The short-term stability and reliability of daily estimates of posttraumatic stress disorder symptoms

Abstract

Intensive longitudinal study designs examine posttraumatic stress disorder (PTSD) symptoms without guidance on how many days of PTSD assessments are sufficient to capture reliable and stable estimates of intraindividual mean (iM) and variability (intraindividual standard deviations [iSD]). Thus, the current study examined the reliability and short-term stability of daily PTSD symptom endorsement measured with the Primary Care PTSD Screen for DSM-5 (PC-PTSD-5). Seventy participants (M_{age} = 30.44±12.78; 72.9% female) completed the PC-PTSD-5 for 21 consecutive days before and after attending four intervention sessions. To examine reliability, generalizability coefficients assessing systematic consistency across multiple days (Rc) and single-day reliability (R1F) were calculated. To examine short-term stability in each phase, we calculated reference iMs and iSDs from 21 days for both pre- and post-intervention phases and used (1) correlation coefficients (r >.80 as "stable") and mean absolute differences (MAD; MAD < .25 as "stable") to compare these reference estimates with estimated values ranging from 2–21 days per participant; and (2) bias and agreement using Bland-Altman analyses. Results indicated that the PC-PTSD-5 yielded varying estimates of intraindividual variability in the short term (pre-intervention Rcs=.45–.67; post-intervention Rcs=.40-.55) but demonstrated good single-day reliability (pre-intervention R1Fs = .72–.78; post-intervention R1Fs=.77-.82). Seven to eleven days of PC-PTSD-5 assessments could produce iM and iSD estimates comparable to 21 days. Overall, the PC-PTSD-5 is more reliable for capturing between-person differences than within-person fluctuations. Intensive longitudinal studies could use 7-11 days of daily PC-PTSD-5 assessments to capture stable estimates of average and variable PTSD symptoms.

Keywords: Primary Care PTSD Screen for DSM-5; posttraumatic stress disorder; intraindividual variability; short-term stability; reliability

The short-term stability and reliability of daily estimates of posttraumatic stress disorder symptoms

Intensive longitudinal study designs, including ecological momentary assessments (EMA), experience sampling method (ESM) and daily diary methods, have been increasingly applied to measure short-term changes in posttraumatic stress disorder (PTSD) symptoms (e.g., Hall et al., 2021; Hruska et al., 2025). Such designs are very useful for PTSD research, given their potential to minimize recall bias, to capture rapid symptom fluctuations, and to examine momentary and temporal associations between symptoms (Chun, 2016; Hruska et al., 2025). Another key advantage of these designs is their ability to monitor within-person symptom variability across time (Chun, 2016). Specifically, they allow researchers to measure intraindividual means (iM) and variability (IIV), which represent a person's average symptom level and within-person symptom fluctuations across assessments, respectively (Estabrook et al., 2012). Such data provide valuable information about how individuals differ from one another on PTSD symptom changes across time and how PTSD symptoms change within individuals over time, which can potentially inform personalized interventions and clinical decision-making.

Indeed, studies using EMA, ESM, or daily diary methods to examine PTSD symptom changes over short time periods or in response to contextual factors (e.g., Greene et al., 2022; Grinapol et al., 2022) have demonstrated utility in helping to better understand the implications of daily PTSD variability for clinical practice. Monitoring fluctuations in PTSD symptoms may enable clinicians to detect early signs of symptom deterioration, refine intervention targets, and deliver more personalized treatments for trauma survivors. For example, Possemato et al. (2015) found that real-time monitoring of PTSD symptoms and alcohol use in combat veterans can help identify high-risk periods, suggesting opportunities for timely interventions that reduce

avoidance-based coping and strengthen self-efficacy. Preliminary evidence also suggests that frequent assessments of PTSD symptoms itself can contribute to symptom reduction. For instance, Pollmann and colleagues (2024) found that a 2-week EMA protocol assessing trauma-related intrusive memories led to significant reductions in PTSD's intrusion severity among trauma-exposed adults.

Despite growing interest in within-person variability of PTSD symptoms, there remains little to no consensus on the conditions necessary to obtain reliable estimates of PTSD symptom IIV. This measurement property is referred to as within-person reliability, which reflects the extent to which a measure can consistently capture true fluctuations of symptoms within the same individual over time (Bolger & Laurenceau, 2013). In contrast, between-person reliability suggests how consistently a measure distinguishes differences in average symptom levels amongst individuals. Reliably estimating PTSD symptom IIV is particularly important in clinical settings, as it can provide critical insights into symptom dynamics and may serve as key indicator for treatment progress, risk assessments, and treatment adjustments (Badawi et al., 2025; Shalom et al., 2018). Further, reliably estimating PTSD symptom iM can more accurately identify high-risk individuals requiring more intensive care or early intervention (Warner et al., 2013). Despite the plethora of psychometric work in the extant literature verifying the between-person reliability of PTSD measures (Lane et al., 2019), there is limited work on within-person reliability of PTSD measures. Notably, existing study findings may be biased if PTSD measures lack sufficient reliability to accurately reflect true symptom fluctuations. To our knowledge, only one study has examined the within-person reliability of an abbreviated 8-item version of the PTSD Checklist for DSM-5 (PCL-5) that was administered over 7 consecutive days (Schuler et al., 2021). The authors reported moderate within-person reliability (r = .78) and excellent between-person

reliability (r = .99), but this remains one of the few to assess the ability of PTSD measures to capture IIV. Thus, more research is needed to establish reliable methods for measuring PTSD symptom IIV.

Another important, yet unresolved question in intensive longitudinal studies on mental health is how many repeated measurement occasions are needed to reliably obtain the symptom iM and IIV estimates (hereafter referred to as short-term stability). For example, is the estimate of PTSD symptom IIV derived from 3 measurement occasions as reliable as the estimate derived from 14 measurement occasions? To date, there is no clear guidance on how many measurement occasions of repeated PTSD symptom assessments are needed to achieve reliable estimates of PTSD symptom IIV (Biggs et al., 2019; Greene et al., 2018). Greater clarity on the optimal measurement occasions needed to achieve short-term stability for PTSD symptom iM and IIV estimates would inform more efficient intensive longitudinal study designs that minimize participant burden (Hasselhorn et al., 2022) and reduce logistic costs for researchers and other stakeholders (Pullenayegum et al., 2021).

To address these critical research gaps, the current study evaluated both the between- and within-person reliability of 21-days of daily administration of the Primary Care PTSD Screen for DSM-5 (PC-PTSD-5; Prins et al., 2016) and the short-term stability of PTSD symptom iM and IIV estimates. Although the PC-PTSD-5 has been demonstrated as a reliable tool for assessing between-person differences in PTSD symptoms (e.g., Bovin et al., 2021), its ability to reliably capture within-person variability in PTSD symptoms remains unknown. Our central hypothesis was that as the number of repeated measurement occasions of the PC-PTSD-5 increased, the reliability and short-term stability of PTSD symptom estimates would also increase. Specifically, we hypothesized that as the number of consecutive recorded days increased, estimates of

between- and within-person reliability of PTSD symptoms derived from the PC-PTSD-5 would increase (**Hypothesis 1**). Regarding the short-term stability, we hypothesized that, as the number of consecutive recorded days increased, (1) correlations between the estimated PTSD symptom iM and IIV and reference PTSD symptom iM and IIV estimates derived from 21 days of PC-PTSD-5 assessments (hereafter referred to as "reference estimates") would increase (**Hypothesis 2**); (2) the mean absolute differences between the estimated and reference PTSD symptom iM and IIV estimates would decrease (**Hypothesis 3**); and (3) the agreement between estimated and reference PTSD symptom iM and iSD estimates would increase and bias would decrease (**Hypothesis 4**). Lastly, drawing on the findings from the hypotheses above, we aimed to examine the number of days required to obtain stable estimates of PTSD symptom iM and IIV that approximate the reference estimates. A reference estimate is calculated using all available data (i.e., 21 days within each phase), representing the most accurate estimate of an individual's average symptom or variability of symptoms across days, and bias refers to the average difference between the estimated and reference estimates.

Method

Procedure

The current study was approved by [redacted] Institutional Review Board. Details were described in the protocol paper by [redacted]. The current study was a secondary analysis of data from a pilot intervention trial with multiple phases. During the Screening Phase, interested participants provided electronic informed consent and were screened for eligibility. Eligible and consenting participants completed a baseline survey and 21 daily surveys in the pre-intervention phase. During the intervention phase, participants received four Processing of Positive Memories Technique (PPMT) sessions (Contractor et al., 2021) and measures to assess their psychological

symptoms on a weekly basis. PPMT is conceptualized as a potentially novel intervention for PTSD, wherein treatment recipients are guided to recall and process aspects of salient positive autobiographical memories, which is purported to help them improve mood, cognitions, and posttrauma symptoms (Contractor et al., 2021). In the post-intervention phase, participants received an outcome survey (similar to the baseline survey) and another 21 daily surveys. Compensation was provided based on the total surveys completed and the number of PPMT sessions attended with a maximum of $150 per participant. The current study analyzed data from the pre- and post-intervention daily surveys separately to eliminate intervention effects from the analyses.

Participants

Participants were deemed eligible if they a) were 18-65 years old; b) had access to internet; c) were fluent in English; d) experienced at least one traumatic event assessed by the Life Event Checklist-5 (LEC-5; Weathers, Blake, et al., 2013) and posttraumatic symptoms indicated by a total score of ≥3 on the Primary Care PTSD Screen for DSM-5 (PC-PTSD-5; Prins et al., 2016); e) reported no suicidal/homicidal plans and attempts in the past three months; f) were not currently in therapy; g) resided in the United States; h) expressed willingness and availability to participate for ~10 weeks of the study; and i) agreed to be video-recorded during PPMT sessions. From an initial 1372 participants who attempted the screening survey, the final analytic sample for the current study included 70 participants (M_{age} = 30.44±12.78; 72.9% woman; 52.9% probable PTSD as determined by the PTSD Checklist for DSM-5). See [redacted] for information on sample size truncation. Demographic information of the current sample is provided in Supplemental Table 1.

Measures Relevant to the Current Study

The Life Event Checklist –5 (LEC-5; Weathers, Blake, et al., 2013) is a 17-item measure that was administered to screen participants' lifetime traumatic experiences. The first 16 items referenced specific traumas, and the last item enabled participants to describe an event not covered in the checklist. The 6-point response scale represented different levels of exposure: directly experienced, witnessed, learned about the incident, being exposed to aversive details as part of one's occupation, not sure, or does not apply. For the current study's inclusion criteria, endorsement of any of the first four responses on the first 16 items of the LEC-5 was considered indicative of trauma exposure.

The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5; Weathers, Litz, et al., 2013) is a 20-item measure to assess PTSD symptom severity. In the baseline survey, participants rated symptoms in reference to their worst traumatic event on the LEC-5 over the past month using a 5-point Likert scale (0 = not at all, 1 = a little bit, 2 = moderately, 3 = quite a bit, 4 = extremely). Research showed that the PCL-5 demonstrates good psychometric properties across studies (Forkus et al., 2023). In the current study, total PCL-5 scores equal to or greater than 33 indicate probable diagnostic PTSD (Forkus et al., 2023). Cronbach's alpha for the PCL-5 total score was .93 at baseline.

The Primary Care PTSD Screen for DSM-5 (PC-PTSD-5; Prins et al., 2016) was used to assess daily PTSD symptoms during the pre-intervention and post-intervention phases. The five items correspond to PTSD criteria based on the DSM-5: one item assessed intrusions, one item assessed avoidance of trauma reminders, two items assessed negative alterations in cognition and mood, and one item assessed alterations in arousal and reactivity. Reponses were coded dichotomously (i.e., yes = 1 or no = 0). A cutoff score of 3 or higher indicates positive screening for PTSD that requires further assessment (Prins et al., 2016). Literature showed that the PC-

PTSD-5 scores demonstrate good between-person reliability and validity in primary care samples (Williamson et al., 2022). In the current study, the total PC-PTSD-5 score for each participant on each day was calculated. The reference iM and iSD were calculated using the mean and standard deviation of the total PC-PTSD-5 scores across all available days (21 days pre- or post-intervention separately) for each participant, respectively. Estimated iM and iSD for each day were calculated as the mean and standard deviation of the PC-PTSD-5 total scores from day 1 to the given day (21 days pre- or post-intervention separately).

Data Analyses

We calculated test-retest reliability between each specified day and day 21 for each phase separately and used the Generalizability Theory (or G Theory) to evaluate both between- and within-person reliability of the PC-PTSD-5 total score (Bolger & Laurenceau, 2013; **Hypothesis 1**). Hereby, we conducted Analysis of Variance (ANOVA) with a random effects model to understand the sources of variability between three dimensions: person, time (i.e., consecutive recorded days), and item. Then we calculated two Generalizability coefficients from variance values extracted from the ANOVA model derived from 2 to 21 days (pre- or post-intervention phase): 1) we calculated R1F (**equation 1)** to examine how consistently individuals could be differentiated (i.e., between-person reliability) based on their PTSD symptoms reported on one fixed day, which can be interpreted as the approximate average of Cronbach's alpha values calculated separately for each day; and 2) we calculated Rc (**equation 2**) to assess the consistency of individuals' reported PTSD symptoms (i.e., within-person reliability) across occasions. This metric indicates the degree to which these repeated measures are adequate (i.e., able to capture meaningful symptom fluctuations) and systematic (i.e., able to reflect consistent

patterns of fluctuations rather than random fluctuations). Both Rc and R1F were interpreted using the same heuristics as Cronbach's alpha (Bolger & Laurenceau, 2013)

$$R1F = \frac{\sigma^2_{person}+[\sigma^2_{person:item}/k]}{\sigma^2_{person}+[\sigma^2_{person:item}/k]+[\sigma^2_{error}/k]} \quad (1)$$

$$Rc = \frac{\sigma^2_{person:time}}{(\sigma^2_{person:time}+\left(\sigma^2_{person:time:item}+\sigma^2_{error}\right)/k)} \quad (2)$$

Where σ^2_{person}, $\sigma^2_{person:item}$, σ^2_{error}, $\sigma^2_{person:time}$, and $\sigma^2_{person:time:item}$ represent variance terms extracted from the ANOVA with random effects model and *k* represents the number of items.

To investigate the short-term stability of daily administration of the PC-PTSD-5, we first (**Hypothesis 2**) calculated reliability correlation coefficients (i.e., reliability index) where the total day iM and iSD estimates for PTSD symptom severity (21 days pre- or post-intervention) were used as the reference estimates – a proxy for the "true scores" – and estimates of iM and iSD for PTSD symptom severity derived from 2 to 21 days (pre- or post- intervention) in 1-unit intervals were used as estimates of the "observed scores." Calculated reliability correlation coefficients were summarized, and the trend was plotted across number of consecutive recorded days. For **Hypothesis 3**, we followed the same approach except that we calculated the mean absolute differences (MAD) instead of reliability correlation coefficients. Finally, we conducted Bland-Altman analyses to address **Hypothesis 4**. Bland-Altman plots were used to probe a more detailed evaluation of differences between the reference estimate and observed value(s) for numbers of consecutive recorded days that are of particular interest for daily diary study designs (e.g., 7-, 14- days; Chun, 2016; Lane et al., 2019). Limits of agreement (LOA) and average bias

were computed from the Bland-Altman plots to assess level of agreements between the reference estimate and observed value(s).

Results

During the pre-intervention phase, participants completed 1,327 daily surveys in total, with an average of 19.0 surveys per participant (range: 14–21). On average, participants ($N = 70$) positively endorsed 1.53 items per day ($SD = 1.62$), with 26.9% of daily responses indicating a positive screen of probable PTSD (i.e., >= 3 items endorsed as a *yes* on the PC-PTSD-5). During the post-intervention phase, 1,253 daily surveys were completed, averaging 17.9 surveys per participant (range: 14–21). On average, participants positively endorsed 1.13 items per day ($SD = 1.48$), with 18.6% of daily responses suggesting a positive screen for probable PTSD. Notably, there were no missing item-level data. Further, Little's MCAR tests indicated that the PC-PTSD-5 data (complete measures) were missing completely at random for both the pre-intervention phase (χ^2 [595] = 420.14, $p = 1.00$) and the post-intervention phase (χ^2 [1255] = 762.93, $p = 1.00$). Missing data were addressed using pairwise deletion, which allowed us to use all available data without relying on within-day information to guide imputation of entire PC-PTSD-5 surveys – an approach that could have introduced inaccuracy and/or artificial within-person variability.

The intraclass correlation coefficients for PC-PTSD-5 were .62 and .66 for the pre- and post-intervention phases, respectively. These results indicated that 62% and 66% of the variance in the PC-PTSD-5 scores is attributed to between-person differences during the pre- and post-intervention phases, respectively. The remaining 38% and 34% of the variance reflect within-person variability during the pre- and post-intervention phases, respectively.

For the pre-intervention phase, the reference estimate of iM was 1.53, and the reference estimate of iSD was 1.47. The estimated iM ranged from 1.19 to 1.86 across 21 days, and the

estimated iSD ranged from 1.42 to 1.73. For the post-intervention phase, the reference estimate of iM was 1.53, and the reference estimate of iSD was 1.48. The estimated iM ranged from .91 to 1.31 across 21 days, and the estimated iSD ranged from 1.38 to 1.31 (Supplemental Table 2).

Reliability

Results for test-retest reliability are presented in Supplemental Table 3. For the pre-intervention phase, the test-retest reliability between any given day and day 21 ranged from .48 to .86. For the post-intervention phase, the test-retest reliability between any given day and day 21 ranged from .33 to .80. Regarding **Hypothesis 1**, generalizability coefficients broadly suggested that the PC-PTSD-5 had good between-person reliability, but poorer within-person reliability across days for both the pre- and post-intervention phases (Table 1). For the pre-intervention phase, estimates of R1F ranged between .72 to .78, indicating consistently acceptable between-person reliability. Conversely, estimates of Rc showed less consistency and ranged between .45-.67, indicating generally poor within-person reliability. Similarly, for the post-intervention phase, estimates of R1F ranged between .77 to .82, indicating consistently acceptable between-person reliability. Estimates of Rc were consistently low (.40-.55), indicating consistently poor within-person reliability.

Short-term Stability

We evaluated the short-term stability of daily administration of the PC-PTSD-5 in several ways (Supplemental Table 4). Regarding the evaluation of correlation coefficients (**Hypothesis 2**), plotted trends (Figure 1) suggested that 3 days of daily PC-PTSD-5 assessments may be needed to provide a convergent estimate of iM ($r > .80$) with the reference estimate during the pre-intervention phase, with correlation coefficients continuing to rise before plateauing at approximately 9 days ($r = .96$). It took 7 days to achieve a convergent estimate of iSD with the

reference estimate ($r > .80$) during the pre-intervention phase, and correlation coefficients continued to rise before plateauing at approximately 9 days ($r = .91$). For the post-intervention phase, it took 2 days to achieve a convergent estimate of iM with the reference estimate ($r > .80$), but correlation coefficients continued to rise and started plateauing at approximately 12 days ($r = .99$). It took 6 days to achieve a convergent estimate of iSD with the reference estimate ($r > .80$), but correlation coefficients continued to rise and started plateauing at approximately 12 days ($r = .95$).

Regarding the evaluation of mean absolute differences (**Hypothesis 3**), plotted trends (Figure 2) suggested that 10 days of daily PC-PTSD-5 assessments may be needed to estimate iMs that are minimally different (MAD < .25) from the reference estimates, while 7 days of daily assessments may be needed to estimate iSDs that are minimally different from the reference estimate (MAD < .25). Specifically, for the pre-intervention phase, it took 10 days to achieve a minimally different estimate of iM from the reference estimate, and the mean absolute differences continued to decrease before plateauing at approximately 19 days (MAD = .01). It took 7 days to achieve a minimally different estimate of iSD with the reference estimate, and similarly, the mean absolute differences continued to decrease before plateauing at approximately 9 days (MAD = .15). For the post-intervention phase, it took 8 days to achieve a minimally different estimate of iM with the reference estimate, and the mean absolute differences continued to decrease before plateauing at approximately 18 days (MAD = .02). It took 6 days to achieve a minimally different estimate of iSD with the reference estimate, with the mean absolute differences continuing to decrease before plateauing at approximately 10 days (MAD = .13).

Regarding the Bland-Altman analyses (**Hypothesis 4**), results broadly suggested as the number of consecutive days used to estimate iMs and iSDs increased, agreement between the reference estimate and estimated values increased, and bias decreased for both pre- and post-intervention phases. Specifically, for the pre-intervention phase (Figure 3), the mean difference for iMs between the reference and 7 days was .35 (*SD* = .32), with 95% limits of agreement ranging from -.76 to .67 and the average bias was -.04. Between the reference and 14 days, the mean difference for iMs was .14 (*SD* = .15), with 95% limits of agreement ranging from -.31 to .25 and the average bias was -.03. The mean difference for iSDs between the reference and 7 days was .22 (*SD* = .18), with 95% limits of agreement ranging from -.58 to .39 and the average bias was -.10. The mean difference for iSDs between the reference and 14 days was .09 (*SD* = .10), with 95% limits of agreement ranging from -.30 to .21 and the average bias was -.05. For the post-intervention phase (Figure 4), the mean difference for iMs between the reference and 7 days was .24 (*SD* = .27), with 95% limits of agreement ranging from -.76 to .67 and the average bias was -.04. Between the reference and 14 days, the mean difference for iMs was .09 (*SD* = .11), with 95% limits of agreement ranging from -.31 to .25 and the average bias was -.03. The mean difference for iSDs between the reference and 7 days was .18 (*SD* = .21), with 95% limits of agreement ranging from -.56 to .57 and the average bias was minimal (i.e., .002). The mean difference for iSDs between the reference and 14 days was .07 (*SD* = .12), with 95% limits of agreement ranging from -.26 to .26 and the average bias was minimal (i.e., .0002).

Supplemental Analyses

We conducted supplemental analyses with two subsamples to examine whether the directness of trauma exposure or PTSD symptom severity influenced the observed level of stability in PTSD symptom estimates.

First, we analyzed data from 45 participants who directly experienced or witnessed their index traumas. For the pre-intervention phase, estimates of R1F ranged from .72 to .79 and estimates of Rc ranged from .49 to .60 (Supplemental Table 5). For the post-intervention phase, estimates of R1F ranged from .81 to .85 and estimates of Rc ranged from .36 to .60. Correlation trends (Supplemental Figure 1) suggested that 3 and 8 days of daily PC-PTSD-5 assessments may be needed to provide convergent estimates of iM and iSD ($r > .80$), respectively, with reference estimates for the pre-intervention phase. Eight and six days of daily PC-PTSD-5 assessments may be needed to provide convergent estimates of iM and iSD ($r > .80$), respectively, with reference estimates. Mean absolute difference trends (Supplemental Figure 2) suggested that 8 days of daily PC-PTSD-5 assessments may be needed to obtain minimally different estimates of iM and iSD from the reference estimates during the pre-intervention phase (MAD < .25). For the post-intervention phase, eight and six days of daily PC-PTSD-5 assessments may be needed to provide minimally different estimates of iM and iSD, respectively, from the reference estimates (MAD < .25).

Second, we analyzed data from 37 participants who endorsed probable PTSD (as measured by the PCL-5) at baseline. For the pre-intervention phase, estimates of R1F ranged from .66 to .77 and estimates of Rc ranged from .34 to .54 (Supplemental Table 6). For the post-intervention phase, estimates of R1F ranged from .80 to .86 and estimates of Rc ranged from .35 to .56. Correlation trends (Supplemental Figure 3) suggested that 4 and 7 days of daily PC-PTSD-5 assessments may be needed to provide convergent estimates of iM and iSD ($r > .80$), respectively, with reference estimates for the pre-intervention phase. Two and four days of daily PC-PTSD-5 assessments may be needed to provide convergent estimates of iM and iSD ($r > .80$), respectively, with reference estimates for the post-intervention phase. Mean absolute

difference trends (Supplemental Figure 4) suggested that 11 and 6 days of daily PC-PTSD-5 assessments may be needed to obtain minimally different estimates of iM and iSD, respectively, from the reference estimates during the pre-intervention phase (MAD < .25). For the post-intervention phase, nine and six days of daily PC-PTSD-5 assessments may be needed to obtain minimally different estimates of iM and iSD, respectively, from the reference estimates (MAD < .25).

Discussion

The current study examined the reliability and short-term stability of daily estimates of PTSD symptoms derived from the PC-PTSD-5 within a daily diary design. Overall, results indicated that the PC-PTSD-5 scores demonstrated good between-person reliability but poor within-person reliability across days (including in a subsample of participants with direct exposure to a traumatic event or those with probable PTSD). Further, our findings indicated that approximately 7 to 11 days of daily PC-PTSD-5 assessments may be required to capture stable estimates of average and variable PTSD symptoms in the short term (i.e., comparable to 21 days). We detail the implications of these findings in the context of existing literature.

Consistent with **Hypothesis 1**, findings indicated that both between- and within-person reliability of PC-PTSD-5 improved as the number of days increased. This is unsurprising given that reliability estimates increase as observations increase (Kennedy, 2022). However, results also revealed differences across between- and within-person reliability estimates of the PC-PTSD-5. While results indicated the PC-PTSD-5 demonstrated acceptable between-person reliability in both the pre- and post-intervention phases, the PC-PTSD-5 demonstrated poor within-person reliability. Several factors may have contributed to the poor within-person reliability observed in the current study. First, PTSD symptoms are inherently variable and

susceptible to daily stressors, contextual factors, and individual coping strategies (Biggs et al., 2019; Short et al., 2018). Further, prior research has indicated more within-person fluctuations in PTSD symptoms among individuals with more PTSD severity (Schuler et al., 2021). Given that approximately half of the current sample screened positive for probable PTSD, it is possible that within-person reliability would be even lower in clinical populations with diagnostic PTSD, where symptom fluctuations are often more pronounced. Second, the PC-PTSD-5 was originally designed to screen for probable PTSD, each PC-PTSD-5 item captures more than one PTSD criteria, and it might not be sensitive enough to monitor day-to-day changes in PTSD symptoms (Prins et al., 2016). While the PC-PTSD-5 has demonstrated reliability and validity in cross-sectional studies (e.g., Bovin et al., 2021), more research is needed to support its use in intensive longitudinal designs.

Results also supported **Hypotheses 2-4** in which more consecutive days of daily PC-PTSD-5 measurements were associated with greater short-term stability (i.e., greater convergence and agreement, and decreased differences and bias) of PTSD symptom iM and IIV estimates relative to the reference estimates. These trends reflect the psychometric principle that more observations typically yield more stable estimates (Fisher, 1925). Notably, our results indicated that while it only took three days to differentiate individuals with relatively higher or lower average PTSD symptoms (i.e., observed *rs* > .80), a longer assessment period of up to 10 days could improve the accuracy of estimating individual average PTSD symptoms (i.e., observed MADs < .25). This finding is consistent with the nature of the PC-PTSD-5 as a brief screening tool to identify individuals who might need further assessment for PTSD (Prins et al., 2016). Additionally, while only three days of PC-PTSD-5 assessments were needed to estimate a PTSD symptom iM comparable to the 21-day iM reference estimate, up to seven days of PC-

PTSD-5 assessments were needed to estimate a PTSD symptom IIV comparable to the 21-day IIV reference estimate. Given the dynamic nature of PTSD symptoms (e.g., Greene et al., 2018), a longer observation window may be needed to capture a more meaningful and stable estimate of PTSD symptom IIV. Hence, our study results suggest 7- to 10-days of PC-PTSD-5 assessments may be needed to obtain reasonably stable estimates of both average PTSD symptom levels and PTSD symptom IIV. Notably, for participants who directly experienced or witnessed their index trauma, our findings suggest that approximately 8 days of PC-PTSD-5 assessments may be needed to obtain reliable estimates of both average PTSD symptom levels and symptom IIV, whereas 7 to 11 days may be required for participants with probable PTSD to obtain such reliable estimates. These observations are consistent with prior simulation work that has demonstrated that IIV metrics generally have poorer reliability compared to iM metrics, and more observations are needed to generate stable estimates of IIVs (Estabrook et al., 2012). Taken together, future intensive longitudinal studies and clinicians interested in capturing daily patterns of PTSD symptoms should consider monitoring PTSD symptoms for a period of 7- to 11 days to obtain more reliable and stable information about daily PTSD symptoms derived from daily administrations of the PC-PTSD-5.

Our findings provided preliminary guidance for researchers and clinicians seeking to balance data quality and participant burden. Indeed, our results indicate that between 7- to 11 days of consecutive PC-PTSD-5 assessments produce stability estimates comparable to 21-days of PC-PTSD-5 assessments. As such, these findings can help guide researchers and clinicians select protocol lengths that minimize participant burden while maximizing data quality that may be incorporated in novel mobile technology-based interventions (Heron & Smyth, 2010) or to tailor interventions and monitor treatment response. Nonetheless, studies examining the

reliability of other PTSD measures in estimating PTSD symptom reliability and short-term stability are needed to support our recommendation.

Study Limitations and Future Directions

Although the current study is the first study, to our knowledge, to examine the reliability and short-term stability of the PC-PTSD-5 in an intensive longitudinal study, the current study is not without limitations. Our findings are specific to the PC-PTSD-5, which was originally designed as a screening tool for PTSD. The limited items and dichotomous response format could substantially limit its ability to fully capture within-person symptom variability over time. Furthermore, the once-daily sampling frequency may miss within-daily fluctuations. Future studies should examine whether the discrepancy amongst between- and within-person variability in our study design exists when using lengthier and more comprehensive measures of PTSD such as the PCL-5 (Weathers, Litz, et al., 2013), or when multiple EMA reports are given each day by participants. Second, our study used a community sample, with only 52.9% of participants reporting probable PTSD. This may limit the generalizability of our findings to individuals with more PTSD symptom severity or those with comorbid psychiatric disorders. Third, our sample consisted primarily of individuals who identified as white, non-Hispanic, and women; hence, findings may not apply to more diverse racial, ethnic, and gender-based groups. Research indicates that the severity and presentation of PTSD symptoms can differ across racial and ethnic groups, influenced by factors such as cultural norms and socioeconomic status (e.g., Hall-Clark et al., 2016). Future research would benefit from replicating our findings in more clinically and demographically diverse populations. Lastly, due to the small sample size, our study was underpowered to examine whether types of trauma exposure (e.g., direct or indirect) or baseline PTSD symptom severity would potentially moderate our study findings. Future research with

larger samples should examine these potential effects to better guide intensive longitudinal study designs and inform clinical decision-making.

Conclusion

Overall, our study highlights the importance of collecting multiple days of data to improve reliability and stability of estimates of PTSD symptom patterns derived from the PC-PTSD-5. Regarding reliability, the PC-PTSD-5 demonstrated better between-person reliability than within-person reliability. In terms of short-term stability, 7 to 11 consecutive days of daily PC-PTSD-5 assessments may be sufficient to generate stable estimates of PTSD symptom iM and IIV that are comparable to 21 days of assessment. Notably, PTSD symptom IIV is becoming increasingly recognized as a clinically meaningful metric to describe within-person PTSD symptom change over time. As such, it would be helpful for researchers and clinicians to consider the reliability and stability limitations of the PTSD symptom IIV metrics (e.g., iSD) when designing future intensive longitudinal studies or daily monitoring protocols. Our findings may serve as a preliminary guide for how many daily observations are needed to achieve sufficiently reliable and stable estimates of PTSD symptom IIV when using the PC-PTSD-5. Future studies should seek to identify instruments that are both brief and sensitive to individual PTSD symptom fluctuations, which will be critical for advancing research with intensive longitudinal designs as well as evidence-based practice.

References

Alden, L. E., Regambal, M. J., & Laposa, J. M. (2008). The effects of direct versus witnessed threat on emergency department healthcare workers: implications for PTSD criterion A. *Journal of anxiety disorders*, 22(8), 1337–1346. https://doi.org/10.1016/j.janxdis.2008.01.013

Badawi, A., Steel, Z., Rogers, K., Wijesinghe, N., & Berle, D. (2025). Changes in Trauma-based Intrusive Memory Characteristics Associated with Repetitive Transcranial Magnetic Stimulation (rTMS) for Depression: A Daily Diary Study. *The Psychiatric quarterly*, 10.1007/s11126-025-10156-4. Advance online publication. https://doi.org/10.1007/s11126-025-10156-4

Biggs, Q. M., Ursano, R. J., Wang, J., Krantz, D. S., Carr, R. B., Wynn, G. H., Adams, D. P., Dacuyan, N. M., & Fullerton, C. S. (2019). Daily variation in post traumatic stress symptoms in individuals with and without probable post traumatic stress disorder. *BMC psychiatry*, *19*(1), 56. https://doi.org/10.1186/s12888-019-2041-7

Black, A. C., Cooney, N. L., Sartor, C. E., Arias, A. J., & Rosen, M. I. (2018). Impulsivity interacts with momentary PTSD symptom worsening to predict alcohol use in male veterans. *The American journal of drug and alcohol abuse*, *44*(5), 524–531. https://doi.org/10.1080/00952990.2018.1454935

Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford press.

Bos, F. M., Schoevers, R. A., & aan het Rot, M. (2015). Experience sampling and ecological momentary assessment studies in psychopharmacology: A systematic review. *European neuropsychopharmacology : the journal of the European College of*

Neuropsychopharmacology, *25*(11), 1853–1864. https://doi.org/10.1016/j.euroneuro.2015.08.008

Bovin, M. J., Kimerling, R., Weathers, F. W., Prins, A., Marx, B. P., Post, E. P., & Schnurr, P. P. (2021). Diagnostic Accuracy and Acceptability of the Primary Care Posttraumatic Stress Disorder Screen for the Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition) Among US Veterans. *JAMA network open*, *4*(2), e2036733. https://doi.org/10.1001/jamanetworkopen.2020.36733

Campbell, S. B., Renshaw, K. D., Kashdan, T. B., Curby, T. W., & Carter, S. P. (2017). A Daily Diary Study of Posttraumatic Stress Symptoms and Romantic Partner Accommodation. *Behavior therapy*, *48*(2), 222–234. https://doi.org/10.1016/j.beth.2016.04.006

Chun, C. A. (2016). The expression of posttraumatic stress symptoms in daily life: A review of experience sampling methodology and daily diary studies. *Journal of Psychopathology and Behavioral Assessment, 38*(3), 406–420. https://doi.org/10.1007/s10862-016-9540-3

Contractor, A. A., Weiss, N. H., & Shea, M. T. (2021). Processing of Positive Memories Technique (PPMT) for Posttraumatic Stress Disorder: A Primer. *Journal of psychotherapy integration*, *31*(2), 195–207. https://doi.org/10.1037/int0000239

Decker, S. E., Pavlo, A., Harper, A., Herring, Y., & Black, A. C. (2020). Themes in experiences of PTSD symptoms and relationships among male veterans with risky sexual behavior. *Psychological trauma: theory, research, practice and policy*, *12*(7), 678–686. https://doi.org/10.1037/tra0000569

Dworkin, E. R., Ullman, S. E., Stappenbeck, C., Brill, C. D., & Kaysen, D. (2018). Proximal relationships between social support and PTSD symptom severity: A daily diary study of sexual assault survivors. *Depression and anxiety*, *35*(1), 43–49. https://doi.org/10.1002/da.22679

Estabrook, R., Grimm, K. J., & Bowles, R. P. (2012). A Monte Carlo simulation study of the reliability of intraindividual variability. *Psychology and aging*, *27*(3), 560–576. https://doi.org/10.1037/a0026669

Fisher, R. A. (1925, July). Theory of statistical estimation. In *Mathematical proceedings of the Cambridge philosophical society* (Vol. 22, No. 5, pp. 700-725). Cambridge University Press.

Forkus, S. R., Raudales, A. M., Rafiuddin, H. S., Weiss, N. H., Messman, B. A., & Contractor, A. A. (2023). The Posttraumatic Stress Disorder (PTSD) Checklist for DSM-5: A Systematic Review of Existing Psychometric Evidence. *Clinical psychology: a publication of the Division of Clinical Psychology of the American Psychological Association*, *30*(1), 110–121. https://doi.org/10.1037/cps0000111

Greene, T., Gelkopf, M., Epskamp, S., & Fried, E. (2018). Dynamic networks of PTSD symptoms during conflict. *Psychological medicine*, 48(14), 2409–2417. https://doi.org/10.1017/S0033291718000351

Greene, T., Sznitman, S., Contractor, A. A., Prakash, K., Fried, E. I., & Gelkopf, M. (2022). The memory-experience gap for PTSD symptoms: The correspondence between experience sampling and past month retrospective reports of traumatic stress symptoms. *Psychiatry research*, *307*, 114315. https://doi.org/10.1016/j.psychres.2021.114315

Grinapol, S., Greene, T., Pagorek-Eshel, S., Rodrig, J., Halpern, P., & Gelkopf, M. (2022).

Prospective examination of daily life traumatic stress symptom expression following motor vehicle collisions. *Psychological trauma : theory, research, practice and policy*, *14*(7), 1212–1220. https://doi.org/10.1037/tra0000537

Hall-Clark, B., Sawyer, B., Golik, A., & Asnaani, A. (2016). Racial/ethnic differences in symptoms of posttraumatic stress disorder. *Current Psychiatry Reviews, 12*(2), 124–138. https://doi.org/10.2174/1573400512666160505150257

Hall, M., Scherner, P. V., Kreidel, Y., & Rubel, J. A. (2021). A Systematic Review of Momentary Assessment Designs for Mood and Anxiety Symptoms. *Frontiers in psychology*, *12*, 642044. https://doi.org/10.3389/fpsyg.2021.642044

Hasselhorn, K., Ottenstein, C., & Lischetzke, T. (2022). The effects of assessment intensity on participant burden, compliance, within-person variance, and within-person relationships in ambulatory assessment. Behavior research methods, 54(4), 1541–1558. https://doi.org/10.3758/s13428-021-01683-6

Heitman, R. J., Kovaleski, J. E., & Pugh, S. F. (2009). Application of generalizability theory in estimating the reliability of ankle-complex laxity measurement. *Journal of athletic training*, *44*(1), 48–52. https://doi.org/10.4085/1062-6050-44.1.48

Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *British journal of health psychology*, *15*(Pt 1), 1–39. https://doi.org/10.1348/135910709X466063

Hruska, B., Piccirillo, M. L., Lenferink, L. I. M., Pacella-LaBarbara, M. L., Contractor, A. A., Price, M., & Greene, T. (2025). Making trauma ecological momentary assessment studies FAIR: review of design considerations and data procedures. *European journal of psychotraumatology*, *16*(1), 2477423. https://doi.org/10.1080/20008066.2025.2477423

Kennedy, I. (2022). Sample size determination in test-retest and Cronbach alpha reliability estimates. *British Journal of Contemporary Education*, 2(1), 17-29. https://doi.org/10.52589/BJCE-FY266HK9

Lane, A. R., Waters, A. J., & Black, A. C. (2019). Ecological momentary assessment studies of comorbid PTSD and alcohol use: A narrative review. *Addictive behaviors reports*, 10, 100205. https://doi.org/10.1016/j.abrep.2019.100205

Nahleen, S., Nixon, R. D. V., & Takarangi, M. K. T. (2019). Current PTSD symptomatology distorts memory for past symptoms. *Psychiatry research*, *274*, 330–334. https://doi.org/10.1016/j.psychres.2019.02.052

Naragon-Gainey, K., Simpson, T. L., Moore, S. A., Varra, A. A., & Kaysen, D. L. (2012). The correspondence of daily and retrospective PTSD reports among female victims of sexual assault. *Psychological assessment*, *24*(4), 1041–1047. https://doi.org/10.1037/a0028518

Pollmann, Y., Clancy, K. J., Devignes, Q., Ren, B., Kaufman, M., & Rosso, I. M. (2024). Ecological Momentary Assessments of Trauma-Related Intrusive Memories: Potential Clinical Utility. *medRxiv : the preprint server for health sciences*, 2024.05.15.24307377. https://doi.org/10.1101/2024.05.15.24307377

Possemato, K., Maisto, S. A., Wade, M., Barrie, K., McKenzie, S., Lantinga, L. J., & Ouimette, P. (2015). Ecological momentary assessment of PTSD symptoms and alcohol use in combat veterans. *Psychology of addictive behaviors: journal of the Society of Psychologists in Addictive Behaviors*, 29(4), 894–905. https://doi.org/10.1037/adb0000129

Prins, A., Bovin, M. J., Smolenski, D. J., Marx, B. P., Kimerling, R., Jenkins-Guarnieri, M. A., Kaloupek, D. G., Schnurr, 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

Pullenayegum, E. M., Xi, Y., Lim, L., Levin, J., & Feldman, B. M. (2021). Choosing the frequency of follow-up in longitudinal studies: Is more necessarily better?. *Research Methods in Medicine & Health Sciences*, *2*(2), 61-67. https://doi.org/10.1177/2632084320975260

Ruggero, C. J., Schuler, K., Waszczuk, M. A., Callahan, J. L., Contractor, A. A., Bennett, C. B., Luft, B. J., & Kotov, R. (2021). Posttraumatic stress disorder in daily life among World Trade Center responders: Temporal symptom cascades. *Journal of psychiatric research*, *138*, 240–245. https://doi.org/10.1016/j.jpsychires.2021.04.002

Schuler, K., Ruggero, C. J., Mahaffey, B., Gonzalez, A., L Callahan, J., Boals, A., Waszczuk, M. A., Luft, B. J., & Kotov, R. (2021). When Hindsight Is Not 20/20: Ecological Momentary Assessment of PTSD Symptoms Versus Retrospective Report. *Assessment*, *28*(1), 238–247. https://doi.org/10.1177/1073191119869826

Shalom, J. G., Gilboa-Schechtman, E., Atzil-Slonim, D., Bar-Kalifa, E., Hasson-Ohayon, I., van Oppen, P., van Balkom, A. J. L. M., & Aderka, I. M. (2018). Intraindividual variability in symptoms consistently predicts sudden gains: An examination of three independent datasets. *Journal of consulting and clinical psychology*, *86*(11), 892–902. https://doi.org/10.1037/ccp0000344

Shapiro, D. L., Chen, D., & Duncan, A. (2007). Classical test theory: A critical review of its assumptions and implications for clinical practice. *Psychological Assessment, 19*(2), 134-141. https://doi.org/10.1037/1040-3590.19.2.134

Shavelson, R. J., & Webb, N. M. (1981). Generalizability theory: 1973–1980. *British Journal of Mathematical and Statistical Psychology*, *34*(2), 133-166.

Short, N. A., Boffa, J. W., Clancy, K., & Schmidt, N. B. (2018). Effects of emotion regulation strategy use in response to stressors on PTSD symptoms: An ecological momentary assessment study. *Journal of affective disorders*, *230*, 77–83. https://doi.org/10.1016/j.jad.2017.12.063

Warner, C. H., Warner, C. M., Appenzeller, G. N., & Hoge, C. W. (2013). Identifying and managing posttraumatic stress disorder. *American family physician*, *88*(12), 827–834.

Weathers, F. W., Blake, D. D., Schnurr, P. P., Kaloupek, D. G., Marx, B. P., & Keane, T. M. (2013). The life events checklist for DSM-5 (LEC-5).

Weathers, F.W., Litz, B.T., Keane, T.M., Palmieri, P.A., Marx, B.P., & Schnurr, P.P. (2013). The PTSD Checklist for DSM-5 (PCL-5). Scale available from the National Center for PTSD at www.ptsd.va.gov.

Westermeyer, J., Shiroma, P., Thuras, P., Kattar, K., Johnson, D., & Crosby, R. D. (2015). Daily versus monthly reporting of post-traumatic symptoms: A study of reliability across time and instruments. *Psychiatry research*, *227*(2-3), 309–312. https://doi.org/10.1016/j.psychres.2015.03.017

Table 1

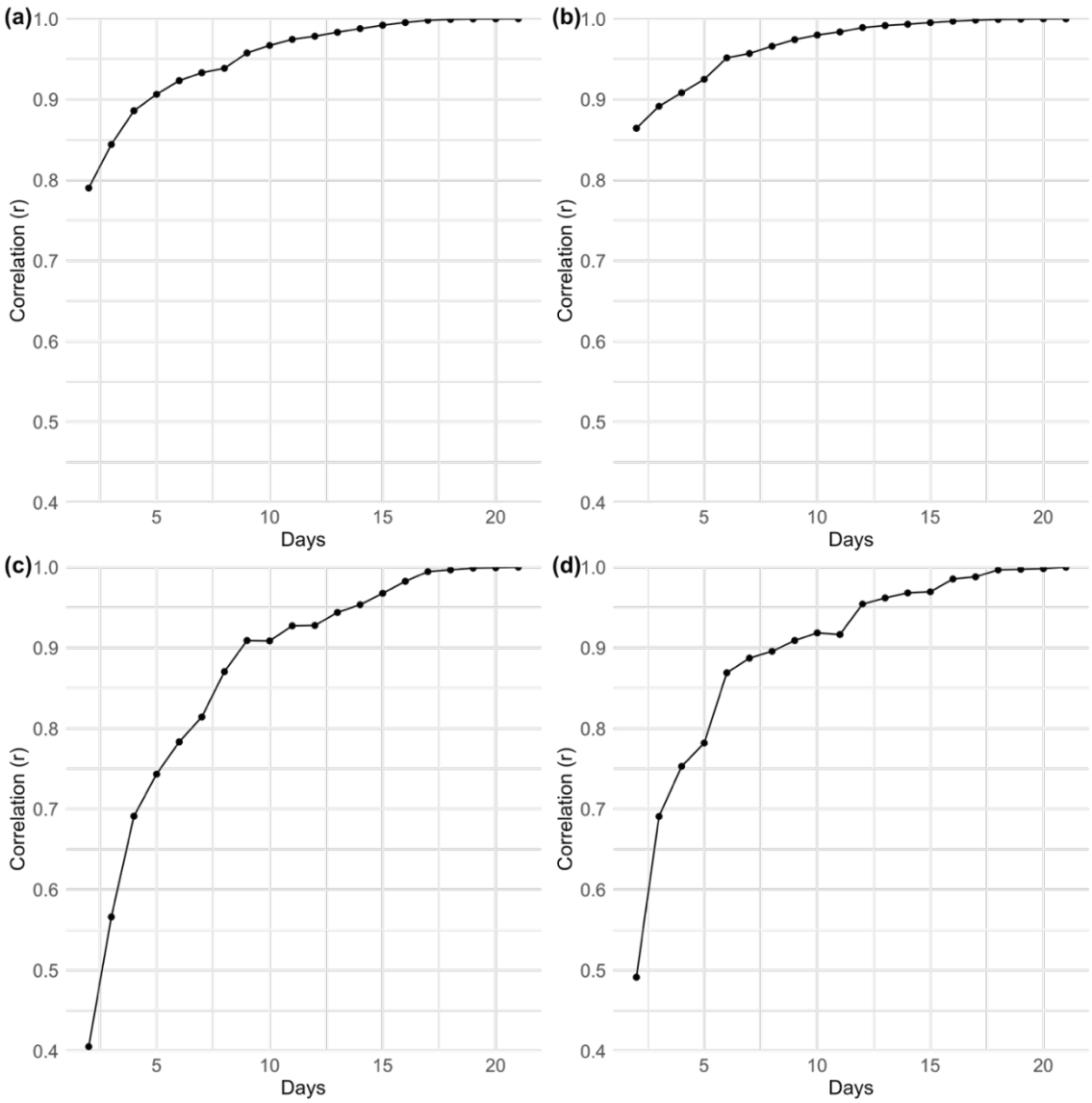
Daily Values of Rc and R1F During the Pre- and Post-Intervention Phases (n = 70)

Days	Pre-Intervention Phase		Post-intervention Phase	
	Rc	*R1F*	*Rc*	*R1F*
2	.45	.73	.47	.82
3	.54	.73	.40	.81
4	.55	.73	.42	.81
5	.51	.72	.44	.82
6	.67	.72	.49	.82
7	.53	.75	.49	.82
8	.54	.76	.48	.81
9	.54	.76	.47	.81
10	.54	.77	.49	.81
11	.55	.77	.55	.77
12	.54	.77	.49	.80
13	.54	.77	.48	.80
14	.53	.78	.49	.80
15	.54	.78	.50	.81
16	.54	.78	.53	.81
17	.56	.78	.53	.81
18	.55	.78	.53	.81
19	.55	.78	.53	.81
20	.55	.78	.53	.81
21	.55	.78	.53	.81

Note. Rc refers to between-person reliability and R1F refers to within-person reliability.

Figure 1

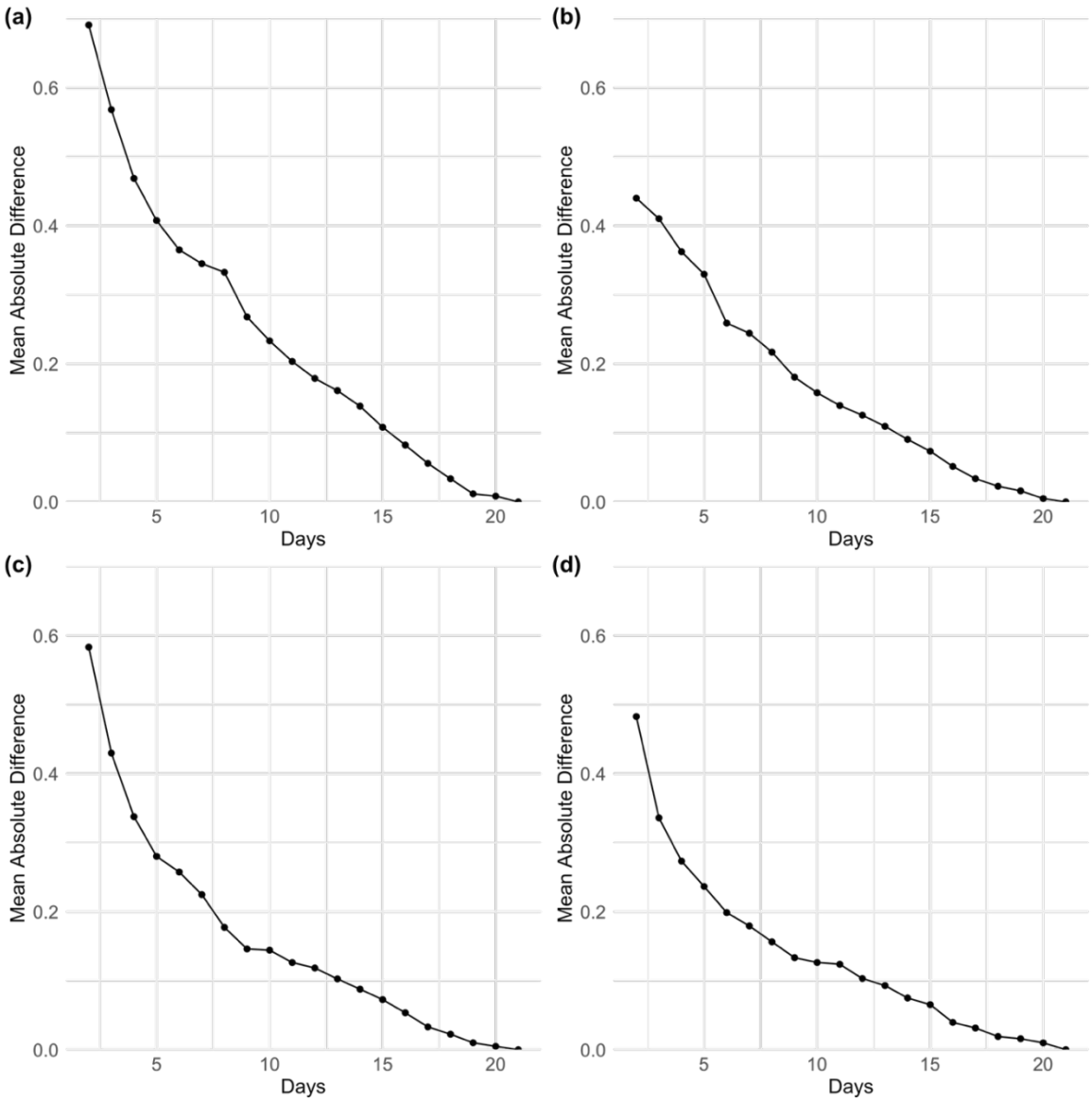
Correlations Between Reference and Estimate Values of PTSD Symptom iMs and iSDs During the Pre- and Post-Intervention Phases (n = 70)



Note. a) correlations between reference and estimate values of PTSD symptom iMs during the pre-intervention phase; b) correlations between reference and estimate values of PTSD symptom iMs during the post-intervention phase; c) correlations between reference and estimate values of PTSD symptom iSDs during the pre-intervention phase; d) correlations between reference and estimate values of PTSD symptom iSDs during the post-intervention phases. iM refers to intraindividual mean and iSD refers to intraindividual standard deviation.

Figure 2

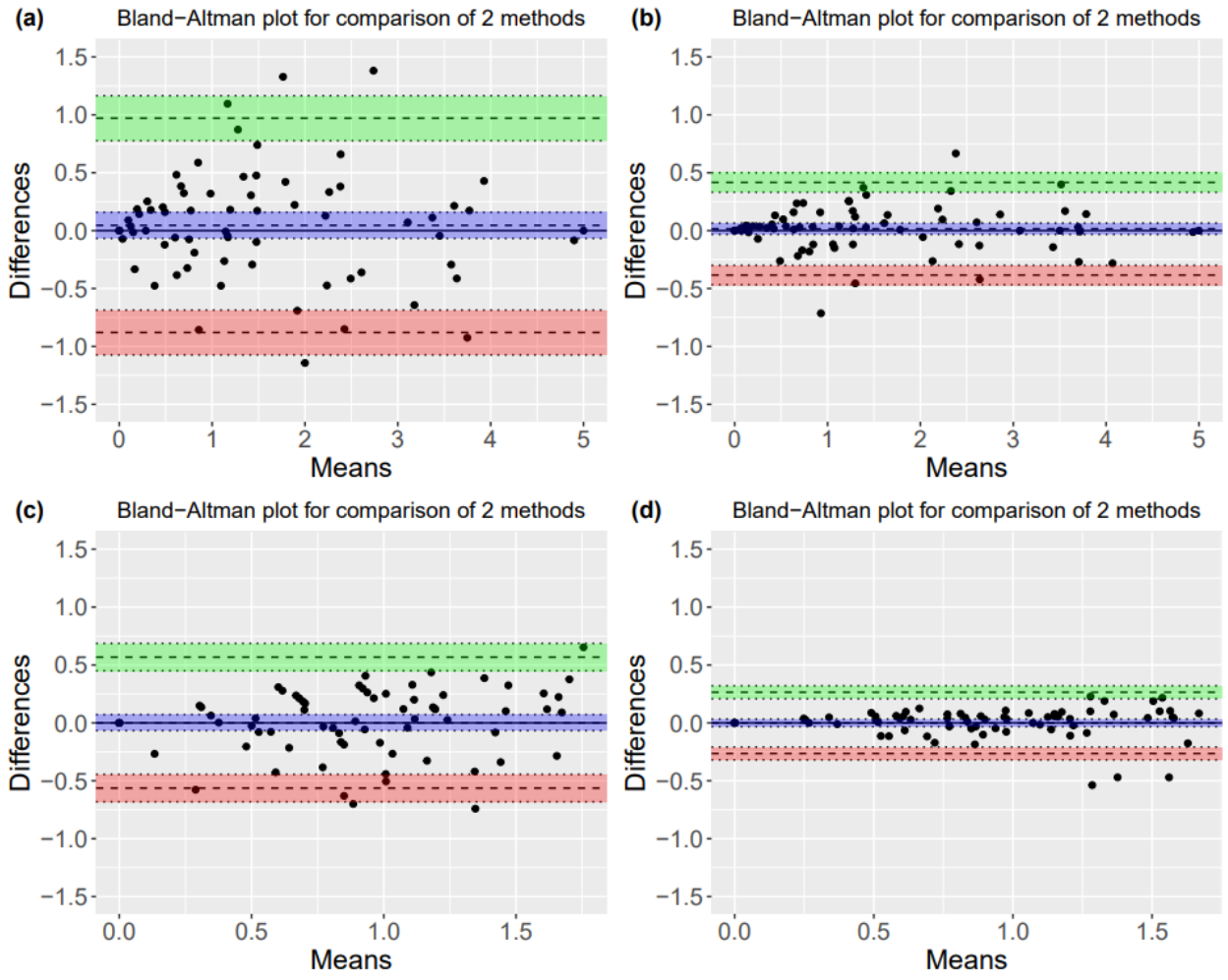
Mean Absolute Differences (MAD) Between Reference and Estimate Values of PTSD Symptom iMs and iSDs During the Pre- and Post-Intervention Phases (n = 70)



Note. a) MAD between reference and estimate values of PTSD symptom iMs during the pre-intervention phase; b) MAD between reference and estimate values of PTSD symptom iMs during the post-intervention phase; c) MAD between reference and estimate values of PTSD symptom iSDs during the pre-intervention phase; d) MAD between reference and estimate values of PTSD symptom iSDs during the post-intervention phases. iM refers to intraindividual mean and iSD refers to intraindividual standard deviation.

Figure 3

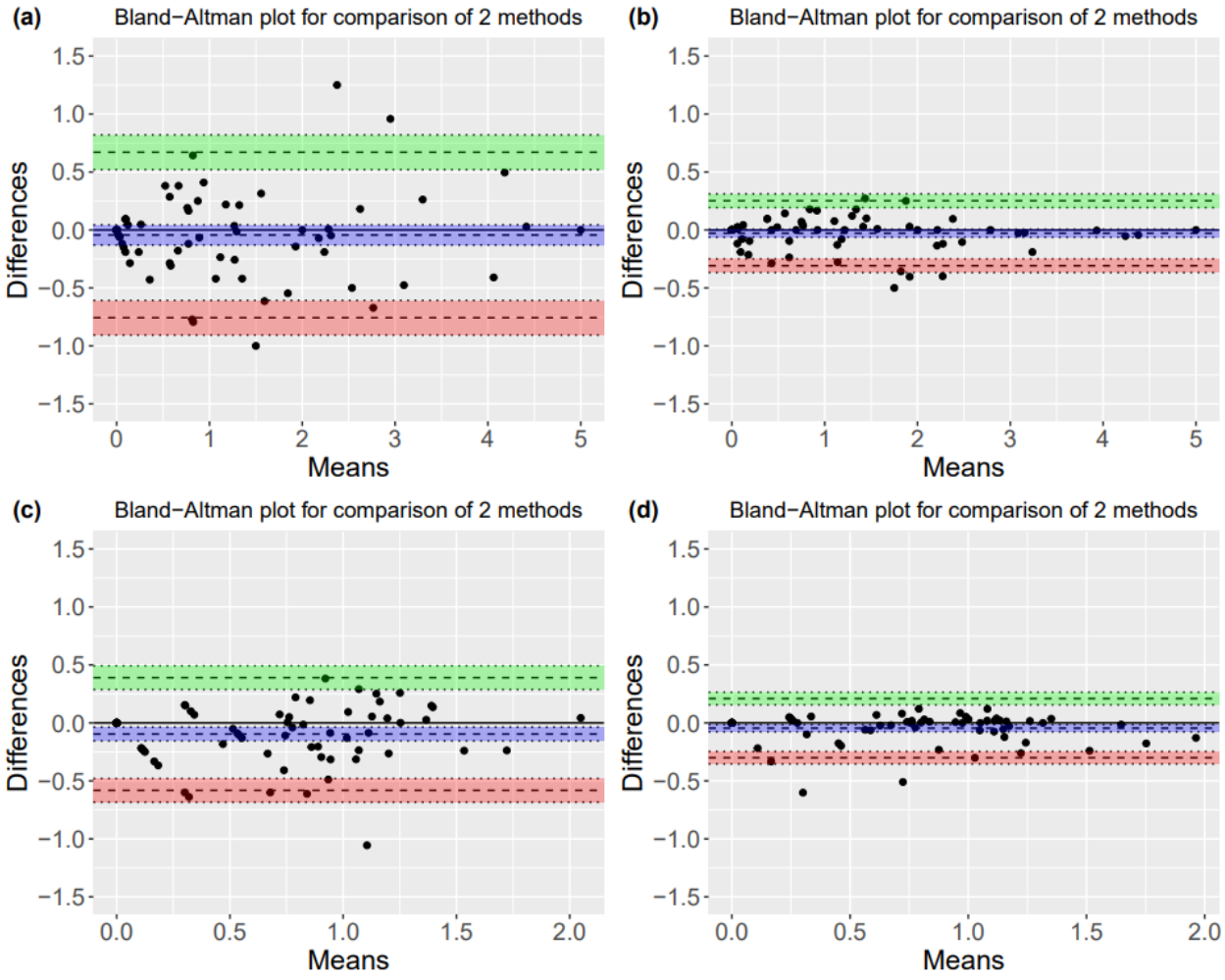
Bland-Altman Plots for 7 and 14 Days During the Pre-Intervention Phase (n = 70)



Note. a) agreement between reference and estimated values of PTSD symptom iMs on day 7; b) agreement between reference and estimated values of PTSD symptom iMs on day 14; c) agreement between reference and estimated values of PTSD symptom iSDs on day 7; d) agreement between reference and estimated values of PTSD symptom iSDs on day 14. iM refers to intraindividual mean and iSD refers to intraindividual standard deviation.

Figure 4

Bland-Altman Plots for 7 and 14 Days During the Post-Intervention Phase (n = 70)



Note. a) agreement between reference and estimated values of PTSD symptom iMs on day 7; b) agreement between reference and estimated values of PTSD symptom iMs on day 14; c) agreement between reference and estimated values of PTSD symptom iSDs on day 7; d) agreement between reference and estimated values of PTSD symptom iSDs on day 14. iM refers to intraindividual mean and iSD refers to intraindividual standard deviation.