Article

Assessment
2021, Vol. 28(3) 829–844

DOI: 10.1177/1073191119888582
journals.sagepub.com/home/asm
SAGE

Do State and Trait Measures Measure States and Traits? The Case of Community-Dwelling Caregivers of Older Adults

Charles E. Lance[1,2], Juliette Christie[3], and Gail M. Williamson[3†]

Abstract

Spielberger's state and trait anxiety and anger scales are widely used and documented, but there is little or no direct evidence that they actually measure their respective state and trait aspects as was intended. We conducted latent state-trait analyses on data collected from 310 community-dwelling caregivers of older adult care recipients and found that (a) both state and trait scales reflected a mixture of state and trait aspects of their latent constructs, (b) state scales reflected more state-like variance than did corresponding trait scales, but (c) both state and trait scales were dominated by stable trait-like variance. Follow-up bivariate latent state-trait analyses indicated that correlations between trait components of anger and anxiety correlated more strongly with trait components of caregiver–care recipient mutually communal behavior and care recipient problem behavior than did state–state component correlations. Implications for the measurement of state and trait components of psychological constructs are discussed.

Keywords

latent state-trait analyses, psychometric assessment, community-dwelling caregivers, older adult care recipients, anxiety, anger

Spielberger's measures of state and trait anxiety (Spielberger, 1983) and anger (Spielberger, Jacobs, Russell, & Crane, 1983) have been used thousands of times in basic and applied research and for diagnostic and classification purposes and have well-documented reliability (Barnes, Harp, & Jung, 2002; Deffenbacher et al., 1996; Eckhart, Norlander, & Deffenbacher, 2004; Therrien & Hunsley, 2013), but almost all the empirical evidence that these scales actually *do* measure state and trait components of anxiety and anger is indirect. The main purpose of the present study was to apply Cole, Martin, and Steiger's (2005) trait-state-occasion (TSO) model to data collected from a sample of community-dwelling caregivers of older adult care recipients to obtain direct assessments of whether these scales actually *do* measure discriminable state and trait components of the constructs they are designed to assess in the first place and, if they do, to determine the *extent to which* they do. As such, this study addresses a critical missing step in the validation of Spielberger's measures. As we explain in more detail later, our sample was from a growing and increasingly important population—informal caregivers of older adult care recipients (Talley et al., 2004). Thus, a secondary aim of this study was to investigate time-varying and time-invariant aspects of interpersonal and environmental factors related to the focal anxiety and anger constructs among members of this population. Specifically, we investigated relations between the focal anxiety and anger constructs and (a) a historically more stable interpersonal factor, communal behavior, and (b) a typically more labile environmental factor, care recipient problem behavior.

Background

Four outcomes of the person–situation interaction debate of the 1970s and 1980s (Epstein & O'Brien, 1985) were (a) the conclusion that both trait and state operationalizations of the same construct are possible depending on the level aggregation of measures across occasions and situations; (b) the understanding that some individual difference

[1]Organizational Research & Development, Lawrenceville, GA, USA
[2]University of the Western Cape, Cape Town, South Africa
[3]University of Georgia, Athens, GA, USA
†Deceased

Corresponding Author:
Charles E. Lance, Organizational Research & Development, 173 Crystal River Drive, Lawrenceville, GA 30043, USA.
Email: clancephd@gmail.com

variables (e.g., intelligence) represent more stable trait-like characteristics, while others (e.g., mood) reflect more labile state-like components; (c) the realization that most, if not all, human psychological attributes exhibit both state-like and trait-like aspects; and (d) a widespread acceptance of some form of person–situation interactionism in human behavior (see also Steyer, Schmitt, & Eid, 1999). As Hertzog and Nesselroade (1987) are often quoted, "Generally it is certainly the case that most psychological attributes will neither be, strictly speaking, traits or states. That is, attributes can have both trait and state components" (p. 95). A common conception is that those constructs that are rather stable temporally are seen to be more trait-like, while others that are relatively unstable are seen to be more state-like (G. Chen, Gully, Whiteman, & Kilcullen, 2000; Cole et al., 2005). More generally, traits and states can be viewed as fuzzy prototype-based categories that individuals use to describe and attempt to understand their own and others' behavior; traits are prototypically more stable, long-lasting, and internally caused, while states are seen as being more unstable, short-term, and caused by external circumstances (Chaplin, John, & Goldberg, 1988). Relatedly, multidimensional, multifaceted structure of personality has been advanced in the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed., *DSM-5*; American Psychiatric Association, 2013) that recognizes both genetic and environmental determinants of personality and supplements the *DSM-IV*'s discrete, categorical classification system.

There have been several direct attempts to develop separate operational measures of trait and state aspects of a number of diverse constructs, including arousal (Wilding & Mohindra, 1982), curiosity (Naylor, 1981), guilt (Saklofske & Schulz, 1975), negative affect (Brondolo et al., 2008), worry (O'Neil, Baker, & Matsuura, 1992), and the Big Five traits (Schutte, Malouff, Segrera, Wolf, & Rodgers, 2003). Among the first, best known, and widely used of these state and trait measures are Spielberger's measures of state and trait anxiety (Spielberger, 1983) and anger (Spielberger et al., 1983).

Spielberger's Measures

State and Trait Anxiety. As summarized succinctly by Grös, Antony, Sims, and McCabe (2007),

> Spielberger (1983) described state anxiety as existing in a transitory emotional state that varies in intensity and fluctuates over time . . . [whereas] . . . trait anxiety refers to a stable susceptibility or a proneness to experience state anxiety frequently. (p. 369)

Speilberger (1983) developed the State-Trait Anxiety Inventory (STAI) to measure these conceptually distinct components of anxiety and is composed of two 20-item scales. Instructions for the 20-item State scale (STAI-S) ask respondents how they feel "right now . . . at this very moment" (e.g., "I am tense," "I feel upset") and to respond on a 4-point scale ranging from 1 = *not at all* to 4 = *very much so*. As such, these instructions direct the respondents' attention to their immediate, present circumstances. In contrast, instructions for the Trait scale (STAI-T) include asking the respondents to use a 4-point scale (1 = *almost never* to 4 = *almost always*) to indicate "how you *generally* feel" (e.g., "I try to avoid facing a crisis or difficulty"),[1] defining a more global temporal frame of reference. Even nearly 30 years ago, the STAI had been used in more than 3,000 studies and had been translated into 30 languages (Spielberger, 1989). Barnes et al.'s (2002) reliability generalization study indicated that mean internal consistency reliability estimates were comparable for the STAI-T ($\bar{r}_{xx'} = .89, k = 51$) and the STAI-S ($\bar{r}_{xx'} = .91, k = 52$). The STAI has also demonstrated good convergent validity with other measures of anxiety and correlates negatively with self-reported health outcomes (Elwood, Wolitzky-Taylor, & Olatunji, 2012). Efforts to distinguish the STAI-T and STAI-S qua trait versus state measures sometimes compare test–retest reliabilities, and Barnes et al. showed these were lower for the STAI-S ($\bar{r}_{xx'} = .70, k = 7$) than for the STAI-T ($\bar{r}_{xx'} = .88, k = 7$) as would be expected, given the definitions of the underlying constructs. To our knowledge, no studies have estimated latent state and trait components of the STAI. As such, there is little or no direct evidence that the STAI actually measures state and trait components of anxiety.

Trait and State Anger. Spielberger et al. (1983) defined anger as "an emotional state that consists of feelings that vary in intensity, from mild irritation or annoyance to intense fury and rage" (p. 16). "State anger refers to a transitory emotional-physiological condition . . . [whereas trait anger] . . . refers to a stable personality dimension of anger proneness or the tendency to experience state anger" (Deffenbacher et al., 1996, p. 131). Developed by Spielberger et al. (1983), the state (SAS) and trait (TAS) anger scales contain 10 items each that ask how respondents have felt "during the past week" (SAS, e.g., "angry," "furious," "mad") or how they "generally feel" (TAS, e.g., "quick tempered," "hotheaded person") using 4-point scales (1 = *almost never* to 4 = *almost always*). Scores on the SAS and TAS have internal consistencies that range between .84 and .93 (Spielberger, 1988). The psychometric properties of SAS and TAS have been evaluated in a number of countries and have been useful in clinical and normal populations (Eckhart et al., 2004). Deffenbacher et al. (1996; Deffenbacher, Lynch, Oetting, & Yingling, 2001; Deffenbacher, Richards, Filetti, & Lynch, 2005) have also demonstrated considerable support for the convergent, predictive, and construct validity of these scales in a number of experimental tests of various propositions derived from state-trait theory. However, we located no studies that attempted direct assessments of whether and the extent to which Speilberger's anger scales actually measure state and trait anger.

The Case of Community-Dwelling Informal Caregivers

Informal, nonpaid, caregivers are considered to be a national health care resource (Schulz & Sherwood, 2008), with 34.2 million Americans, or 14.3% of the adult population, providing care to ill or disabled adult relatives aged 50 years or older (National Alliance for Caregiving & AARP Public Policy Institute, 2015). Community-dwelling older adults compromised by chronic disease and/or disability typically rely on family members or close friends for home-based care and social support. Informal caregivers incur the economic burden equivalent of billions of dollars annually (e.g., Joo, Dunet, Fang, & Wang, 2014), in addition to the psychological and physical health burden of providing care (Pinquart & Sorensen, 2001). Often studied psychological well-being outcomes of caregiving include anxiety, depression, and anger (MacNeil et al., 2009; Schulz & Sherwood, 2008). We focused on caregivers' anxiety and anger incurred as a result of caregiving, as Spielberger's (1983) scales are specifically designed to address the trait and state components of these constructs. We also included measures of two factors known to affect caregivers' psychological functioning: communal behavior (an interpersonal factor) and care recipient problem behavior (an environmental factor) as putative determinants of trait and state components of caregiver anxiety and anger.

Covariates of Caregiver Anxiety and Anger

Communal Behavior. One covariate of caregiver anxiety and anger is caregiver–care recipient preillness relationship quality, or the degree to which caregivers and care recipients historically engaged in mutually communal behavior. As opposed to exchange relationships where benefits are given based on the expectation that a comparable benefit will be returned or *exchanged*, communal relationships rest on the underlying assumption that benefits are given in response to a need or to demonstrate a general concern for each other's welfare (Williamson & Clark, 1989, 1992; Williamson, Clark, Pegalis, & Behan, 1996). Individuals in communal relationships are more likely to perceive caregiving as an extension of general orientations toward reciprocity and relationship reward (Hui, Elliott, Shewchuk, & Rivera, 2007) and thus as more rewarding, satisfying, and stable over time than exchange relationships (Clark, Ouellette, Powell, & Milberg, 1987). As such, we expected that premorbid caregiver–care recipient communal behavior would be characterized as being more trait-like than state-like and that its trait-like component would be more closely related to caregiver anxiety and anger than its state-like component.[2]

Care Recipient Problem Behavior. A second correlate of caregiver anxiety and anger is care recipient problem behavior (Kramer, Gibson, & Teri, 1992). Problem behavior can be described as behavior that the care recipient may exhibit as an attempt to control or manipulate the caregiver to get his or her way (e.g., scream or yell, pout or withdraw) and as troublesome behavior (e.g., agitation, wandering, delusions) displayed by care recipients as a result of cognitive impairment or dementia (Teri et al., 1992). To more broadly apply this construct to other caregiving populations, some caregiving researchers (e.g., Smith, Williamson, Miller, & Schulz, 2011; Williamson et al., 2005) have operationalized care recipient problem behavior as controlling or manipulative behavior of care recipients and examined the association of these behavioral problems with caregivers' maladaptive psychological functioning (e.g., depression and resentment). We expected more occasion- or situation-specific variability in this variable over the course of the caregiving trajectory given the more episodic nature of problem behavior over time.

Method

Sample and Procedure

Three waves of data were collected 1 year apart in Athens, GA, Pittsburgh, PA, Dallas, TX, and the surrounding areas in the early 2000s as part of the Family Relations in Late Life (FRILL) project (Williamson & Shaffer, 2001). Older adult care recipients and informal family caregiver dyads were recruited with eligible caregivers providing at least one basic activity of daily living (ADL, e.g., bathing, toileting) or at least two instrumental ADLs (e.g., paying bills, handling financial matters) to care recipients older than 60 years. Carefully trained pairs of interviewers conducted structured, face-to-face interviews to administer study measures, most typically in the dyad's home and, when possible, in separate rooms. Participating caregivers were paid $20 for each of three interviews. Mean caregiver age was 63 years (N = 310 at Time 1; SD = 14.3), most had some college or trade school education, median income in the $25,000 to $30,000 range, and had been providing care for the care recipient for M = 6.1 years (SD = 6.8). Most caregivers were women (77%), who lived with the care recipient (75%), the care recipient's spouse (54%) or adult children (39%), White (79%) or Black (17%), and assisted with M = 10.5 ADLs (SD = 4.4). For care recipients, mean age was 78 years (SD = 8.7), median education was high school graduate, median income was $10,000 to $15,000, most (53%) were women, 82% White, and 17% Black, and 41% had been diagnosed with Alzheimer's disease or related dementia. Data reported here are FRILL caregivers' responses to Spielberger's state and trait anxiety and anger scales, Williamson and Schulz's (1995) measure of premorbid communal behavior, and Steinmetz's (1988) measure of care recipient problem behavior.

Additional Measures

Communal Behavior. We used the 10-item Mutual Communal Behavior Scale (MCBS; Wlliamson & Schulz, 1995) to measure preillness–disability relationship quality between the caregiver and the older adult care recipient. The MCBS measures the frequency (1 = *never* to 4 = *always*) with which behavioral expressions of communal feelings were displayed between a caregiver and care recipient before the onset of the caregiving relationship. Specifically, caregivers were directed to think about "the kind of interactions you had with (care recipient) *before* he or she became ill." Five items each evaluated caregiver communal behavior toward the care recipient (e.g., "If he or she was feeling bad, I tried to cheer him or her up") and care recipient communal behavior toward the caregiver ("He or she did things just to please me"). The MCBS has good psychometric properties and is stable over time (Williamson & Schulz, 1995). Total scale coefficients alpha for each of the three waves of data collection in this study were the following: Wave 1 = .87, Wave 2 = .89, and Wave 3 = .90.[3]

Care Recipient Problem Behavior. We assessed care recipient problem behavior by asking caregivers to report how often "in the last month" (0 = *never* to 4 = *always*) the care recipient exhibited 14 common behavior problems (e.g., screamed or yelled, refused to eat, refused medical treatment) included in the Steinmetz Control Scale (Steinmetz, 1988). Consistent with previous research (e.g., Smith et al., 2001), total scale internal consistency was high (α = .87) at all three waves of measurement in this study. Also, principal components analysis of the data reported here indicated that a single dominant general factor accounted for 38% of the variance among items at each measurement wave, supporting a unidimensional structure (Ree, Carretta, & Teachout, 2015).

Latent State-Trait Theory

Latent state-trait theory (LST) was developed as an extension of classical test theory (Lord & Novick, 1968) that expands the true-score space beyond the individual to also recognize situational influences on test scores (Steyer et al., 1999). LST analyses involves two fundamental decompositions: (a) first, observed test score variance is decomposed into state true-score variance and nonsystematic measurement error variance (as in classical test theory) and (b) second, state true-score variance is decomposed into stable

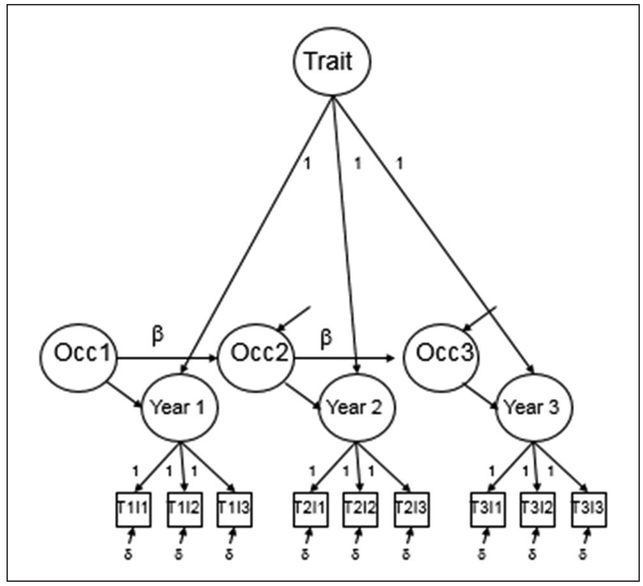

Figure 1. Generic three-wave three-indicator TSO model.
Note. TSO = trait-state-occasion; Occ = occasion.

trait-like variance and state residual variance that represents occasion- or situation-specific components of the construct along with trait × occasion ineractions (Steyer, 1989; Steyer, Ferring, & Schmitt, 1992; Steyer, Mayer, Geiser, & Cole, 2015). As such, LST was developed in recognition that most psychological attributes reflect both stable trait (i.e., time-invariant) and more labile state (time-varying) components (Cattell, 1946; Hertzog & Nesselroade, 1987).

The TSO Model. Since the 1990s (Kenny & Zautra, 1995; Steyer et al., 1992), development of analytic models for LST data has flourished (Geiser, Hintz, Burns, & Servera, 2020). Cole et al.'s (2005) TSO model is a constrained version of a no-change second-order latent curve model with structured residuals (Curran, Howard, Bainter, Lane, & McGinley, 2014; Tisak & Tisak, 2000). We chose it over other available models for four reasons. First, it is faithful to mathematical development of LST (e.g., Steyer et al. 1999). Second, it is a straightforward generalization of Kenny and Zautra's (1995) multiple-indicator state-trait-error model. Third, it contains the multiple-indicator LST model as a special case when autoregressive effects are not present. And finally, it tends to avoid estimation problems associated with more complex models such as Kenny and Zautra's (2001) multiple-indicator state-autoregressive-trait model (Geiser et al., 2020). A generic univariate three-wave, three-indicator TSO model is shown in Figure 1. Conceptually, the TSO model begins as a longitudinal measurement model with a unidimensional factor structure imposed for each measurement wave, but first-order factor loadings are fixed at 1.0 to pass the indicators' covariance structure up to the level of the State first-order factors, which are shown in

Figure 1 as "Year *x*." Loadings for the Trait and Occasion (Occ) second-order factors are also fixed at 1.0 and State first-order factors' variances are fixed at 0 so that they are apportioned entirely into either Trait or Occasion components. An autoregressive function (β) that distinguishes the TSO model from a multiple-indicator LST model (Cole et al., 2005) is sometimes included to account for stable variance between adjacent States that is not stable across all measurement waves and therefore is not accounted for by the general Trait second-order factor. Also, a congeneric measurement structure is often tenable, implying equality constraints on the manifest indicators' uniquenesses both within and across waves (although this is a testable assumption, see below), and stationarity is usually assumed for the Occasion factors, implying equality constraints on β over time, where appropriate (Cole et al., 2005).

Analyses

First, we computed item parcels for each scale by randomly assigning individual scale items to parcels with the constraint that each parcel contained equal or approximately equal numbers of items (Little, Cunningham, Shahar, & Widaman, 2002) and used the same item allocations to parcels for each measurement wave. For all measures, we computed three parcels for each measurement wave so that the measurement model and subsequent TSO models would be identified locally as well as globally. While we acknowledge that the use of item parcels is not without controversy (see, e.g., Hagtvet & Nasser, 2004; Hall, Snell, & Foust, 1999; Marsh, Lüdtke, Nagengast, Morin, & Von Davier, 2013), we adopted an item parcel measurement strategy due to the widely acknowledged benefits of higher indicator reliability, more nearly continuous data distributions, more parsimonious measurement models, fewer dual factor loadings, and less sampling error than use of individual items (Little et al., 2002; Little, Rhemtulla, Gibson, & Schoeman, 2013). Analyses proceeded in four steps using LISREL 8.8 (Jöreskog & Sörbom, 1996): (a) evaluation of the longitudinal measurement invariance (MI) of each scale across the three measurement waves, (b) selection of an appropriate univariate TSO model for each scale, (c) calculation of trait and state variance components for each scale, and (d) development of bivariate TSO models (see below) to assess covariate relationships between Spielberger's scales and caregiver–care recipient communal behavior and care recipient problem behavior. Missing data were treated using full information maximum likelihood procedures (Newman, 2014). We used a combination of the overall χ^2 statistic and two often-reported fit indexes: the root mean squared error of approximation (RMSEA and its 90% confidence intervals) and the comparative fit index (CFI), with guidelines for good fit recommended by Hu and Bentler (1999).

Table 1. Measurement Invariance Omnibus Test Results.

Scale	*df*	χ^2	RMSEA	90% CI	CFI
STAI-S	21	17.87	0	[0, .04]	1.00
STAI-T	21	30.31	.04	[0, .06]	.99
SAS	21	80.72*	.10	[.07, .12]	.96
TAS	21	45.15*	.06	[.04, .09]	.98
MCBS	21	15.87	0	[0, .03]	1.00
CR-PB	21	31.17	.04	[0, .07]	.99

Note. RMSEA = root mean squared error of approximation; CI = confidence interval; CFI = comparative fit index; *df* = degrees of freedom; STAI-S = State-Trait Anxiety Inventory–State scale; STAI-T = State-Trait Anxiety Inventory–Trait scale; SAS = State Anger Scale; TAS = Trait Anger Scale; MCBS = Mutual Communal Behavior Scale; CR-PB = Care Recipient–Problem Behavior.
*$p < .01$.

Results

Measurement Invariance

We first conducted an omnibus test of MI across measurement waves using Vandenberg and Lance's (2000) augmented covariance matrix approach, and as Table 1 shows, this model fit well for the STAI-S, STAI-T, MCBS, and care recipient problem behavior scales, supporting full MI across all three measurement waves and indicating that "further tests of specific aspects of ME/I are neither needed nor warranted" (p. 36). The omnibus test was rejected statistically for the TAS, $\chi^2(21) = 45.15$, $p < .01$, but the RMSEA and CFI indicated good model fit (Hu & Bentler, 1999). However, the omnibus test of MI did not indicate good fit for the SAS, $\chi^2(21) = 80.72$, $p < .001$, RMSEA = .10. Following Yoon and Millsap's (2007) approach using modification indices and Jung and Yoon's (2016) approach based on parameter estimates' confidence intervals to identify specific violations of MI, we found that two factor loadings in the SAS model failed to exhibit strict metric invariance. Freeing these factor loadings in a partial metric invariance model improved model fit, $\Delta\chi^2(2) = 21.70$, ΔCFI = .016 (Cheung & Rensvold, 2002); ΔRMSEA = .02 (F. F. Chen, 2007); so we carried over this partial metric invariance measurement model into the subsequent TSO models for SAS.

TSO Model Results

STAI-S model selection results in Table 2 showed that the basic TSO model (Model 1) provided a poor fit to the data. LaGrange and Cole (2008) suggested that the basic TSO model fit can often be improved on by adding Method factors to account for the repeated administration of the same measures over time, and Table 2 shows that doing so (Model 2, see Figure 2) improved model fit considerably.[4] Allowing heteroscedastic uniquenesses within measurement waves

Table 2. Model Selection Results: STAI-S.

Model	*df*	χ^2	RMSEA	90% CI	CFI
1. Basic TSO model	40	100.57*	.070	[.053, .087]	.970
1 versus 2	3	37.37*			
2. Model 1 with uncorrelated Method factors	37	63.20*	.048	[.027, .068]	.987
2 versus 3	3	3.20			
3. Model 2 with heteroscedastic uniquenesses	35	60.00*	.048	[.026, .068]	.988
2 versus 4	2	1.99			
4. Model 2 without autoregressive effects[a]	39	65.19*	.047	[.025, .066]	.986

Note. STAI-S = State-Trait Anxiety Inventory–State scale; TSO = trait-state-occasion; RMSEA = root mean squared error of approximation; CI = confidence interval; CFI = comparative fit index; *df* = degrees of freedom.
[a]Selected model.
$*p < .01$.

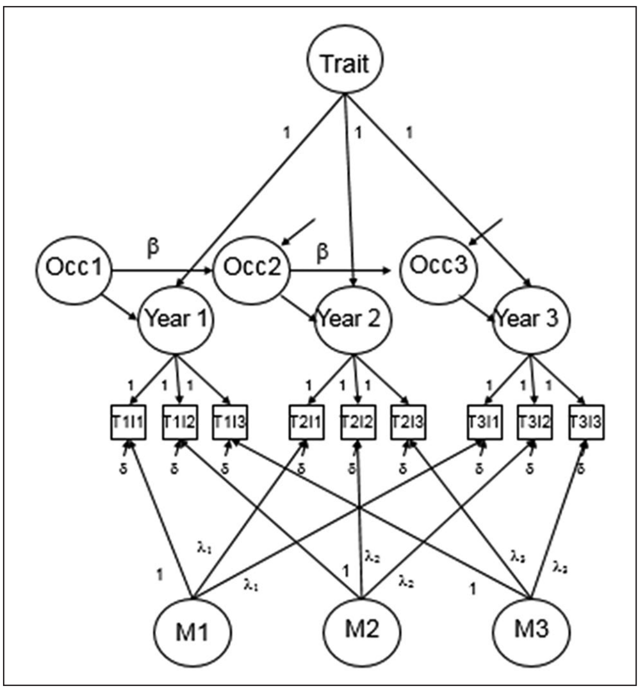

Figure 2. TSO model with orthgonal Method factors.
Note. TSO = trait-state-occasion; Occ = occasion.

(Model 3) did not improve model fit, indicating that the homoscedasticity assumption held both across measurement waves and across parcels within waves. Removing autoregressive effects did not worsen model fit (Model 4) indicating that there was no significant residual stability above and beyond the global trait stability component. As such, Model 4 (what Prenoveau et al., 2011, referred to as a "trait stability model," p. 838) was judged as best representing the data. Similar series of model comparisons were also conducted for the remaining constructs to identify the most appropriate univariate model as a prerequisite for additional analyses. The results for each scale are summarized in the appendix. Results indicated that the same Model 4 also best represented the data for STAI-T, but for the remaining scales, allowing heterogeneous uniquenesses for different parcels within each measurement wave (indicating heteroscedasticity across different parcels within measurement waves, but homogeneous uniquenesses for like parcels across measurement waves) improved model fit and thus was the selected model. Table 3 provides an overview of selected TSO models and their respective goodness-of-fit indices.

The top portion of Table 4 shows the decomposition of State first-order factors' variance into corresponding Trait and Occasion second-order factors and, as such, present this study's key results as they pertain to Spielberger's state and trait anxiety and anger scales. Three findings are noteworthy here. First, *both* trait *and* state scales reflect both trait and state variance components. This comes as no surprise, as time after time LST studies have shown that most measures of psychological constructs exhibit both trait and occasion-specific variance components (Cole et al., 2005; Geiser et al., 2020; Steyer et al., 2015). Often, even constructs thought to be predominantly trait-like still exhibit nonnegligible state components (e.g., cognitive ability, Hermes & Stelling, 2016) and others that are thought to be predominantly state-like, still exhibit nonnegligible trait-like components (e.g., mood, Windle & Dumenci, 1998). As such, the first key finding here, and one that contributes to the accumulating LST literature, is that even measures that are designed specifically to separate trait from state aspects of the underlying constructs still reflect *both* trait and state components of their respective constructs.

The second noteworthy finding in Table 4 is that for both the anxiety and anger constructs, their corresponding trait scales reflected proportionally more stable trait-related variance (84% and 81%, respectively) than did their state counterparts (66% and 70%, respectively) and, conversely, the state scales reflected relatively more occasion-specific variance (34% and 30%, respectively) than did their trait counterparts (16% and 19%, respectively). In isolation, these results signal the psychometric success of distinguishing

Table 3. TSO Model Selection Summary.

Scale	Method effects	Hetero or homo uniquenesses	Autoregressive effects	Hetero or homo SOF loadings	Model goodness-of-fit				
					df	χ^2	RMSEA	90% CI	CFI
STAI-S	Yes	Homo	No	Homo	39	65.19*	.047	[.025, .066]	.986
STAI-T	Yes	Homo	No	Homo	39	75.14*	.055	[.036, .073]	.981
SAS	Yes	Hetero	No	Hetero	29	105.38*	.092	[.074, .112]	.951
TAS	Yes	Hetero	No	Homo	37	98.62*	.073	[.056, .091]	.943
MCBS	Yes	Hetero	No	Homo	37	43.02	.023	[0, .048]	.997
CR-PB	Yes	Hetero	No	Homo	37	111.97*	.081	[.064, .098]	.951

Note. TSO = trait-state-occasion; STAI-S = State-Trait Anxiety Inventory–State scale; STAI-T = State-Trait Anxiety Inventory–Trait scale; SAS = State Anger Scale; TAS = Trait Anger Scale; MCBS = Mutual Communal Behavior Scale; CR-PB = Care Recipient–Problem Behavior; Hetero = heteroscedastic; Homo = homoscedastic; RMSEA = root mean squared error of approximation; CI = confidence interval; CFI = comparative fit index; *df* = degrees of freedom.
*$p < .01$.

Table 4. Trait and Occasion Components of State First-Order Factors' Variances.

	Trait variance	Occasion-specific variance
Focal variables		
STAI-S	.66	.34
STAI-T	.84	.16
SAS	.70	.30
TAS	.81	.19
Covariates		
MCBS	.90	.10
CR-PB	.73	.27

Note. STAI-S = State-Trait Anxiety Inventory–State scale; STAI-T = State-Trait Anxiety Inventory–Trait scale; SAS = State Anger Scale; TAS = Trait Anger Scale; MCBS = Mutual Communal Behavior Scale; CR-PB = Care Recipient–Problem Behaviors. Trait and Occasion–specific variance components = squared standardized second-order factor loadings of Trait and Occasion second-order factors on State first-order factors.

between trait versus state components of the anxiety and anger constructs operationally. However, this conclusion must be tempered by the third key set of results.

This third noteworthy set of findings in Table 4 is that *all* of Spielberger's scales were dominated by trait, versus state variance components, even those that were designed specifically to assess immediate, in-the-moment (vs. temporally stable) feelings of anxiety and anger. Thus, even scales whose instructions to respondents directed them to focus on how they feel "right now . . . at this very moment" (STAI-S) and how they felt specifically "during the past week" (SAS) were dominated by stable trait-like variance components that spanned the entire 3-year data collection course of the FRILL project. In fact, correlations between the trait components for both the state and the trait anxiety scales ($\phi = .86$, $p < .01$) and the state and trait anger scales ($\phi = .72$, $p < .01$) were quite high. In part, these correlations indicate the conceptual overlap between the constructs, but they also reflect statistical corrections for (a) attenuation due measurement error and contamination due to (b) item-wording method effects and (c) transient effects of time-varying influences on the observed measures (Steyer et al., 1999). As such, these correlations represent an extension of Lance, Dawson, Birkelbach, and Hoffman's (2010) simultaneous correction for measurement error and common method variance to correct also for attenuating transient effects (Green, 2003) on estimates of trait–trait correlations. In the next section, we explore some possible reasons for these findings.

Covariates

As a preliminary to covariate analyses, we conducted TSO model variance decompositions on the communal behavior and care recipient problem behavior in order to isolate stable versus situation-specific variance components to relate to their counterparts among the Spielberger measures, and so as to unconfound the stable and unstable variance components that would otherwise comprise a measure's total variance at any single measurement wave. This allows for the estimation of stable aspects of the constructs independent of changes over time and, alternately, changes in the constructs over time independent of their stabilities (Brauchli, Schaufeli, Jenny, Füllemann, & Bauer, 2013).

The bottom portion of Table 4 shows state variance decompositions for this study's covariates. Communal behavior's trait component was the largest of any of the present study's measures at 90%. As was expected, communal behavior was very trait-like, having developed over the course of many years in most cases and, likely, decades of interactions between the caregiver and care recipient. These communal (vs. exchange) relationships extended into the caregiving phase of the caregiver–care recipient relationship and were subject only to minor transitory fluctuations over the course of this study. In contrast, and also as

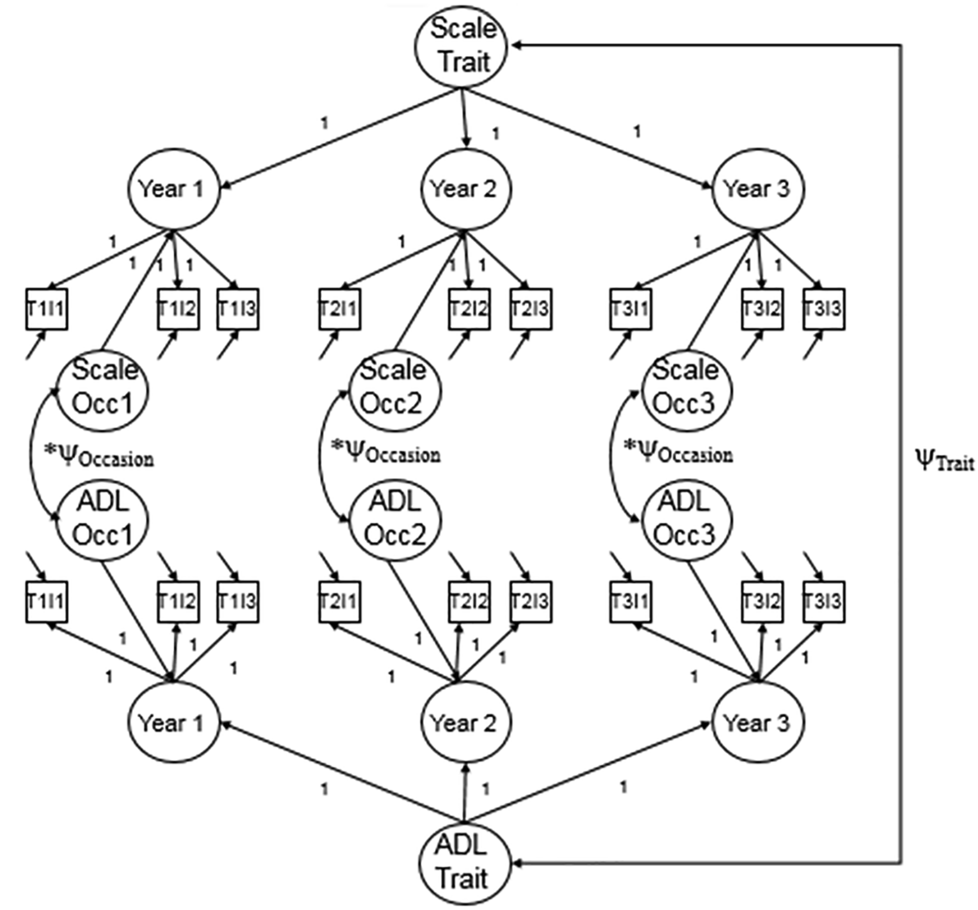

Figure 3. Bivariate TSO model.
Note. TSO = trait-state-occasion; Occ = occasion; ADL = activity of daily living.

was expected, care recipient problem behavior, while still largely a stable construct (73% trait variance), was subject to more situation-specific variability (27%) over the course of the study.

We also estimated bivariate TSO models to assess relationships between the focal variables' and covariates' trait and state components by combining the univariate TSO models selected earlier as best representing the data. A generic version of the bivariate TSO model is shown in Figure 3. Here, the covariance between the two variables' trait components is shown as Ψ_{Trait}, and Ψ_{Occasion} indicates the covariance between the variables' state components. The variable Ψ_{Occasion} was constrained to be equal over the three measurement waves as the univariate TSO models indicated that stationarity between measurement waves was a reasonable assumption. Also, the orthogonal measurement method structures, although present for both the focal variable and the covariate, are omitted from Figure 3 for clarity.

Results for the bivariate TSO analyses are shown in Table 5. First, the trait component of communal behavior was significantly related to the trait components of all the focal variables. This is consistent with previous cross-sectional relationships (Mahoney, Regan, Katona, & Livingston, 2005; Sperberg & Stabb, 1998) and extends them in supporing the ideas (a) that communal behavior qua stable interpersonal trait-like construct at least in part helps explain the stability in anxiety and anger over time and (b) that caregivers in more communal relationships experience less adverse emotional reactions to the caregiving situation than caregivers in less communal relationships. The low and nonsignificant relationships between the state components of communal behavior and focal variables reflect the facts that (a) there was little occasion-specific fluctuation in them (and especially communal behavior) to begin with (i.e., they had restricted ranges) and (b) what over-time variation there was in communal behavior bore little or no relationship to caregivers' variations in emotional reactions to the caregiving situation.

Table 5 also shows that both the trait and the state components of care recipient problem behaviors were

Table 5. Correlations Between Spielberger's Scales and Covariates.

	Mutual Communal Behavior Scale		Care Recipient–Problem Behavior	
	Trait variance	Occasion-specific variance	Trait variance	Occasion-specific variance
STAI-S	−.18**	−.11	.30**	.08
STAI-T	−.23**	−.09	.45**	.15*
SAS	−.21**	−.02	.45**	.30**
TAS	−.13*	−.10	.30**	.24**

Note. STAI-S = State-Trait Anxiety Inventory–State scale; STAI-T = State-Trait Anxiety Inventory–Trait scale; SAS = State Anger Scale; TAS = Trait Anger Scale; trait variance = TI-TI component correlations; occasion-specific variance = TV-TV component correlations.
*$p < .05$. **$p < .01$.

(generally speaking) related to both the trait and the state components of anxiety and anger. These results indicate that (a) stability in caregivers' emotional responses to caregiving was related to *consistency* in care recipients' problem behavior and (b) instability in caregivers' emotional responses to caregiving was related to *variations* in care recipients' problem behavior frequencies over the course of the study.

Discussion

So, do state and trait measures measure states and traits? The answer seems to be (a) yes, they measure both, (b) but not quite as was intended, (c) because stable, trait-like variance dominates both state *and* trait measures, (d) at least among the present sample of community-dwelling caregivers of older adult care recipients. The first two of these findings replicate an accumulating literature that indicates that psychological measures routinely assess both trait *and* state components of the latent construct (Cole et al., 2005; Steyer et al., 2015). The present study extends this research to indicate that this robust finding also generalizes to measures designed specifically to assess state versus trait aspects of their constructs and is a first step in this regard, examining state and trait measures of anxiety and anger. An obvious next step would be to determine the extent to which our findings generalize further to the many state and trait measures of other psychological constructs that have been developed, for example, Spielberger's (1999) State-Trait Anger Expression Inventory–2, the State-Trait Inventory for Cognitive and Somatic Anxiety (Grös et al., 2007), and state and trait aspects of other factors related to caregiving for physically and/or psychologically compromised older adult care recipients (e.g., caregiver burden, Horwitz & Reinhard, 1995; Miller, McFall, & Montgomery, 1991). LST models may be highly informative as research expands in this area to account for purer measures of symptomatology, such as is the case with the State-Trait Inventory for Cognitive and Somatic Anxiety, and to include other dimensions of caregiver burden and personality variables. In fact, conducting LST analyses for newly developed scales should be a routine part of scale development for scales designed specifically to assess enduring (vs. ephemeral) individual characteristics.[5]

Future research should also seek to determine the generalizability of our findings beyond our niche population. Community-residing caregivers of older care recipients represent a somewhat unique population in many respects and may exhibit greater stability in emotional responses to the caregiving situation than is seen in other populations' responses to their environments. For example, Olatunji and Cole (2009) found 52% trait variance in anxiety for grade school children over a 4-year period and Nett, Bieg, and Keller (2017) found 52% trait variance in math anxiety for 9th- and 10th-grade German students (vs. 66% to 84% for the measures studied here). Typically, caregivers of older adult care recipients are spouses (46% in this study) or children (39% in this study) who have known and developed their relationship with care recipients for many years and provide care for care recipients for up to 4 years or longer (National Alliance for Caregiving & AARP Public Policy Institute, 2015). As such, caregivers' anxiety and anger reactions likely reflect on both the inherent stability in their relations with care recipients and the caregiving environment. Future quantitative studies of LST research should help establish population-type and other boundary conditions for the prevalence of trait versus state variance in psychological measures.

One issue we encountered during the course of this study was the wide variety of analytic models that were presented as LST models. Many of these were first-order factor models, for example, single trait–multiple state (e.g., Kaczmarek, Bujacz, & Eid, 2015) and multitrait–multistate models (e.g., Bonnefon, Vautier, & Eid, 2007) that resemble multitrait–multimethod models (Campbell & Fiske, 1959). Perhaps the most often reported model was Cole et al.'s (2005) TSO model, but Steyer et al.'s (1999) LST model and Kenny and Zautra's (1995, 2001) stable trait autoregressive state (STARTS) model is also often reported. As Cole et al. (2005) discuss, the TSO model can be viewed as a multiple-indicator generalization of Kenny and Zautra's (1995) single-indicator TSE model or as the equivalent of Steyer and Schmitt's

(1994) latent state-trait autoregressive (LST-AR) model but with autoregressive effects parameterizes between latent state residuals rather than between the latent states themselves. A multiple-indicator LST model could also be seen as arising as a special case of either the TSO or LST-AR model, but with autoregressive effects absent. Geiser et al. (2020) has presented a catalog of many of these and other LST-related models and suggested some decision contingencies for the adoption of one versus another, but more work is needed along these lines to clarify just what it is that these various models offer and what they do not.

We also noted that many LST applications went beyond the fundamental partitioning of measures' variances measured longitudinally. For example, studies often augmented the basic LST analytic model they used by invoking one of LaGrange and Cole's (2006) strategies for the control of method effects, usually arising from the repeated administration of the same manifest indicators over time (e.g., Cole, 2006; Danner, Hagemann, Shankin, Hager, & Funke, 2011; Körner, Silbereisen, & Canter, 2014). Doing so usually results in improved model fit by accounting for the fact that these measurement artifacts are themselves stable over time. As a second example, researchers also introduced exogenous predictors of LST model components (e.g., Hazel & Hankin, 2014; Ormel & Schaufeli, 1991) in attempts to explain observed trait or state facets of the focal construct(s). Still others estimated bivariate LST models (usually the TSO model) either with (Brauchli et al., 2013; Lucas & Donnellan, 2012) or without (Schaufeli, Maassen, Bakker, & Sixma 2011; Wetzel, Lüdtke, Zettler, & Böhnke, 2016) structural parameters linking model components in attempts to understand how trait and state components of different constructs are related over time.

A third category of studies presented extensions of the traditional LST models to accommodate particular study design features that extend beyond the longitudinal random samples of persons and situations over time. For example, Hamaker, Nesselroade, and Molenaar (2007) introduced an integrated and general LST model that allows for theorizing and modeling of idiosyncratic and subgroup differences in trait and state variance components above and beyond the more commonly assumed nomothetic latent state-trait models that assume homogeneous factor structures across individuals. As a second example, Geiser et al. (2014) introduced a generalization of the more traditional LST model that includes only random situations and person × situation interactions to a more general model that also parameterizes fixed (i.e., known) situations. This model allows the estimation of situation-specific trait effects and person × situation effects that are ordinarily confounded with occasion effects in more traditional LST models. Other examples include Courvossier, Eid, and Nussbeck's (2007) mixture distribution LST model that allows the identification of different classes of individuals in an LST framework and Geiser, Bishop, Lockhart, Shiffman, and Grenard's (2013) presentation of LST and growth curve models as multilevel structural equation models. These few examples indicate how LST is continuing to evolve and to provide more complete understandings of persons–situation interactions in human behavior.

Conclusion

Routinely, applications of LST models support the idea that most psychological measures tap into both trait and state components of their underlying constructs. The present findings add to this literature by supporting the idea that mere variations in scale instructions are unlikely to isolate the trait and state components in producing pure trait or state measures of their intended constructs. In fact, our findings suggest that trait components can dominate over state components over extended time periods (here, 3 years) despite scale instructions that focus the respondent on their immediate circumstances. Rather, application of one of a number of available LST models (Geiser et al., 2020) to longitudinal data can accomplish this goal, permitting tests of theory relating specifically to state (vs. trait) variables. Of course, such a research agenda raises both the data collection (requiring longitudinal data) and the analytic (requiring application of reasonably sophisticated structural equation models) bars, but nonetheless supports inferences regarding trait- versus occasion-specific aspects of focal constructs.

Finally, there are several important considerations for future research examining state versus trait variance as it relates to measuring key factors of psychological functioning and well-being. Given the emerging understanding of personality (American Psychiatric Association, 2013) as a multidimensional, multifaceted construct derived from genetic and environmental determinants, future studies examining state versus trait variance associated with emotions and cognitions relevant to psychological health should include additional measures assessing positive aspects of psychological functioning. Research investigations assessing positive emotions and cognitions would provide a more holistic picture of the key determinants of health for community-dwelling caregivers, as well as for other populations that may be disproportionally affected and overburdened by poor health outcomes. In particular, Kruithof et al. (2012) underscore the clinical importance of examining the positive aspects of caregiving. Findings from future studies examining positive, in addition to negative, aspects of psychological health would further broaden our understanding of the person–situation interaction debate within this context, particularly when using advanced statistical methods, such as LST models. Increased knowledge and understanding of this complex interaction could inform the development of clinical and community-based interventions aimed at improving caregiver health and well-being.

Appendix

Detailed Model Selection Results

Table A1. Model Selection Results: STAI-T.

Model	*df*	χ^2	RMSEA	90% CI	CFI
1. Basic TSO model	40	236.04*	.126	[.111, .142]	.894
1 versus 2	3	63.73*			
2. Model 1 with uncorrelated methods	37	72.31*	.056	[.036, .075]	.981
2 versus 3	3	4.51			
3. Model 2 with heteroscedastic uniquenesses	35	67.80*	.055	[.035, .075]	.982
2 versus 4	2	2.83			
4. Model 2 without autoregressive effects[a]	39	75.14*	.055	[.036, .073]	.981

Note. STAI-T = State-Trait Anxiety Inventory–Trait scale; TSO = trait-state-occasion; RMSEA = root mean squared error of approximation; CI = confidence interval; CFI = comparative fit index; *df* = degrees of freedom.
[a]Selected model.
*$p < .01$.

Table A2. Model Selection Results: State Anger.

Model	*df*	χ^2	RMSEA	90% CI	CFI
1. Basic TSO model	38	243.12*	.132	[.117, .148]	.868
1 versus 2	3	95.46*			
2. Model 1 with uncorrelated methods	35	147.66*	.102	[.085, .119]	.928
2 versus 3	3	30.55*			
3. Model 2 with heteroscedastic uniquenesses[a]	33	117.11*	.091	[.073, .109]	.946
3 versus 4	2	8.88			
4. Model 3 without autoregressive effects	35	125.99*	.092	[.075, .109]	.941

Note. TSO = trait-state-occasion; RMSEA = root mean squared error of approximation; CI = confidence interval; CFI = comparative fit index; *df* = degrees of freedom.
*$p < .01$.
[a]Model selected.

Table A3. Model Selection Results: Trait Anger.

Model	*df*	χ^2	RMSEA	90% CI	CFI
1. Basic TSO model	40	364.54*	.162	[.147, .177]	.683
1 versus 2	3	133.76*			
2. Model 1 with uncorrelated Method factors	37	228.73*	.129	[.113, .146]	.777
2 versus 3	2	137.63*			
3. Model 2 with heteroscedastic uniquenesses	35	91.10*	.072	[.054, .090]	.948
3 versus 4	2	7.52			
4. Model 3 without autoregressive effects[a]	37	98.62*	.073	[.056, .091]	.943

Note. TSO = trait-state-occasion; RMSEA = root mean squared error of approximation; CI = confidence interval; CFI = comparative fit index; *df* = degrees of freedom.
*$p < .01$.
[a]Model selected.

Table A4. Model Selection Results: Mutual Communal Behavior.

Model	*df*	χ^2	RMSEA	90% CI	CFI
1. Basic TSO model	40	156.32*	.098	[082, .114]	.939
1 versus 2	3	105.09*			
2. Model 1 with uncorrelated Method factors	37	51.23	.036	[0, .057]	.993
2 versus 3	2	10.08*			
3. Model 2 with heteroscedastic uniquenesses	35	41.15	.024	[0, .050]	.997
3 versus 4	2	1.87			
4. Model 3 without autoregressive effects[a]	37	43.02	.023	[0, .048]	.997

Note. TSO = trait-state-occasion; RMSEA = root mean squared error of approximation; CI = confidence interval; CFI = comparative fit index; *df* = degrees of freedom.
*$p < .01$.
[a]Model selected.

Table A5. Model Selection Results: Problem Behavior.

Model	*df*	χ^2	RMSEA	90% CI	CFI
1. Basic TSO model	40	555.77*	.204	[.189, .219]	.663
1 versus 2	3	296.86*			
2. Model 1 with uncorrelated Method factors	37	238.91*	.133	[.117, .149]	.931
2 versus 3	2	129.92*			
3. Model 2 with heteroscedastic uniquenesses	35	108.99*	.083	[.065, .100]	.952
3 versus 4	2	2.98			
4. Model 3 without autoregressive effects[a]	37	111.97*	.081	[.064, .098]	.951

Note. TSO = trait-state-occasion; RMSEA = root mean squared error of approximation; CI = confidence interval; CFI = comparative fit index; *df* = degrees of freedom.
*$p < .01$.
[a]Model selected.

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.

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 in part by National Institute on Aging Grant AG15321 to Gail M. Williamson, principal investigator. The Family Relationships in Late Life (FRILL) project was conducted at the University of Georgia (Charles E. Lance and L. Stephen Miller co-investigators) in collaboration with the University of Pittsburgh (Richard Schulz, co-investigator) and the University of Texas Southwestern Medical Center at Dallas (Myron F. Weiner, co-investigator).

ORCID iD

Juliette Christie https://orcid.org/0000-0002-6681-837X

Notes

1. It is important to note that the addition of positively worded items to measurement scales may not fully capture positive emotions and cognitions associated with the construct of interest (e.g., Kruithof, Visser-Meily, & Post, 2012).
2. In the caregiving literature, historically communal relationships between caregivers and care recipients have been found to buffer caregivers from poor mental and psychological health outcomes (e.g., Williamson & Shaffer, 2001). That is, caregivers who have had more mutually communal relationships with care recipients are less likely to report depressed affect and resentment in the caregiving role (Williamson & Shaffer, 2001). Thus, examining associations between communal behavior and positive psychological outcomes is an important avenue for research; however, an investigation of this scope is beyond the purpose of this study. The current study aims to validate Speilberger's anxiety and anger measures using a community-dwelling sample of caregivers. Examining communal behavior as a covariate provides a meaningful stable variable in this context by which to evaluate the trait and state components of caregiver anxiety and anger.
3. We do not report the 90% confidence interval for coefficients alpha here because the upper and lower bounds were equal to the point estimate when rounded to the second decimal place (Raykov & Marcoulides, 2015).
4. We parameterized Method effects as uncorrelated Method (UM) factors. LaGrange and Cole also considered a correlated uniqueness model as an alternative to the UM model but in the case of a three indicator–three wave model such as in the present case, the UM and correlated uniqueness models are mathematically equivalent. Also, Geiser and Lockhart

(2012) considered the efficacy of four different approaches to modeling method effects in LST models, and while the UM model was not optimal in every condition of their simulation it performed well in conditions that described our study (i.e., low proportions of method variance, high proportions of true score variance). Because the calibration of Method variance components is also very straightforward under the UM parameterization, we opted for it in the present study.

5. We thank an anonymous reviewer for this suggestion.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author. doi:10.1176/appi.books.9780890425596

Barnes, L. L. B., Harp, D., & Jung, W. S. (2002). Reliability generalization of scores on the Spielberger State-Trait Anxiety Inventory. *Educational and Psychological Measurement*, *62*, 603-618. doi:10.1177/0013164402062004005

Bonnefon, J., Vautier, S., & Eid, M. (2007). Modeling individual differences in contrapositive reasoning with continuous latent state and trait variables. *Personality and Individual Differences*, *42*, 1403-1413. doi:10.1016/j.paid.2006.10.017

Brauchli, R., Schaufeli, W. B., Jenny, G. J., Füllemann, D., & Bauer, G. F. (2013). Disentangling stability and change in job resources, job demands, and employee well-being: A three-wave study on the Job-Demands Resources model. *Journal of Vocaational Behavior*, *83*, 117-129. doi:10.1016/j.jvb.2013.03.003

Brondolo, E., Brady, N., Thompson, S., Tobin, J. N., Cassells, A., MacFarlane, D., & Contrada, R. J. (2008). Perceived racism and negative affect: Analyses of state and trait measures of affect in a community sample. *Journal of Social and Clinical Psychology*, *27*, 150-173. doi:10.1521/jscp.2008.27.2.150

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

Cattell, R. B. (1946). *The description and measurement of personality*. New York, NY: World Book.

Chaplin, W. F., John, O. P., & Goldberg, L. R. (1988). Conceptions of states and traits: Dimensional attributes with ideals as prototypes. *Journal of Personality and Social Psychology*, *54*, 541-557.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464-04. doi:10.1080/10705510701301834

Chen, G., Gully, S. M., Whiteman, J. A., & Kilcullen, R. N. (2000). Examination of relationships among trait-like individual differences, state-like individual differences, and learning performance. *Journal of Applied Psychology*, *85*, 835-847. doi:10.1037/OQ21-9010.85.6.835

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233-255. doi:10.1207/S15328007SEM0902_5

Clark, M., Ouellette, R., Powell, M., & Milberg, S. (1987). Recipient's mood, relationship type, and helping. *Journal of Personality and Social Psychology*, *53*, 94-103.

Cole, D. A. (2006). Coping with longitudinal data in research on psychopathology. *International Journal of Behavioral Development*, *30*, 20-25.

Cole, D. A., Martin, N. C., & Steiger, J. H. (2005). Empirical and conceptual problems with longitudinal trait-state models: Introducing a trait-state-occasion model. *Psychological Methods*, *10*, 3-20. doi:10.1037/1082-989X.10.1.3

Courvossier, D. S., Eid, M., & Nussbeck, F. W. (2007). Mixture distribution latent state-trait analysis: Basic ideas and applications. *Psychological Methods*, *12*, 80-104. doi:10.1037/1082-989X.12.1.80

Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology*, *82*, 879-894. doi:10.1037/a0035297

Danner, D., Hagemann, D., Shankin, A., Hager, M., & Funke, J. (2011). Beyond IQ: A latent state-trait analysis of general intelligence, dynamic decision making, and implicit learning. *Intelligence*, *39*, 332-334. doi:10.1016/j.intell.2011.06.004

Deffenbacher, J. L., Lynch, R. S., Oetting, E. R., & Yingling, D. A. (2001). Driving anger: Correlates and a test of state-trait theory. *Personality and Individual Differences*, *31*, 1321-1331. doi:10.1016/S0191-8869(00)00226-9

Deffenbacher, J. L., Oetting, E. R., Thwaites, G. A., Lynch, R. S., Baker, D. A., Stark, R. S., . . . Eiswerth-Cox, L. (1996). State-trait anger theory and the utility of the trait anger scale. *Journal of Counseling Psychology*, *43*, 131-148.

Deffenbacher, J. L., Richards, T. L., Filetti, L. B., & Lynch, R. S. (2005). Angry drivers: A test of state-trait theory. *Violence and Victims*, *20*, 455-469. doi:10.1891/vivi.2005.20.4.455

Eckhart, C., Norlander, B., & Deffenbacher, J. L. (2004). The assessment of anger and hostility: A critical review. *Aggression and Violent Behavior*, *9*, 17-43. doi:10.1016/S1359-1789(02)001167

Elwood, L. S., Wolitzky-Taylor, K., & Olatunji, B. O. (2012). Measurement of anxious traits: A contemporary review and synthesis. *Anxiety, Stress & Coping*, *25*, 647-666. doi:10.1080/10615806.2011.582949

Epstein, S., & O'Brien, E. J. (1985). The person-situation debate in historical and current perspective. *Psychological Bulletin*, *98*, 513-537.

Geiser, C., Bishop, J., Lockhart, G., Shiffman, S., & Grenard, J. L. (2013). Analyzing latent state-trait and latent growth curve models as multilevel structural equation models. *Frontiers in Psychology*, *4*, 1-23. doi:10.3389/fpsyg.2013.00975

Geiser, C., Hintz, F., Burns, G. L., & Servera, M. (2020). Latent variable modeling of person-situation data. In J. Rauthman, R. S. Sherman & D. Funder (Eds.), *The Oxford handbook of psychological situations*. Oxford, England: Oxford University Press, 230. doi:10.1093/oxfordhb/9780190263348.001.0001

Geiser, C., Litson, K., Bishop, J., Keller, B., Burns, G., Servera, M., & Shiffman, S. (2014). Analyzing person, situation and person × situation interaction effects: Latent state-trait models for the combination of random and fixed situations. *Psychological Methods*, *20*, 165-192. doi:10.1037/met0000026

Geiser, C., & Lockhart, G. (2012). A comparison of four approaches to account for method effects in latent state–trait analyses. *Psychological Methods*, *17*, 255-283. doi:10.1037/a0026977

Green, S. B. (2003). A coefficient alpha for test-retest data. *Psychological Methods*, *8*, 88-101.

Grös, D. F., Antony, M. M., Sims, L. J., & McCabe, R. E. (2007). Psychometric properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State-Trait Anxiety Inventory. *Psychological Assessment*, *19*, 369-381. doi:10.1037/1040-3590.19.4.369

Hagtvet, K. A., & Nasser, F. M. (2004). How well do item parcels represent conceptually defined latent constructs? A two-facet approach. *Structural Equation Modeling*, *11*, 168-193.

Hall, R. J., Snell, A. F., & Foust, M. S. (1999). Item parceling strategies in SEM: Investigating the subtle effects of unmodeled secondary constructs. *Organizational Research Methods*, *2*, 233-256.

Hamaker, E. L., Nesselroade, J. R., & Molenaar, P.C. M. (2007). The integrated trait–state model. *Journal of Research in Personality*, *41*, 295-315. doi:10.1016/j.jrp.2006.04.003

Hazel, N. A., & Hankin, B. L. (2014). A trait-state-error model of adult hassles over two years: Magnitude, sources and predictors of stress continuity. *Journal of Social and Clinical Psychology*, *33*, 103-123.

Hermes, M., & Stelling, D. (2016). Context matters, but how much? Latent state-trait analysis of cognitive ability assessments. *International Journal of Selection and Assessment*, *24*, 285-295. doi:10.1111/ijsa.12147

Hertzog, C., & Nesselroade, J. R. (1987). Beyond autoregressive models: Some implications of the state-trait distinction for the structural modeling of developmental change. *Child Development*, *58*, 93-109.

Horwitz, A. V., & Reinhard, S. C. (1995). Ethnic differences in caregiving duties and burdens among parents and siblings of persons with severe mental illness. *Journal of Health and Social Behavior*, *36*, 138-150.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55. doi:10.1080/10705519909540118

Hui, S. A., Elliott, T. R., Shewchuk, R., & Rivera, P. (2007). Communal behaviors and psychological adjustment of family caregivers and persons with spinal cord injury. *Rehabilitation Psychology*, *52*, 113-119. doi:10.1037/0090-5550.52.1.113

Joo, H., Dunet, D. O., Fang, J., & Wang, G. (2014). Cost of informal caregiving associated with stroke among the elderly in the United States. *Neurology*, *83*, 1831-1837. doi:10.1212/WNL.0000000000000986

Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Retrieved from http://www.ssicentral.com/

Jung, E., & Yoon, M. (2016). Comparisons of three empirical methods for partial factorial invariance: Forward, backward, and factor-ratio tests. *Structural Equation Modeling*, *23*, 577-584. doi:10.1080/10705511.2015.1138092

Kaczmarek, L. D., Bujacz, A., & Eid, M. (2015). Comparative latent state-trait analysis of satisfaction with life measures: The Steen Happiness Index and the Satisfaction with Life Scale. *Journal of Happiness Studies*, *16*, 443-453. doi:10.1007/s10902-014-9517-4

Kenny, D. A., & Zautra, A. (1995). The trait-state-error model multiwave data. *Journal of Consulting and Clinical Psychology*, *63*, 52-59.

Kenny, D. A., & Zautra, A. (2001). Trait-state models for longitudinal data. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 243-263). Washington, DC: American Psychological Association.

Körner, A., Silbereisen, R. K., & Canter, U. (2014). Work-related demands emanating from social change and their relation to trait-like and occasion-specific aspects of subjective well-being. *Social Indicators Research*, *115*, 203-222. doi:10.1007/s11205-012-0215-5

Kramer, B., Gibson, J., & Teri, L. (1992). Interpersonal family stress in Alzheimer's disease: Perceptions of patients and caregivers. *Clinical Gerontologist*, *12*, 57-75.

Kruithof, W. J., Visser-Meily, J. M. A., & Post, M. W. M. (2012). Positive caregiving experiences are associated with life satisfaction in spouses of stroke survivors. *Journal of Stroke & Cerebrovascular Diseases*, *21*, 801-807.

LaGrange, B., & Cole, D. A. (2008). An expansion of the trait-state-occasion model: Accounting for shared method variance. *Structural Equation Modeling*, *15*, 241-271. doi:10.1080/10705510801922381

Lance, C. E., Dawson, B., Birkelbach, D., & Hoffman, B. J. (2010). Method effects, measurement error, and substantive conclusions. *Organizational Research Methods*, *13*, 435-455. doi:10.1177/1094428109352528

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, *9*, 11-173. doi:10.1207/S15328007SEM0902_1

Little, T. D., Rhemtulla, M., Gibson, K., & Schoeman, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods*, *18*, 285-300. doi:10.1037/a0033266

Lord, F. M., & Novick, M. R. (1968). *Statistical theory of mental test scores*. Reading, MA: Addison-Wesley.

Lucas, R. E., & Donnellan, M. B. (2012). Estimating the reliability of single-item life satisfaction measures: Results from four national panel studies. *Social Indicators Research*, *105*, 323-331. doi:10.1007/s11205-011-9783-z

MacNeil, G., Kosberg, J. I., Durkin, D. W., Dooley, K. W., DeCoster, J., & Williamson, G. M. (2009). Caregiver mental health and potentially harmful caregiving behavior: The central role of caregiver anger. *The Gerontologist*, *50*, 76-86. doi:10.1093/geront/gnp099

Mahoney, R., Regan, C., Katona, C., & Livingston, G. (2005). Anxiety and depression in family caregivers of people with Alzheimer disease: The LASER-AD study. *American Journal of Geriatric Psychiatry*, *13*, 795-801. doi:10.1176/appi.ajgp.13.9.795

Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J. S., & Von Davier, M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right—Camouflaging misspecification with item parcels in CFA models. *Psychological Methods*, *18*, 257-284. doi:10.1037/a0032773

Miller, B., McFall, S., & Montgomery, A. (1991). The impact of elder health, caregiver involvement, and global stress on two dimensions of caregiver burden. *Journal of Gerontology: Social Sciences*, *46*, S9-S19.

National Alliance for Caregiving & AARP Public Policy Institute. (2015). *Caregiving in the U.S.* Washington, DC: Author.

Naylor, F. D. (1981). A state-trait curiosity inventory. *Australian Pyschologist*, *16*, 172-183.

Nett, U. E., Bieg, M., & Keller, M. M. (2017). How much trait variance is captures by measures of academic state emotions? A latent state-trait analysis. *European Journal of Psychological Assessment*, *33*, 239-255.

Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods*, *17*, 372-411. doi:10.1177/1094428114548590

Olatunji, B. O., & Cole, D. A. (2009). The longitudinal structure of general and specific anxiety dimensions in children: Testing a latent trait-state-occasion model. *Psychological Assessment*, *21*, 412-424. doi:10.1037/a0016206

O'Neil, H. F., Baker, E. L., & Matsuura, S. (1992). Reliability and validity of Japanese children's trait and state worry and emotionality scales. *Anxiety, Stress & Coping*, *5*, 225-239.

Ormel, J., & Schaufeli, W. B. (1991). Stability and change in psychological distress with self-esteem and locus of control: A dynamic equilibrium model. *Journal of Personality and Social Psychology*, *60*, 288-299.

Pinquart, M., & Sorensen, S. (2001). Gender differences in self-concept and psychological well-being in old age: A meta-analysis. *Journals of Gerontology, Series B*, *568*, 195-213.

Prenoveau, J. M., Craske, M. G., Zinbarg, R. E., Mineka, S., Rose, R. D., & Griffith, J. W. (2011). Are anxiety and depression just as stable as personality during late adolescence? Results from a three-year longitudinal latent variable study. *Journal of Abnormal Psychology*, *120*, 832-843. doi:10.1037/a0023939

Raykov, T., & Marcoulides, G. A. (2015). A direct latent variable modeling based method for point and interval estimation of coefficient alpha. *Educational and Psychological Measurement*, *75*, 146-156.

Ree, M. J., Carretta, T. R., & Teachout, M. S. (2015). Pervasiveness of dominant general factors in organizational measurement. *Industrial and Organizational Psychology*, *8*, 409-427. doi:10.1017/iop.2015.16

Saklofske, D. H., & Schulz, H. W. (1975). Factor analysis of repeated state hostility and guilt measures: Males. *Psychological Reports*, *37*, 756-758.

Schaufeli, W. B., Maassen, G. H., Bakker, A. B., & Sixma, H. J. (2011). Stability and change in burnout: A 10-year follow-up study among primary care physicians. *Journal of Occupational and Organizational Psychology*, *84*, 248-267. doi:10.1111/j.2044-8325.2010.02013.x

Schulz, R., & Sherwood, P.R. (2008). Physical and mental health effects of family caregiving. *American Journal of Nursing*, *108*, 23-27. doi:10.1097/01.NAJ.0000336406.45248.4c

Schutte, N. S., Malouff, J. M., Segrera, E., Wolf, A., & Rodgers, L. (2003). States representing the Big Five dimensions. *Personality and Individual Differences*, *34*, 591-603. doi:10.1016/S0191-8869(02)00031-4

Smith, G. R., Williamson, G. M., Miller, L. S., & Schulz, R. (2011). Depression and informal care: A longitudinal investigation of caregiving stressors. *Psychology and Aging*, *26*, 584-591. doi:10.1037/a0022263

Sperberg, E. D., & Stabb, S. D. (1998). Depression in women as related to anger and mutuality in relationships. *Psychology of Women Quarterly*, *22*, 223-238.

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y)*. Palo Alto, CA: Mind Garden.

Spielberger, C. D. (1988). *State-Trait Anger Expression Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.

Spielberger, C. D. (1989). *State-Trait Anxiety Inventory: A comprehensive bibliography*. Palo Alto, CA: Consulting Psychologists.

Spielberger, C. D. (1999). *STAXI-2: State-Trait Anger Expression Inventory-2: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Spielberger, C. D., Jacobs, G., Russell, J. S., & Crane, R. S. (1983). Assessment of anger: The state-trait anger scale. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment* (Vol. 2). Hillsdale, NJ: Lawrence Erlbaum.

Steinmetz, S. K. (1988). *Duty bound: Elder abuse and family care*. Newbury Park, CA: Sage.

Steyer, R. (1989). Models of classical psychometric test theory as stochastic measurement models: Representation, uniqueness, meaningfulness, identifiability, and testability. *Methodika*, *111*, 25-60.

Steyer, R., Ferring, D., & Schmitt, M. J. (1992). States and traits in psychological assessment. *European Journal of Psychological Assessment*, *8*, 79-98.

Steyer, R., Mayer, A., Geiser, C., & Cole, D. A. (2015). A theory of traits and states—Revisited. *Annual Review of Clinical Psychology*, *11*, 71-98. doi:10.1146/annurev-clinpsy-032813-153719

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality*, *13*, 389-408.

Steyer, R., & Schmitt, T. (1994). The theory of confounding and its application in causal modeling with latent variables. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 36-67). Thousand Oaks, CA: Sage.

Talley, R. C., Crews, J., Dorn, P., Silvernail, J., Hunt, G., & Zeitzer, J. (2004, July). *Caregiving in America as an emerging public health issue: Surveillance and response by the nation's public health system*. Paper presented at the Second National Center on Birth Defects and Developmental Disabilities Conference, Washington, DC.

Teri, L., Truax, P., Logsdon, R., Uomoto, J., Zarit, S., & Vitaliano, P. P. (1992). Assessment of behavioral problems in dementia: The Revised Memory and Behavior Problems Checklist. *Psychology and Aging*, *7*, 622-631.

Therrien, Z., & Hunsley, J. (2013). Assessment of anxiety in older adults: A reliability generalization meta-analysis of commonly used measures. *Clinical Gerontologist*, *36*, 171-194. doi:10.1080/07317115.2013.767871

Tisak, J., & Tisak, M. S. (2000). Permanency and ephemerality of psychological measures with application to organizational commitment. *Psychological Methods*, *5*, 175-198. doi:10.1037/1082-989X.5.2.175

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*, 4-69. doi:10.1177/109442810031002

Wetzel, E., Lüdtke, O., Zettler, I., & Böhnke, J. R. (2016). The stability of extreme response style and acquiescence over 8 years. *Assessment*, *23*, 279-291. doi:10.1177/1073191115583714

Wilding, J. M., & Mohindra, N. (1982). A note on state and trait measures of arousal. *Current Psychological Research*, *2*, 55-88.

Williamson, G. M., & Clark, M. S. (1989). Providing help and desired relationship type as determinants of changes in moods and self-evaluations. *Journal of Personality and Social Psychology*, *56*, 722-734.

Williamson, G. M., & Clark, M. S. (1992). Impact of desired relationship type on affective reactions to choosing and being required to help. *Personality and Social Psychology Bulletin*, *18*, 10-18.

Williamson, G. M., Clark, M. S., Pegalis, L. J., & Behan, A. (1996). Affective consequences of refusing to help in communal and exchange relationship. *Personality and Social Psychology Bulletin*, *22*, 34-47.

Williamson, G. M., Martin-Cook, K., Weiner, M. F., Svetlik, D. A., Saine, K., Hyman, L. S., . . . Shulz, R. (2005). Caregiver resentment: Explaining why care recipients exhibit problem behavior. *Rehabilitation Psychology*, *50*, 215-223. doi:10.1037/0090-5550.50.3.215

Williamson, G. M., & Schulz, R. (1995). Caring for a family member with cancer: Past communal behavior and affective reactions. *Journal of Applied Social Psychology*, *25*, 93-116.

Williamson, G. M., & Shaffer, D. R. (2001). Relationship quality and potentially harmful behaviors by spousal caregivers: How we were then, how we are now. *Psychology and Aging*, *16*, 217-226. doi:10.1037/0882-7974.16.2.217

Windle, M., & Dumenci, L. (1998). An investigation of maternal and adolescent depressed mood using a latent-trait-state model. *Journal of Research on Adolescence*, *8*, 461-484.

Yoon, M., & Millsap, R. E. (2007). Detecting violations of factorial invariance using data-based specification searches: A Monte Carlo study. *Structural Equation Modeling*, *14*, 435-463. doi:10.1080/10705510701301677