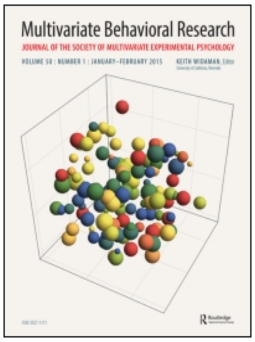

Multivariate Behavioral Research

ISSN: 0027-3171 (Print) 1532-7906 (Online) Journal homepage: http://www.tandfonline.com/loi/hmbr20

Using a Few Snapshots to Distinguish Mountains from Waves: Weak Factorial Invariance in the Context of Trait-State Research

E. L. Hamaker, N. K. Schuurman & E. A. O. Zijlmans

MULTIVARIATE BEHAVIORAL RESEARCH
2017, VOL. 52, NO. 1, 47–60
http://dx.doi.org/10.1080/00273171.2016.1251299

Routledge
Taylor & Francis Group

OPEN ACCESS

Using a Few Snapshots to Distinguish Mountains from Waves: Weak Factorial Invariance in the Context of Trait-State Research

E. L. Hamaker[a], N. K. Schuurman[a], and E. A. O. Zijlmans[b]

[a]Methodology and Statistics, Faculty of Social and Behavioural Sciences, Utrecht University; [b]Department of Methodology and Statistics, TS Social and Behavioral Sciences, Tilburg University

ABSTRACT
In this article, we show that the underlying dimensions obtained when factor analyzing cross-sectional data actually form a mix of within-person state dimensions and between-person trait dimensions. We propose a factor analytical model that distinguishes between four independent sources of variance: common trait, unique trait, common state, and unique state. We show that by testing whether there is weak factorial invariance across the trait and state factor structures, we can tackle the fundamental question first raised by Cattell; that is, are within-person state dimensions qualitatively the same as between-person trait dimensions? Furthermore, we discuss how this model is related to other trait-state factor models, and we illustrate its use with two empirical data sets. We end by discussing the implications for cross-sectional factor analysis and suggest potential future developments.

KEYWORDS
Factorial invariance; longitudinal data analysis; multilevel modeling; trait-state distinction; within-person versus between-person

The Great Wave off Kanagawa is a wood-block print by the Japanese artist Hokusai picturing a massive cresting wave that is towering over three fishing boats, with a snowcapped Mount Fuji in the background. Due to the perspective that is taken, the wave looks much larger and more impressive than the mountain, even though Mount Fuji is the highest peak of Japan. Cattell may have been thinking of an image like this when he used the analogy of a photograph to expose a fundamental problem associated with cross-sectional research: He pointed out that—just as we cannot tell the difference between a mountain and a wave in a frozen snapshot—it is impossible to determine to what extent individual differences observed at a single occasion are due to enduring, trait-like differences between people, and to what extent they reflect transient, state-like fluctuations within people (Cattell, 1978).

In the context of factor analysis, this implies that the common factors (i.e., underlying dimensions) that are obtained from cross-sectional data are partly determined by the between-person covariance structure and partly by the within-person covariance structure. As a result, these factors may not be very meaningful with respect to either of these structures (Cattell, 1967). Put differently, the results from cross-sectional factor analysis may very well form an "uninterpretable blend" of the relationships that exist at the between-person level and the within-person level (cf. Raudenbush & Bryk, 2002). While the danger of mixing within- and between-cluster factor structures has been recognized in the context of multilevel factor analysis when individuals are nested in groups (Muthén, 1994), it has received surprisingly little attention in the context of cross-sectional research, as well as in the context of repeated measures designs (with occasions nested in persons).

In this article, we propose a longitudinal structural equation modeling (SEM) approach that allows us to investigate potential differences between the between-person, trait-like factor structure on the one hand and the within-person, state-like factor structure on the other. In many ways, the model we propose only forms a minor extension of existing factor models; for example, it does not account for some common features of longitudinal data, such as trends or cycles over time, autoregressive relationships, or individual differences in within-person factor structures. However, the unique strength of our model is that it exposes the fundamental and often implicit assumption underlying cross-sectional factor analysis and many longitudinal factor models, that the within-person and the between-person factor structures are the same. Moreover, the model offers a way to test this assumption using as few as two measurement occasions.

The remainder of this article is organized as follows. We begin by discussing how trait-like and state-like sources of variance are contributing to our observations,

CONTACT E. L. Hamaker e.l.hamaker@uu.nl Methods and Statistics, Faculty of Social and Behavioural Sciences, Utrecht University, P.O. Box 80140, 3508 TC, Utrecht, The Netherlands.

how a particular longitudinal factor model can be used to separate them, and how this model can be used to investigate whether the underlying trait-like dimensions that describe enduring differences between individuals coincide with underlying within-person dimensions on which individuals tend to differ from themselves over time. In addition, we discuss how some of the model's restrictions across time can be relaxed (and tested). In the second section, we discuss the connection with other modeling approaches, including multilevel factor analysis and trait-state modeling using SEM. This is followed by two empirical applications that illustrate some of the unique strengths of this approach. We end by discussing the most important findings in our article and some future directions that need to be explored.

Investigating whether trait and state dimensions coincide

We begin by distinguishing between four sources of variance that are likely to contribute to our data: two *time-invariant* and thus *trait-like* sources and two *time-varying* and thus *state-like* sources.[1] Then we show how these four sources affect our cross-sectional factor analysis and how weak factorial invariance across the trait and state factors plays a key role in this context. Subsequently, we present a longitudinal factor model that can be used to separate the four independent sources of variance and allows us to formally investigate whether the underlying between-person dimensions coincide with the underlying within-person dimensions. We end by relaxing some of the assumptions across time that are present in the initial trait-state model we propose.

Distinguishing between four sources of variance

When we have a score on variable j at occasion t for person i, y_{jti}, we can think of this score as consisting of two parts: a trait part that remains stable across time, and a state part that is the temporal deviation from this trait score. Extending this to m variables, and gathering these in an m-variate vector $\boldsymbol{y}_{ti}$, we may write

$$\boldsymbol{y}_{ti} = \boldsymbol{\mu}_i + \boldsymbol{\delta}_{ti}, \tag{1}$$

[1] Throughout, we use rather simplistic definitions of trait and state. Trait is a stable, time-invariant feature, for example, an individual's mean over time, whereas state is something that varies over time, for example, the individual's temporal deviation from this mean. Note that others have used different definitions in the context of SEM models (e.g., Steyer, Mayer, Geiser, & Cole, 2015). Note also that we do not consider the option of trait change here, which is of course an important phenomenon in longitudinal research, especially if we take a life-span perspective.

where $\boldsymbol{\mu}_i$ is an m-variate vector that contains the intraindividual means for person i on the observed variables, which we refer to as the *trait scores* of the individual, and $\boldsymbol{\delta}_{ti}$ is an m-variate vector that contains the temporal deviations from the intraindividual means for individual i at occasion t, which we refer to as the person's *state scores*.

Cattell (1967) proposed that by factor analyzing the trait scores across individuals—an approach he termed the *averaged R-technique*—we can obtain the p underlying trait dimensions; that is,

$$\boldsymbol{\mu}_i = \boldsymbol{c} + \boldsymbol{\Gamma}\boldsymbol{\xi}_i + \boldsymbol{u}_i, \tag{2}$$

where (a) $\boldsymbol{c}$ is an m-variate vector with the grand means; (b) $\boldsymbol{\xi}_i$ is a p-variate vector with *common traits* for individual i, representing an individual's stable position on the underlying trait dimensions; (c) $\boldsymbol{\Gamma}$ is an $m \times p$ matrix with factor loadings relating the m trait scores $\boldsymbol{\mu}_i$ to the p common traits $\boldsymbol{\xi}_i$; and (d) $\boldsymbol{u}_i$ is an m-variate vector that contains the part of the trait scores that could not be accounted for by the common traits and are therefore referred to as *unique traits*. These unique traits are assumed to be multivariate normally distributed with means of zero and an $m \times m$ covariance matrix $\boldsymbol{\Phi}$. The common traits are assumed to be multivariate normally distributed, with means of zero and a $p \times p$ covariance matrix $\boldsymbol{\Omega}$.

The state-part $\boldsymbol{\delta}_{ti}$ from Equation (1) can be factor analyzed in a number of ways. When enough timepoints are available per person, an individual factor analysis can be performed—referred to as Cattell's *P-technique analysis*, and its extension called dynamical factor analysis (Molenaar, 1985)—to obtain truly idiographic results (e.g., Cattell, 1967; Hamaker, Dolan, & Molenaar, 2005). When there are not enough timepoints to do individual factor analysis (especially when the number of timepoints is smaller than the number of variables per wave), we can pool these within-person centered data and factor analyze them simultaneously in what has been called a *pooled P-technique analysis* (Cattell, 1967); that is,

$$\boldsymbol{\delta}_{ti} = \boldsymbol{K}\boldsymbol{\zeta}_{ti} + \boldsymbol{v}_{ti}, \tag{3}$$

where (a) $\boldsymbol{\zeta}_{ti}$ is a q-variate vector with q *common states* for person i at occasion t representing the individual's temporal position on the underlying state dimensions; (b) $\boldsymbol{K}$ is an $m \times q$ matrix with factor loadings relating the m state scores $\boldsymbol{\delta}_{ti}$ to q common states $\boldsymbol{\zeta}_{ti}$; and (c) $\boldsymbol{v}_{ti}$ is an m-variate vector that contains the part of the state scores that cannot be accounted for by the underlying common states and are thus referred to as the *unique states*. These residuals are assumed to be multivariate normally distributed, with means of zero and $m \times m$ covariance matrix $\boldsymbol{A}$. The common states are also assumed to be multivariate normally distributed, with means of zero and $q \times q$ covariance matrix $\boldsymbol{\Pi}$.

Through the preceding reasoning, we have now distinguished between four independent sources of variance (which have been previously identified by others as well; cf. Dumenci & Windle, 1996; Marsh & Grayson, 1994), that is, (a) a *lasting general* source that is invariant over time and variables, which we refer to as the *common trait* ξ; (b) a *lasting-specific* source that is measured by only one variable and is invariant over time, which we refer to as the *unique trait u*; (c) a *temporary-general* source that is measured by multiple variables and varies over time, which we refer to as the common state ζ; and (d) a *temporary-specific* source that is measured by only one variable and that varies over time, which we refer to as the unique state v.

Implications for cross-sectional factor analysis

To see how these four sources of variance contribute to regular cross-sectional factor analysis, we begin by plugging Equations (2) and (3) into Equation (1), which results in

$$\boldsymbol{y}_{ti} = \boldsymbol{c} + \boldsymbol{\Gamma}\boldsymbol{\xi}_i + \boldsymbol{u}_i + \boldsymbol{K}\boldsymbol{\zeta}_{ti} + \boldsymbol{v}_{ti}. \tag{4}$$

Equation (4) reveals something interesting: When we have weak factorial invariance across the two factor structures (i.e., $\boldsymbol{\Gamma} = \boldsymbol{K}$), the model can be expressed as

$$\begin{aligned} \boldsymbol{y}_{ti} &= \boldsymbol{c} + \boldsymbol{\Gamma}\left(\boldsymbol{\xi}_i + \boldsymbol{\zeta}_{ti}\right) + \boldsymbol{u}_i + \boldsymbol{v}_{ti} \\ &= \boldsymbol{c} + \boldsymbol{\Gamma}\boldsymbol{\eta}_{ti} + \boldsymbol{\epsilon}_{ti}, \end{aligned} \tag{5}$$

where the latent variables $\boldsymbol{\eta}_{ti}$ are sums of the common traits $\boldsymbol{\xi}_i$ and common states $\boldsymbol{\zeta}_{ti}$. Furthermore, the residuals $\boldsymbol{\epsilon}_{ti}$ consist of both unique traits $\boldsymbol{u}_i$, which can also be interpreted as *systematic error*, and unique states $\boldsymbol{v}_{ti}$, which can also be interpreted as *random measurement error*. Hence, when we factor analyze cross-sectional data that are generated by the model in Equation (5) (with $\boldsymbol{\Gamma} = \boldsymbol{K}$), the factor solution we obtain adequately represents both the between-person factor structure and the within-person factor structure, as the two are identical. The variance on the common factors $\boldsymbol{\eta}_{ti}$ will be a sum of the trait-like variance of $\boldsymbol{\xi}_i$ and the state-like variance of $\boldsymbol{\zeta}_{ti}$, and the unique variance will be a sum of the variance of systematic error $\boldsymbol{u}_i$ and of the variance of measurement error $\boldsymbol{v}_{ti}$.

By contrast, however, when there is no weak factorial invariance across the common traits and common states (i.e., $\boldsymbol{\Gamma} \neq \boldsymbol{K}$), the model does not simplify to the expression in Equation (5). As a result, factor analyzing cross-sectional data that were generated by the model in Equation (4) results in a factor-loading matrix that forms a blend of $\boldsymbol{\Gamma}$ and $\boldsymbol{K}$, such that the cross-sectionally obtained factors may more closely represent either between-person or within-person factors, depending on the relative contribution of the common traits and common states to the total covariance structure.

It is important to realize that it is impossible to determine whether weak factorial invariance holds across the trait and state structures (i.e., whether data were generated by Equation [4] or Equation [5]) if we only have cross-sectional data. As Cattell (1978) already pointed out with his analogy of a photograph, it is simply impossible to distinguish between lasting and temporal individual differences on the basis of a single measurement occasion. Thus far, the literature focusing on how to tackle this issue has predominantly suggested using intensive longitudinal data, which are then analyzed separately for each person such that the idiographic factor solution could be compared to the cross-sectional factor structure (cf., Zevon & Tellegen, 1982).[2] Although this approach allows individual differences to be maximal, the disadvantage is that it requires a lot of data per person. In the following, we present a longitudinal factor model that requires a minimum of two waves of data and allows us to separate the trait-factor structure from the (average) state-factor structure. In addition, we can use this model to investigate whether weak factorial invariance holds across the trait and state factors (as in Equation [5]): If it does, the underlying trait dimensions coincide with the underlying state dimensions, such that both the enduring differences between individuals (i.e., $\boldsymbol{\xi}_i$) and temporal fluctuations within individuals (i.e., $\boldsymbol{\zeta}_{ti}$) can be situated on identical underlying dimensions. In this case, cross-sectional factor analysis adequately recovers such dimensions. However, when weak factorial invariance does not hold, this implies that the factor structures do not coincide and stable between-person differences are situated on qualitatively different dimensions than within-person fluctuations.

The common and unique trait-state model

The model presented in Equation (4) can be interpreted as a factor model for occasion t (with $t = 1, 2, \ldots, T$). When applying this approach in practice, we can express it as an SEM model for all T occasions simultaneously. To this end, we make use of the stacked vector

[2] Cattell (1967) already pointed out that this was not a clear comparison of state and trait factors because the cross-sectional factor structure is partly determined by the state variance present in the data.

$\boldsymbol{y}_i = [\boldsymbol{y}'_{1i}\ \boldsymbol{y}'_{2i}\ \dots\ \boldsymbol{y}'_{Ti}]'$, such that[3]

$$\begin{bmatrix} \boldsymbol{y}_{1i} \\ \boldsymbol{y}_{2i} \\ \dots \\ \boldsymbol{y}_{Ti} \end{bmatrix} = \begin{bmatrix} \boldsymbol{c} \\ \boldsymbol{c} \\ \dots \\ \boldsymbol{c} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Gamma} & \boldsymbol{I} & \boldsymbol{K} & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{\Gamma} & \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{K} & \dots & \boldsymbol{0} \\ \dots \\ \boldsymbol{\Gamma} & \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{K} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}_i \\ \boldsymbol{u}_i \\ \boldsymbol{\zeta}_{1i} \\ \boldsymbol{\zeta}_{2i} \\ \dots \\ \boldsymbol{\zeta}_{Ti} \end{bmatrix}$$

$$+ \begin{bmatrix} \boldsymbol{v}_{1i} \\ \boldsymbol{v}_{2i} \\ \dots \\ \boldsymbol{v}_{Ti} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Gamma}\boldsymbol{\xi}_i + \boldsymbol{u}_i + \boldsymbol{K}\boldsymbol{\zeta}_{1i} + \boldsymbol{v}_{1i} \\ \boldsymbol{\Gamma}\boldsymbol{\xi}_i + \boldsymbol{u}_i + \boldsymbol{K}\boldsymbol{\zeta}_{2i} + \boldsymbol{v}_{2i} \\ \dots \\ \boldsymbol{\Gamma}\boldsymbol{\xi}_i + \boldsymbol{u}_i + \boldsymbol{K}\boldsymbol{\zeta}_{Ti} + \boldsymbol{v}_{Ti} \end{bmatrix}, \quad (6)$$

where $\boldsymbol{I}$ is an $m \times m$ identity matrix and $\boldsymbol{0}$ is an $m \times q$ zero matrix. In this representation, we have $p + m + q \times T$ latent variables in the vector $\boldsymbol{\eta}$ (that is, p common traits, m unique traits, and for each of the T occasions, q common states). For now, it is assumed the factor-loading matrices for the common traits (i.e., $\boldsymbol{\Gamma}$), and for the common states (i.e., $\boldsymbol{K}$) are invariant over time; we will relax this assumption in the following subsection. The covariance matrix of $\boldsymbol{\eta}$ can be specified as a diagonal block matrix; that is,

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\Omega} \\ \boldsymbol{0}_{(m\times p)} & \boldsymbol{\Phi} \\ \boldsymbol{0}_{(q\times p)} & \boldsymbol{0}_{(q\times m)} & \boldsymbol{\Pi} \\ \boldsymbol{0}_{(q\times p)} & \boldsymbol{0}_{(q\times m)} & \boldsymbol{0}_{(m\times m)} & \boldsymbol{\Pi} \\ \dots \\ \boldsymbol{0}_{(q\times p)} & \boldsymbol{0}_{(q\times m)} & \boldsymbol{0}_{(m\times m)} & \boldsymbol{0}_{(m\times m)} & \dots & \boldsymbol{\Pi} \end{bmatrix}, \quad (7)$$

where the subscripts for the zero matrices indicate their dimensionality. Note that $\boldsymbol{\Omega}$ and $\boldsymbol{\Phi}$ were defined as covariance matrices associated with the trait model (see Equation [2]) and $\boldsymbol{\Pi}$ was defined as a covariance matrix associated with the state model (see Equation [3]). Finally, the covariance matrix of the residuals in Equation (6) is denoted as $\boldsymbol{\Theta}$, and it is the diagonal block matrix

$$\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{0}_{(m\times m)} & \boldsymbol{A} \\ \dots \\ \boldsymbol{0}_{(m\times m)} & \boldsymbol{0}_{(m\times m)} & \dots & \boldsymbol{A} \end{bmatrix}, \quad (8)$$

where $\boldsymbol{A}$ was defined as a covariance matrix associated with the state model (see Equation [3]). We refer to the model in Equations (6)–(8) as the common and unique trait-state (CUTS) model to emphasize that it separates between four sources of variance. In the upper left panel of Figure 1, an example of the CUTS model is given, with four indicators, three measurement occasions, and one common trait and one common state (per occasion).

The CUTS model can be used to investigate whether the trait dimensions coincide with the state dimensions, through constraining the factor loadings; that is, $\boldsymbol{\Gamma} = \boldsymbol{K}$. Since the resulting model is nested under the model in which this constraint is not imposed, we can simply perform a chi-square difference test: If the null hypothesis of no difference between models is not rejected, the researcher may conclude that weak factorial invariance holds across the trait- and state-factor structures.[4]

Relaxing (and testing) the constraints over time in the CUTS model

In the CUTS model presented in the preceding, all the parameters are constrained to be invariant over time. As a result, changes over time in $\boldsymbol{y}_{ti}$ are solely accounted for by changes in the common states $\boldsymbol{\zeta}_{ti}$ and the unique states $\boldsymbol{v}_{ti}$. However, such strict stationarity constraints may not be realistic in practice, and we could actually investigate whether they hold. To this end, we propose to relax these constraints in the CUTS model, thus allowing for the means, the factor loadings, the variances, and covariance to vary over time. In this case, Equation (6) becomes

$$\begin{bmatrix} \boldsymbol{y}_{1i} \\ \boldsymbol{y}_{2i} \\ \dots \\ \boldsymbol{y}_{Ti} \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ \dots \\ \boldsymbol{c}_T \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Gamma}_1 & \boldsymbol{I} & \boldsymbol{K}_1 & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{\Gamma}_2 & \boldsymbol{R}_2 & \boldsymbol{0} & \boldsymbol{K}_2 & \dots & \boldsymbol{0} \\ \dots \\ \boldsymbol{\Gamma}_T & \boldsymbol{R}_T & \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{K}_T \end{bmatrix}$$

$$\times \begin{bmatrix} \boldsymbol{\xi}_i \\ \boldsymbol{u}_i \\ \boldsymbol{\zeta}_{1i} \\ \boldsymbol{\zeta}_{2i} \\ \dots \\ \boldsymbol{\zeta}_{Ti} \end{bmatrix} + \begin{bmatrix} \boldsymbol{v}_{1i} \\ \boldsymbol{v}_{2i} \\ \dots \\ \boldsymbol{v}_{Ti} \end{bmatrix}, \quad (9)$$

where $\boldsymbol{R}_t$ is a diagonal $m \times m$ matrix with the factor loadings for the unique traits at occasion t on its diagonal. Furthermore, the blocks $\boldsymbol{\Pi}$ in Equation (7) will now carry a time index, as will the blocks $\boldsymbol{A}$ in Equation (8).

It should be noted that when there are only two waves of data, the factor loadings of common traits and the common states can be allowed to vary over time (i.e., $\boldsymbol{\Gamma}_t$ and $\boldsymbol{K}_t$), but the factor loadings of the unique traits (i.e.,

[3] We are making use of the well-known measurement equation $\boldsymbol{y} = \boldsymbol{\tau} + \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\epsilon}$ from the LISREL model here (with $\boldsymbol{\eta} \sim N(\boldsymbol{0}, \boldsymbol{\Psi})$ and $\boldsymbol{\epsilon} \sim N(\boldsymbol{0}, \boldsymbol{\Theta})$); alternatively, one could use the RAM notation.

[4] Weak factorial invariance is one of the first steps in investigating measurement invariance when applied to multiple group analysis or longitudinal factor analysis (cf. Meredith, 1993). Subsequent steps typically consist of constraining the intercepts across groups or timepoints (resulting in strong factorial invariance) and constraining the residual variances across groups or timepoints (resulting in strict factorial invariance). In the current setting, however, the factor model for the state part by definition has a mean vector of zero, while the trait part has a mean vector containing the grand means $\boldsymbol{c}$. Hence, the concept of strong factorial invariance does not make sense here.

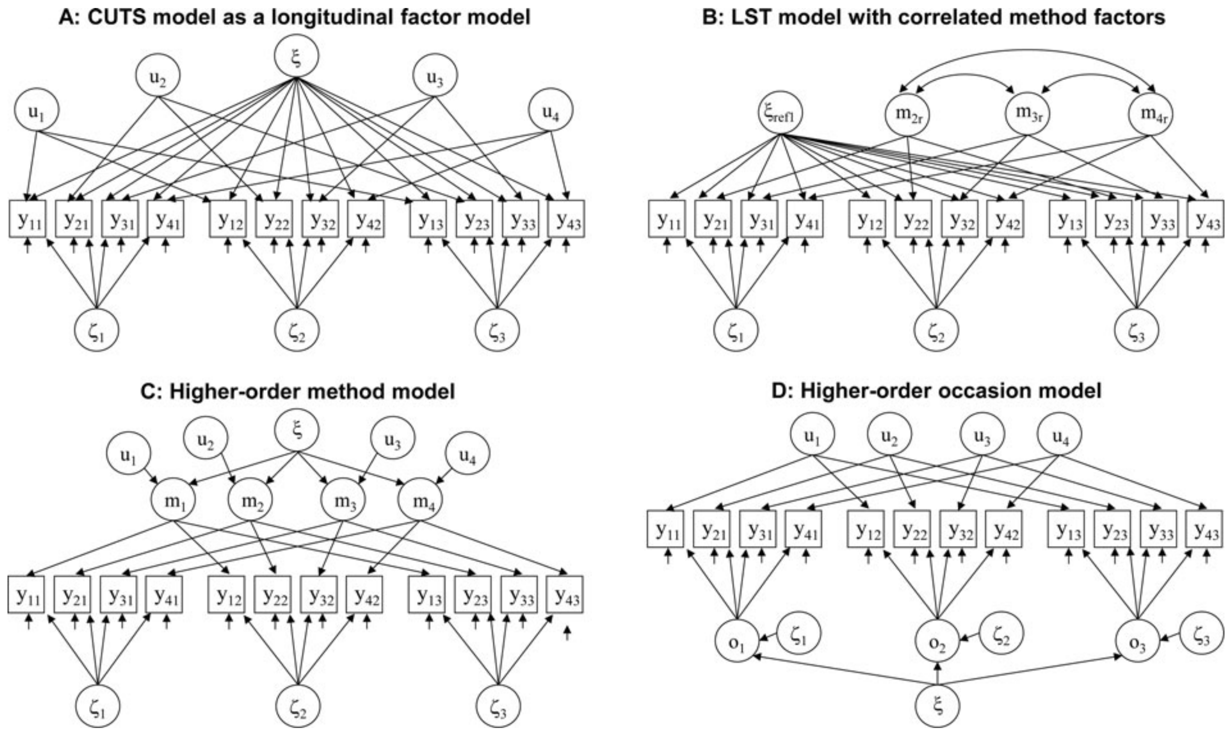

Figure 1. Four models for decomposing observed variances into trait and state components. Panel A contains the common and unique trait-state (CUTS) model presented in the current paper; panel B contains the latent state-trait (LST) model with correlated method factors as discussed by Geiser and Lockhart (2012), with indicator 1 as the reference indicator; panel C contains the higher-order method (HOM) model that includes the common trait as a second-order factor that links the method factors (i.e., the higher-order item model by Marsh and Grayson, 1994); panel D contains the higher-order occasion (HOO) model, which includes the common trait as a second-order factor that links the occasion factors (i.e., the higher-order time model by Marsh and Grayson, 1994).

$\boldsymbol{R}_t$) cannot. The reason for this is that when there are only two waves, these unique traits have only two indicators each, such that additional constraints are needed to ensure the model is identified (i.e., $\boldsymbol{R}_t = \boldsymbol{I}$). When three or more waves are available, the factor loadings for the unique traits no longer have to be constrained over time, as there will be three or more indicators for each of these factors.

In practice, we propose to start with this least restricted version of the CUTS model, followed by a model in which the factor loadings are constrained over time (i.e., weak factorial invariance across time for the common traits and the common states as in Equation [6]); if this constraint is tenable, we may proceed with constraining the factor loadings across the trait- and the state-factor structures (i.e., weak factorial invariance across traits and states through $\boldsymbol{\Gamma} = \boldsymbol{K}$). Finally, additional time-invariant constraints on the (co-)variances of the residuals (in $\boldsymbol{\Theta}$) and the latent variables (in $\boldsymbol{\Psi}$) as initially included in Equations (7) and (8) can be of interest, because—if these hold—the contribution of the four distinct sources of variance to our observations is invariant over time. However, these constraints are not crucial when the interest is in whether the underlying dimensions coincide (i.e., they are only of interest when we want to know whether the reliabilities of the indicators are invariant across trait and state structures); therefore, we do not emphasize these latter constraints here.

Connection with other longitudinal factor models

The CUTS model presented here can be thought of as a missing link between three different strands of literature. First, the work by Cattell and his followers on isolating the within-person factor structure and how to distill the between-person factor structure forms an important source of inspiration for the work presented here (Borsboom, Mellenbergh, & Van Heerden, 2003; Cattell, 1967; Hamaker, Nesselroade, & Molenaar, 2007; Kievit, Frankenhuis, Waldorp, & Borsboom, 2013; Molenaar, 2004; Molenaar, Huizenga, & Nesselroade, 2003; Voelkle, Brose, Schmiedek, & Lindenberger, 2014). Many of the contributions in this area have focused on the idiographic nature of within-person factor structures, thus requiring a single-subject approach based on a large number of repeated measures. In some of these articles, there has also been a focus on the theoretical and sometimes also empirical comparison between within-person factor structures and the cross-sectional factor structure. However, most of these contributions have overlooked the fact that the cross-sectional factor structure is a weighted sum of the within-person factor structure and the between-person factor structure. The integrated trait-state (ITS) model proposed by Hamaker et al. (2007) actually considers this between-person trait factor model and combines it with individual latent first-order vector autoregressive models (to allow for a purely idiographic factor structure as well as relationships between common states over time). The advantage of the CUTS model proposed here is that it does not require extensive time series data per person; however, this comes at the cost of not being able to model truly idiographic factor structures.

Second, the strictly stationary CUTS model can also be represented as a multilevel factor model using the within-person state-factor model given in Equation (3) as the Level 1 equation, while the between-person trait-factor model given in Equation (2) is used as the Level 2 equation. A detailed comparison between the CUTS model conceived of as a multilevel factor model and the longitudinal factor model approach presented in Equations (6) to (8) is included in the online supporting material for the current article: It shows that these two approaches lead to the exact same parameter estimates and can thus be considered equivalent, even though the data are organized in different ways.

Third, there are numerous longitudinal factor modeling approaches in SEM that have been proposed to separate trait-like and state-like sources of variance in multiple indicator models. In the following, we will focus on two issues in particular in this context: (a) the diverse ways in which other multiple indicator longitudinal trait-state models have separated lasting-specific variance from the other sources of variance and (b) the effect of including the lasting-general factor as a second-order factor in the model, rather than as a first-order factor.[5]

Modeling lasting-specific variance: Unique traits versus other method factors

In the CUTS model presented in the preceding, we model lasting-specific variance as the unique traits, which are not allowed to be correlated with each other or the common trait (see also panel A in Figure 1). The rationale for this is that it represents variance that is specific to a particular indicator, whereas the common trait represents all lasting variance that the indicators share.

There are a number of other ways in which one can account for variance that is particular to a specific indicator. Geiser and Lockhart (2012) discussed and compared several approaches, indicating that the uncorrelated factors approach such as we discussed in the preceding is most popular in the literature. However, they criticized this approach on both substantial and practical grounds. Substantially, they indicate that the uncorrelated factors are not compatible with the latent state-trait (LST) theory, which makes the interpretation of these factors problematic in their opinion.[6] Practically, they showed through simulations that the model often results in inadmissable solutions, convergence problems, or a nonsignificant variance of one of these uncorrelated factors. Hence, although the model with uncorrelated factors is *theoretically identified*, it is often *empirically unidentified*, meaning that there is not enough information in the data to obtain a unique and acceptable solution. Clearly, this hampers the application of this model in practice, as we will also see in our empirical illustrations. We revisit the issue of identification in the Discussion section.

An alternative approach, which has been advocated in the LST literature, is to allow the method factors, which capture what is specific to a particular indicator, to be correlated to each other (Eid, Schneider, & Schwenkmezger, 1999). However, this model is only theoretically identified if one of the method factors is omitted, such that Geiser and Lockhart (2012) referred to this as the $M - 1$ correlated methods approach. An example of this model is included in panel B of Figure 1. Two issues associated with this approach are worth noting here.

First, the use of a reference indicator and the inclusion of the correlations between the method factors implies that variance is redistributed across the different trait-like factors in this model. As a result, the general trait factor can no longer be interpreted as capturing all that the indicators have in common, as some of what the nonreference indicators have in common is now captured by the correlations between the method factors. Hence, it is not possible for this model to exclusively link particular model aspects to the four separate sources we distinguished (see also Marsh & Grayson, 1994).

[5] Note that the diverse strands of literature that we pull from use different terminology. Since we aim to use terminology in a consistent manner throughout this article, our language necessarily deviates at times from the way others have used the terms *trait* versus *state*, *common* versus *unique*, and *occasion* and *method* in the literature.

[6] The reason for this, as explained by Geiser and Lockhart (2012), is that LST theory (which is heavily based on classical test theory) is concerned with a *random experiment* in which individuals and situations (i.e., occasions) are sampled. Subsequently, all latent variables are defined in terms of expectations that are conditional on person or on the person × occasion interaction, or they are functions of these expectations. Uncorrelated method factors (which we refer to as unique traits here) cannot be defined in this way and are therefore not compatible with LST theory (see Geiser and Lockhart [2012] for a more detailed discussion). It would be interesting to see, however, whether extending the random experiment in LST theory with the random sampling of variables (i.e., items) would overcome this limitation. The notion that we are also sampling variables is congruent with Cattell's data box (which consists of three dimensions: individuals, timepoints, and variables) and has been used in classical test theory (e.g., Lord, 1965).

Table 1. Common and unique trait-state (CUTS) model compared to four versions of the correlated-method (CM) model, which uses one indicator as the reference indicator through omitting the method-specific factor for that indicator (denoted as minus M1 to M4). The left part contains the chi-square test for the models without any constraints; the right part contains the results for models with factor loadings to be invariant over time (i.e., for the common trait or general trait factor, the common states or occasion factors, and the unique traits or method factors).

	Nonconstrained factor loadings			Time-invariant factor loadings		
	χ^2	df	p	χ^2	df	p
CUTS model	30.45	30	.44	49.15	52	.59
CM minus M1 model	32.46	30	.35	48.78	50	.52
CM minus M2 model	34.14	30	.28	48.78	50	.52
CM minus M3 model	34.46	30	.26	48.78	50	.52
CM minus M4 model	26.35	30	.66	48.78	50	.52

Second, the results of this $M - 1$ correlated methods approach depend on the choice of reference indicator: Specifically, models based on a different reference indicator are not statistically equivalent unless the factor loadings are constrained to be equal over time. To illustrate this, we simulate data based on a strictly stationary CUTS model with four indicators, three waves, and 500 cases. First, we fit the CUTS model, as well as all the four versions of the $M - 1$ correlated methods models (using the first indicator as the reference indicator, then the second, and so on), without constraints on the factor loadings over time. The results for the chi-square test are included in Table 1, showing that none of these models are statistically equivalent. Subsequently, we impose weak factorial invariance across time in all five of these models; the results are presented in the right part of Table 1. It is clear that when the factor loadings are constrained to be invariant over time, the choice of the reference indicator no longer matters; however, the $M - 1$ correlated method model still provides a different solution than the CUTS model in this case.

Higher-order trait-state models

Many multiple-indicator trait-state models include the general lasting factor (i.e., the common trait factor) as a second-order factor. We highlight two possibilities here. First, we consider a higher-order methods (HOM) model (also referred to as the higher-order item model, see Marsh & Grayson, 1994), which is graphically represented in panel C of Figure 1. In this model, the common trait is included as a second-order factor that connects the first-order method factors. The residuals of these method factors can be considered to represent the unique traits, as these are the components that do not vary over time and that are specific to a particular indicator. Second, we consider a higher-order occasion (HOO) model (also referred to as the higher-order time model by Marsh & Grayson, 1994), which is represented in panel D of Figure 1. In this model, the common trait is included as a second-order factor connecting the first-order occasion factors; the residuals of these occasion factors can be considered to represent the common states, as they capture what the indicators have in common at a particular occasion, but not with indicators at other occasions.

We are particularly interested in the circumstances under which the higher-order models and the first-order CUTS model are statistically equivalent. Yung, Thissen, and McLeod (1999) have shown that second-order models are nested under first-order models that contain a general factor such as our common traits. To illustrate this, we make use of the same simulated data set that we used before (i.e., four indicators, three waves, and 500 cases) and compare the chi-square and *df* of the diverse models. Specifically, we consider two versions of the CUTS model: one in which there are no constraints imposed across the factor loadings of the common trait and common state (referred to as the *free CUTS model*) and one in which these factor loadings are constrained to be identical, implying we have weak factorial invariance across the common trait and common state (referred to as the *WFTS CUTS model*). In addition, we consider the HOM model and the HOO model. For all four models, we consider two versions: the first one in which there are no constraints imposed on the factor loadings across time and the second one in which the factor loadings are constrained to be invariant over time.[7]

The results presented in Table 2 show that (a) the HOM model with time-invariant factor loadings is equivalent to the CUTS model with time-invariant factor loadings (i.e., both models have a χ^2 of 49.15 with $df = 52$) and (b) the HOO model with time-invariant factor loadings is equivalent to the CUTS model with time-invariant factor loadings *and* weak factorial invariance across the common trait and common state (i.e., both models have a χ^2 of 50.33 with $df = 55$). Furthermore, the results show that when we compare the models without constraints across time, the free CUTS model is the least restrictive, whereas the WFTS CUTS model (with weak factorial invariance across common trait and common state, but not across time, such that $\boldsymbol{\Gamma}_t = \boldsymbol{K}_t$) is actually the most restrictive model. Without going into the details, we point out that it is also possible to impose certain proportionality constraints on the factor loadings of the CUTS

[7] Note that for the free CUTS model, this constraint implies we have $\boldsymbol{\Gamma}_t = \boldsymbol{\Gamma}$, $\boldsymbol{R}_t = \boldsymbol{I}$, and $\boldsymbol{K}_t = \boldsymbol{K}$; for the WFTS CUTS model, we start with the constraint $\boldsymbol{\Gamma}_t = \boldsymbol{K}_t$, and the additional constraint of weak factorial invariance across time implies we get $\boldsymbol{\Gamma} = \boldsymbol{K}$.

Table 2. Comparison between four models without and with weak factorial invariance over time using simulated data based on four indicators measured at three occasions. Two versions of the common and unique trait-state (CUTS) model are included: without weak factorial invariance across the common trait and common state factors (free CUTS model) and with weak factorial invariance across the common trait and common state factors (WFTS CUTS model). The other two models are the higher-order occasion (HOO) model and the higher-order method (HOM) model. The left part of the table contains results for models without constraints over time; the right part contains the results when weak factorial invariance over time is imposed.

	No constraints			Time-invariant		
	χ^2	df	p	χ^2	df	p
free CUTS model	30.45	30	.44	49.15	52	.59
WFTS CUTS model	45.86	41	.28	50.33	55	.65
HOO model	44.15	39	.26	50.33	55	.65
HOM model	36.88	38	.52	49.15	52	.59

model with time-varying factor loadings, such that this model becomes equivalent to the HOO or HOM model with time-varying factor loadings, implying that the latter two are special cases of the former.

These comparisons lead to an interesting observation: If the within-person and the between-person factor structures differ, using a HOO model will result in factor loadings that are a weighted mix of these two factor structures. This means that the problem we identified before with respect to cross-sectional research actually *also arises in the context of longitudinal trait-state models* if the within-person covariance structure is not properly separated from the between-person covariance structure.

Empirical applications

We consider two empirical applications to illustrate the use of the CUTS model. In the first application we have four waves of data in two groups, and we illustrate how the hypotheses of weak factorial invariance over time and weak factorial invariance across the common trait structure and common state structure can be investigated. The second application is based on two waves of data, and the central questions are whether the underlying factor structure is characterized by one or two factors and whether this is different for the trait-like and the state-like parts of the model.

Depression in adolescents

Dumenci and Windle (1996) studied the degree of stability in adolescents' depression through applying the HOO model to four waves of CES-D data from adolescents. Specifically, they had data for 372 males and 433 females, which they analyzed separately. The indicators they used were the four subscales of the CES-D, that is, *Depressed Affect* (DA), *Positive Affect* (PA), *Somatic Complaints* (SC), and *Interpersonal* (I). In their models, they distinguished between a common trait factor, a common state factor per occasion, uncorrelated method factors (i.e., unique traits), and residuals (i.e., unique states). However, the common trait was included as a second-order factor that connected the occasion factors, which implies it was (implicitly) assumed that the common trait and common state dimensions coincide.

By contrast, we begin with the CUTS model that is characterized by (a) one common trait that directly influences all four indicators at all four measurement occasions; (b) four unique traits that directly influence the same indicator at different occasions; (c) one common state per occasion (i.e., four in total), which directly influences the indicators within the same occasion; and (d) a unique state for each indicator at each occasion (note that a graphical representation of this model would be like the upper-left panel of Figure 1, with one more wave added). Even though this model is theoretically identified (as we established analytically using the approach described in Bekker, Merckens, & Wansbeek, 1993),[8] it did not converge for either males or females, which implies the model is empirically unidentified. Therefore, we omitted the unique traits for DA in both groups, as these resulted in nonsignificant negative variance estimates. As a result, Model 1 contains one common trait, four common states, three unique traits, and sixteen unique states (see the online supporting material for a detailed representation of the models used here for males). The model fit for both males and females is reported in Table 3.

In Model 2, we constrained all factor loadings across time (i.e., $\boldsymbol{\Gamma}_t = \boldsymbol{\Gamma}$ for the common traits; $\boldsymbol{R}_t = \boldsymbol{I}$ for the unique traits; and $\boldsymbol{K}_t = \boldsymbol{K}$ for the common states). The chi-square difference test comparing Model 2 to Model 1 was not significant for males, indicating that we can assume weak factorial invariance across time for this group (which, as we have seen in the previous section, is actually equivalent to a HOM model with time-invariant factor loadings). However, for females, the chi-square difference test was significant, which implies that weak factorial invariance over time is not tenable in this group.

[8] To determine whether a model is theoretically identified, one can use the approach proposed by Bekker et al. (1993). In a first step, the covariance matrix of the model is determined in terms of the unknown parameters based on the LISREL model $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda}' + \boldsymbol{\Theta}$. Subsequently, all unique elements of this matrix are placed in a vector, and a second vector is created with all the unknown parameters. Then, the derivative of the first vector toward the elements of the second is taken; this results in the Jacobian matrix. Finally, the null space of this matrix is considered. If it is empty, the model is identified; if it is not empty, the model is not identified (cf. Bekker et al., 1993). We performed this procedure using Maple (version 18.01), which is a symbolic (and numeric) computing environment.

Table 3. Fit measures for the common and unique trait-state (CUTS) model in a sample of males and a sample of females observed at four waves with the CES-D (from Dumenci & Windle, 1996). $\mathbf{\Gamma}_t$ contains the factor loadings by which the indicators at occasion t load on the common trait; $\boldsymbol{K}_t$ contains the factor loadings by which the indicators at occasion t load on the common state of occasion t.

Model description	χ^2	*df*	*p*	$\Delta\chi^2$	Δdf	*p*	RMSEA	CFI	SRMR	AIC	BIC
Males											
Model 1: CUTS model	129.03	76	.01				.04	.98	.04	25420	25655
Model 2: Model 1 with $\mathbf{\Gamma}_t = \mathbf{\Gamma}$ and $\boldsymbol{K}_t = \boldsymbol{K}$	162.39	106	<.01	33.36	30	.31	.04	.98	.05	25393	25511
Model 3: Model 2 with $\mathbf{\Gamma} = \boldsymbol{K}$	203.46	109	<.01	41.07	3	<.01	.05	.97	.07	25428	25534
Females											
Model 1: CUTS model	106.09	76	.01			<.01	.03	.99	.03	31364	31608
Model 2: Model 1 $\mathbf{\Gamma}_t = \mathbf{\Gamma}$ and $\boldsymbol{K}_t = \boldsymbol{K}$	158.50	106	<.01	52.41	30	<.01	.03	.99	.07	31357	31479
Model 3: Model 1 with $\mathbf{\Gamma}_t = \boldsymbol{K}_t$	160.84	91	<.01	54.75	15	<.01	.04	.99	.05	31370	31566

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

Model 3 is based on imposing weak factorial invariance constraints across the common trait and common states. For males, we add this constraint to Model 2 (i.e., we have weak factorial invariance across time and across common trait and common states). To show how to proceed when factor loadings are not invariant over time, as was the case for the sample of women, we added the constraint across trait and state loadings to Model 1 for this group (i.e., we had $\mathbf{\Gamma}_t = \boldsymbol{K}_t$ for each t). In both cases, the chi-square difference test was significant, indicating that identical factor loadings across common trait and common states cannot be assumed. Hence, the underlying latent depression dimension on which individuals *differ from each other structurally over time* is not the same dimension as the underlying latent depression dimension on which individuals *differ from themselves over time*.

For males, these results imply that the common trait depression cannot be included as a second-order factor connecting the depression factors at the different occasions (i.e., the HOO model, which was originally used by Dumenci and Windle [1996], does not hold). However, it could be included as a second-order factor connecting the method factors (i.e., the HOM model does hold). For females, we can further investigate whether the common trait depression can be included as a second-order factor connecting the occasion factors (i.e., the HOO model), even though there is no weak factorial invariance over time. This model is slightly less restrictive than the model in which we constrain the factor loadings to be identical across the common trait and common states, which we discussed in the preceding. However, this time-varying HOO model also fitted significantly worse than the first model (i.e., the chi-square difference test was $136.27 - 106.09 = 30.18$, $df = 88 - 76 = 12$, $p < .01$). The factor loadings for the best-fitting models for males and females are given in Table 4.

In summary, we can conclude that for both males and females, the trait depression cannot be included as a second-order factor that connects the occasion factors. This implies that the way the common trait depression influences the observations is qualitatively different from the way the common states influence the observations, both for males and for females. It also implies that a factor structure that is obtained with cross-sectional factor analysis or a longitudinal higher-order factor model (like the HOO model) represents neither the trait nor the state factor structure.

Togetherness in elderly

Tiikkainen, Heikkinen, and Leskinen (2004) investigated the way that elderly people experience various aspects of their social life. To this end, they obtained data from 111 participants who were 80 years old at the first wave and

Table 4. Factor loadings for males and females in the CES-D data of Dumenci and Windle (1996) based on the common and unique trait-state (CUTS) model with weak factorial over time for males and time-varying factor loadings for females.

Group	Variable	$\hat{\gamma}$	(*SE*)	$\hat{\kappa}$	(*SE*)
Males	DA	1.00	—	1.00	—
	PA	0.54	(0.05)	0.42	(.03)
	SO	0.81	(0.05)	0.62	(.04)
	IN	0.31	(0.02)	0.15	(.01)
Females	DA1	1.00	—	1.00	—
	PA1	0.50	(0.04)	0.30	(0.04)
	SO1	0.54	(0.05)	0.53	(0.06)
	IN1	0.21	(0.02)	0.15	(0.02)
	DA2	1.17	(0.09)	1.00	—
	PA2	0.53	(0.05)	0.34	(0.05)
	SO2	0.69	(0.07)	0.58	(0.07)
	IN2	0.22	(0.02)	0.16	(0.03)
	DA3	1.02	(0.08)	1.00	—
	PA3	0.53	(0.05)	0.39	(0.05)
	SO3	0.59	(0.06)	0.78	(0.09)
	IN3	0.23	(0.03)	0.18	(0.03)
	DA4	0.77	(0.08)	1.00	—
	PA4	0.42	(0.05)	0.42	(0.04)
	SO4	0.50	(0.06)	0.63	(0.05)
	IN4	0.16	(0.02)	0.14	(0.02)

Note. $\hat{\gamma}$ = the estimated factor loading by which the indicator loads on the common trait; $\hat{\kappa}$ = the factor loading by which the indicator loads on the common state; *SE* = the standard error of the estimate; DA = *Depressed Affect*; PA = *Positive Affect* (reversely coded); SO = *Somatic Complaints*; IN = *Interpersonal*.

85 years old at the second wave. They used the Social Provision Scale, which results in six scores (i.e., *Attachment, Reliable alliance, Guidance, Social integration, Opportunity of nurturance*, and *Reassurance of worth*).

The researchers had two main research questions. First, they wanted to know whether these six scores measure a single underlying construct, *Togetherness*, or two separate constructs, *Emotional Togetherness* (with *Attachment, Reliable alliance*, and *Guidance* as its indicators) and *Social Togetherness* (with *Social integration, Opportunity of nurturance*, and *Reassurance of worth* as its indicators). To tackle this question, they analyzed their data cross-sectionally (i.e., for each occasion separately) and concluded that at both occasions a two-factor model provided a better fit than a one-factor model. Second, they wanted to determine whether the experience of togetherness was stable over time. To address this question, they used longitudinal factor models with autoregressive relationships between the common factors at the subsequent timepoints and concluded that there was a substantial amount of stable variance across time.

The approach taken by Tiikkainen et al. (2004) is less than ideal in several ways. Most important, they did not decompose the total variance into stable, trait-like between-person variation and temporal, state-like within-person variation, and as a result it is not clear whether the two-factor solution they found represents both the between-person and the within-person factor models, or predominantly one of them. We specify a series of models that allow us to investigate whether the dimensions underlying these structures actually coincide, specifically focussing on the question whether togetherness should be perceived of as one or two (correlated) dimensions at the within-person and the between-person levels.

We began with a model in which we have two common traits (which are allowed to correlate), and at each occasion two common states (which are allowed to correlate within the same occasion but not across occasions). Furthermore, instead of including six unique traits (i.e., one for each indicator), we included these components as covariations between the unique parts of the indicators. This approach is identical to including unique trait factors when only two waves of data are available. Although this model is theoretically identified (as we established analytically, using Bekker et al., 1993), it did not converge, implying it is not empirically identified. After fixing the factor loading for the first common trait to the first variable at the *second* occasion to 1 in addition to the usual identification constraints, the model did converge (note that this additional constraint is part of the constraints that are imposed when testing whether the factor loadings are invariant over time). We refer to this model as Model 1, and the fit measures are included in Table 5 (see the online supporting material for a detailed presentation of the models used in this application).

Because all covariances between the unique parts of the indicators (representing the presence of unique traits) were not significantly different from zero, we omitted them in Model 2. The chi-square difference test indicates that this did not lead to a significant decrease in fit (see Table 5), and therefore we conclude that in this data set there was no unique trait variance. A graphical representation of this model is given in the upper half of Figure 2. In Model 3, we constrained all the factor loadings over time (i.e., $\boldsymbol{\Gamma}_t = \boldsymbol{\Gamma}$ and $\boldsymbol{K}_t = \boldsymbol{K}$); the chi-square difference test (compared to Model 2) was not significant, implying that time-invariant factor loadings are tenable.

Model 4 is based on constraining the factor loadings across the common traits and common states, implying weak factorial invariance across the two common traits and two common state factors. The chi-square difference test for this model (in comparison to Model 3) was significant, implying that weak factorial invariance across the trait and state structures does not hold. Hence, within-person changes in togetherness occur on *qualitatively different* dimensions from the dimensions on which stable between-person differences in togetherness are located. This also implies that cross-sectional analyses result in some mix of the within-person and between-person factor structures, while it is unclear to what extent it represents either. Specifically, it is not clear whether the two-factor structure obtained by Tiikkainen et al. (2004) actually represents both the within-person and the between-person factor structure, or primarily one of them.

To further investigate this matter, we specified a model with only one common trait, while keeping the two-factor model for the common states. We refer to this model, which is represented in the lower part of Figure 2, as Model 5. The chi-square difference for this model (in comparison to Model 3), was not significant, such that we can conclude that the trait part is characterized by only one underlying dimension. Subsequently, we replaced the two common states per occasion by a single common state per occasion in Model 6. The chi-square difference test for this model (in comparison to Model 5) was significant, indicating that the common state structure was actually characterized by two underlying dimensions rather than one.

In summary, the current approach has shown that the within-person fluctuations occur on two underlying dimensions (i.e., *Emotional Togetherness* and *Social Togetherness*), whereas the between-person differences are located on a single underlying dimension (i.e., *Togetherness*). Hence, it seems that the cross-sectional two-factor solution found earlier by Tiikkainen et al. (2004) was

Table 5. Fit measures for six versions of the common and unique trait-state (CUTS) model, applied to the togetherness ratings from Tiikkainen et al. (2004). $\boldsymbol{\Gamma}_t$ contains the factor loadings by which the indicators at occasion t load on the common traits; $\boldsymbol{K}_t$ contains the factor loadings by which the indicators at occasion t load on the common states of occasion t.

Model description	χ^2	*df*	*p*	$\Delta\chi^2$	Δdf	*p*	RMSEA	CFI	SRMR	AIC	BIC
Model 1: CUTS model with two common traits and two common states	65.29	34	<.01				.09	.94	.07	5,152	5,271
Model 2: Model 1 without unique traits	69.88	40	<.01	4.59	6	.60	.08	.94	.08	5,144	5,247
Model 3: Model 2 with with $\boldsymbol{\Gamma}_t = \boldsymbol{\Gamma}$ and $\boldsymbol{K}_t = \boldsymbol{K}$	78.10	49	<.01	8.22	9	.51	.07	.94	.08	5,135	5,213
Model 4: Model 3 with $\boldsymbol{\Gamma} = \boldsymbol{K}$	90.20	53	<.01	12.01	4	.02	.08	.93	.07	5,139	5,206
Model 5: Model 3 with a single common trait	78.13	50	<.01	0.03	1	.86	.07	.94	.08	5,133	5,208
Model 6: Model 5 with a single common state per occasion	89.97	53	<.01	11.84	3	.01	.08	.93	.08	5,138	5,206

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

dominated by the within-person state structure. The factor loadings for the final model are presented in Table 6.

Discussion

Factor analysis as developed by Spearman (1904) was originally intended to extract a general underlying trait (i.e., intelligence) from a set of correlated tests, allowing the separate indicators to have residuals that contain both systematic and random measurement error. In line with this tradition, it is often assumed that the common factors that are obtained with cross-sectional analysis represent trait-like underlying dimensions, while the measurement errors can contain both time-invariant and time-varying components. However, as has become clear in this article, this is not correct when there are common states that contribute to the observed variance, as these sources will also affect the factor structure. Furthermore, those who have been concerned about the idiographic nature of within-person factor structures have often (implicitly)

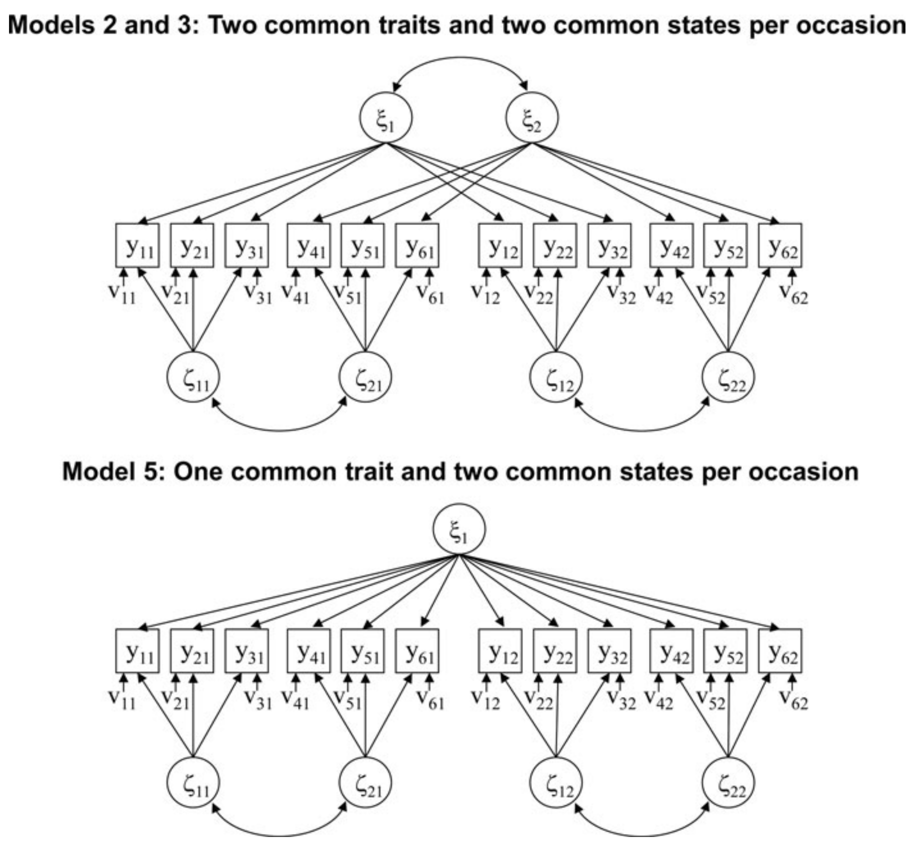

Figure 2. Two longitudinal factor models for the two-wave, six-indicator data from Tiikkainen et al. (2004) to separate the common traits ξ, common states ζ, and unique states (i.e., measurement error) v from each other. There was no evidence for unique traits (i.e., systematic error) u in these data. The top model is based on assuming a two-factor structure for both the common traits and the common states; the bottom model consists of a one-factor model for the between-person part combined with a two-factor structure for the within-person part.

Table 6. Factor loadings and proportions of explained variance for the one-common-trait two-common-states version of the common and unique trait-state (CUTS) model selected for the togetherness data measured by Tiikkainen et al. (2004).

Variable	T trait $\hat{\gamma}$ (SE)	ET state $\hat{\kappa}$ (SE)	ST state κ (SE)
ATT	1.00(—)	1.00(—)	—
REL	0.58(.13)	1.31(.18)	—
GUI	0.80(.11)	1.65(.22)	—
REA	0.55(.14)	—	1.00(—)
SOC	0.94(.16)	—	1.88(.75)
OPN	0.98(.20)	—	0.94(.48)

Note. T = the common trait *Togetherness*; ET = the common state *Emotional Togetherness*; ST = the common state *Social Togetherness*; $\hat{\gamma}$ = the estimated factor loading by which the indicator loads on the common trait; $\hat{\kappa}$ = the factor loading by which the indicator loads on the common state; *SE* = standard error of the estimate; ATT = *Attachment*; REL = *Reliable alliance*; GUI = *Guidance*; REA = *Reassurance of worth*; SOC = *Social integration*; OPN = *Opportunity of nurturance*.

assumed that the common factors obtained with cross-sectional analysis reflect an average (across individuals) of underlying within-person dimensions (cf. Borsboom et al., 2003; Molenaar et al., 2003; Voelkle et al., 2014; Zevon & Tellegen, 1982). This is equally untrue, however, because when there are common traits that contribute to the observed variance, the cross-sectional factor structure will also represent this trait factor structure to some extent (cf. Cattell, 1967).

In the current article, we presented a longitudinal factor analytical approach that can be used to separate the (average) within-person factor structure from the between-person factor structure, and we explained how testing for weak factorial invariance across these two factor structures is of key interest in this context; if it holds, this implies that (a) these two kinds of underlying dimensions coincide such that the enduring individual difference can be situated on the same underlying dimensions as the transient within-person fluctuations over time; (b) cross-sectionally obtained factors adequately represent both trait-like and state-like dimensions in the data; and (c) the general trait factor can be included as a second-order factor that relates the first-order state factors as is customary in many longitudinal factor models (i.e., the HOO model). In all our empirical samples, weak factorial invariance across the trait and state structures was proved to be absent, such that these structures should not be merged as is done in cross-sectional research or certain popular longitudinal models. This illustrates the need for more scrutiny when investigating factor structures.

We want to raise three concerns here with respect to the current approach. First, although our analytical checks based on Bekker et al. (1993) showed that the initial models in our empirical applications are *theoretically identified*, they both led to improper solutions that are illustrative of *empirical unidentification*. The problems of empirical underidentification of models with M uncorrelated unique trait factors were already noted in a simulation study by Geiser and Lockhart (2012); they indicated that most problems disappeared when sample size was 300 or larger. However, for the first application, we had samples sizes of 372 and 433 individuals, but this still resulted in negative variance estimates for one of the unique traits in both groups. Clearly, such empirical underidentification is likely to hamper the practical application of the CUTS model presented here. It implies that researchers should carefully check the results for signs of empirical underidentification and, if necessary, decide what could be considered reasonable restrictions to identify the model. Note that Geiser and Lockhart (2012) also indicated that the convergence problems and improper solutions were most common when the variance of the unique traits was either relatively small or large; one could argue that having to fix the variance of a unique trait to zero to obtain a proper solution is an acceptable sacrifice if this variance is actually small in comparison to the other sources of variance in the model. The fact that in the first application it was the same unique trait in both groups that led to a nonsignificant negative variance estmate may imply that indeed the variance for this factor is very small. Future research could focus on whether nonconvergence and improper solutions point the user to the “right” variance (i.e., a variance that in the data-generating model is indeed relatively small).

Second, in the current article, we limited ourselves to a rather restrictive dichotomy, that is, traits versus states. However, when considering longitudinal data, often there is some form of development or decline present as well, which is sometimes referred to as *trait change* or simply as *change* (Nesselroade, 1991). There have been several hybrid approaches based on combining trait-state models with latent growth curve modeling (Geiser et al., 2015; Tisak & Tisak, 2000). In such extensions, the common trait can be interpeted as the (common) random intercept, while trait change is modeled through the inclusion of a (common) random slope. However, there may also be other ways to conceptualize trait changes, which are not necessarily concerned with (individual differences in) smooth developmental trajectories over time. For example, if the number of common traits and/or common states may change over time as a result of differentiation, this could also be considered a form of trait change. Hence, while including trait change into our models is a logical next step, the way in which we account for this source of variability is not straightforward and should be based on theory as well as the research questions at hand. Another way in which the model could be extended is by the inclusion of autoregressive relationships over time between the common states (and potentially also between the unique states) in order to account for potential carryover between subsequent observations at the within-person level (e.g., Cole, Martin, & Steiger, 2005; Hamaker et al., 2005; Hamaker et al., 2007).

Third, although the current framework allows us to address the question of whether the within-person factor structure coincides with the between-person factor structure, there may be a more fundamental question we need to ask: Are individuals actually characterized by the same within-person factor structure? This question has received considerable attention from Cattell (1978) and others, specifically through factor analyzing large numbers of repeated measurements per person using P-technique analysis, multivariate time series analysis Hamaker et al. (2005), or dynamic factor analysis (Molenaar, 1985). Subsequently, through comparing this idiographic within-person factor structure across individuals, one can begin to answer the question whether there is a general within-person structure or whether there are important differences between individuals (cf. Hamaker et al., 2005; Lebo & Nesselroade, 1978; Nesselroade & Molenaar, 1999). If we are simultaneously interested in the trait factor structure and the idiosyncracies in the state factor structure, we may consider the integrated trait-state model proposed by Hamaker et al. (2007), but this requires a large number of repeated measurements from a large number of individuals.

In conclusion, while the point we are making here is closely related to what has been known for decades already in multilevel regression analysis (cf., Curran & Bauer, 2011; Enders & Tofighi, 2007; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012; Wang & Maxwell, 2015), and to a lesser extent in the context of multilevel factor analysis for multiple groups (e.g., Hox & Maas, 2001; Muthén, 1994), for some reason it has not been recognized well in the broader context of factor analysis or in the more specific context of longitudinal factor analysis. We hope that the explicit comparison of Equations (4) and (5) convincingly exposes the problem that is both fundamental and ubiquitous; that is, cross-sectional factor analysis may be of very limited value because it fails to disentangle within-person from between-person sources of variance. The ramification of this is that, in general, cross-sectional factor analysis will be influenced by both the trait factor structure and the state factor structure, and unless we systematically investigate whether these two structures coincide and/or to what degree they contribute to our data, it is impossible to determine the value and substantive meaning of results obtained with cross-sectional factor analysis.

Article information

Conflict of interest disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: This work was supported by Grant VIDI 452-10-007 from the Netherlands Organization for Scientific Research (NWO).

Role of the funders/sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Acknowledgments: The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institutions or the Netherlands Organization for Scientific Research (NWO) is not intended and should not be inferred.

References

Bekker, P. A., Merckens, A., & Wansbeek, T. J. (1993). *Identification, equivalent models, and computer algebra*. Boston, MA: Academic Press.

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203–219. doi:10.1037/0033-295X.110.2.203

Cattell, R. B. (1967). The structuring of change by P-technique and incremental R-technique. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 167–198). Madison, WI: The University of Wisconsin Press.

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York, NY: Plenum Press. doi:10.1007/978-1-4684-2262-7

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*(1), 3–20. doi:10.1037/1082-989x.10.1.3

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*(1), 583–619. doi:10.1146/annurev.psych.093008.100356

Dumenci, L., & Windle, M. (1996). A latent trait-state model of adolescent depression using the Centre for Epidemiological

Studies-Depression Scale. *Multivariate Behavioral Research*, *31*(3), 313–330. doi:10.1207/s15327906mbr3103_3

Eid, M., Schneider, C., & Schwenkmezger, P. (1999). Do you feel better or worse? The validity of perceived deviations of mood states from mood traits. *European Journal of Personality*, *13*(4), 283–306. doi:10.1002/(SICI)1099-0984(199907/08)13:4<283::AID-PER341>3.0.CO;2-0

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138. doi:10.1037/1082-989x.12.2.121

Geiser, C., Keller, B. T., Lockhart, G., Eid, M., Cole, D. A., & Koch, T. (2015). Distinguishing state variability from trait change in longitudinal data: The role of measurement (non)invariance in latent state-trait analyses. *Behavioral Research Methods*, *47*(1), 172–203. doi:10.3758/s13428-014-0457-z

Geiser, C., & Lockhart, G. (2012). A comparison of four approaches to account for method factors in latent state-trait analyses. *Psychological Methods*, *17*(2), 255–283. doi:10.1037/a0026977

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical modeling of the individual: Rationale and application of multivariate time series analysis. *Multivariate Behavioral Research*, *40*(2), 207–233. doi:10.1207/s15327906mbr4002_3

Hamaker, E. L., Nesselroade, J. R., & Molenaar, P. C. M. (2007). The integrated trait-state model. *Journal of Research in Personality*, *41*(2), 295–315. doi:10.1016/j.jrp.2006.04.003

Hox, J. J., & Maas, C. J. M. (2001). The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. *Structural Equation Modeling*, *8*, 157–174. doi: 10.1207/S15328007SEM0802_1

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*, 513. doi:10.3389/fpsyg.2013.00513

Lebo, M. A., & Nesselroade, J. R. (1978). Intraindividual difference dimensions of mood change during pregnancy identified in five P-technique factor analyses. *Journal of Research in Personality*, *12*(2), 205–224. doi:10.1016/0092-6566(78)90098-3

Lord, F. M. (1965). *Item sampling in test theory and in research design*. Princeton, NJ: ETS Research Bulletin Series.

Marsh, H. W., & Grayson, D. (1994). Longitudinal confirmatory factor analysis: Common, time-specific, item-specific, and residual-error components of variance. *Structural Equation Modeling*, *1*(2), 116–145. doi:10.1080/10705519409539968

Meredith, W. M. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. doi:10.1007/bf02294825

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika*, *50*(2), 181–202. doi:10.1007/bf02294246

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology—this time forever. *Measurement: Interdisciplinary Research and Perspectives*, *2*(4), 201–218. doi:10.1207/s15366359mea0204_1

Molenaar, P. C. M., Huizenga, H. M., & Nesselroade, J. R. (2003). The relationship between the structure of interindividual and intraindividual variability: A theoretical and empirical vindication of developmental systems theory. In U. M. Staudinger & U. Lindenberger (Eds.), *Understanding human development: Dialogues with lifespan psychology* (pp. 339–360). Norwell, MA: Kluwer Academic. doi:10.1007/978-1-4615-0357-6_15

Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods and Research*, *22*, 376–398. doi:10.1177/0049124194022003006

Nesselroade, J. R. (1991). Interindividual differences in intraindividual change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 92–105). Washington, DC: American Psychological Association.

Nesselroade, J. R., & Molenaar, P. C. M. (1999). Pooling lagged covariance structures based on short, multivariate time series for dynamic factor analysis. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 223–250). Thousand Oaks, CA: Sage.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London, UK: Sage.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, *15*(2), 201–292. doi:10.2307/1412107

Steyer, R., Mayer, A., Geiser, C., & Cole, D. (2015). A theory of states and traits—revised. *Annual Review of Clinical Psychology*, *11*(1), 71–98. doi:10.1146/annurev-clinpsy-032813-153719

Tiikkainen, P., Heikkinen, R.-L., & Leskinen, E. (2004). The structure and stability of perceived togetherness in elderly people during a 5-year follow-up. *The Journal of Applied Gerontology*, *23*(3), 279–294. doi:10.1177/0733464804267582

Tisak, J., & Tisak, M. S. (2000). Permanency and ephemerality of psychological measures with application to organizational commitment. *Psychological Methods*, *5*(2), 175–198. doi:10.1037/1082-989x.5.2.175

Voelkle, M., Brose, A., Schmiedek, F., & Lindenberger, U. (2014). Toward a unified framework for the study of between-person and within-person strcutures: Building a bridge between two research paradigms. *Multivariate Behavioral Research*, *49*(3), 193–213. doi:10.1080/00273171.2014.889593

Wang, L., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Multivariate Behavioral Research*, *20*(1), 63–83. doi:10.1037/met0000030

Yung, Y.-F., Thissen, D., & McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, *64*(2), 113–128. doi:10.1007/bf02294531

Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Jounral of Personality and Social Psychology*, *43*(1), 111–122. doi:10.1037/0022-3514.43.1.111