



Published in final edited form as:

Assessment. 2022 September ; 29(6): 1301–1319. doi:10.1177/10731911211008254.

## Psychometric Evaluation of a Big Five Personality State Scale for Intensive Longitudinal Studies

Whitney R. Ringwald<sup>1</sup>, Stephen B. Manuck<sup>1</sup>, Anna L. Marsland<sup>1</sup>, Aidan G. C. Wright<sup>1</sup>

<sup>1</sup>University of Pittsburgh, Pittsburgh, PA, USA

### Abstract

Despite enthusiasm for using intensive longitudinal designs to measure day-to-day manifestations of personality underlying differences between people, the validity of personality state scales has yet to be established. In this study, we evaluated the psychometrics of 20-item and 10-item daily, Big Five personality state scales in three independent samples ( $N = 1,041$ ). We used multilevel models to separately examine the validity of the scales for assessing personality variation at the between- and within-person levels. Results showed that a five-factor structure at both levels fits the data well, the scales had good convergent and discriminative associations with external variables, and personality states captured similar nomological nets as established global, self-report personality inventories. Limitations of the scales were identified (e.g., low reliability, low correlations with external criterion) that point to a need for more, systematic psychometric work. Our findings provide initial support for the use of personality state scales in intensive longitudinal designs to study between-person traits, within-person processes, and their interrelationship.

### Keywords

personality assessment; personality states; intensive longitudinal design; multilevel modeling

A comprehensive model of personality must account for relatively stable traits that vary between people as well as trait-relevant states that vary within a person across time and situations (Baumert et al., 2017; Cervone, 2005). Although most personality research has focused on the *between-person* trait structure, contemporary personality theories emphasize *within-person* processes that give rise to differences between people (e.g., DeYoung, 2015; Fleeson & Jayawickreme, 2015; Pincus & Ansell, 2013; Read et al., 2017). Understanding how personality traits manifest in day-to-day life, including the precipitants and consequences of shifts in personality states (i.e., short-term expressions of personality traits), is also key to explaining the well-established links between traits and life outcomes (Hampson, 2012).

**Corresponding Author:** Whitney R. Ringwald, Department of Psychology, University of Pittsburgh, 4305 Sennott Square, 210 South Bouquet Street, Pittsburgh, PA 15260, USA., wringwald@pitt.edu.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Supplementary Material

Supplementary material for this article is available online.

## Measuring Personality States

The gap between measurement and theory has prompted growing interest in empirical methods capable of capturing variance in personality at both the individual differences (i.e., between-persons) and occasion differences (i.e., within-person) levels of analysis. One promising approach is the use of intensive longitudinal designs (ILDs), encompassing more specific techniques like ambulatory assessment, ecological momentary assessment, or daily diaries, which involve repeated measurement of personality states over the course of days or weeks. ILD research has shown there is considerable within-person variability of self-reported personality states and that the aggregation of an individual's states approximates their personality trait standing (Fleeson, 2001; Fleeson & Gallagher, 2009; Jones et al., 2017). These methods have also been used to relate fluctuations in personality states to proximal contexts, affording insight into questions about the dynamics of the person and situation that have eluded cross-sectional methods (Bleidorn, 2009; Fleeson, 2007; Fleeson & Law, 2015; Matz & Harari, 2020; Rauthmann et al., 2016; Sherman et al., 2015).

Although ILDs are the vanguard of personality measurement, there has yet to be a thorough psychometric evaluation of personality state scales used in these studies. In their recent review of the ILD personality state literature, Horstmann and Ziegler (2020) identified major limitations in previous efforts to validate these measures. Notably, nearly all studies that tested scale psychometrics examined the reliability and validity of average state scores (i.e., between-person level), but not individual state scores (i.e., within-person level). Horstmann and Ziegler reported that the most common validity evidence has been indexing convergence between average states and personality traits measured by global, self-report inventories, and reliability has typically only been estimated for the average states. These studies only offer psychometric evidence for measuring trait-like, between-person variance—not within-person variance—despite the importance of each source of variability for understanding personality.

In one exception to this pattern, Zimmerman et al. (2019) examined both the within- and between-person levels in developing their ILD scales of personality states, with the primary goal of measuring maladaptive processes (see also Wright & Simms, 2016). The scales introduced by that study were constructed in an exploratory manner and were not intended to align with a particular model of personality. This approach was well-suited for the scale's purpose of clinical assessment, but not for testing personality theories built on the Big Five model of personality, which is arguably the dominant framework for describing individual differences. Because of the Big Five's eminence in personality research, it is unsurprising that many ILD personality state measures have been informed by this model (e.g., Aschwanden et al., 2019; Bleidorn, 2009; Borkenau & Ostendorf, 1998; Fleeson, 2001, 2007; Fleeson & Law, 2015; McCabe & Fleeson, 2016). Tethering personality state measures to the Big Five permits researchers to leverage the model's robust structure and nomological net for conceptualizing trait expressions in daily life. For personality state measures to be validly used for studying the Big Five between-person traits, within-person processes, and their interrelationship, the psychometrics for both levels of analysis must be established.

## The Need for Brief Measures

A unique advantage of ILD over traditional cross-sectional measures of personality is that it affords naturalistic, temporally sensitive data. The resolution of personality process models is limited, however, by the burden placed on participants of completing lengthy surveys in the stream of everyday life. Furthermore, the potential for participant reactivity to a tedious ILD protocol threatens the validity of their responses (Barta et al., 2012). Thus, the assessment demands of ILD protocols exacerbate the traditional tensions of the bandwidth-fidelity trade-off. That is, researchers are faced with a trade-off between content coverage or temporal granularity when both components are essential to modeling personality processes. Short scales offer one solution, and there are several such scales for personality traits with good psychometrics (e.g., Gosling et al., 2003; Thalmayer et al., 2011), but none have been systematically psychometrically validated for assessing personality states with ILDs. For these reasons, there is great potential value in validating brief measures of personality states that can be administered multiple times per day with minimal interference in participant's daily activities.

## Current Study

The purpose of this study was to assess the reliability and validity of a 20-item and 10-item personality state scale for measuring both between- and within-person variance. Our goal was not to construct a definitive instrument from scratch; rather, we aimed to evaluate a Big Five ILD scale akin to those used in a number of published studies. This study is an initial effort that can inform a broader, ongoing validation process needed to comprehensively test the psychometrics of personality state measures for ILD.

Our analyses were guided by expectations of what constitutes a valid measure at each, conceptually and statistically distinct level of personality (Horstmann & Ziegler, 2020). Because a personality state is the contextualized, short-term enactment of a personality trait, each personality state score is presumed to be influenced by the corresponding trait as well as the situation in which it manifests. In this study, we separated these sources of variance to examine the structure and correlates of personality traits and states.

At the between-person level, individual differences reflect variation in how people tend to behave, think, and feel across situations (i.e., personality traits). In ILDs, these characteristic patterns can be estimated by averaging each person's personality state scores over the study period. For a Big Five personality state scale to be valid for studying between-person differences we expected that (a) covariation of average personality states would conform to the expected five-factor trait structure (i.e., structural validity); (b) there would be adequate variance in average states between people to justify studying average states as individual differences; (c) average states would correlate with corresponding traits measured by established global self-report inventories (i.e., convergent validity), and (d) average states and global self-report trait measures would share nomological nets evidenced by a similar set of associations with an individual's typical behaviors, thoughts, and affect assessed in the ILD protocol (i.e., construct validity or nomological homomorphy; Rauthmann et al., 2019).

At the within-person level, deviation of personality states from an individual's average level (i.e., trait) are elicited by and lead to other behavioral, cognitive, and affective states. States and situations that tend to covary with shifts in personality states over time suggest characteristic processes underlying between-person differences. For personality states and corresponding traits to be functionally related via these within-person processes, they would presumably share a nomological net. Thus, to be a valid measure at the within-person level, we expected that (a) the same personality states that covary between people would also covary from day-to-day and so would form five, within-person factors resembling the between-person structure; (b) there would be adequate within-person variance to justify studying within-person processes; and (c) personality states would co-occur with other behavioral, cognitive, and affective states measured in the ILD protocol that map onto the nomological net of corresponding traits assessed by global self-report inventories.

## Method

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

### Participants

**Sample 1 (S1).**—S1 consisted of undergraduate students enrolled in introductory psychology courses at the University of Pittsburgh. The sample was mostly White (86%; 10% Asian, 6.5% Black or African American)<sup>1</sup> and mostly female (62%) with a mean age of 18.6 years ( $SD = 0.96$ ). The total sample size was 330. Sample size selection was influenced by the goal of arriving at stable effect sizes for analyses based on covariance/correlation matrices (e.g., structural equation modeling, multilevel modeling). Based on recent empirical work suggesting correlation estimates stabilize at  $n = 250$  (Schönbrodt & Perugini, 2013), we over-sampled with a target of  $n > 300$  to ensure the minimum sample size was achieved. After removing outliers, the final sample size for S1 was 294. Criteria for outliers is described in the Analytic Plan section.

**Sample 2 (S2).**—Participants in S2 were community members recruited from posted flyers and online postings for a study on personality and daily life. For inclusion, participants had to be between 18 and 40 years of age. Participants also had to be users of a smartphone running iOS or Android software. To recruit a distinct community sample, individuals were not eligible if they were enrolled in a full-time undergraduate program. All participants in S2 were prescreened to ensure a gender-balanced sample as well as adequate representation of personality traits of interest for the parent study on narcissism. Namely, the modesty facet scale of the NEO Personality Inventory–Revised (NEO-PI-R; Costa & McCrae, 1992) was used to assess participants during the prescreening interview. Low modesty, a core feature of narcissism, was oversampled such that a 2:1:1 ratio of low, moderate, and high levels of modesty within each gender were recruited. S2 was mostly White (89%; 8.5% Asian, 4.4% Black or African American), balanced on gender (female = 52%), with a mean age

---

<sup>1</sup>Percentages add up to >100% because participants in S1 and S2 could identify with more than one race

of 27.6 years ( $SD = 4.9$ ). The total sample size was 342 and was determined by the same considerations as S1. After removing outliers, the final sample size for S2 was 316.

**Sample 3 (S3).**—S3 was drawn from the University of Pittsburgh Adult Health and Behavior Project (AHAB) project, which comprises a registry of behavioral and biological measurements for the study of individual differences. Participants 30 to 54 years of age were recruited via mass-mail solicitation from communities of Southwestern Pennsylvania in two periods of data collection (2001–2005; 2008–2011). At enrollment, participants were in good general health, without reported history of atherosclerotic cardiovascular disease, chronic kidney or liver disease, recent treatment for cancer, major neurological disorders, or psychotic illness (e.g., Manuck et al., 2010; John-Henderson et al., 2016). Data used in the present analyses derive from an ongoing, second wave of AHAB data collection, occurring roughly 11 to 16 years following participants' initial participation. The sample available for analysis is mostly White (85%, 14% African American, <.01% Asian, <.01% biracial), 54% females, with an average age of 59.5 years ( $SD = 7.2$ ). Total projected enrollment for the second wave of the AHAB registry is  $N \sim 800$ . However, due to the COVID-19 global pandemic, participant recruitment and data accrual was suspended in March of 2020. This study uses those data that were available at that time. The total sample size was 458. After removing outliers, the final sample size was 431.

## Procedure

**Sample 1 and Sample 2.**—Study procedures were nearly identical for S1 and S2. For both samples, participants completed baseline self-report questionnaires then could elect to participate in the ILD portion of the study. Orientation to the protocol and participation were conducted entirely online without direct contact with study staff. Participants in S1 received course credit for completing the study. In S2, participants who completed baseline questionnaires were entered into prize drawings for \$75 Amazon gift cards. For the ILD portion, they received a \$100 Amazon gift card for completing 90% or greater of the total surveys administered during the study period. Gift cards of prorated value (e.g., \$75 was given for 75% participation) were given to those who completed less than 90% of surveys.

Baseline assessments included self-report questionnaires on personality and psychological functioning. Baseline data were collected and managed using Research Electronic Data Capture (REDCap) hosted at the University of Pittsburgh (Harris et al., 2009; 2019). REDCap is a secure, web-based software platform designed to support data capture for research studies supported by the Department of Biomedical Informatics (Clinical and Translational Sciences Institute at the University of Pittsburgh Grant Number UL1-TR-001857).

When participants chose to participate in the ILD protocol, they viewed a video training presentation explaining the ILD procedures and instructions for downloading the MetricWire smartphone application (MetricWire, Inc., 2019) used to administer surveys. A short comprehension quiz was given following the training to check for understanding. Failure to show adequate comprehension lead to exclusion from further participation.

The ILD protocol involved randomly initiated surveys delivered via push notifications between 9 am and 9 pm. Participants in S1 received five surveys per day, and those in the S2 received seven surveys per day. For this study, only self-reported positive and negative affect from the random surveys were used to index average daily affect. A longer survey was administered at 9 pm each day asking about thoughts, feelings, and events occurring in the past 24 hours. Participants had 3 hours after the initial push notification to complete the end of day survey.

**Sample 3.**—The study procedures for S3 involved several in-person lab sessions and 8 days of ILD monitoring outside of the lab. Participants were compensated \$150 to \$250, depending on their participation in an ancillary brain imaging protocol. Other components of data collection included neurocognitive testing, assessments of social and demographic factors, and instrumented measurements of physical functioning and cardiometabolic risk factors. Only data from self-report questionnaires were used for this study, so the procedures from other sessions will not be detailed further.

For the ILD portion of the study, participants were provided electronic tablets to complete end-of-day surveys and received in-person training about the protocol. Daily surveys were completed in Qualtrics and included the same questions administered to S1 and S2.

## Measures

Items used to assess personality states and other ILD variables are shown in Table 1. Descriptive statistics for every measure and omega coefficients for scales are included in the online supplementary material. Unless otherwise noted, items were administered to all three samples in the same format.

**Personality States.**—Personality states were measured at the end of the day with 20 bipolar adjective items informed by those used in Fleeson (2001). For each item, participants indicated the extent to which two adjectives best described them in the past 24 hours. Items were rated on a 7-point Likert-type scale. For example, to assess state Extraversion, participants would rate whether *talkative describes extremely well* (1) to *silent describes extremely well* (7), with the mid-point indicating *both describe equally well*. Two (for the brief scale) or four items (for the full scale) were averaged to calculate personality state subscale scores. Seven items were reverse scored.

**Daily Behaviors.**—Occurrence of personality-relevant behaviors were recorded at the end of each day. Participants used binary scales to indicate whether a behavior did or did not occur in the past 24 hours.

**Daily Positive and Negative Affect.**—Emotion adjectives from the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) were adapted to assess affect. In S1 and S2, affect was rated at each randomly initiated prompted time-point throughout the day. Items were reworded from the PANAS in the surveys to read “How ADJECTIVE do you feel right now?” Ratings were made on a slider scale from *not at all* (0) to *extremely* (100) for each adjective. Each individual’s mean affect reported across surveys within a day was used to measure daily affect. In S3, affect was rated at the end of day surveys.



Items were reworded to read “During the past 24 hours to what extent did you feel . . .” and adjectives were rated on a Likert-type scale from *not at all* (1) to *very much* (5). The mean of corresponding items indexed daily positive and negative affect.

**Daily Energy.**—Level of energy was assessed with two items in which participants rated how often they felt tired or out of energy in the past 24 hours on a Likert-type scale from *almost never* (1) to *very often* (4).

**Daily Loneliness and Stress.**—Daily loneliness and stress were each measured by three items. Loneliness items were adapted from the UCLA Loneliness Scale (Russell, 1996). Stress items were adapted from the Perceived Stress Scale (Cohen et al., 1983). Participants rated the extent to which they felt lonely and stressed in the past 24 hours on a 4-point (S1 and S2) or 5-point (S3) Likert-type scale from *never* to *very often*.

**Global Personality Traits.**—Big Five personality traits and facets were self-reported at baseline in S2 and S3. The 60-item Big Five Inventory–2 (BFI-2; Soto & John, 2017) was used in S2, and the 240-item NEO-PI-R was used in S3. For each item of the BFI-2 and NEO-PI-R, participants rated the extent to which a characteristic applies them (e.g., “I am someone who is outgoing”). In both instruments, items were rated on a Likert-type scale from *disagree strongly* (0) to *agree strongly* (4). Items from both instruments produce personality facet scores that are averaged to produce trait scores for Extraversion, Neuroticism (NEO-PI-R)/Negative Emotionality (BFI-2), Conscientiousness, Agreeableness, and Openness to Experience (NEO-PI-R)/Open-Mindedness (BFI-2).

## Analytic Plan

All analyses were conducted with *Mplus* Version 8.4 (Muthén & Muthén, 2017) and RStudio Version 3.6.3 (Rstudio Team, 2020). We used multilevel models to disaggregate personality traits and states. In these models, each personality state score was decomposed into statistically independent sources of variance using latent decomposition. Variance at the between-person level reflects an individual’s average reported personality state expression across study days, indicating their relatively stable, cross-situational tendencies. Within-person variance represents how much an individual’s personality state deviates from their average level on a given day and suggests situational influences on personality expression. Thus, between-person variance approximates personality traits and variability at the within-person level are the day-to-day fluctuations in personality states.

**Twenty-Item Scale Structural Validity.**—To examine the within- and between-person structures of the personality states, we conducted a series of multilevel exploratory factor analyses (MEFAs) of the 20 items from our ILD scale. For these analyses, we pooled the three samples to maximize generalizability. In MEFA, a separate factor structure is estimated at each level. The structure at the between-person level represents how average personality states covary from *person to person*, whereas the within-person structure represents how personality states covary from *day to day*.

Factors were extracted using maximum likelihood estimation and oblique Geomin rotation. Models were evaluated in terms of factor interpretability and by using standard criteria for

single-level factor analysis (i.e., comparative fit index [CFI] = .95, root mean square error of approximation (RMSEA) < .06, and standardized root mean square residual (SRMR) < .08; Hu & Bentler, 1999). We considered factor loadings  $\geq |.30|$  to be markers for a factor. Level-specific reliabilities for the subscales were estimated using omega coefficients, which has been shown to be preferable to other internal consistency measures in multilevel settings (Geldhof et al., 2014).

**Twenty-Item Scale Construct Validity.**—Construct validity of the subscales was assessed by examining the amount of variability in personality states subscales, convergent and discriminant relations with other trait and state variables, and by indexing their nomological homomorphy with commonly used global self-report personality trait measures. Because the samples had differences in criterion variables, we conducted these analyses separately for each sample. We used Bayes estimation with noninformative priors for all the following multilevel models. The Bayes estimator does not provide two-tailed  $p$  values, so we considered point estimates with 95% credibility intervals that did not contain zero to be statistically significant.

**Ten-Item Scale Structural and Construct Validity.**—Candidate items for a brief, 10-item scale were selected based on the MEFA results. We considered items for inclusion with primary factor loadings on the target factor at the between- and within-person levels, and items having the lowest cross-loadings on nontarget factors at both levels. We then fit models using the pooled sample with different combinations of the selected items with confirmatory factor analysis (CFA). Overall model fit was evaluated with the same criteria used for the MEFA. The final model was chosen to balance model fit and breadth of item content. After identifying two items per personality state, we conducted the same construct validation analyses described above in each sample with the reduced subscales. Additionally, we compared the content coverage of the 10-item and 20-item scales by indexing their nomological homomorphy.

## Results

### Twenty-Item Scale Structural Validity

For this stage of analysis, we pooled the samples to maximize generalizability. Prior to evaluating structural validity, 1,438 observations (~16% of total observations) were removed due to unusually high or low personality state item variability within a day suggestive of invalid responding (i.e., flatline or flip-flopping between extreme values). To determine variance thresholds, we visually inspected distribution plots of within-day item standard deviations for outliers. Based on these distributions, we removed observations with item-level standard deviations  $\leq 0.80$  and  $\geq 2.60$ . This resulted in removing 50 participants for a total pooled sample size of 1,041 participants and 7,664 days.

*Mplus* output and table of model fit indices for the MEFA are available in the online supplementary materials. Although we expected five factors resembling the Big Five traits within- and between people, we estimated all permutations of one to six factors at each level as a basis of comparison for the hypothesized model. The model with five factors at the between- and within-person levels achieved good fit (CFI = .96; RMSEA = .03;



SRMR<sub>within</sub>/SRMR<sub>between</sub> = .02/.03). Model fit indices improved slightly if a sixth factor was extracted at either level, but the added factors had uniformly low factor loadings (i.e., <|.30|) or were marked by secondary factor loadings, so we proceeded with analysis of the model with five factors between- and within-person.

Factor loadings for this model can be found in Table 2. At the between-person level, all but two items loaded most strongly on the target factor and only two items had cross loadings |.30| on nontarget factors. The factor structure was less consistent at the within-person level, but all factors except for Agreeableness had at least two primary loadings on the target factor.

After confirming good overall model fit and factor interpretability, we calculated personality state subscale scores. We then removed 39 participants with zero variance in every personality state subscale. Most participants removed for this reason were those who completed only one or two daily surveys, thus did not contribute enough data for reliable estimates, otherwise the responses were indicative of invalid responding (e.g., flatlining). Table 3 shows the means and standard deviations for the subscales in the pooled sample and sample specific descriptive statistics can be found in the online supplementary material. Intercorrelations among subscales at each level are presented in Table 4. The pattern of associations was similar across levels and ranged from modest to strong. The intercorrelations suggest a considerable amount of shared variance indicating a general factor in these scales.

### Reliability.

Omeegas for the state subscales in the pooled sample are shown in Table 3, and sample-specific reliability estimates are provided in the online supplementary material. Average reliability of the subscale scores was higher at the between-person ( $\omega = .86$ , range = .81–.87) than within-person level ( $\omega = .51$ , range = .44–.56).

### Twenty-Item Scale Construct Validity

**Variability.**—To determine whether there was sufficient within- and between-person variance in personality states, we calculated intraclass correlations (ICCs) using an unconditional multilevel model (i.e., no predictors). ICCs index the proportion of variance that is between-persons and 1—ICC is the proportion that is within-persons. ICCs for the personality state subscales in the pooled sample are shown in Table 3. The average ICC was .61 (range = .57–.67) indicating 61% of the variance in personality states is between people and 39% is within-persons from day-to-day.

**Convergent and Discriminative Validity.**—To determine the degree of convergence between average personality states and personality traits, we correlated personality states with the BFI-2 (S2) and NEO-PI-R (S3) trait scales at the between-person level (Table 5). Correlations with BFI-2/NEO-PI-R facets are available in the online supplementary material. In both samples, average personality states had the highest correlations with the corresponding trait with one exception: average daily Agreeableness was associated (negatively) just as strongly with NEO-PI-R Neuroticism as with Agreeableness. At the

same time, with few exceptions, average personality states were also significantly associated with each other trait consistent with the sizeable amount of shared variance across scales observed in Table 4. However, after adjusting for their shared variance using multiple regression, associations among the same average personality states and traits remained relatively strong, but cross-trait associations were attenuated and became non-significant in most cases (see Table 6) consistent with a strong general factor.

Next, we examined correlations among personality state subscales and external state variables from the ILD protocol at the within- and between-person levels. Like the personality states, the other daily state variables were decomposed into between- and within-person variance allowing us to investigate the nomological nets at each level. Correlations at the between-person level show the associations among average personality states and average behavior, thoughts, and feelings, and correlations at the within-person level indicate how shifts in personality states relate to shifts in behavior, thoughts, and feelings across days. Of the 220 associations tested per sample (22 external variables  $\times$  five traits at between- and within-person levels), ~64% replicated (i.e., the correlation's statistical significance and sign were the same) in all three samples, and an additional ~15% replicated in two samples. Furthermore, these patterns of associations were in line with theory and previous research on personality correlates. We summarize the replicated effects here, and full correlation tables are in Table 7.

Correlations at the between-person level indicated that individuals who reported more Extraversion on average were more likely to report exercising over the study period. Those who reported more Conscientiousness were more likely to report exercising, going to work, and doing housework and were less likely to report having an argument. Individuals who reported more Extraversion, Conscientiousness, Agreeableness, or Openness on average reported more positive affect and energy, and less negative affect, loneliness, and stress. In contrast to the positive valence of the other average personality states, individuals who reported more Neuroticism overall reported less positive affect and energy, and more negative affect, loneliness, and stress. Those who reported more Neuroticism were also more likely to report having an argument and receiving emotional support and were less likely to report engaging in leisure activities.

Turning to correlations at the within-person level, we found that on days people reported more Extraversion than was typical of them, they were more likely than their usual to report drinking alcohol, socializing, running errands, going to work, and exercising. When people reported more Agreeableness, they were less likely to report having an argument and more likely to report drinking alcohol, engaging in leisure activities, and exercising. On days people reported higher Conscientiousness, they were also more likely to report going to work and exercising but were less likely to report drinking alcohol than their usual. On days people reported more Neuroticism, they were more likely than their usual to report having an argument and working and were less likely to report engaging in leisure activities and drinking alcohol. When people reported more Openness, they were more likely to report exercising and socializing the same day.

Personality states were valenced at the within-person level like at the between-person level; on days when people reported more Extraversion, Agreeableness, or Openness than their usual, they reported more positive affect and energy, and less negative affect, loneliness, and stress whereas on days people reported higher Neuroticism they reported more negative affect, stress, and loneliness. In contrast to the between-person level, however, when people reported more Conscientiousness than is typical of them, they reported somewhat more positive affect, but not less negative affect or stress.

**Nomological Homomorphy With the BFI-2/NEO-PI-R.**—For the final construct validity analyses, we compared the nomological nets of our personality state subscales and BFI-2 and NEO-PI-R trait scales. These analyses provide distinct, psychometric information from convergent validity (i.e., correlations between personality state scales and global trait assessments) by indexing whether our personality state scale and two global trait measures have a comparable pattern of associations with another set of variables. This was done by first estimating a vector of correlations between a personality state subscale and all external variables, and a vector of correlations between the corresponding BFI-2/NEO-PI-R trait and those same external variables. Then, we correlated these resulting vector “profiles” to evaluate their nomological homomorphy. Higher correlations indicate that the state and trait measures of personality are tapping similar constructs.

Table 8 shows the vector correlations between BFI-2/NEO-PI-R traits and both levels of personality states. The nomological profiles between corresponding average personality states and traits had a mean correlation of .90 (range = .74-.98) with the BFI-2 in S2 and .81 (range = .47-.99) with the NEO-PI-R in S3. Within-person personality state variance had lower nomological homomorphy with the trait measures, but with the exception of state Agreeableness and Openness and their NEO-PI-R counterparts, all vector correlations were significant ( $ps < .001$ ). The average vector correlation at the within-person level was .77 (range = .64-.88) with the BFI-2 traits and .66 (range = .20-.90) with the NEO-PI-R. At both levels, Extraversion, Neuroticism, and Conscientiousness showed the most nomological homomorphy, whereas Agreeableness and Openness showed relatively less.

### Ten-Item Scale Structural and Construct Validity

*Mplus* output and a table of model fit indices for CFAs using different combinations of candidate items for the 10-item scale are available in the online supplementary files. Approximately half of the models tested were non-positive definite, likely due to collinearity of certain items (e.g., Relaxed to Tense and At Ease to Nervous). Of models that were positive definite, all achieved adequate model fit, so the final items were selected to maximize content coverage. These items are marked with superscript in Table 1. Before calculating subscale scores, we removed outliers following the same criteria used for the 20-item scale, for a total of 1,039 participants and 7,486 days.<sup>2</sup> Intercorrelations between subscales were similar to the 20-item scale as shown in Table 4.

---

<sup>2</sup>The sample size differs from the 20-item scale analyses because we removed outliers based on variability, and the item-level variability was slightly different using the 10-item scale.

Descriptive statistics, omegas, and ICCs for the brief personality state subscales are shown in Table 2. The average reliabilities of these scales were lower than the 20-item scales ( $\omega_{\text{between}} = .79$ ,  $\omega_{\text{within}} = .43$ ). Within- and between-person variability of the 10-item scales was comparable to the 20-item scales with an average ICC of .57 (range = .43-.79).

Table 5 shows the correlations between the brief subscales and the BFI-2/NEO-PI-R. Convergence between methods was comparable to the 20-item scale with the highest correlations between daily personality subscales and corresponding traits. Like the 20-item scale, cross-trait correlations and multiple regression results suggested a strong general factor as seen in Table 6.

We compared the content coverage of 10- and 20-item scales by indexing their nomological homomorphy. This was done by applying the same vector correlation approach described previously; that is, we took the vector of correlations between each personality state subscale and external variables in the 10-item and 20-item versions separately then correlated the vector profiles. We found there were only negligible differences in correlations with the set of external variables. The average nomological vector correlation between the 10- and 20-item daily personality subscales was .99 between-persons and .97 within-persons indicating the reduced set of items captured a nearly identical nomological net at both levels. These vector correlations and the 10-item scale correlations with external variables are reported in the online supplementary material.

**Nomological Vector Similarities With the BFI-2/NEO-PI-R.**—Vector correlations between the 10-item scale and the BFI-2 and NEO-PI-R can be found in Table 8. The average vector correlations differed by .01 or less from the 20-item scale.

## Discussion

Although there is much enthusiasm for using ILDs in naturalistic settings to test theories integrating personality trait structure with processes, the reliability and validity of available scales for studying both the between- and within-person levels of analysis has yet to be established. In this study, we evaluated the psychometrics of two versions of a personality state scale for daily assessment like those in active use. Across three samples, we found initial evidence that the ILD scales have the potential to validly measure Big Five traits and states, while also identifying several notable limitations. We found alignment between the Big Five traits and personality states, and the ILD scales captured a similar nomological net as two, widely used personality trait measures. However, we also found that the scales had low reliability, and the magnitude of correlations with criterion variables were small. Because we did not set out to construct an authoritative ILD instrument from the ground-up, these limitations should be addressed with future concerted efforts following psychological scale development best practices in future research (see, e.g., Clark & Watson, 2019, for a general overview on scale construction; Horstmann & Ziegler, 2020, for recommendations on personality state measurement; Wright & Zimmermann, 2019, for ILD scale development). In addition to the general recommendation of more psychometric work, we offer specific suggestions based on our results for the next steps needed to mitigate these limitations when using personality state measures.

## Differences Between Trait and State Levels

Parity in the structures and nomological nets at the between- and within-person levels supports the possibility of using these ILD scales to investigate links between personality traits and state-level processes. Areas of divergence between levels may have theoretical importance as well by revealing the nature of functional trait–state relationships proposed by personality theories. For instance, in line with work showing robust associations between Conscientiousness and positive life outcomes (Roberts et al., 2007), we found that individuals who reported more Conscientiousness on average tended to report less negative affect and stress. However, on days people were more Conscientious than was typical of them, they did not report less negative affect and stress than usual. This suggests that being more Conscientious in general is adaptive, but instances of elevated Conscientiousness are not necessarily pleasant. These results are consistent with theories that propose a willingness/ability to sacrifice immediate rewards for long-term goal pursuit is a core mechanism underlying individual differences in Conscientiousness (Carver, 2016; DeYoung, 2015) and points to the rich opportunity for testing theoretical models like these with ILDs that allow separation of between- and within-person levels. Indeed, Beckmann et al. (2010) have illustrated this aspect of Conscientiousness in the within-person association with negative affect in managers.

Other discrepancies between levels have less apparent theoretical grounding and provoke more questions than they provide answers. In particular, Agreeableness items formed a clear between-person factor but did not cohere within-person. Is this result because instances of warmth, generosity, trust, and politeness follow separate, day-to-day processes despite strongly covarying between people? Or is the issue methodological? Items validated for assessing traits cannot a priori be assumed to validly measure states, so it is possible that the items we used, which were adapted from trait measures, are poor indicators of state Agreeableness. For example, it makes conceptual sense that warmth may be more context-dependent and thus a better state indicator, whereas generosity may be more trait-like and not as effective for measuring states (and indeed warmth had more within-person variance than generosity [ $ICC_{\text{warmth}} = .39$ ,  $ICC_{\text{generosity}} = .53$ ]; see online supplementary materials for all item-level ICCs). Underscoring the possibility that state Agreeableness is especially distinct from its trait counterpart, we found it had low nomological homomorphy with BFI-2/NEO-PI-R Agreeableness, consistent with previous work (Rauthmann et al., 2019). Concerted item construction efforts are needed to go beyond simply changing the instructions for trait measures to determine what items are best for assessing Agreeableness states.

Related to the role of context, our sampling schedule may not have been well-suited for assessing variance (or covariance) in Agreeableness, as shifts in this state may be somewhat limited to interpersonal situations. Although all personality states are thought to fluctuate in response to situational cues, other traits may have more manifestations that extend beyond the duration of an event (e.g., a social interaction) such as state Extraversion and Neuroticism reflected in day-to-day mood changes, and Conscientiousness in day-to-day motivation levels. Researchers could use event-contingent designs to sample social interactions and see whether homing in on a proximal context results in a more

coherent within-person Agreeableness factor. If this were the case, it would suggest that within-person changes in Agreeableness occur primarily in interpersonal contexts and would help explain the weak factor structure we found sampling at the daily level. Alternatively, it may be that state Agreeableness does persist beyond social situations, but has more intrapersonal components (e.g., cognitive or affective components like thinking prosocial thoughts, reflecting warmly on relationships) that are not well captured by the current set of items.

### Low Scale Reliability

Judging reliability of personality state scales is difficult without accepted guidelines for omega coefficients; but, according to conventional standards of internal consistency (i.e., alpha coefficients), reliability for the 20-item scale was good at the between-person level and poor at the within-person level, and reliability for the 10-item scale was moderate to poor at both levels. This shortcoming is not unique to the personality state scales we used, as low reliability is commonly reported in ILD research (e.g., Halvorson et al., 2021; Zimmerman et al., 2019; Zygari et al., 2018). One reason for generally low reliability of ILD instruments is the premium placed on scale brevity. Although ultra-brief scales with comparable reliabilities to what we found in our study have been regularly used in the published individual differences literature (e.g., Donnellan et al., 2006), a scale's validity is inherently limited by its reliability. In addition to the impact of scale length, within-person scale variance tends to have particularly low reliability because substantive day-to-day variability of interest *and* unsystematic residual (i.e., error) variance are incorporated into this part of the model. The possible consequences of error variance on internal consistency becomes even greater for scales with a small number of items like those used in ILDs.

To maximize the potential of personality state measurement without sacrificing the defining advantage of sampling frequency in ILDs, researchers can improve reliability by assessing fewer constructs in each survey but with more items per construct. For instance, studies may need to focus on measuring one or two personality states over the ILD protocol or randomly assess a different trait at each survey. Another consideration is that the items we administered were not selected following best practices of scale construction. Future scale development efforts can increase reliability by following standard procedures of auditioning a large pool of items for each personality state scale and narrowing them down to those that achieve good internal consistency.

### Do the Magnitude of Correlations Represent True Effects or Method Effects?

An unresolved question from our analyses is whether the effect sizes we observed, which could be considered small in magnitude (mean significant correlations  $r_{\text{within-person}} = .15$ ,  $r_{\text{between-person}} = .32$ ), reflect the true relationship between personality states and the external criterion or byproducts of methodological choices. Our results are not unprecedented; comparable effect sizes and the finding of relatively smaller associations at the within-versus between-level are often reported in ILD studies of other state constructs (e.g., Giacomini & Jordan, 2016; Halvorson et al., 2021; Ringwald & Wright, 2021), and the correlation between personality and external criterion, regardless of method, tends to be around the .10 to .30 range we found (Gignac & Szodorai, 2016; Richard et al., 2003).



Conceptually, because of the multifinality of personality, we might expect correlations with a particular behavior, thought, or emotional state to be small. For example, at the between-person level, some people's high level of Conscientiousness manifests in keeping a tidy house but for others it manifests in maintaining a strict exercise regimen. At the within-person level, there is even more variability as some *days* a person's Conscientiousness manifests in doing housework, whereas on other days it manifests in exercising. This variability in state expression is also part of the reason the within-person structure of personality tends to be less well-defined than the between-person structure (e.g., Edershile et al., 2019; Wright & Simms, 2016). Thus, despite clear regularities in the types of behavior, thoughts, and emotions that covary with personality states and traits, the true correlation with any specific instantiation may be small in magnitude and the within-person factor loadings may be relatively weak.

However, there needs to be more systematic work comparing different sampling schedules and scale designs to determine whether the effect sizes could be stronger with other methods. We measured personality once per day, but personality states fluctuate within a day, so a more frequent sampling rate could draw a more direct link between personality and other state-level variables and result in stronger correlations. In support of this possibility, there is some evidence that stronger associations would be found with shorter lags between state measurements (Rauthmann et al., 2016). Another potential way to tighten the relationship between states would be to ask participants to rate their personality states in relation to a circumscribed situation (e.g., "how would you describe yourself during the social interaction") rather than over a particular time interval (e.g., "how would you describe yourself in the past 24 hours") as has been common practice in this literature. We suggested event-contingent sampling methods for testing the context-specificity of Agreeableness, but this approach could be broadly applied to strengthen observed correlations by assessing personality states during theoretically trait-specific situations (e.g., Conscientiousness when engaging in a difficult task, or Extraversion when drinking alcohol).

The effect sizes we found might also be attenuated due to the noted low reliability of the personality state scales *and* the low reliability of the external criteria, which were almost all single item measures in our study. In addition to designing studies to accommodate more items per construct as discussed earlier, using analytic approaches that reduce error variance, like latent variable modeling that isolates the reliable shared variance (e.g., multilevel structural equation modeling; Sadikaj et al., 2021), could help detect meaningful effects with brief ILD scales. As a proof-of-concept, we reanalyzed some of the correlations in our study using latent scores and found consistent increases in the size of significant effects (differences ranged from  $|.02|$  to  $|.11|$ ). These supplementary analyses suggest that (a) low reliability indeed contributes to the small magnitude of correlations and (b) latent variable modeling could help address the impact of poor reliability inherent to brief scales.

An issue related to the size and meaningfulness of associations with personality state measures is the question of whether they provide unique information from trait measures. Establishing the incremental validity of state measures for theory-testing purposes requires showing they are better predictors of external criteria than trait measures (i.e., have stronger associations) *because they are tapping the state construct*. Determining if incremental

associations are attributable to the target construct, in turn, means ruling out the possibility that extraneous aspects of the method are driving the associations. For example, if a person completed a trait assessment every day, the average of those reports could incrementally predict outcomes over a single trait measure because of the precision gained from repeated measures, not because it tapped a distinct construct. The approaches discussed previously to unconfound method and construct by isolating the effects of different instructions, sampling schedules, and items can all be used to ensure incremental validity is established for a personality state construct as measured by a given scale.

### The General Personality Factor

Although we found a number of specific associations that distinguish personality states from one another and that align with the nomological nets of corresponding traits (e.g., Neuroticism with arguments, Conscientiousness with working), there was also evidence for a strong general factor. The ILD scales had high intercorrelations, and they had high correlations with nearly all BFI-2/NEO-PI-R traits, not just the personality state's trait equivalent. After adjusting for their shared variance, however, personality states uniquely and almost exclusively associated with the trait equivalent. There is little precedence with which to compare these results as few Big Five personality state studies reported scale intercorrelations or correlations with global trait measures. Two studies that used comparable, adjective-based ILD personality state measures showed more discriminability with BFI traits (Matz & Harari, 2020; Sherman et al., 2015) and among state scales (Matz & Harari, 2020) than we found, but another study reported intercorrelations among state scales similar to ours (Bleidorn, 2009). Our personality state scales also had weak discriminability with some external variables, such that the “adaptive” traits (i.e., Extraversion, Conscientiousness, Agreeableness, and Openness) associated with nearly all positively valenced outcomes and Neuroticism with nearly all negatively valenced ones.

Taken together, the personality state scores seem to measure valid, trait-relevant variance as well as nonspecific but valenced variance. This valenced, general factor of personality is not unique to our ILD scales as a large body of research suggests it is a common feature of self-report personality trait measures (Irwing, 2013; Revelle & Wilt, 2013). There is disagreement about whether personality scales of similar valence tend to cluster together for substantive or artifactual reasons. Some researchers have speculated that the general factor is an indicator of social adjustment (van der Linden et al., 2016), but many studies suggest it is primarily the product of rater-specific variance (Anusic et al., 2009; DeYoung, 2006; Gnambs, 2013; McCrae et al., 2008) or social desirability/evaluative bias (Bäckström, 2007; Pettersson et al., 2012).

The contribution of these sources of bias found in trait assessments of personality has yet to be thoroughly investigated in state assessments. Validation of personality state scales would benefit from undergoing the same empirical scrutiny as trait measures such as collecting state-level informant reports of personality in dyadic ILD studies to isolate rater-specific variance or directly estimating a latent general factor from personality states and examining its correlates. It has been shown for trait assessments that changing the wording of items to be more neutral attenuates their common variance (Bäckström et al., 2009), so testing the

structure of states using items that de-emphasize valence could be another way to achieve better differentiation. A challenge for state assessment in particular is that affect is a core component of some personality states, meaning the strong valence may accurately reflect the construct rather than being a nuisance to correct for.

The temporal dimension of ILDs also adds new information that cannot be gleaned from cross-sectional measures, but that could bear on interpretation of the general factor. In our study, personality state and baseline trait ratings of the middle-aged sample had a stronger general factor, but also less within-person *variability* than the other two samples (see online supplemental material for sample specific results). Previous work has similarly shown older adults are less variable in some personality states than younger adults (Nofhle & Fleeson, 2010). This could be evidence that the general factor reflects social adjustment because, from a developmental view of personality maturation, older adults tend to be more emotionally and behaviorally stable (i.e., less variable from day-to-day) than younger adults (Roberts et al., 2006). On the other hand, there is research showing older adults rate their personality as more “normal” than younger adults (Wood et al., 2007), so the low variability and strong general factor could be evidence for acquiescent responding. The relationship between variability and structure of personality states has psychometric and theoretical implications that should be explored further with ILDs.

## Conclusion

As personality assessment expands beyond descriptive, between-person trait structure to include explanatory, within-person processes it is important to ensure the methods being introduced for this purpose are psychometrically sound. Our study adds to this necessary foundation by showing initial support for the validity of two versions of a personality state scale that generalized across individuals from emerging adulthood to midlife in three independent samples. We examined the scale’s construct validity by testing associations with a wide array of behavioral, cognitive, and affective variables and found preliminary evidence that the 20-item and 10-item scales capture similar content as well-validated, global personality inventories. At the same time, our evaluation of these personality state scales for ILD identifies a pressing need for additional psychometric work using these methods to fulfill their promise of advancing personality science.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by grants from the National Institute on Aging (R01 AG056043), National Institute of Diabetes and Digestive and Kidney Diseases (R01 DK110041), and the University of Pittsburgh’s Clinical and Translational Science Institute, which is funded by the National Institutes of Health Clinical and Translational Science Award program (UL1 TR001857). The opinions expressed are solely those of the authors and not those of the funding source.

## References

- Anusic I, Schimmack U, Pinkus RT, & Lockwood P (2009). The nature and structure of correlations among big five ratings: The halo-alpha-beta model. *Journal of Personality and Social Psychology*, 97(6), 1142–1156. 10.1037/a0017159 [PubMed: 19968424]
- Aschwanden D, Luchetti M, & Allemand M (2019). Are open and neurotic behaviors related to cognitive behaviors in daily life of older adults? *Journal of Personality*, 87(3), 472–484. 10.1111/jopy.12409 [PubMed: 30047126]
- Bäckström M (2007). Higher-order factors in a five-factor personality inventory and its relation to social desirability. *European Journal of Psychological Assessment*, 23(2), 63–70. 10.1027/1015-5759.23.2.63
- Bäckström M, Björklund F, & Larsson MR (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality*, 43(3), 335–344. 10.1016/j.jrp.2008.12.013
- Barta WD, Tennen H, & Litt MD (2012). Measurement reactivity in diary research. In Mehl MR & Conner TS (Eds.), *Handbook of research methods for studying daily life* (pp. 108–123). Guilford Press.
- Baumert A, Schmitt M, Perugini M, Johnson W, Blum G, Borkenau P, Costantini G, Denissen JJA, Fleeson W, Grafton B, Jayawickreme E, Kurzius E, MacLeod C, Miller LC, Read SJ, Roberts B, Robinson MD, Wood D, Wrzus C, & Möttus R (2017). Integrating personality structure, personality process, and personality development. *European Journal of Personality*, 31(5), 503–528. 10.1002/per.2115
- Beckmann N, Wood RE, & Minbashian A (2010). It depends how you look at it: On the relationship between neuroticism and conscientiousness at the within-and the between-person levels of analysis. *Journal of Research in Personality*, 44(5), 593–601. 10.1016/j.jrp.2010.07.004
- Bleidorn W (2009). Linking personality states, current social roles and major life goals. *European Journal of Personality*, 23(6), 509–530. 10.1002/per.731
- Borkenau P, & Ostendorf F (1998). The Big Five as states: How useful is the five-factor model to describe intraindividual variations over time? *Journal of Research in Personality*, 32(2), 202–221. 10.1006/jrpe.1997.2206
- Carver CS (2016). Impulse and constraint: Perspectives from personality psychology, convergence with theory in other areas, and potential for integration. *Personality and Social Psychology Review*, 9(4), 312–333. 10.1207/s15327957pspr0904\_2
- Cervone D (2005). Personality architecture: Within-person structures and processes. *Annual Review of Psychology*, 56(1), 423–452. 10.1146/annurev.psych.56.091103.070133
- Clark LA, & Watson D (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, 31(12), 1412–1427. 10.1037/pas0000626 [PubMed: 30896212]
- Cohen S, Kamarck T, & Mermelstein R (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24(4), 386–396. 10.2307/2136404
- Costa PT, & McCrae RR (1992). Normal personality assessment in clinical practice: The NEO personality inventory. *Psychological Assessment*, 4(1), 5–13. 10.1037/1040-3590.4.1.5
- DeYoung CG (2006). Higher-order factors of the big five in a multi-informant sample. *Journal of Personality and Social Psychology*, 91(6), 1138–1151. 10.1037/0022-3514.91.6.1138 [PubMed: 17144770]
- DeYoung CG (2015). Cybernetic big five theory. *Journal of Research in Personality*, 56(June), 33–58. 10.1016/j.jrp.2014.07.004
- Donnellan MB, Oswald FL, Baird BM, & Lucas RE (2006). The Mini-IPIP Scales: Tiny-Yet-Effective Measures of the Big Five factors of personality. *Psychological Assessment*, 18(2), 192–203. 10.1037/1040-3590.18.2.192 [PubMed: 16768595]
- Edershile EA, Woods WC, Sharpe BM, Crowe ML, Miller JD, & Wright AGC (2019). A day in the life of narcissus: Measuring narcissistic grandiosity and vulnerability in daily life. *Psychological Assessment*, 31(7), 913–924. 10.1037/pas0000717 [PubMed: 30883153]

- Fleeson W (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, 80(6), 1011–1027. 10.1037/0022-3514.80.6.1011 [PubMed: 11414368]
- Fleeson W (2007). Situation-based contingencies underlying trait-content manifestation in behavior. *Journal of Personality*, 75(4), 825–862. 10.1111/j.1467-6494.2007.00458.x [PubMed: 17576360]
- Fleeson W, & Gallagher P (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology*, 97(6), 1097–1114. 10.1037/a0016786 [PubMed: 19968421]
- Fleeson W, & Jayawickreme E (2015). Whole trait theory. *Journal of Research in Personality*, 56(June), 82–92. 10.1016/j.jrp.2014.10.009 [PubMed: 26097268]
- Fleeson W, & Law M (2015). Trait enactments as density distributions: The role of actors, situations, and observers in explaining stability and variability. *Journal of Personality and Social Psychology*, 109(6), 1090–1104. 10.1037/a0039517 [PubMed: 26348598]
- Geldhof GJ, Preacher KJ, & Zyphur MJ (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, 19(1), 72–91. 10.1037/a0032138 [PubMed: 23646988]
- Giacomin M, & Jordan CH (2016). The wax and wane of narcissism: Grandiose narcissism as a process or state. *Journal of personality*, 84(2), 154–164. 10.1111/jopy.12148 [PubMed: 25388437]
- Gignac S, & Szodorai ET (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102(November), 74–78. 10.1016/j.paid.2016.06.069
- Gnams T (2013). The elusive general factor of personality: The acquaintance effect. *European Journal of Personality*, 27(5), 507–520. 10.1002/per.1933
- Gosling SD, Rentfrow PJ, & Swann WB Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), 504–528. 10.1016/S0092-6566(03)00046-1
- Halvorson MA, Pedersen SL, Feil MC, Lengua LJ, Molina B, & King KM (2021). Impulsive states and impulsive traits: A Study of the multilevel structure and validity of a multifaceted measure of impulsive states. *Assessment*, 28(3), 796–812. [PubMed: 32680433]
- Hampson SE (2012). Personality processes: Mechanisms by which personality traits “Get outside the skin.” *Annual Review of Psychology*, 63(1), 315–339. 10.1146/annurev-psych-120710-100419
- Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O’Neal L, McLeod L, Delacqua G, Delacqua F, Kirby J, & Duda SN (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, 95(July), Article 103208. 10.1016/j.jbi.2019.103208
- Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, & Conde JG (2009). Research electronic data capture (REDCap): A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377–381. 10.1016/j.jbi.2008.08.010 [PubMed: 18929686]
- Horstmann KT, & Ziegler M (2020). Assessing personality states: What to consider when constructing personality state measures. *European Journal of Personality*, 34(6), 1037–1059. 10.1002/per.2266
- Hu LT, & Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. 10.1080/10705519909540118
- Irwing P (2013). The general factor of personality: Substance or artefact? *Personality and Individual Differences*, 55(3), 234–242. 10.1016/j.paid.2013.03.002
- John-Henderson NA, Marsland AL, Kamarck TW, Muldoon MF, & Manuck SB (2016). Childhood socioeconomic status and the occurrence of recent negative life events as predictors of circulating and stimulated levels of Interleukin-6. *Psychosomatic Medicine*, 78(1), 91–101. 10.1097/PSY.000000000000262 [PubMed: 26727383]
- Jones AB, Brown NA, Serfass DG, & Sherman RA (2017). Personality and density distributions of behavior, emotions, and situations. *Journal of Research in Personality*, 69(August), 225–236. 10.1016/j.jrp.2016.10.006
- Manuck SB, Phillips JE, Gianaros PJ, Flory JD, & Muldoon MF (2010). Subjective socioeconomic status and presence of the metabolic syndrome in midlife community volunteers. *Psychosomatic Medicine*, 72(1), 35–45. 10.1097/PSY.0b013e3181c484dc [PubMed: 19933505]

- Matz SC, & Harari GM (2020). Personality–place transactions: Mapping the relationships between big five personality traits, states, and daily places. *Journal of Personality and Social Psychology*. Advance online publication. 10.1037/pspp0000297
- McCabe KO, & Fleeson W (2016). Are traits useful? Explaining trait manifestations as tools in the pursuit of goals. *Journal of Personality and Social Psychology*, 110(2), 287–301. 10.1037/a0039490 [PubMed: 26280839]
- McCrae RR, Yamagata S, Jang KL, Riemann R, Ando J, Ono Y, Angleitner A, & Spinath FM (2008). Substance and artifact in the higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, 95(2), 442–455. 10.1037/0022-3514.95.2.442 [PubMed: 18665712]
- MetricWire Inc. (2019). MetricWire (Version 4.2.8). [Mobile application software]. <https://metricwire.com/>
- Muthén LK, & Muthén BO (2017). Mplus: Statistical analysis with latent variables: User’s guide (Version 8). Muthén & Muthén.
- Noftle EE, & Fleeson W (2010). Age differences in big five behavior averages and variabilities across the adult life span: Moving beyond retrospective, global summary accounts of personality. *Psychology and Aging*, 25(1), 95–107. 10.1037/a0018199 [PubMed: 20230131]
- Pettersson E, Turkheimer E, Horn EE, & Menatti AR (2012). The general factor of personality and evaluation. *European Journal of Personality*, 26(3), 292–302. 10.1002/per.839
- Pincus AL, & Ansell EB (2013). Interpersonal theory of personality. In Suls J & Tennen H (Eds.), *Handbook of psychology Vol. 5: Personality and social psychology* (2nd ed., pp. 141–159). Wiley.
- Rauthmann JF, Jones AB, & Sherman RA (2016). Directionality of person–situation transactions: Are there spillovers among and between situation experiences and personality states? *Personality and Social Psychology Bulletin*, 42(7), 893–909. 10.1177/0146167216647360 [PubMed: 27229678]
- Rauthmann JF, Horstmann KT, & Sherman RA (2019). Do self-reported traits and aggregated states capture the same thing? A nomological perspective on trait-state homomorphy. *Social Psychological and Personality Science*, 10(5), 596–611. 10.1177/1948550618774772
- Read SJ, Smith BJ, Droutman V, & Miller LC (2017). Virtual personalities: Using computational modeling to understand within-person variability. *Journal of Research in Personality*, 69(August), 237–249. 10.1016/j.jrp.2016.10.005 [PubMed: 28781390]
- Revelle W, & Wilt J (2013). The general factor of personality: A general critique. *Journal of Research in Personality*, 47(5), 493–504. 10.1016/j.jrp.2013.04.012 [PubMed: 23956474]
- Richard FD, Bond CF, & Stokes-Zoota JJ (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7(4), 331–363. 10.1037/1089-2680.7.4.331
- Ringwald WR, & Wright AGC (2021). The affiliative role of empathy in everyday interpersonal interactions. *European Journal of Personality*, 35(2), 197–211. 10.1002/per.2286 [PubMed: 34970022]
- Roberts BW, Kuncel NR, Shiner R, Caspi A, & Goldberg LR (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2(4), 313–345. 10.1111/j.1745-6916.2007.00047.x [PubMed: 26151971]
- Roberts BW, Walton KE, & Viechtbauer W (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, 132(1), 1–25. 10.1037/0033-2909.132.1.1 [PubMed: 16435954]
- Team RStudio. (2020). RStudio: Integrated development for R. <http://www.rstudio.com/>
- Russell D (1996). UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *Journal of Personality Assessment*, 66(1), 20–40. 10.1207/s15327752jpa6601\_2 [PubMed: 8576833]
- Sadikaj G, Wright AGC, Dunkley D, Zuroff D, & Moskowitz DS (2021). Multilevel structural equation modeling for intensive longitudinal data: A practical guide for personality researchers. In Rauthmann JF (Ed.), *Handbook of personality dynamics and processes* (pp. 856–882). Elsevier.
- Schönbrodt FD, & Perugini M (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609–612. 10.1016/j.jrp.2013.05.009



- Sherman RA, Rauthmann JF, Brown NA, Serfass DG, & Jones AB (2015). The independent effects of personality and situations on real-time expressions of behavior and emotion. *Journal of Personality and Social Psychology*, 109(5), 872–888. 10.1037/pspp0000036 [PubMed: 25915131]
- Soto CJ, & John OP (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, 113(1), 117–143. 10.1037/pspp0000096 [PubMed: 27055049]
- Thalmayer A, Saucier G, & Eigenhuis A (2011). Comparative Validity of Brief to Medium-Length Big Five and Big Six Personality Questionnaires. *Psychological Assessment*, 23(4), 995–1009. 10.1037/a0024165 [PubMed: 21859221]
- van der Linden D, Dunkel CS, & Petrides KV (2016). The general factor of personality (GFP) as social effectiveness: Review of the literature. *Personality and Individual Differences*, 101(October), 98–105. 10.1016/j.paid.2016.05.020
- Watson D, Clark LA, & Tellegen A (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. 10.1037/0022-3514.54.6.1063 [PubMed: 3397865]
- Wood D, Gosling SD, & Potter J (2007). Normality evaluations and their relation to personality traits and well-being. *Journal of Personality and Social Psychology*, 93(5), 861–879. 10.1037/0022-3514.93.5.861 [PubMed: 17983305]
- Wright AGC, & Simms LJ (2016). Stability and fluctuation of personality disorder features in daily life. *Journal of Abnormal Psychology*, 125(5), 641–656. 10.1037/abn0000169 [PubMed: 27196437]
- Wright AGC, & Zimmermann J (2019). Applied ambulatory assessment: Integrating idiographic and nomothetic principles of measurement. *Psychological Assessment*, 31(12), 1467–1480. 10.1037/pas0000685 [PubMed: 30896209]
- Zygar C, Hagemeyer B, Pusch S, & Schönbrodt FD (2018). From motive dispositions to states to outcomes: An intensive experience sampling study on communal motivational dynamics in couples. *European Journal of Personality*, 32(3), 306–324. 10.1002/per.2145

**Table 1.**

Items Used to Assess State Constructs.

State construct	Items
<i>Personality</i>	
Extraversion	Lethargic/Energetic <sup>a</sup> , Talkative/Silent <sup>a</sup> , Bold/Timid, Unassertive/Assertive
Neuroticism	Insecure/Secure <sup>a</sup> , Relaxed/Tense <sup>a</sup> , Nervous/At Ease, Unexcitable/Excitable
Conscientiousness	Careless/Thorough <sup>a</sup> , Hardworking/Lazy <sup>a</sup> , Serious/Frivolous, Inefficient/Efficient
Agreeableness	Stingy/Generous <sup>a</sup> , Distrustful/Trustful <sup>a</sup> , Warm/Cold, Rude/Polite
Openness	Unimaginative/Imaginative <sup>a</sup> , Uninquisitive/Curious <sup>a</sup> , Conventional/Creative, Imperceptive/Perceptive
<i>Behaviors</i>	
	Did you drink any alcoholic beverages today?
	Did you engage in any physical activity, long enough to work up a sweat, get your heart thumping, or get out of breath?
	Did you receive emotional support (e.g., listening to you, comforting you) from anyone?
	Did you have social contact with anyone, either at your home or away from your home?
	Did you engage in any leisure activities today (watching TV, listening to music, playing a game, going to a movie or sporting event, taking a walk, hike or run, etc.)?
	Did you do any work around the house or yard?
	Did you do any family or personal errands away from home (grocery shopping, doctor appointment, etc.)?
	Did you go to work?
	Did you engage in work-related activities that involved other people (e.g., attended a meeting; collaborated on a project, engage in a task-related discussion)?
	Did you socialize with anyone at work or during your commute?
	How often did you feel isolated?
	How often did you feel left out?
<i>Affect</i>	
Positive affect	Happy, Excited, and Relaxed (S1 and S2) Happy, Calm, Full of Pep, Lively, Cheerful, and At Ease (S3)
Negative affect	Nervous, Sad, and Angry (S1 and S2) Sad, Hostile, On Edge, Angry, Tense, and Unhappy (S3)
<i>Energy</i>	
	How often did you feel tired today?
	How often did you run out of energy today?
<i>Loneliness</i>	
	How often did you feel lonely? How often did you feel isolated? How often did you feel left out?
<i>Stress</i>	
	How often did you feel like you were unable to control the important things in your life?
	How often did you feel confident about your ability to handle your personal problems?
	How often did you feel like things were going your way?
	How often did you feel like difficulties were piling up so high that you could not overcome them?

*Note.* All items were assessed at the daily level except for affect in S1 and S2, which was measured multiple times per day and averaged for a daily score. S1 = Sample 1; S2 = Sample 2; S3 = Sample 3.

<sup>a</sup>Indicates the item was used in the 10-item scale.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2.** Factor Loadings for Multilevel Exploratory Factor Analysis of Personality State Items.

Item	Extraversion	Neuroticism	Conscientiousness	Agreeableness	Openness
<i>Between-person</i>					
Energetic	.13	<u>-.35</u>	<u>.44</u>	.05	.08
Bold	<u>.73</u>	-.12	.12	.01	.01
Talkative	<u>.66</u>	.01	.02	.29	-.09
Assertive	<u>.62</u>	-.09	<u>.30</u>	-.04	.10
Insecure	-.18	<u>.56</u>	<u>-.31</u>	-.05	.02
Nervous	.03	<u>.99</u>	-.07	.06	-.04
Tense	.01	<u>.92</u>	.11	-.08	.02
Excitable	<u>.43</u>	.18	-.03	<u>.29</u>	.19
Thorough	-.12	-.03	<u>.85</u>	.15	.04
Hardworking	.09	-.01	<u>.83</u>	.02	-.07
Serious	-.07	.09	<u>.69</u>	-.15	-.13
Efficient	.06	-.18	<u>.72</u>	.02	.08
Warm	.14	-.04	-.03	<u>.58</u>	.02
Polite	-.27	-.02	.12	<u>.88</u>	.00
Generous	.01	.00	.19	<u>.54</u>	.14
Trustful	.04	-.22	.00	<u>.66</u>	-.01
Imaginative	-.01	-.05	.13	.00	<u>.90</u>
Curious	.04	.00	.19	.07	<u>.66</u>
Creative	-.02	.01	-.12	-.04	<u>.91</u>
Perceptive	.09	.00	<u>.38</u>	.17	<u>.37</u>
<i>Within-person</i>					
Energetic	<u>.32</u>	-.08	.26	.07	.07
Bold	<u>.32</u>	-.10	.10	-.02	.12
Talkative	<u>.42</u>	-.13	.01	.10	-.02
Assertive	.26	.04	.02	.21	.08
Insecure	.00	<u>.31</u>	-.17	.02	<u>-.35</u>
Nervous	.02	<u>.60</u>	.01	-.01	-.08

Item	Extraversion	Neuroticism	Conscientiousness	Agreeableness	Openness
Tense	-.12	<b>.68</b>	.06	.00	.07
Excitable	.23	.02	-.05	<b>.36</b>	.03
Thorough	-.02	-.01	<b>.48</b>	.04	.26
Hardworking	.22	.08	<b>.55</b>	.00	-.01
Serious	-.08	.20	.28	-.08	-.01
Efficient	.07	-.01	<b>.39</b>	<b>.37</b>	-.04
Warm	.19	-.21	.12	-.04	.05
Polite	-.14	-.21	.20	.11	.18
Generous	.05	-.12	.00	.22	.08
Trustful	-.16	-.24	.03	<b>.45</b>	.00
Imaginative	.07	.00	-.01	.00	<b>.61</b>
Curious	.01	.00	.00	.21	<b>.32</b>
Creative	.08	.05	-.13	.12	<b>.29</b>
Perceptive	.02	.01	.05	<b>.42</b>	.08

Note. Factor loadings > |.30| are bolded and underlined. Item labels refer to the positively keyed pole of the scale. Comparative fit index = .96; root mean square error of approximation = .03; standardized root mean square residual within/standardized root mean square residual between = .02/.03.

**Table 3.** Descriptive Statistics and Reliability Estimates for 20-Item and 10-Item Personality State Subscales.

Subscale	<i>M</i>	<i>SD</i>	ICC	$\omega$ Within-person	$\omega$ Between-person
Extraversion					
Twenty-item	4.55	1.06	.59	.52	.83
Ten-item	4.59	1.18	.53	.49	.72
Neuroticism					
Twenty-item	3.44	1.00	.58	.56	.83
Ten-item	3.05	1.45	.61	.45	.86
Conscientiousness					
Twenty-item	4.89	1.06	.57	.54	.84
Ten-item	4.98	1.34	.52	.50	.83
Agreeableness					
Twenty-item	5.27	0.98	.62	.44	.81
Ten-item	5.21	1.15	.61	.26	.70
Openness					
Twenty-item	4.73	1.01	.67	.51	.87
Ten-item	4.81	1.18	.60	.46	.84
Average					
Twenty-item	—	—	.61	.51	.84
Ten-item	—	—	.57	.43	.79

*Note.* Results are for the pooled sample. ICC = intraclass correlation coefficient,  $\omega$  = omega coefficient.



**Table 4.**

Correlations Among Personality States at the Within- and Between-Person Levels for the 20-Item and 10-Item Scales.

	Extraversion	Neuroticism	Conscientiousness	Agreeableness	Openness
Twenty-item scale					
Extraversion	—	<b>-.57</b>	<b>.68</b>	<b>.63</b>	<b>.60</b>
Neuroticism	<b>-.22</b>	—	<b>-.55</b>	<b>-.54</b>	<b>-.35</b>
Conscientiousness	<b>.35</b>	<b>-.10</b>	—	<b>.65</b>	<b>.55</b>
Agreeableness	<b>.34</b>	<b>-.29</b>	<b>.27</b>	—	<b>.57</b>
Openness	<b>.39</b>	<b>-.18</b>	<b>.29</b>	<b>.33</b>	—
Ten-item scale					
Extraversion	—	<b>-.61</b>	<b>.57</b>	<b>.57</b>	<b>.55</b>
Neuroticism	<b>-.25</b>	—	<b>-.65</b>	<b>-.64</b>	<b>-.50</b>
Conscientiousness	<b>.23</b>	<b>-.24</b>	—	<b>.62</b>	<b>.54</b>
Agreeableness	<b>.22</b>	<b>-.28</b>	<b>.18</b>	—	<b>.53</b>
Openness	<b>.26</b>	<b>-.31</b>	<b>.28</b>	<b>.25</b>	—

*Note.* Results are from the pooled sample. Between-person correlations are shown above the diagonal; within-person correlations are shown below the diagonal. Bolded correlations were statistically significant (i.e., the 95% credibility intervals for the point estimates did not contain zero).

**Table 5.** Correlations Between Average Personality States and Baseline Traits for the 20-Item/10-Item Scales.

Baseline traits	Average personality state				
	Extraversion	Neuroticism	Conscientiousness	Agreeableness	Openness
<i>BFI-2 Trait subscale</i>					
Extraversion					
Twenty-item	.74	-.25	.34	.29	.36
Ten-item	.74	-.45	.38	.27	.35
Negative emotionality					
Twenty-item	-.32	.63	-.22	-.32	-.23
Ten-item	-.32	.63	-.26	-.29	-.22
Conscientiousness					
Twenty-item	.36	-.19	.65	.31	.15
Ten-item	.33	-.31	.66	.26	.15
Agreeableness					
Twenty-item	.13	-.21	.23	.57	.21
Ten-item	.09	-.25	.28	.51	.17
Open mindedness					
Twenty-item	.31	-.06	.21	.24	.59
Ten-item	.25	-.14	.21	.26	.58
<i>NEO-PI-R Trait subscale</i>					
Extraversion					
Twenty-item	.47	-.15	.22	.28	.29
Ten-item	.47	-.26	.26	.28	.25
Neuroticism					
Twenty-item	-.41	.49	-.35	-.36	-.26
Ten-item	-.33	.49	-.35	-.37	-.26
Conscientiousness					
Twenty-item	.30	-.23	.44	.27	.16
Ten-item	.26	-.28	.42	.28	.17
Agreeableness					

Baseline traits	Average personality state				
	Extraversion	Neuroticism	Conscientiousness	Agreeableness	Openness
Twenty-item	<b>.13</b>	<b>-.22</b>	<b>.20</b>	<b>.36</b>	<b>.10</b>
Ten-item	<b>.11</b>	<b>-.24</b>	<b>.21</b>	<b>.36</b>	<b>.13</b>
Openness					
Twenty-item	<b>.11</b>	<b>-.02</b>	<b>.03</b>	<b>.13</b>	<b>.35</b>
Ten-item	<b>.12</b>	<b>-.05</b>	<b>.06</b>	<b>.11</b>	<b>.33</b>

*Note.* Average personality states were estimated using random intercepts in multilevel models. Values on the diagonal indicate convergent correlations between personality states and corresponding traits. Bolded values are statistically significant (i.e., the 95% credibility intervals for the point estimates do not contain zero). The BFI-2 was only administered in Sample 2, and the NEO-PI-R in Sample 3. BFI-2 = Big Five Inventory-2; NEO-PI-R = NEO Personality Inventory-Revised.

**Table 6.**

Associations Between Average Personality States and Baseline Traits Adjusted for Shared State Variance for the 20-Item/10-Item scales.

Baseline traits	Average personality state				
	Extraversion	Neuroticism	Conscientiousness	Agreeableness	Openness
<i>BFI-2 Trait subscale</i>					
Extraversion					
Twenty-item	<b>.74</b>	.02	-.03	.01	.06
Ten-item	<b>.69</b>	<b>-.12</b>	.06	.02	.06
Negative emotionality					
Twenty-item	-.02	<b>.60</b>	-.06	-.03	-.13
Ten-item	<b>.14</b>	<b>.65</b>	-.04	.02	-.07
Conscientiousness					
Twenty-item	.07	-.05	<b>.62</b>	.08	-.14
Ten-item	.09	-.08	<b>.63</b>	.06	-.11
Agreeableness					
Twenty-item	-.16	-.02	.07	<b>.57</b>	.03
Ten-item	-.13	-.05	.16	<b>.46</b>	.01
Open mindedness					
Twenty-item	.07	.03	-.03	.04	<b>.56</b>
Ten-item	.05	.08	-.02	.11	<b>.56</b>
<i>NEO-PI-R Trait subscale</i>					
Extraversion					
Twenty-item	<b>.40</b>	.05	.04	<b>.12</b>	.11
Ten-item	<b>.44</b>	-.09	.10	<b>.14</b>	.07
Neuroticism					
Twenty-item	<b>-.27</b>	<b>.49</b>	-.19	<b>-.20</b>	<b>-.22</b>
Ten-item	<b>-.19</b>	<b>.45</b>	-.17	<b>-.20</b>	<b>.19</b>
Conscientiousness					
Twenty-item	.06	.01	<b>.34</b>	.04	.01
Ten-item	.04	-.03	<b>.32</b>	.05	.03
Agreeableness					

Baseline traits	Average personality state				
	Extraversion	Neuroticism	Conscientiousness	Agreeableness	Openness
Twenty-item	<b>-.13</b>	-.04	-.01	<b>.23</b>	-.16
Ten-item	<b>-.12</b>	.01	-.01	<b>.23</b>	-.10
Openness					
Twenty-item	-.06	.04	-.05	-.03	<b>.32</b>
Ten-item	-.06	.08	-.07	-.06	<b>.31</b>

*Note.* Associations are standardized regression coefficients from multilevel multiple regression models including all personality state subscales as predictors of BFI-2/NEO-PI-R traits. Average personality states were estimated using random intercepts. Values on the diagonal indicate convergent correlations between personality states and corresponding traits. Bolded values are statistically significant (i.e., the 95% credibility intervals for the point estimates do not contain zero). The BFI-2 was only administered in Sample 2, and the NEO-PI-R in Sample 3. BFI-2 = Big Five Inventory-2; NEO-PI-R = NEO Personality Inventory-Revised.

**Table 7.** Correlations Between the Twenty-Item Personality State Subscales and External Variables.

Behaviors	Sample 1					Sample 2					Sample 3				
	E	N	C	A	O	E	N	C	A	O	E	N	C	A	O
<i>Between-person</i>															
Drank alcohol	.12	-.08	-.13	-.21	.01	.10	-.11	-.08	.04	.04	.19	-.11	.13	.14	.14
Exercised	.27	-.17	.29	.08	-.02	.14	-.06	.19	.08	.00	.23	-.10	.20	.21	.18
Received emotional support	.06	.27	.01	.11	.02	.00	.21	.03	-.1	.04	.12	.01	.07	.11	.17
Had an argument	.12	.15	-.25	-.21	-.07	.03	.22	.02	-.16	-.07	-.19	.39	-.21	-.24	-.16
Had social contact	-.15	.04	.00	.06	-.18	.08	.13	.11	.12	-.03	.25	-.19	.23	.31	.11
Engaged in leisure	-.19	-.16	-.01	.18	.00	.05	-.17	-.04	.14	.05	.07	-.21	.07	.14	.08
Did housework	.01	-.11	.17	-.01	.16	.11	-.05	.18	.07	.09	-.02	-.11	.02	.05	.08
Ran errands	-.04	-.14	.09	-.17	.13	.19	-.03	.12	.06	.02	.12	.04	.05	.10	.15
Worked/volunteered	.06	-.04	.15	.00	-.05	.19	-.24	.17	.15	.05	.11	.05	.15	.01	-.05
Worked on a team	.22	.01	.19	.13	.05	.26	-.10	.26	.04	.09	.17	-.07	.15	.12	.19
Socialized at work	.23	-.10	.21	.10	.03	.27	-.17	.17	.11	.06	.12	-.13	.16	.19	-.03
<i>Affect</i>															
Positive affect	.52	-.47	.39	.27	.25	.41	-.29	.15	.26	.30	.67	-.67	.50	.57	.48
Negative affect	-.21	.47	-.24	-.45	-.18	-.27	.62	-.24	-.4	-.15	-.39	.61	-.31	-.34	-.18
<i>Energy</i>															
Felt tired	-.41	.41	-.36	-.27	-.10	-.31	.46	-.18	-.28	-.20	-.44	.47	-.34	-.31	-.22
Ran out of energy	-.36	.39	-.4	-.39	-.20	-.32	.47	-.24	-.28	-.16	-.4	.42	-.3	-.24	-.18
<i>Loneliness</i>															
Felt lonely	-.49	.49	-.35	-.29	-.10	-.48	.65	-.34	-.43	-.21	-.4	.49	-.28	-.3	-.14
Felt isolated	-.47	.48	-.31	-.35	-.13	-.47	.62	-.33	-.43	-.18	-.41	.48	-.26	-.29	-.13
Felt left out	-.43	.48	-.29	-.36	-.18	-.45	.61	-.34	-.47	-.19	-.39	.49	-.28	-.32	-.14
<i>Stress</i>															
Unable to control things	-.41	.63	-.37	-.22	-.12	-.41	.68	-.32	-.37	-.22	-.45	.58	-.37	-.36	-.16
Confident to handle problems	.13	-.39	.26	.33	.09	.38	-.33	.27	.23	.20	.46	-.58	.40	.41	.23
Things were going your way	.28	-.51	.30	.39	.13	.48	-.51	.34	.37	.31	.58	-.64	.42	.47	.31



Behaviors	Sample 1					Sample 2					Sample 3				
	E	N	C	A	O	E	N	C	A	O	E	N	C	A	O
Difficulties piling high	<b>-.42</b>	<b>.63</b>	<b>-.40</b>	<b>-.20</b>	<b>-.16</b>	<b>-.32</b>	<b>.64</b>	<b>-.23</b>	<b>-.33</b>	<b>-.20</b>	<b>-.4</b>	<b>.57</b>	<b>-.35</b>	<b>-.32</b>	<b>-.18</b>
<i>Within-person</i>															
Drank alcohol	<b>.09</b>	<b>-.15</b>	<b>-.12</b>	<b>.09</b>	<b>.02</b>	<b>.16</b>	<b>-.12</b>	<b>-.07</b>	<b>.07</b>	<b>.07</b>	<b>.14</b>	<b>-.10</b>	<b>-.04</b>	<b>.05</b>	<b>.03</b>
Exercised	<b>.13</b>	<b>.02</b>	<b>.19</b>	<b>.09</b>	<b>.09</b>	<b>.14</b>	<b>-.06</b>	<b>.18</b>	<b>.10</b>	<b>.11</b>	<b>.12</b>	<b>-.03</b>	<b>.10</b>	<b>.05</b>	<b>.04</b>
Received emotional support	<b>-.04</b>	<b>.12</b>	<b>-.02</b>	<b>.01</b>	<b>.08</b>	<b>.08</b>	<b>.09</b>	<b>.02</b>	<b>.09</b>	<b>.00</b>	<b>.09</b>	<b>.03</b>	<b>.01</b>	<b>.02</b>	<b>.05</b>
Had an argument	<b>-.10</b>	<b>.23</b>	<b>.01</b>	<b>-.16</b>	<b>-.01</b>	<b>.01</b>	<b>.15</b>	<b>.03</b>	<b>-.24</b>	<b>-.02</b>	<b>.02</b>	<b>.21</b>	<b>.01</b>	<b>-.10</b>	<b>.01</b>
Had social contact	<b>.06</b>	<b>-.01</b>	<b>.02</b>	<b>.10</b>	<b>.10</b>	<b>.13</b>	<b>.02</b>	<b>.13</b>	<b>.06</b>	<b>.07</b>	<b>.19</b>	<b>.00</b>	<b>.04</b>	<b>.07</b>	<b>.12</b>
Engaged in leisure	<b>.12</b>	<b>-.04</b>	<b>.07</b>	<b>.14</b>	<b>.01</b>	<b>-.02</b>	<b>-.09</b>	<b>-.09</b>	<b>.11</b>	<b>.02</b>	<b>.04</b>	<b>-.10</b>	<b>-.07</b>	<b>.03</b>	<b>.00</b>
Did housework	<b>.04</b>	<b>.03</b>	<b>.06</b>	<b>.07</b>	<b>.05</b>	<b>-.01</b>	<b>-.08</b>	<b>.04</b>	<b>.06</b>	<b>.07</b>	<b>.03</b>	<b>-.02</b>	<b>.06</b>	<b>.02</b>	<b>.03</b>
Ran errands	<b>.11</b>	<b>-.06</b>	<b>.01</b>	<b>.06</b>	<b>.06</b>	<b>.07</b>	<b>-.08</b>	<b>-.04</b>	<b>.10</b>	<b>.03</b>	<b>.03</b>	<b>.03</b>	<b>.01</b>	<b>.03</b>	<b>.03</b>
Worked/volunteered	<b>.02</b>	<b>.10</b>	<b>.17</b>	<b>.00</b>	<b>.01</b>	<b>.09</b>	<b>.16</b>	<b>.32</b>	<b>.00</b>	<b>.05</b>	<b>.11</b>	<b>.06</b>	<b>.25</b>	<b>-.03</b>	<b>.07</b>
Worked on a team	<b>.04</b>	<b>.07</b>	<b>.20</b>	<b>.02</b>	<b>.03</b>	<b>.09</b>	<b>.12</b>	<b>.29</b>	<b>.01</b>	<b>.08</b>	<b>.16</b>	<b>-.05</b>	<b>.09</b>	<b>.09</b>	<b>.10</b>
Socialized at work	<b>.02</b>	<b>.11</b>	<b>.15</b>	<b>.05</b>	<b>.01</b>	<b>.11</b>	<b>.13</b>	<b>.31</b>	<b>.00</b>	<b>.06</b>	<b>.12</b>	<b>-.08</b>	<b>.03</b>	<b>.10</b>	<b>.06</b>
Affect															
Positive affect	<b>.29</b>	<b>-.28</b>	<b>.03</b>	<b>.19</b>	<b>.18</b>	<b>.27</b>	<b>-.25</b>	<b>.10</b>	<b>.26</b>	<b>.21</b>	<b>.40</b>	<b>-.27</b>	<b>.15</b>	<b>.24</b>	<b>.23</b>
Negative affect	<b>-.26</b>	<b>.30</b>	<b>-.02</b>	<b>-.23</b>	<b>-.13</b>	<b>-.14</b>	<b>.31</b>	<b>.01</b>	<b>-.23</b>	<b>-.10</b>	<b>-.17</b>	<b>.38</b>	<b>-.05</b>	<b>-.22</b>	<b>-.14</b>
Energy															
Felt tired	<b>-.21</b>	<b>.10</b>	<b>-.05</b>	<b>-.04</b>	<b>-.11</b>	<b>-.29</b>	<b>.06</b>	<b>-.13</b>	<b>-.10</b>	<b>-.13</b>	<b>-.24</b>	<b>.12</b>	<b>-.12</b>	<b>-.09</b>	<b>-.13</b>
Ran out of energy	<b>-.23</b>	<b>.08</b>	<b>-.07</b>	<b>-.06</b>	<b>-.12</b>	<b>-.29</b>	<b>.11</b>	<b>-.16</b>	<b>-.14</b>	<b>-.15</b>	<b>-.24</b>	<b>.12</b>	<b>-.13</b>	<b>-.10</b>	<b>-.13</b>
Loneliness															
Felt lonely	<b>-.26</b>	<b>.26</b>	<b>-.04</b>	<b>-.16</b>	<b>-.10</b>	<b>-.30</b>	<b>.16</b>	<b>-.12</b>	<b>-.21</b>	<b>-.15</b>	<b>-.17</b>	<b>.14</b>	<b>-.09</b>	<b>-.10</b>	<b>-.13</b>
Felt isolated	<b>-.24</b>	<b>.24</b>	<b>-.05</b>	<b>-.19</b>	<b>-.08</b>	<b>-.25</b>	<b>.15</b>	<b>-.10</b>	<b>-.21</b>	<b>-.12</b>	<b>-.14</b>	<b>.14</b>	<b>-.06</b>	<b>-.09</b>	<b>-.08</b>
Felt left out	<b>-.22</b>	<b>.20</b>	<b>-.06</b>	<b>-.14</b>	<b>-.13</b>	<b>-.17</b>	<b>.13</b>	<b>-.06</b>	<b>-.20</b>	<b>-.11</b>	<b>-.12</b>	<b>.11</b>	<b>-.09</b>	<b>-.09</b>	<b>-.09</b>
Stress															
Unable to control things	<b>-.20</b>	<b>.32</b>	<b>-.04</b>	<b>-.16</b>	<b>-.11</b>	<b>-.14</b>	<b>.21</b>	<b>-.04</b>	<b>-.17</b>	<b>-.11</b>	<b>-.10</b>	<b>.18</b>	<b>-.03</b>	<b>-.07</b>	<b>-.08</b>
Confident to handle problems	<b>.15</b>	<b>-.16</b>	<b>.06</b>	<b>.11</b>	<b>.08</b>	<b>.09</b>	<b>-.13</b>	<b>.02</b>	<b>.10</b>	<b>.08</b>	<b>.06</b>	<b>-.10</b>	<b>.05</b>	<b>.09</b>	<b>.06</b>
Things were going your way	<b>.23</b>	<b>-.28</b>	<b>.09</b>	<b>.21</b>	<b>.13</b>	<b>.18</b>	<b>-.18</b>	<b>.08</b>	<b>.17</b>	<b>.13</b>	<b>.14</b>	<b>-.15</b>	<b>.07</b>	<b>.14</b>	<b>.11</b>
Difficulties piling high	<b>-.13</b>	<b>.25</b>	<b>.02</b>	<b>-.11</b>	<b>-.06</b>	<b>-.20</b>	<b>.29</b>	<b>-.01</b>	<b>-.17</b>	<b>-.10</b>	<b>-.13</b>	<b>.17</b>	<b>-.06</b>	<b>-.11</b>	<b>-.08</b>

Note: All items were assessed at the daily level except for affect in Sample 1 and Sample 2 which was measured multiple times per day and averaged for a daily score. Personality states and external variables were decomposed into between- and within-person variance and correlations were estimated separately at each level.

Bolded correlations were statistically significant (i.e., the credibility intervals do not contain zero). E = Extraversion; N = Neuroticism; C = Conscientiousness; A = Agreeableness; O = Openness.

**Table 8.**

Nomological Vector Correlations Among Between- and Within-Person Variance in Personality States and Baseline Traits for 20-Item and 10-Item scales.

	Between-person, $r_b$	Within-person, $r_w$
<i>BFI-2</i>		
Extraversion		
Twenty-item	<b>.96</b>	<b>.89</b>
Ten-item	<b>.98</b>	<b>.89</b>
Neuroticism		
Twenty-item	<b>.98</b>	<b>.76</b>
Ten-item	<b>.98</b>	<b>.87</b>
Conscientiousness		
Twenty-item	<b>.95</b>	<b>.64</b>
Ten-item	<b>.95</b>	<b>.67</b>
Agreeableness		
Twenty-item	<b>.87</b>	<b>.88</b>
Ten-item	<b>.89</b>	<b>.77</b>
Openness		
Twenty-item	<b>.74</b>	<b>.68</b>
Ten-item	<b>.75</b>	<b>.64</b>
Average		
Twenty-item	.90	.77
Ten-item	.91	.77
<i>NEO-PI-R</i>		
Extraversion		
Twenty-item	<b>.93</b>	<b>.88</b>
Ten-item	<b>.95</b>	<b>.85</b>
Neuroticism		
Twenty-item	<b>.99</b>	<b>.90</b>
Ten-item	<b>.98</b>	<b>.91</b>
Conscientiousness		
Twenty-item	<b>.95</b>	<b>.64</b>
Ten-item	<b>.95</b>	<b>.72</b>
Agreeableness		
Twenty-item	<b>.73</b>	<b>.68</b>
Ten-item	<b>.71</b>	<b>.62</b>
Openness		
Twenty-item	.47	.20
Ten-item	.44	.24
Average		
Twenty-item	.81	.66
Ten-item	.81	.67

*Note.* Vector correlations were calculated by taking the vector of correlations between each personality state subscale and external variables, and between each baseline trait and those same external variables, then correlating the vectors of corresponding personality constructs. Between-person variance reflects each person's average personality state level and within-person variance reflects each person's average daily deviations from their average level. Bolded values indicate correlation is statistically significant ( $p < .001$ ). The BFI-2 was only administered in Sample 2, and the NEO-PI-R in Sample 3.  $r_V$  = vector correlations indexing nomological homomorphy; BFI-2 = Big Five Inventory-2; NEO-PI-R = NEO Personality Inventory-Revised.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript