. The standardized coefficients for the CAPS subscales ranged from a 0.65- to 0.84- $SD$  change for every 1-point standard deviation change in the corresponding PCL subscale. Table 1 contains the total and cluster means and standard deviations for the CAPS and for the PCL for available participants at each time point for each trial.

We expected that the standardized regression coefficients for the PCL would be larger in the symptom-focused treatment conditions of the two trials (i.e., TFGT and CPT) compared with the control conditions. As shown in Table 3, there was a significant difference in the magnitude of associated change between the total PCL scores and the total CAPS scores based on trial condition in CSP 420. The  $R^2$  for the model in the TFGT condition was .61 versus an  $R^2$  of .53 for the model in the PCGT condition. This significant interaction in total scores appears to be attributable to a differential association in ratings of hyperarousal symptoms in this trial because there were not significant interactions in the associated longitudinal changes in reexperiencing and avoidance/numbing symptoms. Figure 1 shows the closer association between changes on the CAPS and the PCL Hyperarousal subscales scores in the TFGT condition compared with the PCGT condition. The  $R^2$  for the association in the TFGT condition was .59 versus an  $R^2$  of .39 for the association in the PCGT condition. Contrary to our hypothesis, there were no significant differences in the magnitude of the associated changes between the CAPS and the PCL found in the CPT trial.

Table 4 shows that in both trials the proportion of agreement in clinically significant improvement and the proportion of agreement in clinically significant exacerbation did not differ: CSP 420,  $\chi^2(1, N = 300) = 0.62, p = .41$ ; CPT,  $\chi^2(1, N = 53) = 1.00, p = .32$ . However, the percentages of agreement on clinically significant improvement and exacerbation differed from the proportion of agreement on no change: CSP 420,  $\chi^2(1, N = 1,361) = 18.06, p < .001$ , for improvement comparison and  $\chi^2(1, N = 1,225) = 14.41, p < .001$ , for exacerbation comparison; CPT,  $\chi^2(1, N = 167) = 6.42, p < .01$ , for improvement comparison and  $\chi^2(1, N = 130) = 3.86, p < .05$  for exacerbation comparison.

## Discussion

Surprisingly little attention has been paid to the issue of correspondence between clinician and patient ratings of symptoms over the course of PTSD treatment, in spite of the ongoing call for

<sup>1</sup> When the number of measurement occasions is relatively small and all individuals are measured at the same set of occasions, it is reasonable to allow the covariance matrix to be arbitrary, with all of its elements unconstrained. With this approach, where no explicit structure is assumed for the covariance among the repeated measures (other than the homogeneity of covariance across different individuals), the resulting covariance is referred to as an *unstructured* covariance. The chief advantage of an unstructured matrix is that no assumptions are made about the variances and covariances (Fitzmaurice, Laird, & Ware, 2004).

Table 1  
*Clinician-Administered PTSD Scale (CAPS) and Posttraumatic Stress Disorder Checklist (PCL)*  
*Scores at Each Time Point for Each Trial*

PTSD	Time 0 <i>M (SD)</i>	Time 1 <i>M (SD)</i>	Time 2 <i>M (SD)</i>	Time 3 <i>M (SD)</i>	Time 4 <i>M (SD)</i>
CSP 420					
Total					
CAPS	81.33 (18.38)	75.15 (22.42)	74.21 (24.31)	75.17 (24.31)	73.21 (25.93)
PCL	62.29 (11.65)	60.41 (13.42)	59.55 (13.85)	59.53 (13.79)	57.21 (15.08)
Reexperiencing					
CAPS	22.18 (7.10)	21.08 (8.17)	20.37 (9.06)	20.48 (8.51)	19.83 (9.67)
PCL	17.50 (4.50)	17.10 (4.63)	16.75 (4.83)	16.77 (4.67)	16.17 (5.01)
Avoidance/Numbing					
CAPS	33.31 (9.16)	29.35 (11.15)	29.46 (11.76)	30.12 (12.34)	28.98 (12.45)
PCL	17.50 (4.20)	17.10 (4.63)	16.75 (4.83)	16.77 (4.67)	16.17 (5.01)
Hypervigilance					
CAPS	25.83 (6.13)	24.73 (6.61)	24.56 (7.11)	24.57 (7.23)	24.40 (7.85)
PCL	19.21 (3.81)	18.63 (4.28)	18.49 (4.27)	18.48 (4.34)	17.68 (4.83)
CPT					
Total					
CAPS	77.92 (16.69)	71.58 (21.71)	64.88 (25.05)	67.08 (25.42)	
PCL	61.25 (9.23)	54.23 (12.30)	50.87 (13.01)	51.96 (13.15)	
Reexperiencing					
CAPS	21.83 (7.58)	19.23 (8.14)	17.26 (9.40)	17.20 (9.56)	
PCL	16.38 (4.88)	11.57 (7.25)	10.53 (6.57)	10.80 (7.21)	
Avoidance/Numbing					
CAPS	30.70 (8.23)	28.23 (12.22)	25.86 (13.20)	26.81 (11.94)	
PCL	24.67 (6.66)	17.10 (10.45)	16.22 (10.35)	16.28 (10.50)	
Hypervigilance					
CAPS	25.37 (5.86)	24.12 (5.93)	21.77 (8.09)	23.06 (7.39)	
PCL	18.77 (4.71)	13.45 (7.99)	12.95 (7.79)	12.83 (7.96)	

*Note.* Means and standard deviations are for all available participants, collapsed across study condition, at each time point. Time 0 represents pretreatment for both trials. The remaining assessments are 7-, 12-, 18-, and 24-months postrandomization for Cooperative Study 420 (CSP 420) and midtreatment or after 3 weeks of waiting, postrandomization or after 6 weeks of waiting, and 1-month follow-up or 10 weeks of waiting for the cognitive processing therapy (CPT) trial conditions, respectively. PTSD = posttraumatic stress disorder.

multimodal assessment (Anthony & Barlow, 2002; Keane, Weathers, & Foa, 2000; Kulka, Schlenger, Fairbank, & Jordan, 1991) and the more general growing emphasis placed on patient-centered outcomes (Freund et al., 1999). We used data from two randomized controlled trials of PTSD treatment with veterans and longitudinal statistical methods that advance prior efforts to evaluate this issue of convergence between clinician and patient outcome ratings.

Overall, the analyses revealed significant longitudinal associations between clinician and patient ratings of PTSD symptoms over the course of treatment, and also over the course of time in patients who are waiting for treatment, when evaluated on a continuous basis. Across the two trials, approximately 60% of the longitudinal variation in total CAPS scores was captured by total PCL scores. Moreover, there was relatively little variability in the magnitude of the associations between clinician and patient ratings across the clusters of PTSD symptoms in the two trials as a whole. Contrary to our hypothesis that the CAPS would reveal relatively larger improvements in symptomatology compared with the PCL, the standardized coefficients that represent change in the CAPS relative to the PCL never reached even a one-to-one relationship with changes on the PCL. Changes on the CAPS were typically

Table 2  
*Prediction of Clinician-Administered PTSD Scale by*  
*Posttraumatic Stress Disorder Checklist (PCL) in Linear Mixed*  
*Effects Model Analyses Overall*

PCL	<i>B</i>	<i>SE</i>	$\beta$	<i>SE</i>	<i>F</i> <sup>a</sup>	<i>dfs</i>	<i>R</i> <sup>2</sup>
CSP 420							
Total	1.28	0.05	0.75	0.03	851.33	1, 320	.57
Reexperiencing	1.17	0.07	0.65	0.04	473.64	1, 316	.45
Avoidance/Numbing	1.26	0.06	0.67	0.03	586.62	1, 321	.52
Hyperarousal	1.08	0.06	0.67	0.04	426.57	1, 314	.49
CPT							
Total	1.50	0.11	0.82	0.07	163.91	1, 46	.65
Reexperiencing	0.96	0.18	0.75	0.14	44.95	1, 52	.25
Avoidance/Numbing	0.96	0.16	0.84	0.14	36.78	1, 52	.35
Hyperarousal	0.66	0.16	0.72	0.17	33.09	1, 52	.21

*Note.* PTSD = posttraumatic stress disorder; CSP 420 = Cooperative Study 420; CPT = cognitive processing therapy.

<sup>a</sup> For all *F*s, *p* < .001.

Table 3

*Longitudinal Associations Between PCL and CAPS Scores in the Trauma-Focused Versus Nontrauma-Focused Conditions of Each Trial*

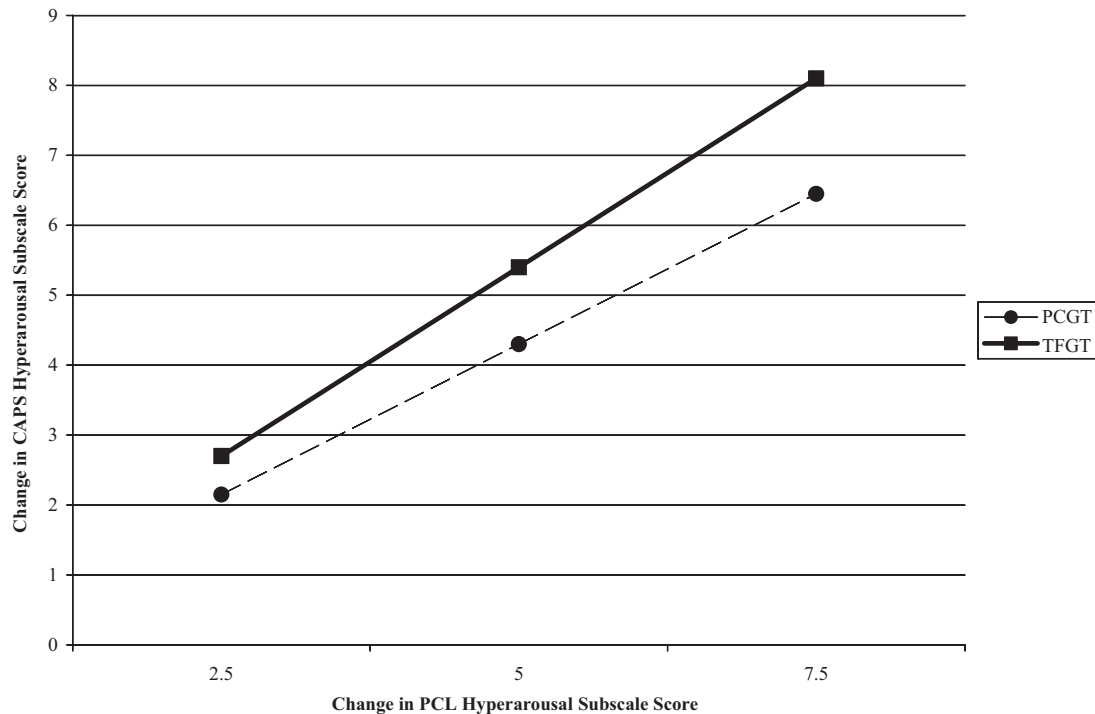
PCL	<i>B</i>	<i>SE</i>	$\beta$	<i>SE</i>	<i>F</i>	<i>dfs</i>
CSP 420						
Total	0.17	0.07	0.10	0.04	5.41*	1, 607
Reexperiencing	0.13	0.09	0.07	0.05	2.00	1, 611
Avoidance/Numbing	0.12	0.09	0.06	0.05	1.84	1, 604
Hyperarousal	0.22	0.08	0.14	0.05	7.20**	1, 611
CPT						
Total	0.27	0.19	0.07	0.11	2.06	1, 71
Reexperiencing	0.18	0.25	0.14	0.20	0.50	1, 98
Avoidance-numbing	0.44	0.23	0.39	0.20	3.55	1, 98
Hyperarousal	-0.01	0.22	-0.01	0.25	0.00	1, 98

*Note.* Significant betas indicate a difference in the strength of the longitudinal association between the Clinician-Administered PTSD Scale (CAPS) and Posttraumatic Stress Disorder Checklist (PCL) in the trauma-focused group therapy condition compared with the referent present-centered group therapy condition in Cooperative Study 420 (CSP 420). PTSD = posttraumatic stress disorder; CPT = cognitive processing therapy.

\* $p < .05$ . \*\* $p < .01$ .

two thirds to three quarters of the changes on the PCL. Similarly, the categorical analyses did not reveal a greater likelihood of clinician-rated clinically significant improvements relative to patient-rated clinically significant improvements.

We found a closer association between patient and clinician ratings of total PTSD symptom changes in the TFGT condition compared with the PCGT condition in CSP 420. This differential association seems to be attributable to more discrepancies in ratings of hyperarousal symptoms. One suspected reason for this difference is that the reexperiencing and behavioral avoidance symptoms of PTSD, compared with hyperarousal symptoms and emotional numbing, are specifically tied to traumatic experiences (e.g., recurrent thoughts of the traumatic experience; avoidance of people, places, and/or things reminiscent of the trauma). The hyperarousal symptoms of PTSD overlap substantially with other mental health conditions. For example, 3 of the 5 hyperarousal symptoms (i.e., sleep disturbance, irritability, concentration difficulties) are more general symptoms of depression and anxiety, and hypervigilance is found in several anxiety disorders. In trauma-focused cognitive-behavioral therapy for PTSD, the rationale for treatment generally includes orienting patients to hyperarousal symptoms as the consequence of reexperiencing symptoms. The treatment also frequently involves exposure exercises to directly increase arousal and facilitate extinction. Given this symptom orientation and direct targeting of hyperarousal in this type of treatment, clinician and patient ratings might be expected to be more similar in these treatments.



*Figure 1.* Longitudinal association between changes on the Clinician-Administered PTSD Scale (CAPS) and Posttraumatic Stress Disorder Checklist (PCL) Hyperarousal subscale scores in Cooperative Study 420. The association is stronger in the trauma-focused group therapy (TFGT) condition compared with the present-centered group therapy (PCGT) condition ( $p < .01$ ). PTSD = posttraumatic stress disorder.

Table 4

*Concordance of Clinically Significant Change on the Clinician-Administered PTSD Scale (CAPS) and Posttraumatic Stress Disorder Checklist (PCL)*

CAPS	PCL					
	Improvement		No change		Exacerbation	
	%	n	%	n	%	n
CSP 420						
Improvement	66 <sub>a</sub>	218	33	111	1	2
No change	13	184	78 <sub>b</sub>	1,143	8	124
Exacerbation	3	4	45	71	52 <sub>a</sub>	82
CPT						
Improvement	79 <sub>a</sub>	45	21	12	0	0
No change	17	28	74 <sub>b</sub>	122	9	15
Exacerbation	7	1	40	6	53 <sub>a</sub>	8

*Note.* Percentages are calculated within CAPS categories of change. In the comparisons of agreement in clinician and patient-rated categories of change, percentages with different subscripts are significantly different from one another within each trial at  $p < .05$ . PTSD = posttraumatic stress disorder;  $n$  = number of assessment interval comparisons in which the clinically significant criteria were met; CSP 420 = Cooperative Study 420; CPT = cognitive processing therapy.

The findings from this study have important theoretical, methodological, and clinical implications. One reason cited for preferring clinician interview over patient report in general is that psychopathology itself may affect patients' ability to reflect on their symptomatology or changes in their symptomatology (e.g., Enns, Larsen, & Cox, 2000). In this vein, the severity and chronicity of PTSD, as represented in the veterans included in the current samples, might be argued to influence the ability to recognize changes in symptomatology. In addition, veterans have been described as potentially less able or willing to note changes in their PTSD symptoms due to secondary gain motivations (e.g., disability entitlements, patient role; Frueh et al., 2003). Our findings argue against these notions and, in fact, indicate that the veterans reported more improvements in PTSD relative to the clinicians. It is important to note that this study was not designed to determine whether the clinicians' or the patients' ratings were "correct" relative to other sources of data. The current findings, however, support the recommended use of brief self-report measures to document the effects of PTSD interventions across treatment in clinical practice, where the administration of independent clinician interviews on multiple occasions is prohibitive (Resick, Monson, & Rizvi, 2007).

The relatively larger changes found on the PCL compared with the CAPS in this study are mostly inconsistent with the previously reviewed treatment studies that found larger treatment effects on the CAPS versus the PCL. Our findings, along with Schnurr et al.'s (2007), showing relatively more change on the PCL than on the CAPS are likely a product of differences in study assessment methods. For example, the differences may be attributable to the manner in which the clinician assessors conducted and scored the CAPS. In addition, there may have been factors influencing the veterans' self-ratings in the trials involved in this study and

Schnurr et al.'s trial that led to greater reported symptom reductions (e.g., opportunity to participate in novel psychotherapy research, effective informed consent procedures that indicated research data would not be included in a participant's medical record) compared with other trials.

There are several limitations of the current study to be considered for future investigations. First, the study samples consisted of veterans receiving psychotherapy within Department of Veterans Affairs Medical Centers in the context of treatment studies. Thus, the current results may or may not represent the convergence between patient- and clinician-rated symptoms in other traumatized populations, treatments, and settings. Moreover, the two studies analyzed here were conducted by overlapping investigators who employed similar training in the CAPS for clinician assessors. Further analyses of other diverse samples by different investigators and with patients receiving other forms of treatment (e.g., drug studies with double-blind assessment) will help ensure the stability of the current results. Additionally, we do not have other assessment information that is less reliant on patient report, such as behavioral assessment, collateral report, or psychophysiological data, to which we can compare the current results. Future investigations should consider the convergence between additional modes of assessment as well.

This study shows that patients with chronic PTSD can and do self-report changes in their symptoms across treatment and time. It also underscores that although there is significant overlap in clinician- and patient-rated outcomes, there is also important independent information that each type of measurement contributes to evaluating if, and how much, symptoms change. Clinicians and researchers are urged to use multimodal methods in assessing outcomes in order to appreciate more fully the effectiveness of PTSD treatment.

## References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Anthony, M. A., & Barlow, D. H. (2002). *Handbook of assessment and treatment planning for psychological disorders*. New York: Guilford Press.
- Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Klauminzer, G., Charney, D. S., et al. (1990). A clinical rating scale for assessing current and lifetime PTSD: The CAPS-1. *Behavior Therapist*, 18, 187-188.
- Blanchard, E. B., Hickling, E. J., Devineni, T., Veazey, C. H., Galovski, T. E., Mundy, E. A., et al. (2003). A controlled evaluation of cognitive behavioral therapy for posttraumatic stress in motor vehicle accident survivors. *Behaviour Research and Therapy*, 41, 79-96.
- Blanchard, E. B., Hickling, E. J., & Forneris, C. A. (1997). Prediction of remission of acute posttraumatic stress disorder in motor vehicle accident victims. *Journal of Traumatic Stress*, 10, 215-234.
- 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.
- Bradley, R., Greene, J., Russ, E., Dutra, L., & Westen, D. (2005). A multidimensional meta-analysis of psychotherapy for PTSD. *American Journal of Psychiatry*, 162, 214-227.
- 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.

- Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis*. Hoboken, NJ: Wiley-Interscience.
- 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.
- Freund, D., Lave, J., Clancy, C., Hawker, G., Hasselblad, V., Keller, R., et al. (1999). Patient outcomes research teams: Contribution to outcomes and effectiveness research. *Annual Review of Public Health*, 20, 337–359.
- Frueh, B. C., Elhai, J. D., Gold, P. B., Monnier, J., Magruder, K. M., Keane, T. M., et al. (2003). Disability compensation seeking among veterans evaluated for posttraumatic stress disorder. *Psychiatric Services*, 54, 84–91.
- Hardin, J. M., & Hilbe, J. W. (2007). *Generalized linear models and extensions* (2nd ed.). College Station, TX: Stata Press.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19.
- Keane, T. M., Weathers, F. W., & Foa, E. B. (2000). Diagnosis and assessment. In E. B. Foa, T. M. Keane, & M. J. Friedman (Eds.), *Effective treatments for PTSD: Practice guidelines from the International Society for Traumatic Stress Studies* (pp. 18–36). New York: Guilford Press.
- Kulka, R. A., Schlenger, W. E., Fairbank, J. A., & Jordan, B. K. (1991). Assessment of posttraumatic stress disorder in the community: Prospects and pitfalls from recent studies of Vietnam veterans. *Psychological Assessment*, 3, 547–560.
- McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12, 153–157.
- Monson, C. M., Schnurr, P. P., Resick, P. A., Friedman, M. J., Young-Xu, Y., & Stevens, S. P. (2006). Cognitive processing therapy for veterans with military-related posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, 74, 898–907.
- Monson, C. M., Schnurr, P. P., Stevens, S. P., & Guthrie, K. A. (2004). Cognitive-behavioral couple's treatment for posttraumatic stress disorder: Initial findings. *Journal of Traumatic Stress*, 17, 341–344.
- Resick, P. A., Monson, C. M., & Chard, K. M. (2007). *Cognitive processing therapy: Veteran/military version*. Washington, DC: Department of Veterans Affairs.
- Resick, P. A., Monson, C. M., & Rizvi, S. L. (2007). Posttraumatic stress disorder. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders* (4th ed., pp. 65–122). New York: Guilford Press.
- Schnurr, P. P., Friedman, M. J., Engel, C. C., Foa, E. B., Shea, T., Chow, B. K., et al. (2007). Cognitive-behavioral therapy for posttraumatic stress disorder in women: A randomized controlled trial. *Journal of the American Medical Association*, 297, 820–830.
- Schnurr, P. P., Friedman, M. J., Foy, D. W., Shea, M. T., Hsieh, F. Y., Lavori, P. W., et al. (2003). Randomized trial of trauma-focused group therapy for posttraumatic stress disorder: Results from a Department of Veterans Affairs cooperative study. *Archives of General Psychiatry*, 60, 481–489.
- Van Etten, M. L., & Taylor, S. (1998). Comparative efficacy of treatment for posttraumatic stress disorder: A meta-analysis. *Clinical Psychology and Psychotherapy*, 5, 126–144.
- Weathers, F. W., Keane, T. M., & Davidson, J. R. (2001). Clinician-Administered PTSD Scale: A review of the first ten years of research. *Depression and Anxiety*, 13, 132–156.
- Weathers, F. W., Litz, B. T., Herman, J. A., Huska, J. A., & Keane, T. M. (1993, November). *The PTSD Checklist (PCL): Reliability, validity and diagnostic utility*. Paper presented at the 9th Annual Conference of the International Society for Traumatic Stress Studies, San Antonio, TX.

Received March 29, 2007

Revision received December 14, 2007

Accepted December 18, 2007 ■

### Instructions to Authors

For Instructions to Authors, please consult the January 2008 issue of the volume or visit [www.apa.org/journals/pas](http://www.apa.org/journals/pas) and click on the "Instructions to authors" link in the Journal Info box on the right.