

Chapman University

## Chapman University Digital Commons

---

Psychology Faculty Articles and Research

Psychology

---

11-30-2021

### The Subcomponents of Affect Scale (SAS): Validating a Widely Used Affect Scale

Brooke N. Jenkins

*Chapman University*, [bjenkins@chapman.edu](mailto:bjenkins@chapman.edu)

Marie P. Cross

*Pennsylvania State University*

Candice Donaldson

*Chapman University*

Sarah D. Pressman

*University of California, Irvine*, [pressman@uci.edu](mailto:pressman@uci.edu)

Michelle A. Fortier

*University of California, Irvine*, [mfortier@hs.uci.edu](mailto:mfortier@hs.uci.edu)

*See next page for additional authors*

Follow this and additional works at: [https://digitalcommons.chapman.edu/psychology\\_articles](https://digitalcommons.chapman.edu/psychology_articles)



Part of the [Health Psychology Commons](#), [Other Psychology Commons](#), and the [Pain Management Commons](#)

---

#### Recommended Citation

Jenkins, B. N., Cross, M. P., Donaldson, C. D., Pressman, S. D., Fortier, M. A., Kain, Z. N., Cohen, S., Martin, L. T., & Farkas, G. (2021). The subcomponents of affect scale (SAS): Validating a widely used affect scale. *Psychology & Health*. <https://doi.org/10.1080/08870446.2021.2000612>

This Article is brought to you for free and open access by the Psychology at Chapman University Digital Commons. It has been accepted for inclusion in Psychology Faculty Articles and Research by an authorized administrator of Chapman University Digital Commons. For more information, please contact [laughtin@chapman.edu](mailto:laughtin@chapman.edu).

---

# The Subcomponents of Affect Scale (SAS): Validating a Widely Used Affect Scale

## Comments

This is an Accepted Manuscript version of the following article, accepted for publication in Psychology & Health in 2021. <https://doi.org/10.1080/08870446.2021.2000612>

It is deposited under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial 4.0 License](https://creativecommons.org/licenses/by-nc/4.0/)

## Copyright

Taylor & Francis

## Authors

Brooke N. Jenkins, Marie P. Cross, Candice Donaldson, Sarah D. Pressman, Michelle A. Fortier, Zeev N. Kain, Sheldon Cohen, Logan T. Martin, and George Farkas

**The Subcomponents of Affect Scale (SAS): Validating a Widely Used Affect Scale**

Brooke N. Jenkins, Ph.D.<sup>a,b,c</sup>, Marie P. Cross, Ph.D.<sup>d</sup>, Candice Donaldson, Ph.D.<sup>a,b</sup>, Sarah D. Pressman, Ph.D.<sup>e</sup>, Michelle A. Fortier, Ph.D.<sup>b,c,e,f,g</sup>, Zeev N. Kain, M.D., M.B.A.<sup>b,c,h,i</sup>, Sheldon Cohen<sup>j</sup>, Ph.D., Logan T. Martin, M.A.<sup>e</sup>, & George Farkas, Ph.D.<sup>k</sup>

<sup>a</sup> Chapman University, Department of Psychology

<sup>b</sup> University of California, Irvine, Center on Stress & Health

<sup>c</sup> University of California, Irvine, Department of Anesthesiology and Perioperative Care

<sup>d</sup> Pennsylvania State University, Department of Biobehavioral Health

<sup>e</sup> University of California, Irvine, Department of Psychological Science

<sup>f</sup> University of California, Irvine, Sue & Bill Gross School of Nursing

<sup>g</sup> CHOC Children's, Department of Pediatric Psychology

<sup>h</sup> CHOC Children's, Department of Pediatrics

<sup>i</sup> Yale University School of Medicine, Child Study Center

<sup>j</sup> Carnegie Mellon University, Department of Psychology

<sup>k</sup> University of California, Irvine, School of Education

Corresponding author contact information: Brooke N. Jenkins, PhD, (bjenkins@chapman.edu)  
Crean College of Health and Behavioral Sciences, Chapman University, One University Drive,  
Orange, CA 92866

Author note: Analyses in this manuscript were presented at the American Psychosomatic Society  
2019 Annual meeting.

### Abstract

Objective: There is a need for a brief affect scale that also encompasses different components of affect relevant for researchers interested in physiological and health outcomes. The Subcomponents of Affect Scale (SAS) meets this need. This 18-item scale has nine positive and nine negative affect items encompassing six subscales (calm, well-being, vigor, depression, anxiety, anger). Previous research using the SAS has demonstrated its predictive validity, but no work has tested its subscale structure or longitudinal validity. Design: Data from the Common Cold Project in which individuals (N = 610) completed the SAS over the course of seven days was used. Results: Confirmatory factor analysis demonstrated the reliability of the subscale structure of the SAS across seven days (positive affect subscale structure: CFIs  $\geq 0.98$ ; negative affect subscale structure: CFIs  $\geq 0.94$  with day 6 CFI = 0.91) and tests of factorial invariance showed the scale is valid to use over time. Conclusions: These results confirm the psychometric validity of the subscale structure of the SAS and imply that the subscales can be used longitudinally, allowing for its use in health research as well as non-health research that can benefit from its subscale structure and longitudinal capabilities.

*Keywords:* affect scale, emotion scale, Subcomponents of Affect Scale, health, measurement

Word count: 7,657

### **The Subcomponents of Affect Scale (SAS): Validating a Widely Used Affect Scale**

Findings from health science overwhelmingly support that affect is related to our physical and mental health (e.g., DeSteno et al., 2013; Pressman & Cohen, 2005; Watson, 1988). Although affect can be measured through a number of techniques, the most popular method is self-report. This involves asking participants how frequently or intensely they feel specific emotions (e.g., happiness, anger), then aggregating items to create positive and negative affect subscales.

Although there are a number of valid and reliable affect scales (e.g., the Profile of Mood States [POMS]; McNair et al., 1971, the Positive and Negative Affect Schedule [PANAS]; Watson et al., 1988), most have limitations when it comes to use in health research. First, health researchers often need to rely on scales that are brief due to the time intensive nature of their study procedures. Participants often have to complete invasive medical/physical examinations (e.g., Kubzansky et al., 2006; Shirom et al., 2010), long questionnaires and symptom checklists assessing other health-relevant variables (e.g., Glejsted Ingstrup et al., 2012; Middelboe et al., 1992), and multiple surveys over time (e.g., Czyz et al., 2019; Mustanski, 2007). For epidemiologic investigations that often contain samples with thousands of participants, the cost of adding just one item to a survey can be substantial (e.g., Montana Department of Public Health and Human Services, 2018). Further, participants often belong to sensitive populations with limited availability. A second limitation is that many instruments are not designed to capture affective subcomponents that are relevant to health outcomes of interest. There is increasing evidence of distinct emotions and affective arousal levels differentially influencing health outcomes (e.g., Kubzansky et al., 2006; Pressman & Cross, 2018; Suls & Bunde, 2005). Hence, a scale representing the key emotions used in health research provides a valuable tool.

A measure that meets the needs of health researchers and has been utilized frequently is the Subcomponents of Affect Scale<sup>1</sup> (SAS; Cohen et al., 2003). The SAS was originally constructed by Cohen and colleagues (2003) based on a factor analysis of 65 emotion adjectives (Usala & Hertzog, 1989), 38 of which were from the POMS (McNair et al., 1971) and 27 of which were from the Pregnancy Mood Checklist (Lebo & Nesselroade, 1978). Cohen et al. (2003) selected 15 items (and added three more), representing a range of affective categories central to theories and empirical work on health. The highest order categories were positive and negative affect. These were further divided into three subscales within each valence: calm, well-being, and vigor under positive affect and depression, anxiety, and anger under negative affect. Both valences included two higher arousal (high body activation; vigor and well-being; anxiety and anger) subscales and one lower arousal (low body activation; calm; depression) subscale.

The SAS has been used in a variety of health studies (e.g., Jenkins et al., 2018) that support its predictive validity in connection with a number of health outcomes. Investigations using the overall negative affect scale find that higher negative affect is associated with greater flu/cold symptoms (Cohen et al., 2003; Cohen et al., 2006), higher disease severity in individuals with Type 2 diabetes (Sultan & Fisher, 2010), and lower sleep quality (Lillis et al., 2018). Similarly, studies using the overall positive affect scale find associations in the expected directions, with greater positive affect being associated with better sleep quality (Lillis et al., 2018) and more physical activity (Poole et al., 2011). Further, studies utilizing the subscales find more nuanced effects. For example, while work has shown that higher scores on positive affect are associated with greater physical activity (Poole et al., 2011), it seems to be the positive affect subscale of calm that most drives this association (Aggio et al., 2017). Similarly, previous work

---

<sup>1</sup>The scale is not named in past publications and is referred to as the State Adjective Questionnaire (18-item version) on the Common Cold Project website. In order to be more descriptive as to the nature of the scale, we have renamed it the Subcomponents of Affect Scale.

found no association between scores on negative affect and physical activity (Poole et al., 2011), but when researchers divided negative affect into subscales, lower depression was associated with greater physical activity (Aggio et al., 2017). Interestingly, subscales of the SAS within a single valence often demonstrate differential effects. For example, individuals with higher vigor and/or well-being were less likely to develop an objectively measured cold, but individuals lower in calm reported more symptoms (Cohen et al., 2003; Cohen et al., 2006). Taken together, this body of work demonstrates the predictive validity of the SAS and suggests that the SAS and its subscales uncover nuanced affect and health associations. However, no investigation has tested the subscale structure of the SAS to confirm that the subscales represent unique constructs. Additionally, while the SAS has been used in longitudinal investigations (e.g., Cohen et al., 2003), no study has confirmed that the subscales have the same structure over time, an important requirement for testing how affect impacts health longitudinally.

Therefore, the first goal of this study is to test the subscale structure of the SAS using confirmatory factor analysis. A confirmatory factor analysis of the subscale structure would provide evidence toward the construct validity of the SAS. Confirmatory factor analysis is an ideal approach as this structural equation modeling framework allows for *a priori* predictions about the measurement model based on theory to be tested. Therefore, we can specify the subscale structure of the SAS to test how closely the specified model accounts for the data. This approach is preferred over other common techniques such as exploratory factor analysis in which no *a priori* hypotheses can be evaluated. A second goal of this paper is to determine the stability of the subscale structure of the SAS over time. One of the strengths of the SAS is that its brevity provides the opportunity for researchers to use the measure in longitudinal investigations. Does the structure hold across days? Is the strength of the association of each item with its factor

similar (or equal) across days? These questions can be answered by using confirmatory factor analysis to test the factorial invariance of the SAS. Confirming the subscale structure and testing the factorial invariance of the SAS will supply needed evidence of its validity in affective-health science research, and provide researchers in several fields with a highly useful scale.

### **Methods**

Data from the Common Cold Project (data are openly available at [www.commoncoldproject.com](http://www.commoncoldproject.com); grant number NCCIH AT006694; Laboratory for the Study of Stress, Immunity, and Disease, 2016) were accessed to examine the subscale structure of the SAS and test its validity longitudinally. Specifically, data from 276 individuals who participated in the Pittsburgh Cold Study 1 (conducted from 1993 to 1996) and 334 individuals who participated in the Pittsburgh Cold Study 2 (conducted from 1997 to 2001) in which affect data from the SAS were collected for seven consecutive days from a total of 610 individuals were used. The 610 participants in this data set had a mean age of 28.98 years ( $SD = 9.82$ ; range 18 to 55) and 53% were female (see Table 1 for more demographic information).

### **Procedure**

Participants for both studies were recruited through newspaper advertisements in the Pittsburgh, Pennsylvania area. Due to the intensive nature of data collection, participants were compensated with \$800. The Pittsburgh Cold Studies 1 and 2 had parallel designs. They were both viral challenge investigations in which participants were quarantined for seven days (six nights) and were exposed to a cold virus (rhinovirus 39 ( $N = 147$ ) or rhinovirus 21 ( $N = 129$ ) in Pittsburgh Cold Study 1; rhinovirus 39 ( $N = 228$ ) or rhinovirus 23 ( $N = 106$ ) in Pittsburgh Cold Study 2) in the evening before the second night of quarantine (i.e., Day 2). In both studies,



participants completed the SAS the evening before the first night of quarantine<sup>2</sup> (referred to from now on as Day 1) and then in the evening on the next 6 days of quarantine (referred to as Days 2 through 7). Thus, this study design allows for a strict test of the SAS in a health-relevant context, providing an excellent opportunity to test the SAS over time. Data from both studies were collected by the Laboratory for the Study of Stress, Immunity, and Disease at Carnegie Mellon University under the directorship of Sheldon Cohen, PhD. The Common Cold Project ([www.commoncoldproject.com](http://www.commoncoldproject.com)) was designed to meticulously combine data from these two studies to allow for analysis of common variables aggregated across the studies (as done previously; e.g., Janicki-Deverts et al., 2016; Prather et al., 2017; Sneed et al., 2012). Data are publicly available on the Common Cold Project website. All participants completed informed consent.

## **Measures**

### ***SAS***

The Subcomponents of Affect Scale (SAS) is not named in past publications and is referred to as the State Adjective Questionnaire (18-item version) on the Common Cold Project website. In order to be more descriptive as to the nature of the scale, we have renamed it the Subcomponents of Affect Scale. The SAS is composed of 18 items, nine positive and nine negative affect adjectives. The nine positive affect adjectives are divided across three subscales: calm (items: calm, at ease, relaxed), well-being (items: happy, cheerful, pleased), and vigor (items: full of pep, lively, energetic). The nine negative affect adjectives are similarly divided across three subscales: depression (items: sad, unhappy, depressed), anxiety (items: on edge, tense, nervous), and anger (items: hostile, angry, resentful). Participants are asked to rate on a scale from 0 (not at all accurate) to 4 (extremely accurate) the extent to which each adjective

---

<sup>2</sup>48 individuals from the Pittsburgh Cold Study 1 were not given the SAS on Day 1.

accurately describes how they felt within a certain time range (e.g., past 24 hours, past hour, current moment). Data used in the present manuscript were collected at 5:30pm each night during quarantine using paper and pencil; participants were instructed to reflect on emotions over the past day with response options ranging from “haven’t felt that way at all since getting up/not at all accurate” (0) to “felt that way a lot since getting up/extremely accurate” (4).

For the current data, the estimated within-day reliability coefficients<sup>3</sup> for the set of nine positive affect adjectives ranged from 0.92 to 0.93 for each of the seven days ( $\omega_{\text{Mean}} = 0.93$ ). The within-day reliability coefficients for the set of nine negative affect adjectives ranged from 0.82 to 0.87 for each of the seven days ( $\omega_{\text{Mean}} = 0.84$ ). The within-day reliability coefficients across the seven days for the subscales had the following: vigor ( $\omega_{\text{Mean}} = 0.93$ ,  $\omega_{\text{Range}} = 0.91, 0.94$ ), well-being ( $\omega_{\text{Mean}} = 0.90$ ,  $\omega_{\text{Range}} = 0.87, 0.92$ ), calm ( $\omega_{\text{Mean}} = 0.85$ ,  $\omega_{\text{Range}} = 0.82, 0.87$ ), depression ( $\omega_{\text{Mean}} = 0.82$ ,  $\omega_{\text{Range}} = 0.70, 0.90$ ), anger ( $\omega_{\text{Mean}} = 0.82$ ,  $\omega_{\text{Range}} = 0.77, 0.87$ ), and anxiety ( $\omega_{\text{Mean}} = 0.63$ ,  $\omega_{\text{Range}} = 0.51, 0.74$ ).

## Statistical Analyses

### *Subscale Structure*

Confirmatory factor analysis was used to examine the subscale structure of the SAS on each of the seven days of data collection using Stata 15 (StataCorp, 2017). First, a positive affect model (Positive Affect Model 1) was built which included one latent variable of positive affect and each of the nine positive affect items as endogenous observed variables. Second, a model with three latent variables (calm, well-being, and vigor) and their corresponding subscale items serving as endogenous observed variables was tested (Model 2). This model was iteratively

---

<sup>3</sup>Because the results of our confirmatory factor analyses detected unequal factor loadings between items and some item error covariance (see results section), we chose to use the reliability estimator McDonald’s Omega ( $\omega$ ) (McDonald, 1999) instead of Cronbach’s alpha. McDonald’s Omega is a robust estimator of reliability under these conditions (see supplemental online material for detailed description of Omega). In contrast, Cronbach’s alpha may either systematically inflate or deflate reliability estimates (Zinbarg et al., 2005).

improved using modification indices to help guide theoretically justified alterations (Acock, 2013; see supplemental online material for detailed description of modification indices), resulting in Positive Affect Models 2a, 2b, 2c, and sometimes 2d. Following theory, we selected changes that specified correlations of measurement residuals within subscales (e.g., correlating the measurement residuals of calm and relaxed) over loading subscale items onto other latent variables (e.g., loading the item at ease onto the latent variable vigor), even when the latter might have had a higher modification index. The models with subscales also allowed us to examine how correlated latent variables were (e.g., correlation between calm and well-being).

Next, this same process was conducted for the negative affect items (i.e., one overall negative affect latent variable [Negative Affect Model 1] was followed by a model including the three subscale latent variables [Negative Affect Model 2] with iterative improvements [Negative Affect Models 2a, 2b, ...]). Lastly, the final positive affect model and negative affect model for each day were merged into one confirmatory factor analysis model (Positive and Negative Affect Model 3) and iterative adjustments were made using the same modification index strategy described above (sometimes resulting in a Positive and Negative Affect Model 3a). Of note, all models testing the subscale structure are in standardized form for ease of interpretation. Additionally, while improving upon the model fit of both the Positive Affect Model 1 (i.e., positive affect model with only one latent variable) and Negative Affect Model 1 (i.e., negative affect model with only one latent variable) was not our main goal, we did examine modification indices to determine if improvements in the models could be made given that researchers sometimes use the overall positive and/or negative affect scores.

Model fit was evaluated using chi-square, with non-significant values reflecting good fit (see supplemental material for detailed description of chi-square). Since chi-square is overly

powered with large samples, we also used goodness of fit tests as recommended by Kline (2015) with the following cutoffs for guidance in model selection: CFI > 0.90 (Bentler, 1990), RMSEA < .08 (Browne & Cudeck, 1993), lower bound of 90% confidence interval of RMSEA below 0.05 and upper bound below 0.10, SRMR < .08, and coefficient of determination (CD) approaching 1. Finally, AIC (Akaike, 1974) and BIC (Schwarz, 1978) were also used to compare nested models, with smaller values reflecting better fit.

Throughout our results, we report 95% confidence intervals when appropriate. Missing data were low and varied across days: Day 1 had 10% missing data, Days 2 through 5 each had 1% missing data, and Days 6 and 7 each had 2% missing data. Confirmatory factor analysis was conducted using the sem command in Stata which follows listwise deletion. Sample sizes are reported throughout our analyses. Sample size requirements in confirmatory factor analysis increase when the number of factors tested is higher, there are fewer endogenous observed variables, the strength of factor loadings is lower, and correlations between factors are weaker. For models in which we had 6 latent variables with 3 overserved variables each (for a total of 18 variables) and factor loadings of 0.5, we conservatively estimated our required sample size at 460 based on simulation analysis in previous work (Wolf et al., 2013). Sample sizes in the analyzed data set varied among days, with the smallest sample size at 507. This provided us with sufficient power across all days. As our models were not predicting outcomes, effect sizes were not produced.

### ***Factorial Invariance***

When testing whether the general subscale structure of the SAS holds across time (i.e., testing for factorial invariance), we elected to test a model that reflected the most common model form across days but still had strong goodness of fit test values. Using this model, we first tested

for configural invariance by specifying a model in which the same group of items loaded on each of the latent variables (same configuration/form) across the days (see supplemental online material for detailed description of factorial invariance). Numerical values of the loadings can be different across days, but if all were significant, we achieved configural invariance. Next, to test for metric invariance, we constrained the factor loadings for each item across days to be equal to one another, which would allow us to conclude that the meaning of the emotion adjectives was invariant over time. Finally, to test scalar invariance, we additionally restricted the intercepts of each of the items on their latent variable to be equal across days. At each step, the model was compared to the previous model using the likelihood ratio chi-squared test. It is important to note that metric invariance is often the model most researchers are satisfied with (Acock, 2013). Scalar invariance is rarely achieved in real data; however, we test it here for completeness.

## Results

### Subscale Structure

#### *Positive Affect*

The positive affect items were first entered into a confirmatory factor analysis model with only the latent variable positive affect for each of the seven days. The initial fit of this model did not reach normative cutoffs for each of the days (Day 1:  $\chi^2(27) = 582.58, p < .001, CFI = 0.84$ ; see Table 2 Positive Affect Model 1; for all other days see Appendix A Tables 1 through 6). Upon examining the modification indices, covariances among the error terms of scale items could be included in the model to obtain CFIs above 0.95 for each of the seven days.

We next ran a model including the subscales. For the first day of data, the model had poor fit,  $\chi^2(27) = 884.94, p < .001, CFI = 0.76$  (see Table 2 Positive Affect Model 2; all other days similarly had poor fit [see Appendix A Tables 1 through 6]). However, after examining the

modification indices and iteratively including covariances among the latent subscales, all CFIs were at or above 0.98, indicating strong fit and other goodness of fit tests similarly met their recommended cutoffs (see Table 2 and Appendix A [Tables 1 through 6] Positive Affect Models 2a, 2b, and 2c). The iterative changes, as indicated by the largest modification index, were identical across days and had the following order: covariance between latent subscales well-being and vigor (Positive Affect Model 2a), covariance between subscales well-being and calm (Positive Affect Model 2b), covariance between subscales vigor and calm (Positive Affect Model 2c). Finally, modification indices of Positive Affect Model 2c suggested including covariances among error terms of the scale items calm and relaxed (days 1, 3, 4, 5, 6) and cheerful and pleased (day 7; see Figure 1; see Positive Affect Model 2d in Table 2 and Appendix A Tables 1 through 6). These additions improved the model fit. Also, the correlations among all latent variables were significant (see Figure 1) with a common pattern across all days such that the strongest correlation was always between well-being and vigor (e.g., Day 1 covariance = 0.89,  $p < .001$ , 95% CI [0.86, 0.91]), followed by the correlation among well-being and calm (e.g., Day 1 covariance = 0.73,  $p < .001$ , 95% CI [0.67, 0.79]), while the covariance between vigor and calm was the least strongly correlated (e.g., Day 1 covariance = 0.64,  $p < .001$ , 95% CI [0.58, 0.71]).

### *Negative Affect*

The negative affect items were first entered into a confirmatory factor analysis model with only the latent variable negative affect for each of the seven days. The initial fit of this model did not reach normative cutoffs for each of the days (Day 1:  $\chi^2(27) = 463.41$ ,  $p < .001$ , CFI = 0.80; see Table 2 Negative Affect Model 1; for all other days see Appendix A Tables 1

through 6). Upon examining the modification indices, covariances among the error terms of scale items could be included in the model to obtain CFIs at or above 0.94.

We next ran a model including the subscales. For the first day of data, the model had poor fit,  $\chi^2(27) = 483.31$ ,  $p < .001$ , CFI = 0.79 (see Table 2 Negative Affect Model 2; all other days similarly had poor fit [see Appendix A Tables 1 through 6]). However, after examining the modification indices and iteratively including covariances among the latent subscales, all CFIs were at or above 0.91, indicating strong fit and other goodness of fit tests similarly met their recommended cutoffs (see Table 2 and Appendix A Tables 1 through 6 Negative Affect Models 2a, 2b, and 2c). While the order of iterative changes of adding the covariances was not the same across days, the final Negative Affect Model 2c for all seven days always ended with an identical pattern which included covariances among the three latent subscales.

Finally, modification indices of Negative Affect Model 2c suggested including covariances among error terms of scale items (see Figure 2; see Table 2 Negative Affect Model 2d; see Appendix A Tables 1 through 6 Negative Affect Model 2d and sometimes 2e and 2f). These additions improved model fit. Also, the correlations among all latent variables were significant (e.g., Day 1 covariance between latent subscales depression and anxiety = 0.68,  $p < .001$ , 95% CI[0.62, 0.75]; Day 1 covariance between latent subscales depression and anger = 0.68,  $p < .001$ , 95% CI[0.61, 0.74]; Day 1 covariance between latent subscales anxiety and anger = 0.62,  $p < .001$ , 95% CI[0.54, 0.71]).

### ***Positive and Negative Affect Model***

Last, we combined the separate positive and negative affect final models described above to create a full Positive and Negative Affect Model. For the first day of data, the model had a strong fit,  $\chi^2(127) = 340.43$ ,  $p < .001$ , CFI = 0.96 (see Figure 3; see Table 2 Positive and

Negative Affect Model 3). Similarly, the CFIs for all the other days for the Positive and Negative Affect Model 3 were at or above 0.96 and other goodness of fit tests similarly met their recommended cutoffs (see Appendix A Tables 1 through 6 Positive and Negative Affect Model 3). Upon examining the modification indices of Model 3, only two days had modification indices. Modification indices for Day 1 indicated that covariances among the latent subscales calm and anxiety would improve the model fit,  $\chi^2(126) = 300.09, p < .001, CFI = 0.97$  (see Table 2 Positive and Negative Affect Model 3a), and modification indices for Day 7 indicated that covariances of error terms of items sad and depressed could be included in the model,  $\chi^2(126) = 396.06, p < .001, CFI = 0.96$ .

While combining the positive and negative affect models into one overall model in this section did have warrant for testing the validity of the SAS in its entirety, we should note that combining them did not drastically improve the model fit. For example, the CFIs for the Positive and Negative Affect Model 3 were 0.96 or 0.97 for each of the seven days. By comparison, the CFIs for the Positive Affect Model 2c were 0.98 or 0.99 for each of the seven days. In essence, the CFIs actually dropped when the positive affect model was combined with the negative affect model. Still, even the Negative Affect Model 2c CFIs were in the 0.94 to 0.97 range with the exception of day 6, which had a CFI of 0.91. Therefore, in testing the factorial invariance of the SAS (see Measurement Model Over Time section below), we test the positive affect model and the negative affect model separately and report the combined model in the Appendix for completeness (Appendix A Table 7). Given that the Positive Affect Model 2c and Negative Affect Model 2c were common across all days and had moderate to strong goodness of fit indices, we selected these models as the final ones to use when testing the factorial invariance across days.



## Measurement Model Over Time

### *Positive Affect*

The configural model resulted in fairly good fit and all the loadings were significant and of the same form,  $\chi^2(1,617) = 3,829.09, p < .001, CFI = 0.94$  (see Table 3). The metric model also resulted in a fairly good fit,  $\chi^2(1,653) = 3,879.35, p < .001, CFI = 0.94$  (see Table 3). The chi-squared difference test between the configural and metric models was not significant,  $\chi^2(36) = 50.52, p = 0.058$ , suggesting that the metric model did not fit significantly worse than the configural model. However, the scalar model fit was significantly worse than the metric model (see Table 3; scalar vs. metric model:  $\chi^2(54) = 313.95, p < .001$ ). Therefore, the model achieved metric invariance, allowing us to conclude that the factor loadings for the subscale structure of the positive affect subscale are equal over time. As mentioned previously, metric invariance is often what most researchers are satisfied with (Acock, 2013). Scalar invariance is often too restrictive, but we tested it here for completeness.

### *Negative Affect*

While the model specifying configural invariance did not meet the 0.90 cutoff for CFI, the RMSEA value, RMSEA lower bound, and coefficient of determination did meet recommended levels (see Table 4). Furthermore, all the item loadings were significant and of the same form, suggesting that the scale did achieve configural invariance. Again, while the model specifying metric invariance did not meet the 0.90 cutoff for CFI, the RMSEA value, RMSEA lower bound, and coefficient of determination did meet recommended levels (see Table 4). However, the chi-squared difference test suggested the metric model fit significantly worse than the configural model,  $\chi^2(36) = 192.68, p < .001$ . Thus, we can only conclude that the negative affect subscale has the same configuration over time. We also tested the scalar invariance of the

negative affect subscale structure, but this model was a significantly worse fit than the configural invariance model (see Table 4).

### **Discussion**

The current study aimed to confirm the subscale dimensionality of the SAS (Cohen et al., 2003) and validate its psychometric properties for use in affect-health research. Confirmatory factor analysis supported the three-factor structure of positive affect as being comprised of vigor, well-being, and calm, and the three-factor dimensionality of negative affect, as represented by the subscales of anger, anxiety, and depression. Tests of measurement invariance across a seven-day time interval supported the validity of the measure for examining changes in affect over time. Further, the SAS had acceptable within-day reliability for the overall scales and subscales. The reliability was particularly strong for positive affect.

The three-factor structure of positive and negative affect has important implications. Although a clear link between general affect and health has been established (Skaff et al., 2009), researchers have advocated for the importance of considering how discrete emotions are differentially related to health outcomes (Consedine & Moskowitz, 2007; Suls & Bunde, 2005). With regard to positive emotion, the issue of conceptualizing affect arousal has received considerable attention (Pressman & Cross, 2018; Pressman et al., 2019). Specifically, emotional experience can take the form of high arousal (e.g., vigor), mid-arousal (e.g., well-being), or low arousal (e.g., calm). Good evidence suggests that arousal level may differentially predict health-relevant outcomes (e.g., Pressman et al., 2017). In some investigations, vigor (high arousal positive affect) has been shown to have beneficial health effects such as increased longevity (Pressman & Cohen, 2012) and lower rates of illness (Cohen et al., 2003; Cohen et al., 2006), but other studies suggest that high arousal positive emotionality is associated with risk of

cardiovascular dysfunction (Armon et al., 2014). As such, correlations between positive affect subscales were evaluated in the current investigation to give light to possible differences in levels of affective arousal. In line with the theory that vigor may be high arousal and calm may be low arousal, with well-being being more mid-arousal, the correlation between calm and vigor (covariance = .64; see Figure 1) was much weaker than that of the association between well-being and the other subscales. This demonstrates that well-being may be more “in the middle” and, thus, more strongly associated with the other arousal levels. However, well-being was consistently more strongly correlated with vigor (covariance = .89) than with calm (covariance = .73), suggesting that well-being may be a little more distinct from the concept of calm as compared to the concept of vigor.

Of note, we could not infer affective arousal differences in the subscales of the negative affect scale as cleanly because the correlations between measures of negative affect were similar across all subscales (ranging from .62 to .68; see Figure 2) across all days. This does seem surprising, as anger and anxiety are typically conceptualized as arousing states with depression being characterized as a lower arousal state. Nevertheless, the subscales of negative affect still represent important health-relevant domains. For example, anger is strongly linked with the development of angina (i.e., chest pain), whereas anxiety is related to the development of both fatal and nonfatal myocardial infarction (Kubzansky et al., 2006). Thus, each specific negative affect subscale is likely to continue to predict different health patterns and outcomes as seen in previous uses of the SAS (e.g., Aggio et al., 2017).

While the subscales do provide utility, some researchers may still opt to use the aggregate positive or negative affect subscales as has been done in past research (e.g., Cohen et al., 2003; Cohen et al., 2006; Jenkins et al., 2018; Lillis et al., 2018; Poole et al., 2011; Sultan & Fisher,

2010). The findings suggest that the items from the SAS do load on to two unique factors (negative and positive affect). For completeness, the overall positive and negative affect subscales were examined and the model fit of aggregate positive and negative affect was poor prior to covarying error terms. As a result, researchers might consider employing a structural equation modeling framework so as to allow error terms to be correlated.

Temporal equality is an important assumption for longitudinal research in the study of affective health science (Meredith, 1993; Putnick & Bornstein, 2016), as health researchers are often examining change in health and health-relevant constructs over time. In other words, scales that have temporal equality indicate that the measure is being interpreted similarly by participants over time. In the case of the SAS, this would mean that the meaning of each adjective would hold constant over time. While the participants in the current data set were expected to have changes in levels of affect given that they were subjected to a health-relevant and stressful situation, findings from the factorial invariance tests supported that participants still interpreted items the same over time even as they potentially got sick, distressed, or homesick as the seven-day quarantine period went on. This analysis demonstrates that both the positive and negative affect subscales could be used over time in a diverse sample. Positive affect exhibited configural and metric invariance, signifying that both the factor structure and loadings of each positive emotion adjective were equivalent throughout the seven days of measurement. Negative affect achieved configural invariance, indicating that the factor structure of the negative affect subscales was the same longitudinally. However, the strength of the factor loadings of the negative affect scale were time variant. In other words, while we can assume the same negative affect theoretical constructs are being measured across days, the relative importance of each emotion adjective over time may not be the same.

The lack of metric equivalence for negative affect is not surprising, as several studies have shown violations of temporal measurement invariance when examining changes in symptoms of depression (Fried et al., 2016; Uher et al., 2008; Wetherwell et al., 2001). As such, changes in level of negative emotionality might influence the participant interpretation of each emotion adjective. Further, data used in the present analysis were collected while participants were quarantined in a hotel and exposed to a cold virus. This context provides a strict, if not overly conservative, assessment of the scale over time in an extreme environment that is of interest to health psychologists.

There are limitations of the current study. Participants were instructed to retrospectively estimate their emotions over the past day. Although retrospective self-report is the standard measurement paradigm for examining affect (Cohen et al., 2003; Cohen et al., 2006; Lac & Donaldson, 2018; Watson et al., 1988), this type of assessment can introduce recall bias. Future investigations might employ smartphones and other electronic devices to record self-reports of emotionality in real time (e.g., Sherman et al., 2015). It is also important to note that affect was not manipulated in this study; thus, we could not assess whether the SAS subscales were sensitive to experimental manipulations of mood. The sample of this study also consisted only of adults, with the majority of participants identifying as White (74%) and African American (23%), limiting the overall generalizability of the results to populations of other racial ethnicities. That said, this study also included diversity in age, education, and employment, improving generalizability in some ways.

Finally, the attentive reader will notice the weaker reliability of the anxiety subscale. A similar result has been observed in past studies of the short form of the POMS (Curran et al., 1995, Table 1). We recommend researchers particularly interested in anxiety consider increasing

the number of anxiety adjectives (e.g., Watson & Clark, 1994), observations per day, or sample size to achieve a higher reliability. Despite this, all of the overall and other subscale scales demonstrated acceptable to strong reliability in our data. Given that the intended use of the SAS is for repeated measures health psychology research, where the researcher's choice is often to use a very short measure or none at all, these reliabilities (including that of anxiety) seem sufficient.

The influential role of global affect, as represented as aggregate emotional experiences comprised of specific emotional states, has been central in the study of health (Consedine & Moskowitz, 2007). The SAS (Cohen et al., 2003; Usala & Hertzog, 1989) is one measurement tool that has been argued to capture both positive and negative affect with different subscales of each affective valence. Results of a confirmatory factor analysis supported the three-factor dimensionality of both positive and negative affect, with positive affect being composed of calm, well-being, and vigor, and negative affect being comprised of anger, anxiety, and depression. Further, analyses demonstrated support for the validity of studying changes in the affect subscales over time. Findings confirm the structure of the SAS and imply that the subscales can be used as a valid longitudinal tool in the study of affective health science.

Acknowledgments: We would like to acknowledge that the data in this article were collected by the Laboratory for the Study of Stress, Immunity, and Disease at Carnegie Mellon University under the directorship of Sheldon Cohen, PhD, and were accessed via the Common Cold Project website ([www.commoncoldproject.com](http://www.commoncoldproject.com)). We would also like to acknowledge the grant from the National Center for Complementary and Integrative Health (NCCIH AT006694) that supported the Common Cold Project as well as the grant from the National Institute of Mental Health (NIMH MH50429) that provided primary funding for the research studies whose data are used in this article and the grant from the National Institutes of Health to the University of Pittsburgh Medical Center General Clinic Research Center (NIH 5M01 RR00056) that provided secondary support for data collection.

Disclosure of Interest: The authors report no conflict of interest. Grant support included funding from the National Center for Complementary and Integrative Health (NCCIH AT006694), the National Institute of Mental Health (NIMH MH50429), and the National Institutes of Health to the University of Pittsburgh Medical Center General Clinic Research Center (NIH 5M01 RR00056).

### References

- Acock, A. C. (2013). *Discovering Structural Equation Modeling Using Stata, Revised Edition*. College Station, TX: Stata Press.
- Aggio, D., Wallace, K., Boreham, N., Shankar, A., Steptoe, A., & Hamer, M. (2017). Objectively measured daily physical activity and postural changes as related to positive and negative affect using ambulatory monitoring assessments. *Psychosomatic Medicine, 79*, 792–797. <https://doi.org/10.1097/PSY.0000000000000485>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716–723.
- Armon, G., Melamed, S., Berliner, S., & Shapira, I. (2014). High arousal and low arousal work-related positive affects and basal cardiovascular activity. *Journal of Positive Psychology, 9*, 146–154. <https://doi.org/10.1080/17439760.2013.848375>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Browne, M. W., & Cudeck, R. (1993). Alternative Ways of Assessing Model Fit. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 136–162). Newbury Park, CA: Sage.
- Cohen, S., Alper, C. M., Doyle, W. J., Treanor, J. J., & Turner, R. B. (2006). Positive emotional style predicts resistance to illness after experimental exposure to rhinovirus or influenza A virus. *Psychosomatic Medicine, 68*, 809–815. <https://doi.org/10.1097/01.psy.0000245867.92364.3c>
- Cohen, S., Doyle, W. J., Turner, R., Alper, C. M., & Skoner, D. P. (2003). Emotional style and susceptibility to the common cold. *Psychosomatic Medicine, 65*, 652–657.



<https://doi.org/10.1111/1467-9280.01452>

Cohen, S., & Pressman, S. D. (2006). Positive affect and health. *Current Directions in Psychological Science*, *15*, 122-125. <https://doi.org/10.1111/j.0963-7214.2006.00420.x>

Consedine, N. S., & Moskowitz, J. T. (2007). The role of discrete emotions in health outcomes: A critical review. *Applied and Preventive Psychology*, *12*, 59–75.

<https://doi.org/10.1016/j.appsy.2007.09.001>

Curran, S. L., Andrykowski, M. A., & Studts, J. L. (1995). Short Form of the Profile of Mood States (POMS-SF): Psychometric information. *Psychological Assessment*, *7*(1), 80-83.

<https://doi.org/10.1037/1040-3590.7.1.80>

Czyz, E. K., Horwitz, A. G., Arango, A., & King, C. A. (2019). Short-term change and prediction of suicidal ideation among adolescents: A daily diary study following psychiatric hospitalization. *The Journal of Child Psychology and Psychiatry*, *60*, 732-741.

doi:10.1111/jcpp.12974

DeSteno, D., Gross, J. J., & Kubzansky, L. (2013). Affective science and health: The importance of emotion and emotion regulation. *Health Psychology*, *32*, 474–486.

<https://doi.org/10.1037/a0030259>

Fried, E. I., van Barkulo, C. D., Epskamp, S., Schoevers, R. A., Tuerlinckx, F., & Borsboom, D. (2016). Measuring depression over time...or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression.

*Psychological Assessment*, *28*, 1354–1367.

Glejsted Ingstrup, K., Schou Andersen, C., Adeltoft Ajslev, T., Pedersen, P., Sørensen, T. I. A., & Nohr, E. A. (2012). Maternal distress during pregnancy and offspring childhood overweight. *Journal of Obesity*, Article ID 462845. doi:10.1155/2012/462845

- Janicki-Deverts, D., Cohen, S., Turner, R. B., & Doyle, W. J. (2016). Basal salivary cortisol secretion and susceptibility to upper respiratory infection. *Brain, Behavior, and Immunity*, *53*, 255-261. <https://doi.org/10.1016/j.bbi.2016.01.013>
- Jenkins, B. N., Hunter, J. F., Cross, M. P., Acevedo, A. M., & Pressman, S. D. (2018). When is affect variability bad for health? The association between affect variability and immune response to the influenza vaccination. *Journal of Psychosomatic Research*, *104*, 41–47. <https://doi.org/10.1016/j.jpsychores.2017.11.002>
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.
- Kubzansky, L. D., Cole, S. R., Kawachi, I., Vokonas, P., & Sparrow, D. (2006). Shared and unique contributions of anger, anxiety, and depression to coronary heart disease: A prospective study in the normative aging study. *Annals of Behavioral Medicine*, *31*, 21–29. [https://doi.org/10.1207/s15324796abm3101\\_5](https://doi.org/10.1207/s15324796abm3101_5)
- Laboratory for the Study of Stress, Immunity, and Disease. (2016). *Common Cold Project*. Retrieved from <http://www.commoncoldproject.com>.
- Lac, A., & Donaldson, C. D. (2018). Validation and psychometric properties of the Alcohol Positive and Negative Affect Schedule: Are drinking emotions distinct from general emotions? *Psychology of Addictive Behaviors*, *32*, 40–51. <https://doi.org/10.1037/adb0000343>.Validation
- Lebo, M. A., & Nesselroade, J. R. (1978). Intraindividual differences dimensions of mood change during pregnancy identified in five P-technique factor analyses. *Journal of Research in Personality*, *12*, 205–224. [https://doi.org/10.1016/0092-6566\(78\)90098-3](https://doi.org/10.1016/0092-6566(78)90098-3)
- Lillis, T. A., Hamilton, N. A., Pressman, S. D., Ziadni, M. S., Khou, C. S., Boddy, L. E., &

- Wagner, L. M. (2018). Sleep quality buffers the effects of negative social interactions on maternal mood in the 3–6 month postpartum period: A daily diary study. *Journal of Behavioral Medicine, 41*, 733–746. <https://doi.org/10.1007/s10865-018-9967-y>
- McDonald, R. P. (1999). *Test Theory: A Unified Treatment*. Routledge.
- McNair, D. M., Lorr, M., & Droppelman, L. F. (1971). *Manual for the Profile of Mood States*. San Diego: Educational and Industrial Testing Services.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525–543.
- Middelboe, T., Andersen, H. S., Birket-Smith, M., & Friis, M. L. (1992). Minor head injury: Impact on general health after 1 year. A prospective follow-up study. *Acta Neurologica Scandinavica, 85*, 5-9.
- Montana Department of Public Health and Human Services (2018). Criteria and Procedures for Addition of Survey Questions Overview. Retrieved on October 9, 2019, from [https://dphhs.mt.gov/Portals/85/publichealth/documents/BRFSS/BRFSS\\_Criteria\\_for\\_Selection\\_2018.pdf](https://dphhs.mt.gov/Portals/85/publichealth/documents/BRFSS/BRFSS_Criteria_for_Selection_2018.pdf).
- Mustanski, B. (2007). The influence of state and trait affect on HIV risk behaviors: A daily diary study of MSM. *Health Psychology, 26*, 618-626. doi:10.1037/0278-6133.26.5.618
- Poole, L., Steptoe, A., Wawrzyniak, A. J., Bostock, S., Mitchell, E. S., & Hamer, M. (2011). Associations of objectively measured physical activity with daily mood ratings and psychophysiological stress responses in women. *Psychophysiology, 48*, 1165–1172. <https://doi.org/10.1111/j.1469-8986.2011.01184.x>
- Prather, A. A., Janicki-Deverts, D., Adler, N. E., Hall, M., & Cohen, S. (2017). Sleep habits and susceptibility to upper respiratory illness: The moderating role of subjective socioeconomic

- status. *Annals of Behavioral Medicine*, 51(1), 137-146. <https://doi.org/10.1007/s12160-016-9835-3>
- Pressman, S. D., & Cohen, S. (2005). Does positive affect influence health? *Psychological Bulletin*, 131, 925–971. <https://doi.org/10.1037/0033-2909.131.6.925>
- Pressman, S. D., & Cohen, S. (2012). Positive emotion word use and longevity in famous deceased psychologists. *Health Psychology*, 31, 297–305. <https://doi.org/10.1037/a0026071>
- Pressman, S. D., & Cross, M. P. (2018). Moving beyond a one-size-fits-all view of positive affect in health research. *Current Directions in Psychological Science*, 27, 339–344. <https://doi.org/10.1177/0963721418760214>
- Pressman, S. D., Jenkins, B. N., Kraft-Feil, T. L., Rasmussen, H., & Scheier, M. F. (2017). The whole is not the sum of its parts: Specific types of positive affect influence sleep differentially. *Emotion*, 17, 778–793. <https://doi.org/10.1037/emo0000256>
- Pressman, S. D., Jenkins, B. N., & Moskowitz, J. T. (2019). Positive affect and health: What do we know and where next should we go? *Annual Review of Psychology*, 70, 627–650. <https://doi.org/10.1146/annurev-psych-010418-102955>
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90. <https://doi.org/10.1111/mec.13536>.Application
- Richman, L. S., Kubzansky, L., Maselko, J., Kawachi, I., Choo, P., & Bauer, M. (2005). Positive emotion and health: Going beyond the negative. *Health Psychology*, 24, 422-429. <https://doi.org/10.1037/0278-6133.24.4.422>
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464.
- Sherman, R. A., Rauthmann, J. F., Brown, N. A., Serfass, D. G., & Jones, A. B. (2015). The

independent effects of personality and situations on real-time expressions of behavior and emotion. *Journal of Personality and Social Psychology*, *109*, 872–888.

<https://doi.org/10.1037/pspp0000036>

Shirom, A., Toker, S., Jacobson, O., & Balicer, R. D. (2010). Feeling vigorous and the risks of all-cause mortality, ischemic heart disease, and diabetes: A 20-year follow-up of healthy employees. *Psychosomatic Medicine*, *72*, 727-733. doi:10.1097/PSY.0b013e3181eeb643

Skaff, M. M., Mullan, J. T., Almeida, D. M., Hoffman, L., Masharani, U., Mohr, D., & Fisher, L. (2009). Daily negative mood affects fasting glucose in Type 2 diabetes. *Health Psychology*, *28*, 265-272. <https://doi.org/10.1037/a0014429>

Sneed, R. S., Cohen, S., Turner, R. B., & Doyle, W. J. (2012). Parenthood and host resistance to the common cold. *Psychosomatic Medicine*, *74*, 567-573.

<https://doi.org/10.1097/PSY.0b013e31825941ff>

StataCorp. (2017). *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.

Suls, J., & Bunde, J. (2005). Anger, anxiety, and depression as risk factors for cardiovascular disease: The problems and implications of overlapping affective dispositions. *Psychological Bulletin*, *131*, 260–300. <https://doi.org/10.1037/0033-2909.131.2.260>

Sultan, S., & Fisher, L. (2010). Depression as a proxy of negative affect? A critical examination of the use of the CES-D in type 2 diabetes. *European Review of Applied Psychology*, *60*, 97–104. <https://doi.org/10.1016/j.erap.2009.12.002>

Uher, R., Farmer, A., Maier, W., Rietschel, M., Hauser, J., Marusic, A., Mors, O., Elkin, A., Williamson, R. J., Schmael, C., Henigsberg, N., Perez, J., Mendlewicz, J., Janzing, J. G. E., Zobel, A., Skibinska, M., Kozel, D., Stamp, A. S.,...Aitchison, K. J. (2008). Measuring depression: Comparison and integration of three scales in the GENDEP study.

*Psychological Medicine*, 38, 289–300. doi:10.1017/S0033291707001730

Usala, P. D., & Hertzog, C. (1989). Measurement of affective states in adults. *Research on Aging*, 11, 403–426.

Watson, D. (1988). Intraindividual and interindividual analyses of positive and negative affect: Their relation to health complaints, perceived stress, and daily activities. *Journal of Personality and Social Psychology*, 54, 1020–1030. <https://doi.org/10.1037/0022-3514.54.6.1020>

Watson, D., & Clark, L. A. (1994). *The PANAS-X Manual for the Positive and Negative Affect Schedule - Expanded Form*. University of Iowa. <https://doi.org/10.17077/48vt-m4t2>

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>

Wetherwell, J. L., Gatz, M., & Pedersen, N. L. (2001). A longitudinal analysis of anxiety and depressive symptoms. *Psychology and Aging*, 16, 187–195. <https://doi.org/10.1037/0882-7974.16.2.187>

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 73, 913-934. <https://doi.org/10.1177/0013164413495237p>

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's  $\alpha$ , Revelle's  $\beta$ , and McDonald's  $\omega_H$ : Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70 (1), 123-133. <https://doi.org/10.1007/s11336-003-0974-7>

Table 1. Demographics

Variable	Mean (SD)	N (%)
Age	28.98 (9.82)	
Sex		
Female		326 (53%)
Male		284 (47%)
Race/Ethnicity		
Caucasian		450 (73.8%)
African American		141 (23.1%)
Native American, Eskimo, or Aleut		3 (0.5%)
Asian or Pacific Islander		10 (1.6%)
Hispanic		6 (1.0%)
Education		
High School Graduate or Lower		153 (25.1%)
Some College but Less than 2 years		195 (32.0%)
2+ years of College and Degree		141 (23.1%)
Bachelor's Degree or Higher		121 (19.8%)
Employed (full or part-time)		362 (59.3%)

Table 2.  
*Model Statistics for Day 1*

Model	Description	N	$\chi^2$ (df)	CFI	RMSEA (90% CI)	SRMR	CD	AIC	BIC
<u>Positive Affect</u>									
1.	Positive Affect with One Factor	555	582.58(27), $p < .001$	0.84	0.19 (0.18, 0.21)	0.09	0.94	12,179	12,295
2.	Subscales with No Covariance	555	884.94(27), $p < .001$	0.76	0.24 (0.23, 0.25)	0.41	1.00	12,481	12,598
2a.	Covariance of Subscales: Well-being & Vigor	555	321.10(26), $p < .001$	0.92	0.14 (0.13, 0.16)	0.29	1.00	11,919	12,040
2b.	Covariance of Subscales: Well-being & Vigor; Well-being & Calm	555	254.91(25), $p < .001$	0.94	0.13 (0.12, 0.14)	0.25	1.00	11,855	11,980
2c.	Covariance of Subscales: all	555	72.44(24), $p < .001$	0.99	0.06 (0.05, 0.08)	0.04	0.99	11,675	11,804
2d.	Covariance of Error Terms: Calm & Relaxed	555	33.44(23), $p = .074$	1.00	0.03 (0.00, 0.05)	0.02	0.99	11,638	11,772
<u>Negative Affect</u>									
1.	Negative Affect with One Factor	512	463.41(27), $p < .001$	0.80	0.18 (0.16, 0.19)	0.09	0.89	7,546	7,661
2.	Subscales with No Covariance	512	483.31(27), $p < .001$	0.79	0.18 (0.17, 0.20)	0.30	1.00	7,566	7,681
2a.	Covariance of Subscales: Anger & Depression	512	288.49(26), $p < .001$	0.88	0.14 (0.13, 0.16)	0.24	0.99	7,373	7,492
2b.	Covariance of Subscales: Anger & Depression; Anxiety & Depression	512	221.41(25), $p < .001$	0.91	0.12 (0.11, 0.14)	0.18	0.99	7,308	7,431
2c.	Covariance of Subscales: all	512	82.741(24), $p < .001$	0.97	0.07 (0.05, 0.09)	0.04	0.99	7,172	7,299
2d.	Covariance of Error Terms: Tense & Nervous	512	53.38(23), $p < .001$	0.99	0.05 (0.03, 0.07)	0.03	0.99	7,144	7,276
<u>Positive and Negative Affect</u>									
3.	Final Positive (2d) and Negative Affect (2d) Models	507	340.43(127), $p < .001$	0.96	0.06 (0.05, 0.07)	0.13	1.00	17,587	17,850
3a.	Covariance of Subscales: Calm & Anxiety	507	300.09(126), $p < .001$	0.97	0.05 (0.05, 0.06)	0.12	1.00	17,549	17,815

*Note.* CFI = comparative fit index; RMSEA = root mean squared error of approximation; SRMR = standardized root mean squared residual; CD = coefficient of determination; AIC = Akaike's information criterion; BIC = Bayesian information criterion.



Table 3.  
*Comparison of Invariance Models for Positive Affect*

Model	N	$\chi^2$ (df)	Comparison	$\chi^2$ (df) difference	CFI	RMSEA (90% CI)	SRMR	CD	AIC	BIC
1. Configural Invariance (Same Form)	528	3,829.09(1,617), $p < .001$	Not applicable	Not applicable	0.94	0.05 (0.05, 0.05)	0.36	1.00	70,022	71,995
2. Metric Invariance (Equal Loadings)	528	3,879.34(1,653), $p < .001$	2 vs. 1	50.52(36), $p = 0.058$	0.94	0.05 (0.05, 0.05)	0.36	1.00	70,000	71,819
3. Scalar Invariance (Equal Intercepts)	528	4,193.29(1,707), $p < .001$	3 vs. 2	313.95(54), $p < .001$	0.93	0.05(0.05, 0.06)	0.36	1.00	70,206	71,795

Table 4.  
*Comparison of Invariance Models for Negative Affect*

Model	N	$\chi^2$ (df)	Comparison	$\chi^2$ (df) difference	CFI	RMSEA (90% CI)	SRMR	CD	AIC	BIC
1. Configural Invariance (Same Form)	488	4,353.49(1,617), $p < .001$	Not applicable	Not applicable	0.86	0.06 (0.06, 0.06)	0.198	1.00	31,410	33,346
2. Metric Invariance (Equal Loadings)	488	4,546.17(1,653), $p < .001$	2 vs. 1	192.68(36), $p < .001$	0.85	0.06 (0.06, 0.06)	0.20	1.00	31,531	33,316
3. Scalar Invariance (Equal Intercepts)	488	4,865.40(1,707), $p < .001$	3 vs. 1	511.91(90), $p < .001$	0.84	0.06 (0.06, 0.06)	0.20	1.00	31,742	33,301

Figure 1

Day 1 Positive Affect Model 2d

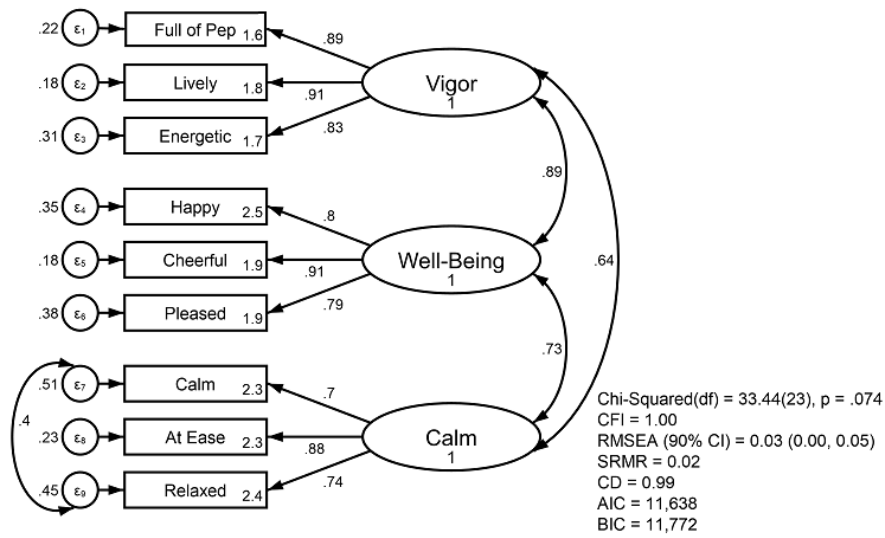


Figure 2

Day 1 Negative Affect Model 2d

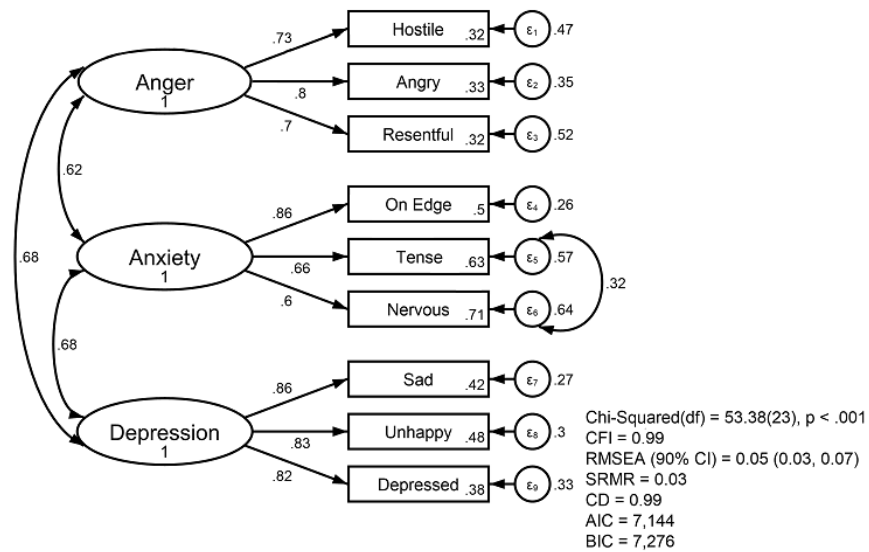


Figure 3

Day 1 Positive and Negative Affect Model 3

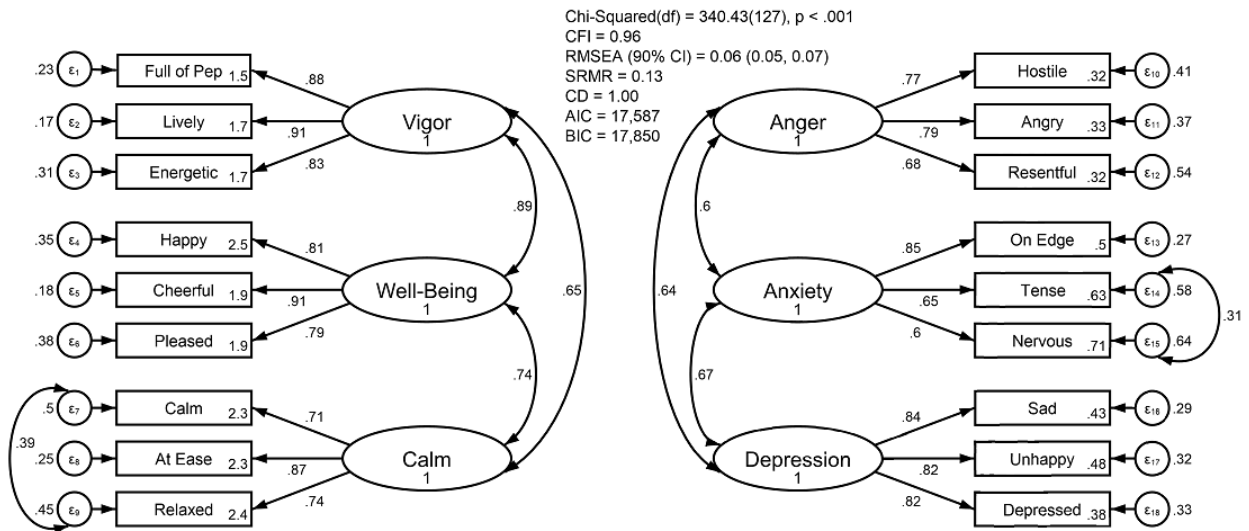


Figure Captions

Figure 1. Day 1 Positive Affect Model 2d.

Figure 2. Day 1 Negative Affect Model 2d.

Figure 3. Day 1 Positive and Negative Affect Model 3.

Appendix A

Table 1.  
*Model Statistics for Day 2*

Model	Description	N	$\chi^2$ (df)	CFI	RMSEA (90% CI)	SRMR	CD	AIC	BIC
<u>Positive Affect</u>									
1.	Positive Affect with One Factor	605	835.18(27), $p < .001$	0.81	0.22 (0.21, 0.24)	0.11	0.94	13,068	13,187
2.	Subscales with No Covariance	605	930.20(27), $p < .001$	0.79	0.24 (0.22, 0.25)	0.41	1.00	13,163	13,282
2a.	Covariance of Subscales: Wellbeing & Vigor	605	339.66(26), $p < .001$	0.93	0.14 (0.13, 0.16)	0.29	1.00	12,575	12,698
2b.	Covariance of Subscales: Wellbeing & Vigor; Wellbeing & Calm	605	229.97(25), $p < .001$	0.95	0.12 (0.10, 0.13)	0.23	1.00	12,467	12,595
2c.	Covariance of Subscales: all	605	69.42(24), $p < .001$	0.99	0.06 (0.04, 0.07)	0.03	1.00	12,309	12,441
<u>Negative Affect</u>									
1.	Negative Affect with One Factor	604	476.10(27), $p < .001$	0.79	0.17 (0.15, 0.18)	0.08	0.87	7,493	7,612
2.	Subscales with No Covariance	604	609.65(27), $p < .001$	0.73	0.19 (0.18, 0.20)	0.27	0.99	7,627	7,746
2a.	Covariance of Subscales: Anger & Depression	604	400.34(26), $p < .001$	0.82	0.15 (0.14, 0.17)	0.21	0.99	7,420	7,543
2b.	Covariance of Subscales: Anger & Depression; Anger & Anxiety	604	283.82(25), $p < .001$	0.88	0.13 (0.12, 0.15)	0.17	1.00	7,305	7,433
2c.	Covariance of Subscales: all	604	124.98(24), $p < .001$	0.95	0.08 (0.07, 0.10)	0.05	0.99	7,148	7,280
2d.	Covariance of Error Terms: Unhappy & Depressed	604	86.86(23), $p < .001$	0.97	0.07 (0.05, 0.08)	0.04	1.00	7,112	7,249
2e.	Covariance of Error Terms: Unhappy & Depressed; On Edge & Tense		Nonconvergent						
<u>Positive and Negative Affect</u>									
3.	Final Positive (2c) and Negative Affect (2d) Models	603	367.63(128), $p < .001$	0.96	0.06 (0.05, 0.06)	0.12	1.00	19,261	19,529
3a.	Covariance of Error Terms: On Edge & Tense		Nonconvergent						

*Note.* CFI = comparative fit index; RMSEA = root mean squared error of approximation; SRMR = standardized root mean squared residual; CD = coefficient of determination; AIC = Akaike's information criterion; BIC = Bayesian information criterion.

Table 2.  
*Model Statistics for Day 3*

Model	Description	N	$\chi^2$ (df)	CFI	RMSEA (90% CI)	SRMR	CD	AIC	BIC
<u>Positive Affect</u>									
1.	Positive Affect with One Factor	605	805.03(27), $p < .001$	0.83	0.22 (0.21, 0.23)	0.10	0.95	13,097	13,216
2.	Subscales with No Covariance	605	1018.04(27), $p < .001$	0.78	0.25 (0.23, 0.26)	0.43	1.00	13,310	13,429
2a.	Covariance of Subscales: Wellbeing & Vigor	605	347.05(26), $p < .001$	0.93	0.14 (0.13, 0.16)	0.30	1.00	12,641	12,764
2b.	Covariance of Subscales: Wellbeing & Vigor; Wellbeing & Calm	605	253.895(25), $p < .001$	0.95	0.12 (0.11, 0.14)	0.25	1.00	12,550	12,678
2c.	Covariance of Subscales: all	605	68.90(24), $p < .001$	0.99	0.06 (0.04, 0.07)	0.04	1.00	12,367	12,499
2d.	Covariance of Error Terms: Calm & Relaxed	605	33.54(23), $p < .001$	1.00	0.03 (0.00, 0.05)	0.02	1.00	12,333	12,470
<u>Negative Affect</u>									
1.	Negative Affect with One Factor	607	327.35(27), $p < .001$	0.87	0.14 (0.12, 0.15)	0.06	0.89	6,365	6,484
2.	Subscales with No Covariance	607	642.97(27), $p < .001$	0.72	0.19 (0.18, 0.21)	0.30	0.99	6,681	6,800
2a.	Covariance of Subscales: Anger & Depression	607	335.59(26), $p < .001$	0.86	0.14 (0.13, 0.15)	0.21	0.98	6,376	6,499
2b.	Covariance of Subscales: Anger & Depression; Anxiety & Depression	607	282.83(25), $p < .001$	0.89	0.13 (0.12, 0.14)	0.17	0.98	6,325	6,453
2c.	Covariance of Subscales: all	607	104.02(24), $p < .001$	0.96	0.07 (0.06, 0.09)	0.03	0.98	6,148	6,280
2d.	Covariance of Error Terms: Hostile & Sad	607	70.71(23), $p < .001$	0.98	0.06 (0.04, 0.07)	0.03	0.98	6,117	6,253
<u>Positive and Negative Affect</u>									
3.	Final Positive (2d) and Negative Affect (2d) Models	604	329.53(127), $p < .001$	0.97	0.05 (0.05, 0.06)	0.13	1.00	18,327	18,600

*Note.* CFI = comparative fit index; RMSEA = root mean squared error of approximation; SRMR = standardized root mean squared residual; CD = coefficient of determination; AIC = Akaike's information criterion; BIC = Bayesian information criterion.

Table 3.  
*Model Statistics for Day 4*

Model	Description	N	$\chi^2$ (df)	CFI	RMSEA (90% CI)	SRMR	CD	AIC	BIC
<u>Positive Affect</u>									
1.	Positive Affect with One Factor	605	678.15(27), $p < .001$	0.86	0.20 (0.19, 0.21)	0.08	0.95	13,074	13,193
2.	Subscales with No Covariance	605	1192.75(27), $p < .001$	0.76	0.27 (0.25, 0.28)	0.47	1.00	13,589	13,708
2a.	Covariance of Subscales: Wellbeing & Vigor	605	471.53(26), $p < .001$	0.91	0.17 (0.16, 0.18)	0.34	1.00	12,870	12,993
2b.	Covariance of Subscales: Wellbeing & Vigor; Wellbeing & Calm		Nonconvergent						
2c.	Covariance of Subscales: all	605	96.26(24), $p < .001$	0.99	0.07 (0.06, 0.09)	0.04	1.00	12,498	12,630
2d.	Covariance of Error Terms: Calm & Relaxed	605	67.75(23), $p < .001$	0.99	0.06 (0.04, 0.07)	0.02	0.99	12,472	12,608
<u>Negative Affect</u>									
1.	Negative Affect with One Factor	607	373.25(27), $p < .001$	0.82	0.15 (0.13, 0.16)	0.08	0.86	6,965	7,084
2.	Subscales with No Covariance	607	586.93(27), $p < .001$	0.71	0.19 (0.17, 0.20)	0.27	0.99	7,178	7,297
2a.	Covariance of Subscales: Anger & Depression	607	384.55(26), $p < .001$	0.81	0.15 (0.14, 0.16)	0.20	0.98	6,978	7,101
2b.	Covariance of Subscales: Anger & Depression; Anger & Anxiety		Nonconvergent						
2c.	Covariance of Subscales: all	607	119.41(24), $p < .001$	0.95	0.08 (0.07, 0.10)	0.05	0.98	6,717	6,849
2d.	Covariance of Error Terms: Sad & Depressed	607	91.64(23), $p < .001$	0.96	0.07 (0.06, 0.09)	0.04	0.98	6,691	6,828
<u>Positive and Negative Affect</u>									
3.	Final Positive (2d) and Negative Affect (2d) Models	605	381.75(127), $p < .001$	0.96	0.06 (0.05, 0.06)	0.13	1.00	19,158	19,431

*Note.* CFI = comparative fit index; RMSEA = root mean squared error of approximation; SRMR = standardized root mean squared residual; CD = coefficient of determination; AIC = Akaike's information criterion; BIC = Bayesian information criterion.

Table 4.  
*Model Statistics for Day 5*

Model	Description	N	$\chi^2$ (df)	CFI	RMSEA (90% CI)	SRMR	CD	AIC	BIC
<u>Positive Affect</u>									
1.	Positive Affect with One Factor	606	935.02(27), $p < .001$	0.82	0.24 (0.22, 0.25)	0.10	0.95	13,307	13,426
2.	Subscales with No Covariance	606	1111.35(27), $p < .001$	0.78	0.26 (0.25, 0.27)	0.46	1.00	13,483	13,602
2a.	Covariance of Subscales: Wellbeing & Vigor	606	412.44(26), $p < .001$	0.92	0.16 (0.14, 0.17)	0.32	1.00	12,786	12,909
2b.	Covariance of Subscales: Wellbeing & Vigor; Wellbeing & Calm		Nonconvergent						
2c.	Covariance of Subscales: all	606	99.81(24), $p < .001$	0.99	0.07 (0.06, 0.09)	0.04	1.00	12,477	12,610
2d.	Covariance of Error Terms: Calm & Relaxed	606	70.84(23), $p < .001$	0.99	0.06 (0.04, 0.07)	0.02	1.00	12,450	12,587
<u>Negative Affect</u>									
1.	Negative Affect with One Factor	607	523.87(27), $p < .001$	0.81	0.17 (0.16, 0.19)	0.08	0.90	6,950	7,069
2.	Subscales with No Covariance	607	761.89(27), $p < .001$	0.72	0.21 (0.20, 0.23)	0.32	0.99	7,188	7,307
2a.	Covariance of Subscales: Anger & Anxiety	607	477.33(26), $p < .001$	0.83	0.17 (0.16, 0.18)	0.27	0.99	6,906	7,029
2b.	Covariance of Subscales: Anger & Anxiety; Anger & Depression	607	390.29(25), $p < .001$	0.86	0.16 (0.14, 0.17)	0.21	0.99	6,821	6,949
2c.	Covariance of Subscales: all	607	169.14(24), $p < .001$	0.94	0.10 (0.09, 0.11)	0.05	0.99	6,602	6,734
2d.	Covariance of Error Terms: Unhappy & Depressed	607	107.16(23), $p < .001$	0.97	0.08 (0.06, 0.09)	0.04	1.02	6,542	6,678
2e.	Covariance of Error Terms: Unhappy & Depressed; Hostile & Nervous	607	72.01(22), $p < .001$	0.98	0.06 (0.05, 0.08)	0.03	1.02	6,508	6,649
<u>Positive and Negative Affect</u>									
3.	Final Positive (2d) and Negative Affect (2e) Models	603	375.16(126), $p < .001$	0.97	0.06 (0.05, 0.06)	0.13	1.00	18,691	18,968

*Note.* CFI = comparative fit index; RMSEA = root mean squared error of approximation; SRMR = standardized root mean squared residual; CD = coefficient of determination; AIC = Akaike's information criterion; BIC = Bayesian information criterion.

Table 5.  
*Model Statistics for Day 6*

Model	Description	N	$\chi^2(df)$	CFI	RMSEA (90% CI)	SRMR	CD	AIC	BIC
<u>Positive Affect</u>									
1.	Positive Affect with One Factor	601	930.47(27), $p < .001$	0.82	0.24 (0.22, 0.25)	0.10	0.95	13,440	13,559
2.	Subscales with No Covariance	601	1131.01(27), $p < .001$	0.77	0.26 (0.25, 0.27)	0.45	1.00	13,641	13,759
2a.	Covariance of Subscales: Wellbeing & Vigor	601	389.80(26), $p < .001$	0.93	0.15 (0.14, 0.17)	0.32	1.00	12,901	13,025
2b.	Covariance of Subscales: Wellbeing & Vigor; Wellbeing & Calm		Nonconvergent						
2c.	Covariance of Subscales: all	601	108.53(24), $p < .001$	0.98	0.08 (0.06, 0.09)	0.04	1.00	12,624	12,756
2d.	Covariance of Error Terms: Calm & Relaxed	601	81.55(23), $p < .001$	0.99	0.07 (0.05, 0.08)	0.02	1.00	12,599	12,736
<u>Negative Affect</u>									
1.	Negative Affect with One Factor	606	672.44(27), $p < .001$	0.73	0.20 (0.19, 0.21)	0.10	0.88	6,084	6,203
2.	Subscales with No Covariance	606	650.70(27), $p < .001$	0.74	0.20 (0.18, 0.21)	0.28	0.99	6,062	6,181
2a.	Covariance of Subscales: Anger & Anxiety	606	461.45(26), $p < .001$	0.82	0.17 (0.15, 0.18)	0.24	0.99	5,875	5,998
2b.	Covariance of Subscales: Anger & Anxiety; Anger & Depression	606	387.99(25), $p < .001$	0.85	0.16 (0.14, 0.17)	0.18	0.99	5,804	5,931
2c.	Covariance of Subscales: all	606	237.79(24), $p < .001$	0.91	0.12 (0.11, 0.14)	0.06	0.99	5,655	5,788
2d.	Covariance of Error Terms: On Edge & Sad	606	191.97(23), $p < .001$	0.93	0.11 (0.10, 0.13)	0.06	0.99	5,612	5,748
2e.	Covariance of Error Terms: On Edge & Sad; On Edge & Nervous	606	156.71(22), $p < .001$	0.94	0.10 (0.09, 0.12)	0.05	0.98	5,578	5,719
2f.	Covariance of Error Terms: On Edge & Sad; On Edge & Nervous; Hostile & Sad	606	125.14(21), $p < .001$	0.96	0.09 (0.08, 0.11)	0.04	0.98	5,549	5,694
<u>Positive and Negative Affect</u>									
3.	Final Positive (2d) and Negative Affect (2f) Models	600	429.26(125), $p < .001$	0.96	0.06 (0.06, 0.07)	0.13	1.00	18,116	18,397

*Note.* CFI = comparative fit index; RMSEA = root mean squared error of approximation; SRMR = standardized root mean squared residual; CD = coefficient of determination; AIC = Akaike's information criterion; BIC = Bayesian information criterion.



Table 6.  
*Model Statistics for Day 7*

Model	Description	N	$\chi^2$ (df)	CFI	RMSEA (90% CI)	SRMR	CD	AIC	BIC
<u>Positive Affect</u>									
1.	Positive Affect with One Factor	600	667.55(27), $p < .001$	0.86	0.20 (0.19, 0.21)	0.08	0.95	14,081	14,200
2.	Subscales with No Covariance	600	1131.68(27), $p < .001$	0.75	0.26 (0.25, 0.27)	0.45	1.00	14,545	14,664
2a.	Covariance of Subscales: Wellbeing & Vigor	600	480.16(26), $p < .001$	0.90	0.17 (0.16, 0.18)	0.33	1.00	13,895	14,019
2b.	Covariance of Subscales: Wellbeing & Vigor; Wellbeing & Calm		Nonconvergent						
2c.	Covariance of Subscales: all	600	106.44(24), $p < .001$	0.98	0.08 (0.06, 0.09)	0.03	1.00	13,526	13,658
2d.	Covariance of Error Terms: Cheerful & Pleased	600	78.84(23), $p < .001$	0.99	0.06 (0.05, 0.08)	0.03	1.00	13,500	13,636
<u>Negative Affect</u>									
1.	Negative Affect with One Factor	602	484.87(27), $p < .001$	0.81	0.17 (0.16, 0.18)	0.08	0.90	6,711	6,830
2.	Subscales with No Covariance	602	759.41(27), $p < .001$	0.70	0.21 (0.20, 0.23)	0.30	0.99	6,986	7,104
2a.	Covariance of Subscales: Anger & Depression	602	486.96(26), $p < .001$	0.81	0.17 (0.16, 0.19)	0.23	0.99	6,715	6,838
2b.	Covariance of Subscales: Anger & Depression; Anger & Anxiety		Nonconvergent						
2c.	Covariance of Subscales: all	602	178.59(24), $p < .001$	0.94	0.10 (0.09, 0.12)	0.05	0.99	6,411	6,543
2d.	Covariance of Error Terms: Resentful & Depressed	602	152.20(23), $p < .001$	0.95	0.10 (0.08, 0.11)	0.05	0.99	6,386	6,523
<u>Positive and Negative Affect</u>									
3.	Final Positive (2d) and Negative Affect (2d) Models	595	426.33(127), $p < .001$	0.96	0.06 (0.06, 0.07)	0.13	1.00	19,625	19,898
3a.	Covariance of Error Terms: Sad & Depressed	595	396.06(126), $p < .001$	0.96	0.06 (0.05, 0.07)	0.13	1.00	19,597	19,874

*Note.* CFI = comparative fit index; RMSEA = root mean squared error of approximation; SRMR = standardized root mean squared residual; CD = coefficient of determination; AIC = Akaike's information criterion; BIC = Bayesian information criterion.

Table 7.

*Comparison of Invariance Models for Positive and Negative Affect Model*

Model	N	$\chi^2$ (df)	Comparison	$\chi^2$ (df) difference	CFI	RMSEA (90% CI)	SRMR	CD	AIC	BIC
1. Configural Invariance (Same Form)	472	14,248.78(7,203), $p < .001$	Not applicable	Not applicable	0.87	0.05 (0.04, 0.05)	0.22	1.00	92,353	96,194
2. Metric Invariance (Equal Loadings)	472	14,473.62(7,275), $p < .001$	2 vs. 1	224.83(72), $p < .001$	0.87	0.05 (0.05, 0.05)	0.22	1.00	92,434	95,976
3. Scalar Invariance (Equal Intercepts)	472	15,053.53(7,383), $p < .001$	3 vs. 1	804.74(180), $p < .001$	0.86	0.05(0.05, 0.05)	0.22	1.00	92,798	95,891

*Note.* CFI = comparative fit index; RMSEA = root mean squared error of approximation; SRMR = standardized root mean squared residual; CD = coefficient of determination; AIC = Akaike's information criterion; BIC = Bayesian information criterion.

## Online Supplemental Material

### Modification Indices Details

Modification indices estimate the extent to which the model chi-square will decrease for various parameter changes (i.e., specifications) within the model. This is beneficial as a reduction in chi-square among nested models (i.e., chi-square decreasing when going from a reduced model to one with more estimated parameters) reflects a better model fit (at the cost of estimating additional parameters). However, theory should guide the selection of model changes based on the modification indices. Therefore, in our analyses, modification indices were examined to help aid theoretically justified alterations to the models. Following theory, we selected changes that specified correlations of measurement residuals within subscales (e.g., correlating the measurement residuals of calm and relaxed) over loading subscale items onto other latent variables (e.g., loading the item at ease [which is part of the subscale calm] onto the latent variable vigor) even when the latter might have had a higher modification index. To assess which modifications would have a small to medium sized effect change or larger while also considering sample size, we used the following formula based on Cohen's effect size measure  $w$  (Cohen, 1988; Newsom, 2017):  $w = \sqrt{\frac{\Delta\chi^2}{N*\Delta df}}$ . In this formula,  $w$  represents the effect size,  $\Delta\chi^2$  indicates the magnitude of change in the chi square,  $N$  reflects sample size, and  $\Delta df$  is the degrees of freedom change. All modification indices reflect a  $\Delta df$  of 1 and we only changed one parameter at a time. To determine which modifications would have a small to medium effect size change ( $w = .2$ ) or greater, we rearrange the formula to solve for  $\Delta\chi^2$ :  $\Delta\chi^2 = (N) * (w^2) * (\Delta df)$ . For example, in a model with  $N = 600$ , a small to medium effect size change would result from a modification index of 24:  $(600) * (0.2)^2 * (1) = 24$ . Because the models based on changes suggested by the modification indices are nested, and in this case the difference of two chi-squared statistics is itself distributed

as chi-squared, successive models can also be compared to one another to determine whether a model with more parameters estimated provides a significantly improved fit to the data. Therefore, using the cutoff of a small-medium effect also provided the additional benefit of ensuring that the chi-square improvement on nested models was statistically significant as the modification cutoff value for samples of 100 or larger is well above 3.84 (as 3.84 is the chi-square critical value for  $df = 1$  at the 0.05 alpha level).

### **Chi-Square Details**

For all of the specified models, chi-squared values were calculated to compare each specified model to a saturated one with the null hypothesis that the specified model perfectly reproduces the covariance matrix of the items in the specified model (standard practice within confirmatory factor analysis). A null result signifies that our specified model represents the data well (i.e., a null result is desired and reflective of a good fitting model). However, chi-squared statistics in confirmatory factor analyses are often statistically significant even though the model may represent a relatively good fit of the data. Thus, goodness of fit tests were used.

### **Factorial Invariance Details**

One question that needs to be answered is whether the general subscale structure of the SAS holds across time. This question can be answered by testing the factorial invariance of the measure. There are different levels of factorial invariance: 1. Configural invariance, 2. Metric invariance, and 3. Scalar invariance. A scale has configural invariance if its items load on the same latent factors across time. In other words, if the items happy, cheerful, and pleased load on to the latent variable (i.e., factor) well-being at time 1 and they continue to do so at times 2 through 7, the subscale would have configural invariance. To test for configural invariance, we specified a model in which the same group of items load on each of the latent variables (same configuration

of variables) across the days. Numerical values of the loadings can be different across days, but if all were significant, we achieved configural invariance.

Metric invariance builds upon configural invariance and assumes that in addition to having the same form across days, the factor loading values are the same across days. Therefore, in order to test for metric invariance, we set the factor loadings for each item across days equal to one another (e.g., the factor loading of the item happy on the latent variable well-being for Day 1 is set equal to the factor loading of the item happy on the latent variable well-being for Day 2 and so on for all 7 days). If the model specifying this equality of factor loadings is not a significantly worse fitting model as compared to the configural invariance model, we achieve metric invariance. Because the models are nested within one another (i.e., the metric invariance model is nested within the configural invariance model), we can test whether they are significantly different using a likelihood ratio chi-squared test. When this test is not significant, we can claim that the model with more restrictions (e.g., equality of factor loadings) is not worse than the model with fewer restrictions. The more restricted model then becomes the model for comparison among even more restricted models.

Next, scalar invariance builds upon metric invariance and requires that the intercepts of the items on each factor are also the same. To test this model, in addition to having the same form and equal factor loadings, we restrict the intercepts of each of the items on their latent variable to be equal across days (e.g., the intercept for the item happy on the latent variable well-being for Day 1 is set equal to the intercept for the item happy on the latent variable well-being for Day 2 and so on for all 7 days).

Each of these invariance models (configural, metric, scalar) are tested iteratively and become more restricted. At each step, the model is compared to the previous model using the

likelihood ratio chi-squared test as described above. Once a model becomes significantly different than the previous model, we use that previous model as the comparison point for all other more restricted models. Throughout the factorial invariance results, we specify which models are being compared. As is common practice, we test the factorial invariance on the unstandardized models (Acock, 2013).

### **Coefficient Omega Details**

Although Coefficient Alpha is by far the most possible estimator of reliability in psychological research, methodologists tend to believe that the assumptions of Alpha: each item reflects the underlying latent trait to the same extent (item factor loadings are equal) and that the items are clearly unidimensional (item error variances do not correlate) are rarely met in psychological data. An alternative estimator that is robust under these conditions is McDonald's Omega ( $\omega$ ) (McDonald, 1999). Omega uses CFA parameter estimates (factor loadings, item error variances, item error covariances) to estimate a reliability coefficient for a scale. Alpha is equal to Omega when its strict assumptions are met. When they are not, Alpha is biased and may underestimate or overestimate reliability depending on the interaction between factor loadings, error variances, and error covariances (Zinbarg et al., 2005).