Annals of Behavioral Medicine OXFORD UNIVERSITY PRESS

Ann Behav Med. 2023 Oct; 57(10): 846–854. PMCID: PMC10498818
Published online 2023 Apr 21. doi: 10.1093/abm/kaad015 PMID: 37084792

The Perceived Stress Scale as a Measure of Stress: Decomposing Score Variance in Longitudinal Behavioral Medicine Studies

Kristie M Harris, PhD,[✉] Allison E Gaffey, PhD, Joseph E Schwartz, PhD, David S Krantz, PhD, and Matthew M Burg, PhD

Abstract

Background

The Perceived Stress Scale (PSS) is a widely used measure designed to assess perceptions of recent stress. However, it is unclear to what extent the construct assessed by the PSS represents factors that are stable versus variable within individuals, and how these components might vary over time.

Purpose

Determine the degree to which variability in repeated PSS assessments is attributable to between-person versus within-person variance in two different studies and populations.

Methods

Secondary analyses utilized data from two studies with up to 13 PSS assessments: An observational study of 127 patients with heart failure followed over 39 months (Study 1), and an experimental study of 73 younger, healthy adults followed over 12 months (Study 2). Multilevel linear mixed modeling was used to estimate sources of variance in the PSS total and subscale scores across assessments.

Results

Between-person variance accounted for a large proportion of the total variance in PSS total scores in Study 1 (42.3%) and Study 2 (51.1%); within-person variance comprised the remainder. Between-person variance was higher for shorter assessment periods (e.g., 1 week), and was comparable when examining only the first 12 months of assessments in each study (52.9% vs. 51.1%).

Conclusions

Within two samples differing in age and health status, between-person variance accounted for approximately half of the total variation in PSS scores over time. While within-person variance was observed, the construct assessed by the PSS may substantially reflect a more stable characteristic of how an individual perceives stressful life circumstances than previously appreciated.

Keywords: Perceived stress scale, Psychological, Stress, Reliability, Longitudinal, Test–retest correlation, Psychometric properties

---

Over repeat administrations, variability in scores on the Perceived Stress Scale results more from differences between participants than from within-person changes over time.

Introduction

The Perceived Stress Scale (PSS) is widely considered the gold standard instrument for measuring stress perception [1–3]. Informed by Lazarus and Folkman's transactional model of stress [4], this self-report questionnaire is used to assess the degree to which a respondent finds circumstances in their life to be unpredictable, uncontrollable, and/or overwhelming. In contrast to measures that assess environmental stress exposures (e.g., life events checklists [5, 6]), the PSS captures a respondent's subjective appraisal of whether life circumstances and experienced events exceed their adaptive capacity [1, 7]. Questions on the PSS ask the respondent to rate their feelings and thoughts for the past month [1], indicating that a relatively recent time frame is assessed.

Since 1983, the PSS has been cited >30,000 times and used to demonstrate the wide-ranging effects of stress in both healthy and patient populations. Research using the PSS has advanced our understanding of the associations between stress and health outcomes (e.g., delayed wound healing, susceptibility to infection, cardiovascular disease) [8–10] and the physiological and behavioral pathways through which stress affects health [7, 11–15]. In 2015, the PSS was added to the National Institutes of Health's Toolbox for the Assessment of Neurological and Behavioral Function (i.e., NIH Toolbox) due to the measure's broad evidence base, strong psychometric properties, brevity, and applicability among diverse populations and across the lifespan [2].

The psychometric properties of the PSS have been extensively examined. Evidence supports the construct and predictive validity of the original 14-item PSS (PSS-14) and the abridged 4- and 10-item versions (i.e., PSS-4 and PSS-10) [6–12, 15–17]. While current scoring guidelines suggest calculating one total stress score, factor analytic studies of the PSS-10 and PSS-14 have indicated that either a unidimensional or two-factor structure (i.e., perceived helplessness and perceived self-efficacy) can be derived from the items [7, 18–22]. Internal consistency reliability is adequate-to-high for the entire PSS-10 and PSS-14 scales (Cronbach's α ≥ 0.75), and for each subscale when using the two-factor structure (Cronbach's α ≥ 0.81) [1, 7, 17, 18, 23, 24]. Due to the limited number of items, the PSS-4 has lower reliability (Cronbach's α = 0.60) [7]. Further, differential item analyses of the PSS-10 have demonstrated the relative lack of measurement variance of items across subgroups defined by sex, race, and education, suggesting that the measure is appropriate for widespread use in the general population [25, 26].

Despite the strength of this psychometric evidence, data on the temporal stability of the PSS remain equivocal. Some studies, including the original validation studies, suggest that there is a decline in the predictive validity and test–retest reliability of the PSS after just 4 weeks [1, 7, 27–32], while two other recent studies, both in medical populations, show a relative stability of scores over 6-month and 1-year intervals [33, 34]. Indeed, as the PSS assesses the perceived balance (or imbalance) between stress exposures and an individual's ability to respond to those stressors [1, 7], responses are likely influenced by both factors that are relatively stable within individuals but differ between persons (e.g., personality, coping style, response style, etc.), and those that vary within an individual over time (e.g., number of stressors, type of stressors, availability of resources, etc.). It remains unclear, however, to what extent the PSS is assessing these types of factors, and how this might change over repeat administrations.

To address this question, we conducted secondary analyses of data from two studies. Multilevel linear mixed models (MLMM) were used to investigate the temporal stability of the PSS by decomposing the total variance of numerous repeat assessments into within-person and between-person sources of variability. Within-person variance captures score variability that fluctuates within the same individual over time. While between-person variance captures score variability that is relatively stable within an individual over time but differs between individuals. Of note, traditional test–retest reliability statistics only provide information about the stability of scores over the specific interval that is assessed and cannot parse out differing sources of score variance. MLMM methodology, however, has the advantage of being able to generate estimates of score stability over any interval, as well as to provide a breakdown of variance by source (within-person and between-person), and to show how the contributions of these sources change as you vary the time interval between assessments.

To increase generalizability of the present analyses, the stability of the PSS was investigated in two disparate samples, each of which included up to 13 administrations of the PSS: (Study 1) 146 patients with heart failure, a chronic, life-threatening cardiovascular disease, who were followed over 39 months, and (Study 2) 80 younger, healthy adults followed over 13 months. These studies differ on several key factors, including study design, the period over which the PSS was administered, the version of the PSS used, and characteristics of the study populations. In this paper we report on the stability of the PSS scores over time and the proportion of variability that is attributable to within-person and between-person sources of variance.

Methods

Participants and Procedures: Study 1

The Biobehavioral Triggers in Heart Failure (i.e., BETRHEART) study was a prospective observational cohort study of 146 patients with a diagnosis of heart failure with reduced ejection fraction who were recruited from the University of Maryland Medical Center [35, 36]. Data were collected from June 2008 to

March 2015. Enrolled participants were diagnosed with heart failure for ≥3 months, symptomatic (New York Heart Association Class II–IV), and had a documented left ventricular ejection fraction ≤40% within the past year. After enrollment, participants completed a baseline assessment followed by 5 bi-weekly assessments, a 3-month assessment, and 6 semi-annual assessments that included the PSS-10, for a total of 13 assessments per participant over a 39-month period. For each PSS-10 assessment, participants were asked to report about their feelings and thoughts during the previous 2 weeks.

Participants and Procedures: Study 2

The Stress and Exercise Study was a single cohort, randomized controlled experiment of personal daily stress and exercise among 80 healthy individuals aged 18 and older who were recruited at Columbia University Medical Center [37, 38]. Data were collected from January 2014 to May 2015. As part of a larger study protocol [37, 38], participants completed a baseline assessment followed by 12 consecutive, monthly, follow-up assessments that included the PSS-14, for a total of 13 assessments per participant over a 12-month period. For each PSS-14 assessment, participants were asked to report about their feelings and thoughts during the previous month.

Both studies received local Institutional Review Board approval and were performed in accordance with the ethical standards in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Written informed consent was obtained from all individual participants included in each study.

Measures

PSS As described above, the PSS-14 and PSS-10 are 14-item and 10-item questionnaires that assess an individual's subjective perception of recent stressful life circumstances [1]. For each version of the PSS, two subscales can be calculated to assess Perceived Helplessness and Perceived Self-Efficacy in response to life events. A list of the items and score ranges for each PSS version and subscale is included in Supplementary Material 1.

Cronbach's alpha, a measure of internal consistency reliability, was calculated using the baseline PSS administration from each study and values were comparable with previous reports in the literature [26]. In Study 1, $\alpha = 0.89$ for the total PSS-10 score, $\alpha = 0.82$ for the Perceived Helplessness subscale, and $\alpha = 0.89$ for the Perceived Self-Efficacy subscale. In Study 2, $\alpha = 0.79$ for the total PSS-14 score and $\alpha = 0.84$ for the Perceived Helplessness subscale. For the Perceived Self-Efficacy subscale, however, α was lower at 0.66, perhaps because this scale includes three of the four items found to have lower factor loadings on the PSS-14 [5].

Demographics and health status In each study, information about participant demographics and health status was collected at baseline via participant interview and query of their electronic health records (Study 1) or via a self-report questionnaire (Study 2).

Statistical Analyses

Analyses were performed using SAS software version 9.4 (SAS Institute Inc., Cary, NC) and $p < .05$ was used to determine statistical significance (1-tailed test for variances, 2-tailed otherwise). All analyses were restricted to participants who completed at least six administrations of the PSS. Demographic characteristics of each sample were summarized, and total and subscale scores were calculated for all PSS data. Missing data were estimated using Markov chain Monte Carlo imputation with a single chain and 100 iterations following 200 burn-in iterations when respondents answered at least seven items on the PSS-10 or PSS-14. PROC MIXED was then used to estimate repeated measures MLMMs, providing a variance decomposition of the PSS over time. Separately for each study, five models were fit (one for each covariance structure considered, see below), to evaluate the sources of variance for the—(a) total PSS score, (b) Perceived Helplessness subscale score, and (c) Perceived Self-Efficacy subscale score. In each case, the intercept was treated as a random effect (capturing the between-person variance) and one of five within-person error covariance structures were considered: Compound symmetry (no autocorrelation, that is, no correlations among the within-person residuals of the 13 assessments), first-order serial correlation (i.e., autocorrelated within-person residuals), first-order serial correlation with unique variance (i.e., the combination of a first-order autocorrelated component and a non-autocorrelated fluctuations component), Toeplitz, and unstructured [39]. All models were estimated by maximum likelihood and model fit was evaluated using the Bayesian information criterion (BIC). The final models used the covariance structure with the overall lowest BIC across models, and we report the estimated total variance, the amounts and percentages attributable to between-person and within-person sources of variance, and the 1-month autocorrelation parameter.

For the Study 2 data, where assessments were obtained at fixed intervals (i.e., monthly), the Toeplitz model generates separate estimates of the 1-, 2-, 3-, … 12-month test–retest correlations for each scale. We created variograms for the PSS-14 total and subscale scores to visually compare these "observed" correlations to those predicted from the parameter estimates of the three more parsimonious models (compound symmetry, first-order serial correlation, and first-order serial correlation with unique variance).

Using the same models described above, again with five error covariance structures considered, a series of sensitivity analyses were also performed. First, to match the time frame over which the Study 2 data were collected and allow for a more direct comparison of score stability between the two studies, MLMMs were fit to only the Study 1 PSS-10 total score data collected during the first 12 months of the study (i.e., up to 8 administrations from baseline through the 9-month follow-up). Second, in addition to the 13 primary Study 1 assessments, the larger study protocol included 1 week of daily assessments with the PSS-10 between the fifth/last bi-weekly assessment and the 3-month assessment. For each daily assessment, participants were instructed to report about their experiences over the prior 24 hr. Again, MLMMs were used to examine score stability over the more condensed 1-week period among participants completing at least 6 of the 7 daily PSS-10 assessments.

Results

Of the 146 Study 1 participants, 127 completed the requisite number of 6 PSS assessments, with a median of 10 (lower, upper quartiles: 8, 12) assessments completed. The baseline average PSS-10 score was 13.08 (*SD* = 8.09, range = 0–32), which is consistent with nationwide averages for a similar age group (ages 55–64) [7]. Of the 80 Study 2 participants, 73 completed the requisite number of 6 PSS assessments, with a median of 13 (lower, upper quartiles: 13, 13) assessments completed. The baseline average PSS-14 score was 22.00 (*SD* = 6.71, range = 10–41), which is also consistent with nationwide averages for a similar age group (ages 30–44) [7]. In total, these studies included data from 2,880 PSS assessments (Study 1: 1268 primary assessments and 704 daily assessments; Study 2: 908 assessments), of which missing data were imputed for 26 assessments (0.9% of total) where the participant answered ≥7 items.

The two samples differed on all assessed demographic characteristics (see Table 1). Compared to the healthy Study 2 sample, the proportions of men, African Americans, and non-Hispanics were higher in Study 1. Participants in Study 1 were also older, less educated, and more likely to be currently retired/on disability, and separated, divorced, or widowed.

Table 1

Participant Demographic Characteristics by Study

Characteristic	Study 1 (n = 127)	Study 2 (n = 73)
Age, years	56.5 ± 11.3	31.9 ± 9.6
Age range, years	23–83	20–58
Gender		
Men	97 (76.4)	30 (41.1)
Women	30 (23.6)	43 (58.9)
Race		
African American	88 (69.3)	2 (2.7)
Asian/Pacific Islander	0 (0.0)	25 (34.3)
White	38 (29.9)	30 (41.1)
Other/Unknown/Declined	1 (0.8)	13 (21.9)
Ethnicity		
Hispanic	1 (0.8)	21 (28.8)
Non-Hispanic	126 (99.2)	51 (69.9)
Unknown/Declined	0 (0.0)	1 (1.3)
Body mass index, kg/m^2	31.4 ± 7.6	26.6 ± 5.5
Highest Education Level		
≤ High school	64 (50.4)	1 (1.3)
Some college/college degree	57 (44.9)	41 (56.2)
Graduate training/degree	6 (4.7)	31 (42.5)
Work Status		
Employed, full or part-time	26 (20.5)	59 (80.8)
Retired/Disability	96 (75.6)	0 (0.0)
Unemployed	5 (3.9)	6 (8.2)
Unknown/Declined	0 (0.0)	8 (11.0)
Partner Status		
Single, never married	29 (22.8)	42 (57.5)
Married/Partnered	36 (28.4)	30 (41.1)
Separated/Divorced/Widowed	62 (48.8)	1 (1.4)

Values represent mean ± standard deviation or n (%). All characteristics differed significantly between the two studies at the $p < .05$ level (2-tailed).

Results of the MLMMs examining between-person and within-person variability of the PSS total and subscale scores in each study are detailed in Table 2. Among the five error covariance structures considered, the first-order serial correlation with unique variance provided the overall best fit across scores in both studies according to the BIC statistic (see Supplementary Material 2). This model provides estimates of the between-person variance and two components of within-person variance—a first-order serially "autocorrelated" component (i.e., associated with an individual's score at the prior assessment) and a "non-autocorrelated fluctuations" component. The non-autocorrelated fluctuations component can include both random measurement error and true within-person variability ("white noise") that fluctuates on a time scale substantially shorter than the interval between assessments. These estimates and the autocorrelation parameter are reported in Table 2.

Table 2

Multilevel Linear Mixed Model Results for PSS Total and Subscale Scores Across Both Studies

Source of variance	Study 1 (n = 127) PSS-10						Study 2 (n = 73) PSS-14			
	Total score		Perceived helplessness		Perceived self-efficacy		Total score		Perceived helplessness	
	Estimate (SE)	%	Estimate (SE)	%	Estimate (SE)	%	Estimate (SE)	%	Estimate (SE)	%
Between-person	29.11 (7.91)	44.6	17.05 (2.74)	51.5	4.42 (1.09)	35.8	27.16 (6.24)	51.1	11.62 (2.30)	53.0
Within-person	31.80	55.4	16.06	48.5	7.92	64.2	25.94	48.9	10.32	47.0
Autocorrelated	15.64 (5.71)	24.0	6.89 (1.73)	20.8	3.04 (0.72)	24.6	17.22 (2.84)	32.4	7.97 (1.35)	36.3
Non-Autocorrelated Fluctuations	20.54 (1.86)	31.5	9.18 (2.36)	27.7	4.88 (0.56)	39.5	8.72 (1.57)	16.4	2.35 (1.38)	10.7
Total	65.30	100.0	33.12	100.0	12.34	100.0	53.10	100.0	21.94	100.0
1-Month Autocorrelation	0.83 (0.12)	—	0.61 (0.23)	—	0.74 (0.14)	—	0.73 (0.09)	—	0.52 (0.11)	—

SE = standard error; % = percentage of total variance.

Each model used a first-order serial correlation with unique variance within-person covariance structure. Within-person variance is sub-divided into a first-order serially "autocorrelated" variance component and a "non-autocorrelated fluctuations" variance component. Except for the estimate of the between-person variance of perceived self-efficacy in Study 2, all estimates are significant at the $p < .05$ level (1-tailed for variances, 2-tailed for the autocorrelation parameter).

In Study 1, the total variance in the PSS-10 total scores (65.30) was primarily attributable to between-person variance (44.6% of total variance), followed by within-person sources of variance including non-autocorrelated fluctuations (31.5%), and autocorrelated variance (24.0%). The estimate of the 1-month autocorrelation for this portion of the within-person variance was 0.83, indicating that change in this component of the PSS was relatively gradual. The Perceived Helplessness subscale scores followed the same pattern, with the total variance (33.12) mostly attributable to between-person variance (51.5%), followed by non-autocorrelated fluctuations (27.7%), and autocorrelated within-person variance (20.8%). For the Perceived Self-Efficacy subscale, between-person variance again accounted for more (35.8%) of the total variance (12.34) than autocorrelated within-person variance (24.6%), but non-autocorrelated fluctuations (39.5%) were the largest source of variance.

Similar findings emerged for Study 2. The total variance (53.10) in the PSS-14 total scores across the 12-month assessment period was also primarily attributable to between-person variance (51.1%) compared to autocorrelated within-person variance (32.4%). The estimated 1-month autocorrelation for this last source of variance was 0.73, somewhat lower, but not significantly lower than that found in Study 1. However, in contrast to the Study 1 data which were collected over a longer assessment period, the variance of non-autocorrelated fluctuations in Study 2 was lower and accounted for the smallest percentage of the total variance (16.4%) among the three sources. Again, the same pattern emerged for the Perceived Helplessness subscale scores, with the total variance (21.94) being mostly due to between-person variance (53.0%), followed by autocorrelated within-person variance (36.3%), and finally non-autocorrelated fluctuations (10.7%). For the Perceived Self-Efficacy subscale however, 100.0% of the total variance (14.87) was attributable to within-person sources of variance, with 62.9% due to autocorrelated within-person variance and the remaining 37.1% attributable to non-autocorrelated fluctuations.

Variograms created using the Study 2 data for the PSS total and subscale scores are included as Fig. 1 and Supplementary Material 3A and B, respectively. Each variogram includes estimates of the 1-, 2-, 3-, …, 12-month test–retest correlations and their 95% confidence intervals, and depicts the portions of the test–retest correlation attributable to stable, trait-like variance and autocorrelated within-person (state-like) variance. These variograms also show that the test–retest correlations change over different intervals of time between assessments, a pattern that is not easily detected by traditional test–retest correlation methodology.

Fig. 1.

Variogram for 13 monthly assessments of the 14-item perceived stress scale. The variogram shows the predicted test–retest correlation for three parsimonious variance decomposition models created using data from Study 2: (1) Compound symmetry (black horizontal line)—assumes that the within-person variance is uncorrelated from one assessment to the next, and thus the test–retest correlation does not depend on the time interval between assessments; (2) First-order serial correlation (i.e., autocorrelated within-person residuals; brown curve): the test–retest correlation of the within-person component of variance declines exponentially as a function of the time interval between assessments; and (3) First-order serial correlation with unique variance (i.e., the combination of a first-order serially autocorrelated component and a non-autocorrelated fluctuations component; blue curve): The non-autocorrelated component consists of "rapidly changing" (so as to not be correlated from 1 month to the next) situation-dependent factors and/or random measurement error. Estimates of the 1-, 2-, 3-, …, 12-month test–retest correlations and their 95% confidence intervals, obtained by fitting a Toeplitz model to the repeated measures data, are shown as red circles with vertical 95% confidence error bars. The variogram reveals that the third model (blue curve) provides the best fit to the 12 estimated test–retest correlations, consistent with the lower BIC statistics reported in Supplementary Material 2. Based on the estimates from the third model, the light green shaded area represents the portion of the test–retest correlation attributable to stable, trait-like variance; the light blue shaded area represents the portion of the test–retest correlation attributable to autocorrelated within-person (state-like) variance. Corresponding variograms for the Perceived Helpless subscale and the Perceived Self-Efficacy subscale are provided in Supplementary Material 3A and B.

Sensitivity Analyses

Table 3 provides a side-by-side comparison of the MLMM results from the 12-month Study 2 assessment period (same as in Table 2) and the MLMM results examining only the first 12 months of data collected in Study 1. The total variance (63.77) in this subset of the Study 1 data was largely accounted for by between-person variance (52.9%), with the variance of non-autocorrelated fluctuations (25.8%), and autocorrelated within-person variance (21.3%) responsible for smaller percentages. The between-person variance was similar in each study (52.9% in Study 1 and 51.1% in Study 2) and again accounted for the largest percentage of the total variance. The percentage of autocorrelated within-person variance (32.4%) was larger in Study 2 than in Study 1 (21.3%), while Study 1 had the larger percentage of non-autocorrelated fluctuations variance among the two studies (25.8% compared to 16.4%).

Table 3

Multilevel Linear Mixed Model Results for PSS Total Score from First 12 Months of Study 1 and Study 2

Source of variance	**Study 1 (*n* = 127)**		**Study 2 (*n* = 73)**	
	PSS-10 Total (First 12 Months ONLY)		**PSS-14 Total**	
	Estimate (SE)	**%**	**Estimate (SE)**	**%**
Between-person	33.75 (7.61)	52.9	27.16 (6.24)	51.1
Within-person	30.02	47.1	25.94	48.9
Autocorrelated	13.60 (3.61)	21.3	17.22 (2.84)	32.4
Non-Autocorrelated Fluctuations	16.42 (4.12)	25.8	8.72 (1.57)	16.4
Total	63.77	100.0	53.10	100.0
1-Month Autocorrelation	0.70 (0.31)	—	0.73 (0.09)	—

SE = standard error; % = percentage of total variance.

Within-person variance is sub-divided into a first-order serially "autocorrelated" variance component and a "non-autocorrelated fluctuations" variance component. All estimates are significant at the $p < .05$ level (1-tailed for variances, 2-tailed for the autocorrelation parameter).

A total of 103 Study 1 participants completed 6 or 7 days of daily ratings in the week between the fifth/last bi-weekly assessment and the 3-month assessment. Again, the total variance (58.65) was largely due to between-person variance (74.3%), followed by autocorrelated within-person variance (20.6%) and non-autocorrelated fluctuations (5.1%). These results are shown in Table 4 alongside the MLMM results for the first 12 months of Study 1 data, and data collected over the entire 39-month assessment period for these n = 103 participants only. As the total assessment period decreased (i.e., 39 months, 12 months, 1 week), the percentage of variance attributable to non-autocorrelated fluctuations also decreased (from 28.6% to 5.1%) while the between-person variance percentage increased from 49.0% to 74.3% of the total variance.

Table 4

Multilevel Linear Mixed Model Results for PSS Total Score Using Study 1 39-Month Data, First 12 Months Only, and 7-Day Daily Assessment

Source of variance	39-Month Data (n = 103; 6–13 assessments)		First 12 Months Only (n = 103; up to 8 assessments)		1 Week (n = 103; 6–7 daily assessments)	
	Estimate (SE)	%	Estimate (SE)	%	Estimate (SE)	%
Between-person	32.73 (8.10)	49.0	38.67 (6.70)	59.3	43.57 (7.42)	74.3
Within-person	34.01	51.0	26.55	40.7	15.08	25.7
Autocorrelated	14.92 (4.98)	22.4	12.18 (4.54)	18.7	12.06 (2.18)	20.6
Non-Autocorrelated Fluctuations	19.09 (2.19)	28.6	14.37 (6.35)	22.0	3.02 (2.17)	5.1
Total	66.74	100.0	65.22	100.0	58.65	100.0
1-Month Autocorrelation	0.80 (0.14)	—	0.54 (0.36)	—	0.62 (0.19)	—

SE = standard error; % = percentage of estimate to total variance.

Within-person variance is sub-divided into a first-order serially “autocorrelated” variance component and a “non-autocorrelated fluctuations” variance component. For the first 12 months only (center panel) 1-Month Autocorrelation p = .13 (2-tailed) and for the 1 Week (right panel) Non-Autocorrelated Fluctuations p = .08 (1-tailed). All other estimates are significant at the $p < .05$ level (1-tailed for variances, 2-tailed for the autocorrelation parameter).

Discussion

In two studies differing widely in methodology and in the demographics and health status of the samples, each including up to 13 longitudinal repeated assessments of the PSS, more of the variation in the PSS total and subscale scores was the result of differences between participants than from within-person changes in scores over time. This pattern held over brief (i.e., 1 week) and longer-term (i.e., 39 months) assessment periods, and was strikingly similar (52.9% vs. 51.1%) across the studies when the assessment period was held constant at 12 months. Additionally, in Study 1, as the period of assessment decreased from 39 months to 12 months to 1 week, the proportion of between-person variance increased from 49.0% to 59.3% to 74.3%, respectively. These results suggest that the construct measured by the PSS may, to a substantial degree, reflect a relatively stable characteristic. Furthermore, the degree of score stability appears to vary by the length of the assessment period, with scores appearing more stable over shorter periods. While not initially predicted, this finding is to be expected as there is likely less real change in people’s circumstances over shorter versus longer intervals.

PSS responses are thought to reflect the dynamic balance between an individuals’ exposure to life circumstances/events, and their ability to respond to those stressors [1, 7]. Data from the present study did reveal a substantial degree of intra-individual (within-person) variation in PSS scores, and we know from earlier work in Study 1 that changes in PSS scores can predict health outcomes [36]. What this study adds, however, is a more nuanced understanding of how PSS scores behave over time when administered repeatedly to a group of individuals—that is, as more stable than previously appreciated. Our novel analytic approach revealed that PSS score variation within the same individuals across assessments *contributed less* to the overall score variance than differences observed between participants. Furthermore, within-person variation was, to a degree, predictable based on an individual's prior assessment (i.e., the autocorrelation parameter) with only a portion of the variation in scores reported by individuals over time seemingly "random" (i.e., non-autocorrelated fluctuations). Even at our longest assessment period—39 months in Study 1—random variation (i.e., non-autocorrelated fluctuations) accounted for less than one-third of the total variance across the 1,268 PSS assessments.

The determination of a large between-person component of variance suggests that PSS scores are influenced, to a large extent, by factors that are relatively stable. These factors could be an individual's personality, or response style, or other characteristics that we think of as more trait-like, in that they exhibit little fluctuation across many years, or even across a person's lifetime. For instance, work in the domain of perceived control suggests that certain individuals are predisposed to feel confident in their ability to effectively respond to life demands, regardless of the specific circumstances [40]. These factors could also be those that are not semi-permanent/trait-like but are stable over a significant period of time, including chronic stressors that are experienced in an individual's lived environment (e.g., crime, caregiver stress, financial instability), and factors that influence how individuals perceive stressors and their ability to respond (e.g., symptoms of anxiety and depression, coping style) [41, 42]. Indeed, it is likely that for at least some participants, certain circumstances/stressors were, to a large degree, constant throughout the periods of observation. It is unlikely however, that the large between-person component observed is due entirely to stable participant circumstances/stress exposure, especially since Study 1 participants underwent frequent hospitalizations—a well-documented major life stressor that is known to disrupt daily life circumstances [36, 43]—after which increases in PSS scores were observed [36]. Thus, it seems more likely that our results suggest that the perception of stress exposure and of coping resources, as assessed by the PSS, is a relatively stable characteristic that varies only slightly over short intervals and only moderately over longer intervals.

The consistency of findings across the two samples, and across our sensitivity analyses, suggests that the results are highly generalizable. Yet, there are limitations. In both samples, approximately 10% of participants were excluded for failing to complete a sufficient number of assessments. It is also possible that the limited within-person variability observed was because there was little variability over time in the circumstances and stressfulness of events experienced by participants. As previously discussed, however, this is unlikely as Study 1 participants experienced the major life stressor of hospitalization [36, 43] frequently throughout the assessment period. Furthermore, it should be noted that with MLMM, less reliable measures are likely to have greater within-person non-autocorrelated fluctuations in variance. Additional discussion of this and other statistical considerations relevant to our modeling approach is included as Supplementary Material 4.

Implications and Future Directions

Longitudinal within-person methodologies with repeated PSS assessments have been used to measure dynamic, intra-individual fluctuations in perceived stress and/or to capture a more global picture of between-person differences in perceived stress over time [9, 12, 33–36]. Our findings suggest, however, that with this methodology, the construct assessed by the PSS behaves as a more stable characteristic than previously appreciated. A substantial portion of score variability over time is due to between-person differences and, among the portion that varies within persons, a significant proportion (i.e., the autocorrelated variance which is largely predictable based on the prior assessment) evolves rather gradually over time. As the period between assessments decreases, the variability of change also decreases.

These findings advance our understanding of the temporal stability of the PSS and how much of the measured construct appears to be the result of stable versus variable factors within individuals. Furthermore, if our gold standard measure of perceived stress is behaving mostly stable over time, we may need to re-think how we conceptualize, and measure, stress as an experience/perception. It is possible that the present findings are a product of the retrospective nature of the PSS, and that by asking individuals to reflect on their perceptions over general time frames (e.g., the last month) their responses are heavily influenced by their individual disposition. However, it is unclear if eliminating the retrospective reporting element—such as when using ecological momentary assessment (EMA) to capture perceptions of stress in the moment—would result in greater within-person score variance [44]. Variance decompensation analyses with EMA reports of anger and anxiety have revealed comparable results to the present study, with EMA methodology also resulting in larger between-person versus within-person components of variance over time [45].

It is important that the present findings are not construed as diminishing the validity of the PSS, but rather as indicating the need for additional study of how the PSS and other measures can best be used to capture the complexity of the human stress experience. Inevitably there will still be research questions that warrant repeat assessments using the PSS (e.g., before and after an intervention). Researchers are nevertheless encouraged to consider the time frame between assessments and the added value of repeat versus single assessments when the latter may provide a general sense of how an individual perceives life stress and their ability to mount a sufficient response. Our study also has implications for revisiting the interpretation of prior work based on the PSS. Specifically, we must reflect on the extent to which differences in PSS scores associated with health outcomes over time reflect stable and chronic differences between individuals as opposed to intra-individual changes precipitated by stressful circumstances. Finally, the results of the present analysis have implications for the statistical power of both cross-sectional and longitudinal studies with the PSS. A detailed description of these implications is available in Supplementary Material 4.

Conclusion

This secondary analysis of two studies that included repeat assessments of the PSS revealed that over time PSS scores are more stable than previously appreciated. Of the total variance in scores obtained over repeat assessments, the largest proportion was attributable to stable between-person differences, and much of the within-person variation was predictable based on an individual's prior assessment. These findings expand our understanding of the psychometric properties of the PSS and suggest opportunities to further study what drives changes in individual perceptions of stress over time and how the PSS or comparable measures can be optimally used to measure such changes.

Supplementary Material

kaad015_suppl_Supplementary_Material

Click here for additional data file.(189K, docx)

Acknowledgments

This work was supported by National Heart Lung and Blood Institute grants R01HL085730 and R01HL115941. Dr. Gaffey was supported by a VA Advanced Fellowship in Women's Health and a VA VISN1 Career Development Award. The opinions and assertions expressed herein are those of the authors and do not necessarily reflect those of the Uniformed Services University, the U.S. Department of Defense, or the National Institutes of Health.

Contributor Information

Kristie M Harris, Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA. VA Connecticut Healthcare System, West Haven, CT, USA.

Allison E Gaffey, Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA. VA Connecticut Healthcare System, West Haven, CT, USA.

Joseph E Schwartz, Department of Psychiatry, Stony Brook University, Stony Brook, NY, USA. Department of Medicine, Columbia University, New York, NY, USA.

David S Krantz, Department of Medical and Clinical Psychology, Uniformed Services University of the Health Sciences, Bethesda, MD, USA.

Matthew M Burg, Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA. VA Connecticut Healthcare System, West Haven, CT, USA. Department of Anesthesiology, Yale School of Medicine, New Haven, CT, USA.

Compliance with Ethical Standards

Authors' Statement of Conflict of Interest and Adherence to Ethical Standards Authors Kristie M. Harris, Allison E. Gaffey, Joseph E. Schwartz, David S. Krantz, Matthew M. Burg declare that they have no conflict of interest. All procedures, including the informed consent process, were conducted in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000.

Authors' Contributions Kristie M. Harris, PhD (Conceptualization: Lead; Formal analysis: Supporting; Investigation: Supporting; Methodology: Equal; Visualization: Supporting; Writing – original draft: Lead; Writing – review & editing: Lead), Allison E. Gaffey, PhD (Conceptualization: Supporting; Writing – review & editing: Supporting), Joseph E. Schwartz, PhD (Conceptualization: Equal; Data curation: Lead; Formal analysis: Lead; Methodology: Lead; Software: Lead; Supervision: Lead; Validation: Lead; Visualization: Lead; Writing – original draft: Supporting; Writing – review & editing: Supporting), David S. Krantz, PhD (Conceptualization: Supporting; Funding acquisition: Lead; Investigation: Lead; Writing – review & editing: Supporting), and Matthew M. Burg, PhD (Conceptualization: Equal; Funding acquisition: Lead; Investigation: Lead; Writing – review & editing: Supporting)

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Transparency Statements (1) *Study Registration*: Neither of the included studies were formally registered. (2) *Analytic Plan Pre-Registration*: The analysis plan was not formally pre-registered. (3) *Analytic Code Availability*: Analytic code used to conduct the analyses presented in this study are not available in a public archive. They may be available by emailing the corresponding author. (4) *Materials Availability*: Materials used to conduct the study are not publicly available.

Data Availability

De-identified data from this study are not available in a public archive. De-identified data from this study will be made available (as allowable according to institutional IRB standards) by emailing the corresponding author.

References

1. Cohen S, Kamarck T, Mermelstein R.. A global measure of perceived stress. *J Health Soc Behav.* 1983; 24(4):385–396. [PubMed] [Google Scholar]

2. Kupst MJ, Butt Z, Stoney CM, et al.. Assessment of stress and self-efficacy for the NIH toolbox for neurological and behavioral function. *Anxiety Stress Coping*. 2015; 28(5):531–544. [PMC free article] [PubMed] [Google Scholar]

3. Monroe SM. Modern approaches to conceptualizing and measuring human life stress. *Annu Rev Clin Psychol.* 2008; 4:33–52. [PubMed] [Google Scholar]

4. Lazarus RS, Folkman S.. *Stress, Appraisal, and Coping*. New York, NY: Springer; 1984. [Google Scholar]

5. Holmes TH, Rahe RH.. The social readjustment rating scale. *J Psychosom Res.* 1967; 11(2):213–218. [PubMed] [Google Scholar]

6. Berry C, Shalowitz M, Quinn K, Wolf R.. Validation of the crisis in family systems-revised, a contemporary measure of life stressors. *Psychol Rep.* 2001; 88(3 Pt 1):713–724. [PubMed] [Google Scholar]

7. Cohen S, Williamson G.. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, eds. *The Social Psychology of Health: Claremont Symposium on Applied Social Psychology*. Newbury Park, CA: Sage;1988:31–67. [Google Scholar]

8. Cohen S, Tyrrell DA, Smith AP.. Psychological stress and susceptibility to the common cold. *N Engl J Med.* 1991; 325(9):606–612. [PubMed] [Google Scholar]

9. Kiecolt-Glaser JK, Marucha PT, Malarkey WB, Mercado AM, Glaser R.. Slowing of wound healing by psychological stress. *Lancet.* 1995; 346(8984):1194–1196. [PubMed] [Google Scholar]

10. Strodl E, Kenardy J, Aroney C.. Perceived stress as a predictor of the self-reported new diagnosis of symptomatic CHD in older women. *Int J Behav Med.* 2003; 10(3):205–220. [PubMed] [Google Scholar]

11. Cohen S, Doyle WJ, Skoner DP.. Psychological stress, cytokine production, and severity of upper respiratory illness. *Psychosom Med.* 1999; 61(2):175–180. [PubMed] [Google Scholar]

12. Ebrecht M, Hextall J, Kirtley LG, Taylor A, Dyson M, Weinman J.. Perceived stress an cortisol levels predict speed of wound healing in healthy male adults. *Psychoneuroendocrinology.* 2004; 29(6):798–809. [PubMed] [Google Scholar]

13. Labbate LA, Fava M, Oleshansky M, Zoltec J, Littman A, Harig P.. Physical fitness and perceived stress. Relationships with coronary artery disease risk factors. *Psychosomatics.* 1995; 36(6):555–560. [PubMed] [Google Scholar]

14. Epel ES, Blackburn EH, Lin J, et al.. Accelerated telomere shortening in response to life stress. *Proc Natl Acad Sci U S A*. 2004; 101(49):17312–17315. [PMC free article] [PubMed] [Google Scholar]

15. Lawless MH, Harrison KA, Grandits GA, Eberly LE, Allen SS.. Perceived stress and smoking-related behaviors and symptomatology in male and female smokers. *Addict Behav.* 2015; 51:80–83. [PMC free article] [PubMed] [Google Scholar]

16. Roberti JW, Harrington LN, Storch EA.. Further psychometric support for the 10-item version of the perceived stress scale. *J Coll Couns*. 2006; 9(2):135–147. [Google Scholar]

17. Ezzati A, Jiang J, Katz MJ, Sliwinski MJ, Zimmerman ME, Lipton RB.. Validation of the perceived stress scale in a community sample of older adults. *Int J Geriatr Psychiatry.* 2014; 29(6):645–652. [PMC free article] [PubMed] [Google Scholar]

18. Taylor JM. Psychometric analysis of the ten-item perceived stress scale. *Psychol Assess.* 2015; 27(1):90–101. [PubMed] [Google Scholar]

19. Golden-Kreutz DM, Browne MW, Frierson GM, Andersen BL.. Assessing stress in cancer patients: a second-order factor analysis model for the perceived stress scale. *Assessment*. 2004; 11(3):216–223. [PMC free article] [PubMed] [Google Scholar]

20. Hewitt PL, Flett GL, Mosher SW.. The perceived stress scale: Factor structure and relation to depression symptoms in a psychiatric sample. *J Psychopathol Behav Assess.* 1992; 14(3):247–257. [Google Scholar]

21. Nielsen MG, Ørnbøl E, Vestergaard M, et al.. The construct validity of the perceived stress scale. *J Psychosom Res.* 2016; 84:22–30. [PubMed] [Google Scholar]

22. Mitchell AM, Crane PA, Kim Y.. Perceived stress in survivors of suicide: psychometric properties of the Perceived Stress Scale. *Res Nurs Health.* 2008; 31(6):576–585. [PubMed] [Google Scholar]

23. Cohen S, Janicki-Deverts D.. Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. *J Appl Soc Psychol*. 2012; 42(6):1320–1334. [Google Scholar]

24. Jiang JM, Seng EK, Zimmerman ME, Sliwinski M, Kim M, Lipton RB.. Evaluation of the reliability, validity, and predictive validity of the subscales of the perceived stress scale in older adults. *J Alzheimers Dis.* 2017; 59(3):987–996. [PMC free article] [PubMed] [Google Scholar]

25. Cole SR. Assessment of differential item functioning in the perceived stress scale-10. *J Epidemiol Community Health.* 1999; 53(5):319–320. [PMC free article] [PubMed] [Google Scholar]

26. Lee E. Review of the psychometric evidence of the perceived stress scale. *Asian Nurs Res*. 2012; 6(4):121–127. [PubMed] [Google Scholar]

27. Teresi JA, Ocepek-Welikson K, Ramirez M, et al.. Evaluation of the measurement properties of the perceived stress scale (PSS) in Hispanic caregivers to patients with alzheimer's disease and related disorders. *Int Psychogeriatr.* 2020; 32(9):1073–1084. [PMC free article] [PubMed] [Google Scholar]

28. Remor E. Psychometric properties of a European Spanish version of the perceived stress scale (PSS). *Span J Psychol.* 2006; 9(1):86–93. [PubMed] [Google Scholar]

29. Reis RS, Hino AA, Añez CR.. Perceived stress scale: reliability and validity study in Brazil. *J Health Psychol.* 2010; 15(1):107–114. [PubMed] [Google Scholar]

30. Chaaya M, Osman H, Naassan G, Mahfoud Z.. Validation of the Arabic version of the cohen perceived stress scale (PSS-10) among pregnant and postpartum women. *BMC Psychiatry*. 2010; 10:111. [PMC free article] [PubMed] [Google Scholar]

31. Wongpakaran N, Wongpakaran T.. The Thai version of the PSS-10: An investigation of its psychometric properties. *Biopsychosoc Med*. 2010; 4:6. [PMC free article] [PubMed] [Google Scholar]

32. Almadi T, Cathers I, Hamdan Mansour AM, Chow CM.. An Arabic version of the perceived stress scale: translation and validation study. *Int J Nurs Stud.* 2012; 49(1):84–89. [PubMed] [Google Scholar]

33. Sexton KA, Walker JR, Graff LA, et al.. Evidence of bidirectional associations between perceived stress and symptom activity: A prospective longitudinal investigation in inflammatory bowel disease. *Inflamm Bowel Dis.* 2017; 23(3):473–483. [PubMed] [Google Scholar]

34. Lehman KA, Burns MN, Gagen EC, Mohr DC.. Development of the brief inventory of perceived stress. *J Clin Psychol.* 2012; 68(6):631–644. [PMC free article] [PubMed] [Google Scholar]

35. Endrighi R, Dimond AJ, Waters AJ, et al.. Associations of perceived stress and state anger with symptom burden and functional status in patients with heart failure. *Psychol Health*. 2019; 34(10):1250–1266. [PMC free article] [PubMed] [Google Scholar]

36. Endrighi R, Waters AJ, Gottlieb SS, et al.. Psychological stress and short-term hospitalisations or death in patients with heart failure. *Heart.* 2016; 102(22):1820–1825. [PMC free article] [PubMed] [Google Scholar]

37. Burg MM, Schwartz JE, Kronish IM, et al.. Does stress result in you exercising less? Or does exercising result in you being less stressed? Or is it both? Testing the bi-directional stress-exercise association at the group and person (n of 1) level. *Ann Behav Med.* 2017; 51(6):799–809. [PMC free article] [PubMed] [Google Scholar]

38. Diaz KM, Thanataveerat A, Parsons FE, et al.. The influence of daily stress on sedentary behavior: Group and person (n of 1) level results of a 1-year observational study. *Psychosom Med.* 2018; 80(7):620–627. [PMC free article] [PubMed] [Google Scholar]

39. Schwartz JE, Stone AA.. The analysis of real-time momentary data: A practical guide. In: Stone AA, Shiffman S, Atienza AA, Nebeling L (eds). *The Science of Real-time Data Capture: Self-report in Health Research*. Oxford, Oxford University Press, 2007, pp 76–113. [Google Scholar]

40. Wallston KA, Wallston BS, Smith S, Dobbins CJ.. Perceived control and health. *Curr Psychol Res Rev.* 1987; 6:5–25. [Google Scholar]

41. Nielsen MB, Knardahl S.. Coping strategies: a prospective study of patterns, stability, and relationships with psychological distress. *Scand J Psychol.* 2014; 55(2):142–150. [PubMed] [Google Scholar]

42. Regier NG, Parmelee PA.. The stability of coping strategies in older adults with osteoarthritis and the ability of these strategies to predict changes in depression, disability, and pain. *Aging Ment Health.* 2015; 19(12):1113–1122. [PMC free article] [PubMed] [Google Scholar]

43. Krumholz HM. Post-hospital syndrome – an acquired, transient condition of generalized risk. *N Engl J Med.* 2013; 368(2):100–102. [PMC free article] [PubMed] [Google Scholar]

44. Murray AL, Xiao Z, Zhu X, et al.. Psychometric evaluation of an adapted version of the perceived stress scale for ecological momentary assessment research. *Stress Health.* 2023; doi:10.1002/smi.3229, Epub ahead of print. [PubMed] [Google Scholar]

45. Edmondson D, Shaffer JA, Chaplin WF, Burg MM, Stone AA, Schwartz JE.. Trait anxiety and trait anger measured by ecological momentary assessment and their correspondence with traditional trait questionnaires. *J Res Pers*. 2013; 47(6): doi: 10.1016/j.jrp.2013.08.005. [PMC free article] [PubMed] [CrossRef] [Google Scholar]