Is it possible to feel good and bad at the same time?
New evidence on the bipolarity of mood-state dimensions

Rolf Steyer and Katrin Riedl
Friedrich Schiller University, Jena

Positively and negatively formulated items that are constructed as indicators of the same psychological construct oftentimes correlate lower than expected. In exploratory factor analyses they sometimes even have their highest loadings on separate factors and, therefore, seem to measure two different monopolar dimensions. However, a longitudinal analysis of positively and negatively formulated items of well-being, and two other dimensions of mood state with a two-construct latent state-trait model reveals that only the trait factors are monopolar, though highly negatively correlated; the state-residual, i.e., the deviations of the actual from habitual mood states are perfectly bipolar, i.e., they are correlated −1, if exact antonymous items are considered. It is concluded that unsystematic measurement errors, the inappropriate use of Pearson correlations instead of polychoric correlations, and systematic response styles mask the deterministic relationship between antonymous mood state self-ratings.

1 Introduction

Is it possible to feel good and bad at the same time? A simple question, it seems, which should have a simple answer. However, this question is neither simple nor clear as the intensive and ongoing discussions at least since Wundt (1896) reveal (see, e.g., Beebe-Center, 1932; Diener, 1999; Diener & Emmons, 1985, Eid, Notz, Schwenkmezger & Steyer, 1994; Egloff, 1998; Green, Goldman & Salovey, 1993; Russell & Carroll, 1999; Tellegen, Watson & Clark, 1999; Schimmack, 2001; Vautier & Raufaste, 2002; Watson, 1988). One reason why this question is not simple is because it is not clear enough what it refers to. Is it feelings or is it mood states we are considering? We distinguish *feelings* from *mood states* by the fact that feelings have an object. For example, I can feel happy, because someone said that he or she likes my work, and I can feel sad, because my dear aunt recently died. In this context, “happy” and “sad” refer to feelings. Hence, I can feel happy with respect to one object and sad with respect to another one, more or less at the same time, although the intensity of both feelings cannot be large in such case. Bradburn (1969), for instance, assessed feelings in this sense.

In contrast, *mood states,* by definition do not have an object. We just feel happy or sad without any obvious or conscious reason. Using the figure-ground metaphor, feelings may be compared to figures, whereas mood states are the ground against which figures are perceived. Note that in this context the same words “happy” and “sad” refer to psychologically different phenomena: mood states instead of feelings. *Considering these mood states*, we believe that it is not possible to feel good and bad at the same time. In more theoretical terms this means that we can construct a single (“bipolar”) dimension with two opposite poles, extremely bad on one side and extremely good on the other.

Even if it is clarified which empirical and psychological phenomena are considered, there is still a major obstacle preventing clear and unanimous empirical findings: using inappropriate designs and inappropriate models used in data analysis. First of all, we believe that a cross-sectional design is not appropriate for answering our question about bipolarity. One reason is that mood state self-reports are not easy to compare between different persons. The same rating on the same item may mean different things for different persons. This problem is due to response styles with respect to which persons differ. Although some attempts have been made to control for acquiescence (see, e.g., Green et al., 1993; Tellegen et al., 1999), other response styles may still mask the true dimensional structure in cross-sectional studies. We believe that response styles can more easily and more completely be controlled for in longitudinal studies. We can control response styles in repeated longitudinal designs, because they are part of the stable individual differences over time (see, e.g., Schimmack, 2001, p. 83). In this paper, we will show how this can be achieved.

The second obstacle mentioned above are inappropriate models used in data analysis. We explicitly mention two points. One is using Pearson correlations for rating scales which are of an ordinal nature. Pearson correlations do not represent the true relationships between ordinal variables and can lead to wrong factorial structures having no reasonable substantive meaning (see the findings on “difficulty factors”, e.g., Moosbrugger & Hartig, 2002). The other point is ignoring the state-trait distinction in the models used in data analysis. Since response styles are trait components, response styles can be controlled for if state and trait components are disentangled.

In the present paper we will show in a longitudinal study with ordinal mood state data that mood states are in fact bipolar. More specifically, we will show that the *deviation of the actual mood state* from the *habitual mood state* (i.e., from the corresponding trait component) is a bipolar dimension that can be measured both by negatively and/or by positively formulated items. Negatively and positively formulated items only make a difference in the measurement of the *habitual* mood state, i.e., the trait. Hence, a trait assessed by negatively formulated items and a trait assessed by positively formulated items may correlate to some degree, but not perfectly. However, the *deviations* of the actual mood state from the habitual mood state have a perfect negative correlation (-1) within each occasion of measurement.

2 Method

2.1 Description of the Data

2.1.1 Sample

A sample of 291 females and 212 males between 17 and 77 years of age (mean age: 31.2 years) responded to a number of questionnaires on four occasions of measurement, each of them three weeks apart. Among others, a mood state questionnaire, a mood trait questionnaire, a personality questionnaire, and a scale of daily hassles and uplifts were administered.[1] The subjects were paid DM 50 for completing the tests on all four occasions of measurement. About half of the subjects were assessed in group sessions in a lecture room at the University of Trier (Germany). The other half of the subjects was recruited via a snowball system and filled in their questionnaires at home Steyer, Schwenkmezger, Eid & Notz, 1991.

The sample analyzed here consists of those among the 548 original subjects who delivered their questionnaires on all four occasions and who had no missing values on the items analyzed. Hence, depending on the item set (see next paragraph), the sample sizes varied between 470 and 501.

---

[1] The complete data set is available in the internet at http://www.uni-jena.de/svw/metheval/daten/start.html.

2.1.2 Items

The analyses to be presented refer to the items and scales of the German version of the Multidimensional Mood Questionnaire (Steyer, Schwenkmezger, Notz & Eid, 1997). This questionnaire has been designed to assess three bipolar dimensions of mood: feeling well vs. feeling bad (GS scale), being awake vs. being tired (WM scale), feeling calm vs. feeling tense (RU scale). Table 1 contains a list of the German items for each of the three scales and their English translation. The subjects rated their actual mood state on a 5-point Likert scale ranging from 1, labeled “überhaupt nicht” (“not at all”), to 5, labeled “sehr” (“very much so”). The response categories between 1 and 5 were only labeled 2, 3, and 4 but had no verbal label.

2.2 Data Analysis

2.2.1 Methodological Background: Latent State-Trait Models

As mentioned in the introduction, our hypothesis is that the deviation of the actual mood state from the habitual mood state (i.e., from the trait) is a bipolar dimension that can be measured by negatively and/or by positively formulated items. How to measure these deviations of the actual from the habitual mood state? The answer is easy in the framework of latent state-trait theory (LST theory) (see, e.g., Steyer, Ferring & Schmitt, 1992; Steyer, Schmitt & Eid, 1999). A typical model of LST theory is the multistate-singletrait model (MSST model; Steyer et al., 1992) depicted in Figure 1. According to this model, a manifest variable Y_{it} observed on occasion t of measurement can be decomposed into a (linear function of a) *latent state variable* η_t that is common to all variables observed on time t and a *measurement error variable* ε_{it}. Each latent state variable η_t can itself be decomposed into a (linear function of a) *latent trait variable* ξ that is assumed to be constant for all occasions of measurement considered, and a *latent state residual* ζ_t. Hence, this latent state residual ζ_t is the deviation of the actual mood state from its associated trait (or “habitual mood state”).

Insert Figure 1 about here

In many applications, the manifest variables Y_{it} are not perfectly parallel (or homogeneous) measures of the same latent state variable. In these cases a more realistic model also includes a “method factor” representing those method effects that are different between subjects (Figure 2). The scores on these method factors can be interpreted as the effects of the interaction between subject and method. The variances of these variables will reflect the degree of heterogeneity of the methods used to measure the latent state variable η_t. In these models, we need just one method factor less than there are different methods: one method factor for two methods, two (correlating) method factors for three methods, etc. Note that “method” in the present context usually means “parallel form” or “item”. Hence we will treat the self-ratings with respect to two items such as feeling “good” and feeling “well” as two different methods of assessing the actual mood state of well-being. Details and more about the rationale of the models with method factors can be found in Eid (2000) and Eid, et al. (2003).

Insert Figure 2 about here

Other LST models, involving several constructs and several traits can be found in Steyer et al. (1989) and Eid et al. (1994). Note that traits, in principle, can also change across time (see, e.g., Eid & Hoffmann, 1998) and that there are models with correlated test-specific (or item-specific) traits avoiding methods factors (Steyer et al., 1999). Furthermore, these models have also been developed for ordinal variables (Eid, 1996; 1997), a methodology that will also be used in the present paper.

2.2.2 Item Sets

Within the GS scale (good vs. bad), there are several items, which, semantically, are antonymmous, i.e., they have opposite meanings: “gut” (good) and “schlecht” (bad), “zufrieden” (content) and “unzufrieden” (discontent), “wohl” (well) and “unwohl” (unwell), “glücklich” (happy) and “unglücklich” (unhappy). Within the WM scale (awake vs. tired) it is more difficult to find exact antonyms. However, the item pairs “wach” (awake) and “müde” (tired) as well as “munter” (lively) and “ermattet” (exhausted) may be rather close to being antonyms. Finally, in the RU scale we may choose “entspannt” (relaxed) and “angespannt” (tense) as well as “ruhig” (calm) and “unruhig” (restless) as antonymous adjectives.

The strategy is now to choose two positive items such as “gut-zufrieden” (good-content) and their antonyms such as “schlecht-unzufrieden” (bad and discontent). Four such items will be referred to as an *item set* in the sequel. We will consider the six item sets displayed in Table 2.

Insert Table 2 about her

From a substantive point of view, our hypothesis (that the deviations of the actual mood states from the habitual mood states correlate −1) should hold perfectly for item sets 1, 2, 3, and 5, whereas it should clearly not hold for item set 6 which contains items from different scales: “gut” and “wohl” belong to the GS scale (good vs. bad), whereas “angespannt” and “unruhig” belong to the RU scale (calm vs. nervous). Although these scales are correlated, we do not expect a perfect negative correlation between the deviations of the actual mood states from the habitual mood states. Item set 6 is included only in order to demonstrate that the high negative correlation between the deviations of the actual mood states from the habitual mood states is not an artefact but an important substantive finding. Finally, for item set 4 we expect slight deviations from our hypothesized −1 correlation, because the items are not perfect antonyms, as is the case for the item sets 1, 2, 3 and 5.

2.2.3 Models

For each of these six item sets we consider four occasions of measurement and test the following models, the input files of which are given in the Appendix. *Model A* is the least restrictive model allowing, within each occasion of measurement, *any correlation* between the deviations of the actual mood state from the habitual mood state. *Model B* is the *most restrictive model postulating a perfect negative correlation (i.e., equal to –1) between these deviations* within each occasion of measurement. Between occasions the deviations of the actual mood state from the habitual mood state are assumed to be uncorrelated in all models considered. Finally, in *Model C* we allow for a nonperfect negative correlation between these deviations within the first occasion of measurement, while postulating perfect negative correlations for times two to four. This exception is substantively meaningful, because the first occasion may be considered a “warming up” occasion of measurement in which subjects learn using the questionnaire (cf. Jagodzinski, Kühnel & Schmidt, 1987).

Insert Figures 3 to 5 about here

Figures 3 to 5 present the models more explicitly. In all three models we assume that there is a trait for the positive items and a separate trait for the negative items. These traits have a high negative correlation which, however, is less than perfect. These two trait variables represent the habitual *self-reported* mood states. Our interpretation of this less than perfect correlation is that the subjects tend to react differently to positive and negative items and that these differences are subject-specific. That is, there are inter-individual differences (response styles) in responding to positive items and negative items, but these inter-individual differences are stable over time. Hence, the less than perfect correlation between the two latent trait variables is due to stable response styles biasing the response to the items.

The deviation of the actual mood state from this habitual mood state, however, is not biased any more. It veridically reflects the true deviation of the actual mood state from the habitual mood state. Therefore, it should not matter any more if this deviation is assessed via a positive or via a negative item.

The models in Figures 3 to 5 also contain two “method factors” or, more specifically, two “item-specific factors”. They are necessary to account for the semantic differences between the items such “gut” (good) and “zufrieden” (content) or “schlecht” (bad) and “unzufrieden” (discontent). The smaller the proportion of variances determined by these item-specific factors, the more homogeneous are these item pairs. The details of the model specifications are given in the three input files printed in the Appendix.

2.2.4 Estimation method

For each of the six item sets mentioned in the previous section, each consisting of four items repeated at four time points, we computed a polychoric correlation matrix and the matrix of their asymptotic covariances using PRELIS 2 (Jöreskog & Sörbom, 1996). The polychoric correlation matrix was then analyzed with LISREL 8.52 (Jöreskog, Sörbom, du Toit & du Toit, 2001) using the matrix of asymptotic covariances with WLS-estimation. In this way we take into account that the items have ordered response categories and allow for different thresholds and difficulties of the items and also different thresholds for the response categories. With this methodology we can circumvent methodological artifacts such as difficulty factors (see, e.g., Moosbrugger & Hartig, 2002), which would occur if we would treat the items as continuous (or metric) variables and use the Pearson correlation. Furthermore, it gives as meaningful information about the items and their response categories.

3 Results

We start looking at Table 3 containing the thresholds for each item of item set 2 at each of the four occasions of measurement. These thresholds characterize the items and their response categories. In order to understand their meaning we have to remember that for each item it is assumed that there is an underlying standard normal latent variable. The first threshold of item “wohl” (well) at time 1 of measurement, -1.737, is the *z*-value cutting off those 4.1% of all persons in the normal distribution which answer 1 (“not at all”). (These percentages can also be found in Table 4 displaying the marginal distributions of all four items of item set 2 at time 1 of measurement.) The second threshold, -0.793, cuts off those 4.1% plus the additional 17.3% of the all subjects answering in response category 2 of this item at the first occasion of measurement. The third threshold, -0.015, cuts off those 4.1% + 17.3% plus the additional 28.0% of all subjects answering, at the first occasion of measurement, in response category 3 of this item. Finally, the fourth and last threshold, 1.118, cuts off those 4.1% + 17.3% + 28.0% plus the additional 37.4% of the all subjects answering in response category 4. The remaining 13.2% of the subjects answer in response category 5 (“very much so”).

Insert Table 3 and Table 4 about here

Looking at Table 3 we can say that “wohl” (well) is the easiest item. “Difficulty” or “easiness” of an item can be defined in different ways such as the average of the thresholds of an item or as the size of the last threshold indicating the point in the standard normal distribution at which all subjects exceeding this point answer in the last response category, or as the size of the first threshold indicating the point in the standard normal distribution at which all subjects below this point answer in the first response category. The item “glücklich” is somewhat more difficult which is in full coincidence with the meaning and the German usage of these words. Much more dramatic is the difference between the positively and negatively formulated items. The last threshold of “unwohl” (unwell) is almost one unit (standard deviation) higher than the last threshold of “wohl”. Almost 50% of all subjects answer in response category 1 (“not at all”) of the item “unwohl” and more than 50% answer in this response category of the item “unglücklich” (unhappy). In principle it would be possible to impose equality constraints on the threshold between items and/or across time. However, in this application we did not use this possibility.

Next we look at the estimated correlations between the deviations of the actual mood states from the habitual mood states. According to our theory, these correlations should be −1 for item sets 1, 2, 3 and 5, close to −1 for item set 4, and distinctly not −1 for item set 6. These correlations which are estimated (via weighted least squares) under Model A described above are shown in the first four columns and their average in column 5 of Table 5. Note that, although true correlations cannot be smaller than −1, their WLS estimates can be smaller than −1. Considering the average correlations displayed in column 5 shows that, from a descriptive point of view, our hypothesis holds close to perfectly for item sets 1, 2, 5, and it is doubtful in item sets 3 and 4. For item set 6 the correlations between the deviations of the actual from the habitual mood states are, in the average -0.699, which meets our expectation that perfect negative correlations only occur for antonyms, i.e., for items with semantically opposite meanings. Also the correlations between the trait components pertaining to positive and negative items are as expected. They are negative and high, but not close to −1. Furthermore, as expected, the correlation between the traits is only −.596 for item set 6 (consisting of items from different scales), whereas it is between −.734 and −.846 for the other item sets.

Insert Table 5 about here

Table 6 gives the model fit statistics for Models A, B and C for each of the six item sets. Looking at the χ^2-goodness of fit statistics reveals that the model fits are close to perfect for item set 1 and 5, while the fits are not perfect for models 2, 3 and 4. The root mean square error of residuals (RMSEA) are all smaller than .04 indicating that all models are quite acceptable. However, what is more important for our hypothesis is *comparing* the fit statistics between the three models, specifically looking at the χ^2-differences. For a test of our hypothesis we have to compare the χ^2-values between models A and B and between models A and C. Remember, *Model A* is the least restrictive model allowing, within each occasion of measurement, *any correlation* between the deviations of the actual mood state from the habitual mood state. *Model B* is the most restrictive model postulating a perfect negative correlation (i.e., equal to −1) between these deviations within each of the four occasions of measurement. In *Model C* we allow for a nonperfect negative correlation between these deviations within the first occasion of measurement, while postulating perfect negative correlations for occasions two to four.

Insert Table 6 about here

Comparing models A to B shows that only item set 5 yields a χ^2-difference (105.40 – 104.23 = 1.17; *df* = 4) that is not significant at the .05-level. However, comparing models A to C yields much smaller χ^2-differences. The results indicate that we have to prefer Model A over Model C only for item sets 4 and 6. For item set 6, the difference (182.65 - 119.06 = 63.59; *df* = 3) is big as expected, because this item set consists of items from different scales. Item set 4 containing the items “wach-munter vs. müde-ermattet” (awake-lively vs. tired-exhausted) is the one consisting of item pairs which are opposite in meaning but not as perfectly as the item sets 1, 2, 3 and 5 for which we do not have to reject the hypothesis that the deviations of the actual mood states from the habitual mood state correlate –1 within occasions 2 to 4.

4 Discussion

The results presented in Table 6 indicate that indeed the deviations of the actual mood states from the habitual mood states are *bipolar*. However, this applies perfectly only for occasions of measurements in which the subjects are already acquainted with using such a questionnaire (in our case: occasions two to four). Furthermore, it only holds for item pairs which can be considered exact antonyms, i.e., exact semantic opposites of each other, such as “good” (gut) and “bad” (schlecht) or “well” (wohl) and “unwell” (unwohl). The more the meanings of the adjectives differ from each other, the lower the negative correlations between the deviations of the actual mood states from the habitual mood states within an occasion of measurement. This fact is most obvious if we look at the analysis of item set 6 which contains items pertaining to different dimensions and scales (GS scale and RU scale) of the mood state questionnaire.

Which are the conclusions? The measurement of habitual mood states via repeated measurement of the actual mood states is biased by response sets. These measurements are specific for the items used and depend on whether or not items are formulated positively or negatively. Even if semantically exactly opposite items are used, such as good and bad, the measurements will differ. Although they correlate, they are not perfectly functionally related, even if unsystematic measurement error is taken into account. Instead there is an interaction (in the ANOVA sense) between the person factor and the factor “positivity vs. negativity” of the items. Hence, the highest reliability of the measurements can be reached if the scales contain either only positive or only negative items. The question about their validity is open.

When it comes to measuring the *deviations of the actual from the habitual mood state*, positive and negative items do equally well. Our results indicate that these deviations *are functionally* related. This also implies that the *changes* in mood states as assessed by a positive item and its negative counterpart are functionally related, i.e., a change rated on a positive item goes perfectly along with a change rated on a negative item. This can be used in experiments designed to manipulate the actual mood states (see, e.g., Vautier & Raufaste, 2002).

A more general substantive conclusion is that unsystematic measurement errors and systematic answer styles mask the deterministic relationship between antonymous mood state self-ratings. A methodological conclusion is that multi-construct latent state-trait models are useful to control for unsystematic measurement error and systematic errors due to response styles and to help revealing the underlying dimensionality in related psychological areas such as the measurement of affect (Diener & Emmons, 1985; Schmukle, Egloff & Burns, 1548, in press; Tellegen et al., 1999), or self-esteem (Marsh, 1996).

Reference List

Beebe-Center, J. G. (1932). *Psychology of pleasantness and unpleasantness*. New York: Van Nostrand.

Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago: Aldine.

Diener, E. (1999). Introduction to the special section on the structure of emotion. *Journal of Personality and Social Psychology*, *76*, 803-804.

Diener, E. & Emmons, R. A. (1985). The independence of positive and negative affect. *Journal of Personality and Social Psychology*, *47*, 1105-1117.

Egloff, B. (1998). The independence of positive and negative affect depends on the affect measure. *Personality and Individual Differences*, *25*, 1101-1109.

Eid, M. (1996). Longitudinal confirmatory factor analysis for polytomous item responses: Model definition and model selection on the basis of stochastic measurement theory. *Methods of Psychological Research Online*, *1*, 69-91.

Eid, M. (1997). Happiness and satisfaction: An application of a latent state-trait model for ordinal variables. In J. Rost & R. Langeheine (Eds.), *Applications of latent trait and latent class models in the social sciences* (pp. 145-151). Münster: Waxmann.

Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika*, *65*, 241-261.

Eid, M. & Hoffmann, L. (1998). Measuring variability and change with an item response model for polytomous variables. *Journal of Educational and Behavioral Statistics*, *23*, 193-215.

Eid, M., Lischetzke, T., Trierweiler, L. I. & Nußbeck, F. W. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple indicator CTC(M-1) model. *Psychological Methods*.

Eid, M., Notz, P., Schwenkmezger, P. & Steyer, R. (1994). Sind Stimmungsdimensionen monopolar? Ein Überblick über empirische Befunde und Untersuchungen faktorenanalytischen Modellen für kontinuierliche und kategoriale Variablen sowie neuere Ergebnisse [Are mood dimensions monopolar? A review of empirical results and investigations with factor analyses of continuous and categorical variables as well as recent outcomes]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 15*, 211-233.

Eid, M., Notz, P., Steyer, R. & Schwenkmezger, P. (1994). Validating scales for the assessment of mood level and variability by latent state-trait analyses. *Personality and Individual Differences*, *16,* 63-76.

Green, D. P., Goldman, S. L. & Salovey, P. (1993). Measurement error masks bipolarity in affect ratings. *Journal of Personality and Social Psychology*, *64*, 1029-1041.

Jagodzinski, W., Kühnel, S. & Schmidt, P. (1987). Is there a "Socratic effect" in nonexperimental panel studies? Consistency of an attitude toward guestworkers. *Sociological Methods & Research*, *15*, 259-302.

Jöreskog, K. G. & Sörbom, D. (1996). *PRELIS 2 user's reference guide: A program for multivariate data screening and data summarization*. Chicago: SSI.

Jöreskog, K. G., Sörbom, D., du Toit, S. & du Toit, M. (2001). *LISREL 8: New statistical features*. Chicago: SSI.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology*, *70,* 810-819.

Moosbrugger, H. & Hartig, J. (2002). Factor analysis in personality research: Some artefacts and their consequences for psychological assessment. *Psychologische Beiträge, 44*, 136-158.

Russell, J. A. & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin*, *125,* 3-30.

Schimmack, U. (2001). Pleasure, displeasure, and mixed feelings: Are semantic opposites mutually exclusive? *Cognition and Emotion*, *15*, 81-97.

Schmukle, S. C., Egloff, B. & Burns, L. R. (1548). The relationship between positive and negative affect schedule (PANAS) (unpublished work). *Journal of Research in Personality*.

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

Steyer, R., Majcen, A. M., Schwenkmezger, P. & Buchner, A. (1989). A latent state-trait anxiety model and its application to determine consistency and specificity coefficients. *Anxiety Research*, *1,* 281-299.

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., Schwenkmezger, P., Eid & Notz, P. (1991). *Befindlichkeitsmessung und Latent-State-Trait-Modelle*. (Arbeitsbericht zum DFG-Projekt "Ste 411/3-1"). Trier: Universität Trier.

Steyer, R., Schwenkmezger, P., Notz, P. & Eid, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF)*. Göttingen: Hogrefe.

Tellegen, A., Watson, D. & Clark, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science, 10*, 297-303.

Vautier, S. & Raufaste, E. (2002). *Measuring dynamic bipolarity in positive and negative activation (unpublished work).*

Watson, D. (1988). The vicissitudes of mood measurement: Effects of varying descriptors, time frames, and response formats on measures of positive and negative affect. *Journal of Personality and Social Psychology*, *55*, 128-141.

Wundt, W. (1896). *Grundrisse der Psychologie [Outlines of psychology]*. Leipzig: Engelmann.

Figure Captions

Figure 1
Conceptual path diagram of a multistate-singletrait model (MSST model) for two measures (tests or items) on each of four occasions of measurement.

Figure 2
Conceptual path diagram of a multistate-singletrait model (MSST model) for two measures (tests or items) on each of four occasions of measurement with a method factor.

Figure 3
Conceptual path diagram of Model A: a two-construct model with a method factor and correlated latent state residuals.

Figure 4
Conceptual path diagram of Model B: a two-construct model with a method factor and perfectly correlated latent state residuals.

Figure 5
Conceptual path diagram of Model C: a two-construct model with a method factor and perfectly correlated latent state residuals for occasions two to four. At occasion one, latent state residuals correlate less than perfect.

Table 1. MDBF scales and their items

Scale	Short form A	Short form B
GS	zufrieden (content) gut (good) schlecht (bad) unwohl (unwell)	wohl (well) glücklich (happy) unglücklich (unhappy) unzufrieden (discontent)
WM	ausgeruht (rested) munter (lively) schlapp (floppy) müde (tired)	wach (awake) frisch (fresh) schläfrig (sleepy) ermattet (exhausted)
RU	gelassen (composed) entspannt (relaxed) ruhelos (restless) unruhig (uneasy)	ruhig (calm) ausgeglichen (well-balanced) angespannt (tense) nervös (nervous)

Table 2. Item sets

Item set 1: “gut-zufrieden“ vs. „schlecht-unzufrieden” ; *N* = 470

Item set 2: “wohl-glücklich“ vs. „unwohl-unglücklich”; *N* = 486

Item set 3: “ruhig-entspannt“ vs. „unruhig-angespannt”; *N* = 501

Item set 4: “wach-munter“ vs. „müde-ermattet”; *N* = 488

Item set 5: “gut-wohl“ vs. „schlecht-unwohl”; *N* = 490

Item set 6: “gut-wohl“ vs. „angespannt-unruhig”; *N* = 499

Note: see Table1 for English translations of the adjectives.

Table 3
Thresholds for item set 2 (Sample Size = 486)

Time	Item	Thresholds				Average Threshold (Location)
1	wohl (well)	-1.737	-0.793	-0.015	1.118	-0.357
1	glücklich (happy)	-1.492	-0.772	0.224	1.254	-0.197
2	wohl (well)	-1.737	-0.967	-0.026	1.081	-0.412
2	glücklich (happy)	-1.540	-0.807	0.114	1.221	-0.253
3	wohl (well)	-1.812	-0.828	0.036	1.138	-0.367
3	glücklich (happy)	-1.432	-0.658	0.240	1.313	-0.134
4	wohl (well)	-1.593	-0.935	-0.010	1.118	-0.355
4	glücklich (happy)	-1.524	-0.793	0.140	1.243	-0.234
1	unwohl (unwell)	0.000	0.658	1.221	2.042	0.980
1	unglücklich (unhappy)	0.103	0.658	1.277	1.931	0.992
2	unwohl (unwell)	-0.088	0.704	1.265	2.042	0.981
2	unglücklich (unhappy)	0.176	0.737	1.350	1.868	1.033
3	unwohl (unwell)	-0.088	0.620	1.221	2.186	0.985
3	unglücklich (unhappy)	0.161	0.731	1.446	1.899	1.059
4	unwohl (unwell)	-0.005	0.717	1.337	2.246	1.074
4	unglücklich (unhappy)	0.315	0.904	1.507	2.042	1.192

Table 4
Marginal distributions of the items of item set 2 on occasion 1.

	wohl (well)		glücklich (happy)		unwohl (unwell)		unglücklich (unhappy)	
Response Category	Frequency	%	Frequency	%	Frequency	%	Frequency	%
1	20	4.1	33	6.8	243	50.0	263	54.1
2	84	17.3	74	15.2	119	24.5	99	20.4
3	136	28.0	179	36.8	70	14.4	75	15.4
4	182	37.4	149	30.7	44	9.1	36	7.4
5	64	13.2	51	10.5	10	2.1	13	2.7

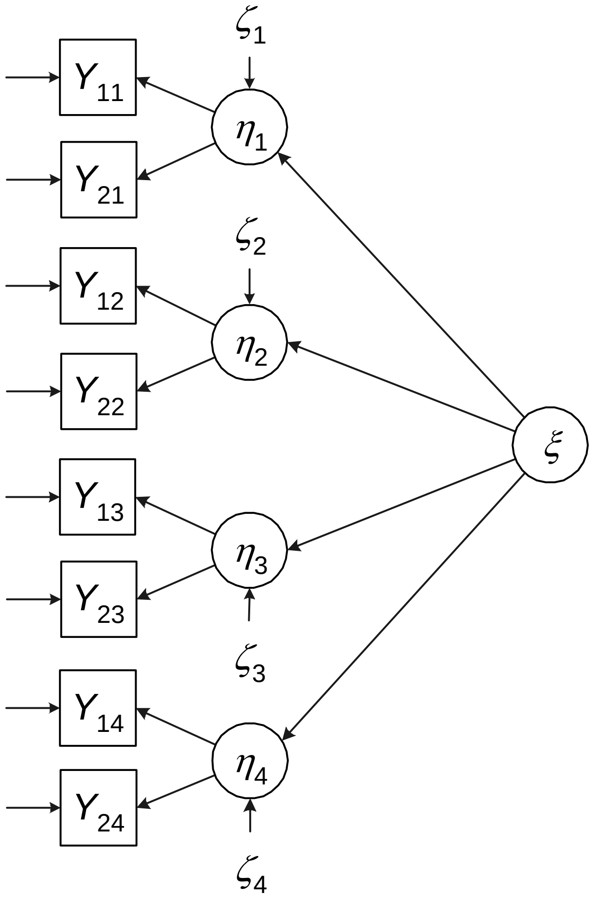
Table 5. Trait correlations and correlations between latent state residuals within each occasion of measurement estimated in Model A

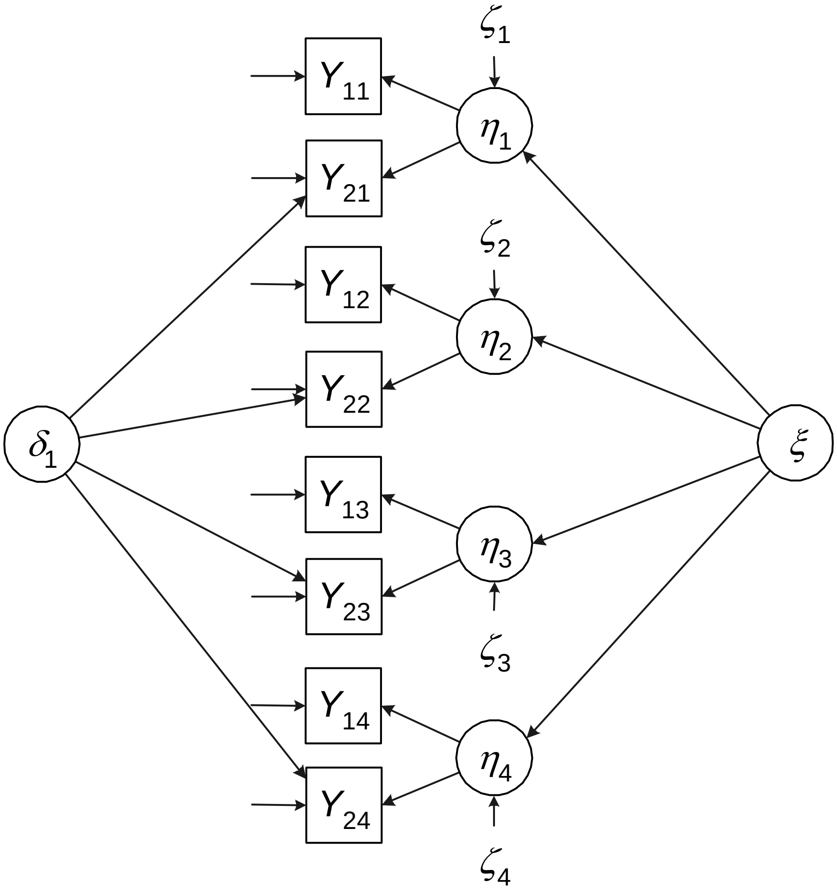
	t_1	t_2	t_3	t_4	average correlation of state residuals	trait correlation
Item set 1	-0.888	-0.999	-1.052	-1.086	-1.006	-0.839
Item set 2	-0.924	-0.996	-1.004	-0.914	-0.960	-0.846
Item set 3	-0.878	-0.883	-0.941	-0.950	-0.913	-0.734
Item set 4	-0.954	-0.977	-0.963	-0.893	-0.947	-0.776
Item set 5	-0.987	-1.009	-1.024	-1.017	-1.009	-0.821
Item set 6	-0.682	-0.661	-0.791	-0.661	-0.699	-0.596

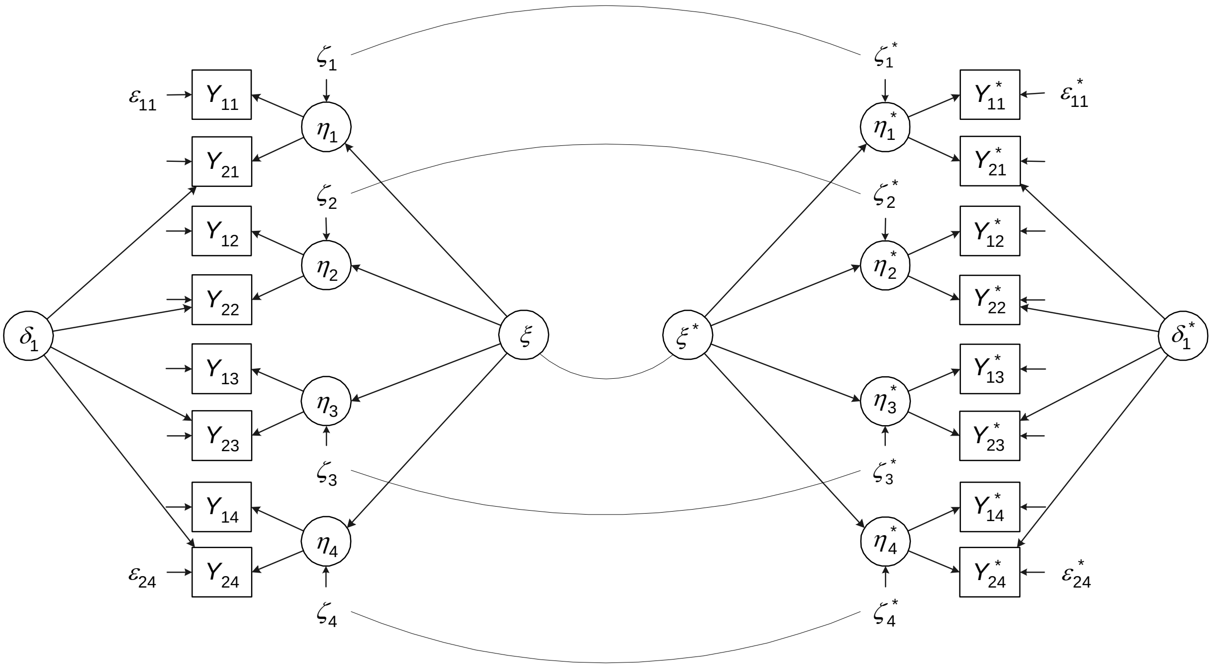
Table 6. Comparing the three models via χ^2-difference tests

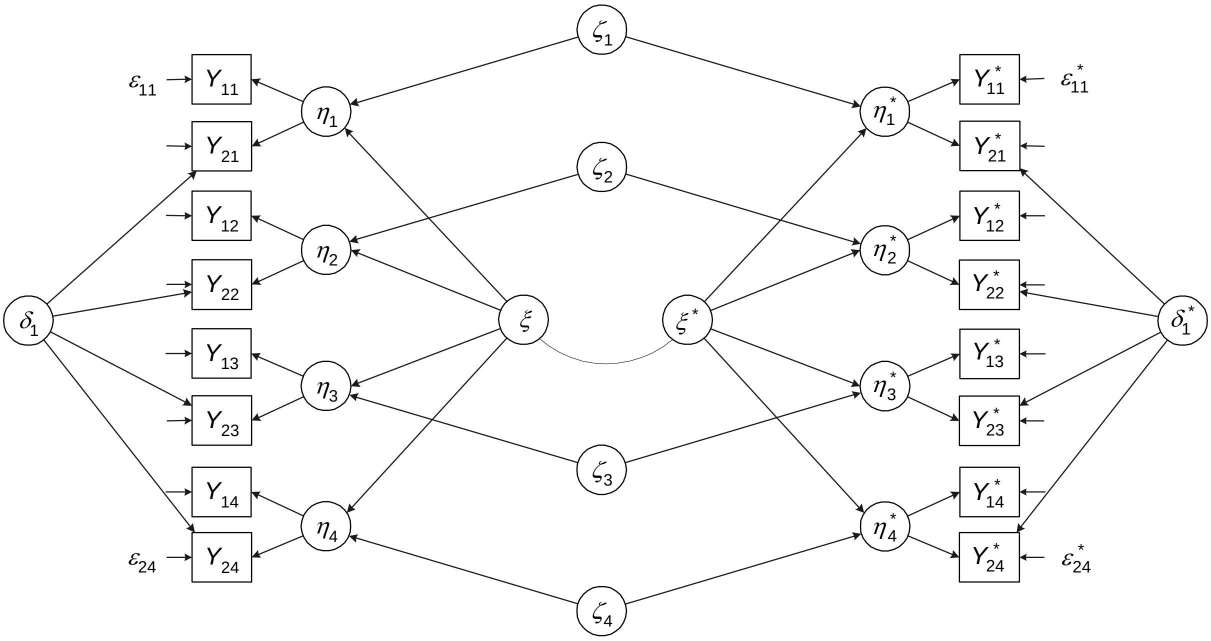
	Model A; *df* = 94			Model B; *df* = 98			Model C; *df* = 97			A vs. C
	χ^2	*p*	RMSEA	χ^2	*p*	RMSEA	χ^2	*p*	RMSEA	
Item set 1	104.71	.21	.016	116.15	.10	.020	110.35	.17	.017	
Item set 2	155.07	.00	.037	162.05	.00	.037	156.88	.00	.036	
Item set 3	145.15	.00	.033	158.99	.00	.035	152.22	.00	.034	
Item set 4	133.83	.00	.029	165.22	.00	.038	161.98	.00	.037	*
Item set 5	104.23	.22	.015	105.40	.29	.012	105.22	.27	.013	
Item set 6	119.06	.04	.023	233.26	.00	.053	182.65	.00	.042	**

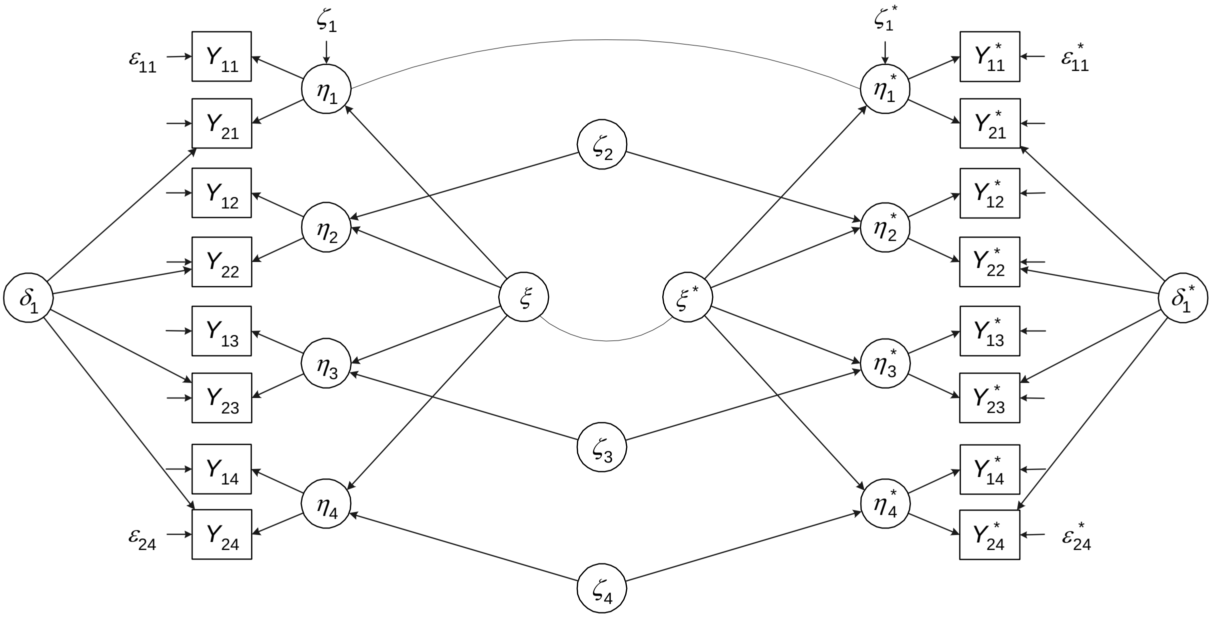
Note: The column headed A vs. C contains an asterisk if the corresponding χ^2-difference test is significant at the .05-level. The critical χ^2 for 3 degrees of freedom at the .05-level is 7.81.


ζ_1
Y_{11}
η_1
Y_{21}
ζ_2
Y_{12}
η_2
Y_{22}
ξ
Y_{13}
η_3
Y_{23}
ζ_3
Y_{14}
η_4
Y_{24}
ζ_4


ζ_1
Y_{11}
η_1
Y_{21}
ζ_2
Y_{12}
η_2
Y_{22}
δ_1
ξ
Y_{13}
η_3
Y_{23}
ζ_3
Y_{14}
η_4
Y_{24}
ζ_4


ζ_1
ζ_1^*
ε_{11}
Y_{11}
η_1
η_1^*
Y_{11}^*
ε_{11}^*
Y_{21}
Y_{21}^*
ζ_2
ζ_2^*
Y_{12}
η_2
η_2^*
Y_{12}^*
Y_{22}
Y_{22}^*
δ_1
ξ
ξ^*
δ_1^*
Y_{13}
η_3
η_3^*
Y_{13}^*
Y_{23}
Y_{23}^*
ζ_3
ζ_3^*
Y_{14}
η_4
η_4^*
Y_{14}^*
ε_{24}
Y_{24}
Y_{24}^*
ε_{24}^*
ζ_4
ζ_4^*


ζ1
ζ2
ζ3
ζ4
ε11
ε24
Y11
Y21
Y12
Y22
Y13
Y23
Y14
Y24
η1
η2
η3
η4
δ1
ξ
ξ*
η1*
η2*
η3*
η4*
Y11*
Y21*
Y12*
Y22*
Y13*
Y23*
Y14*
Y24*
ε11*
ε24*
δ1*


ζ_1
ζ_1^*
ε_{11}
Y_{11}
η_1
η_1^*
Y_{11}^*
ε_{11}^*
Y_{21}
Y_{21}^*
ζ_2
Y_{12}
Y_{12}^*
η_2
η_2^*
Y_{22}
Y_{22}^*
δ_1
ξ
ξ^*
δ_1^*
Y_{13}
Y_{13}^*
η_3
η_3^*
Y_{23}
Y_{23}^*
ζ_3
Y_{14}
Y_{14}^*
η_4
η_4^*
ε_{24}
Y_{24}
Y_{24}^*
ε_{24}^*
ζ_4

Appendix: LISREL input files

```
TI Model A described in the paper

DA  NI = 16  NO = 470  MA=PM

LA
Good1 Cont1 Good2 cont2 Good3 Cont3 Good4 Cont4
Bad1 Discon1 Bad2 Discon2 Bad3 Discon3 Bad4 Discon4

PM  FI = set1.pm
AC  FI = set1.acm

MO NY=16  NE=12  LY=FU,FI  PS=SY,FI  TE=DI,FR  BE=FU,FI

LE
ETA1  ETA2  ETA3  ETA4
ETA*1 ETA*2 ETA*3 ETA*4
KSI   KSI*
MetF1  MetF*1

! Fixing the scales of the latent state variables
VALUE  1.0  LY(1,1)  LY(3,2)   LY(5,3)  LY(7,4)
VALUE  1.0  LY(9,5)  LY(11,6)  LY(13,7) LY(15,8)

! Set free the other loadings on the latent state variables
FREE LY(2,1)   LY(4,2)   LY(6,3)  LY(8,4)
FREE LY(10,5)  LY(12,6)  LY(14,7) LY(16,8)

! Set free the variances of the method factors
FREE PS(11,11) PS(12,12)

! Fixing the scales of the method factors
VALUE  1.0  LY(2,11)  LY(10,12)

! Set free the loadings of the method factors
FREE LY(4,11) LY(6,11) LY(8,11)
FREE LY(12,12) LY(14,12) LY(16,12)
START 1.0 LY(4,11) LY(6,11) LY(8,11)
START 1.0 LY(12,12) LY(14,12) LY(16,12)

! Set free the variances of the latent trait variables
FREE PS(9,9) PS(10,10)

!Fixing the scales of the Traits
VALUE 1.0 BE(1,9) BE(5,10)

! Set free the effects of the Traits
FREE  BE(2,9) BE(3,9) BE(4,9)
FREE  BE(6,10) BE(7,10) BE(8,10)

ST 1.0 BE(2,9) BE(3,9) BE(4,9)
ST 1.0 BE(6,10) BE(7,10) BE(8,10)
ST 0.7 PS(9,10)

! Set free the covariances between the latent traits
FREE  PS(9,10)

! Set free the covariances between the method factors
FREE  PS(11,12)

! Set free the variances of the latent state residuals
```

```
FREE PS(1,1)  PS(2,2)  PS(3,3)  PS(4,4)
FREE PS(5,5)  PS(6,6)  PS(7,7)  PS(8,8)

! Set free the covariances between the latent state residuals
FREE PS(1,5)  PS(2,6)  PS(3,7)  PS(4,8)

EQUAL   TE(1,1)   TE(3,3)    TE(5,5)    TE(7,7)
EQUAL   TE(2,2)   TE(4,4)    TE(6,6)    TE(8,8)
EQUAL   TE(9,9)   TE(11,11)  TE(13,13)  TE(15,15)
EQUAL   TE(10,10) TE(12,12)  TE(14,14)  TE(16,16)

PD
OU ND=3 WP SI=LISOUT.MAT AD=OFF MI SE SC
```

TI Model B described in the paper

```
DA  NI = 16  NO = 470  MA=PM

LA
Good1 Cont1 Good2 cont2 Good3 Cont3 Good4 Cont4
Bad1 Discon1 Bad2 Discon2 Bad3 Discon3 Bad4 Discon4

PM  FI = set1.pm
AC  FI = set1.acm

MO NY=16  NE=16  LY=FU,FI  PS=SY,FI  TE=DI,FR  BE=FU,FI

LE
ETA1  ETA2  ETA3  ETA4
ETA*1 ETA*2 ETA*3 ETA*4
KSI   KSI*
MetF1  MetF*1
Zeta1 Zeta2 Zeta3 Zeta4

!Fixing the scales of the latent state variables
VALUE  1.0  LY(1,1)  LY(3,2)  LY(5,3)  LY(7,4)
VALUE  1.0  LY(9,5)  LY(11,6)  LY(13,7) LY(15,8)

! Set free the other loadings on the latent state variables
FREE LY(2,1)  LY(4,2)  LY(6,3)  LY(8,4)
FREE LY(10,5)  LY(12,6)  LY(14,7) LY(16,8)

! Set free the variances of the method factors
FREE PS(11,11) PS(12,12)

! Fixing the scales of the method factors
VALUE  1.0  LY(2,11)  LY(10,12)

! Set free the loadings of the method factors
FREE LY(4,11) LY(6,11) LY(8,11)
FREE LY(12,12) LY(14,12) LY(16,12)
START 1.0 LY(4,11) LY(6,11) LY(8,11)
START 1.0 LY(12,12) LY(14,12) LY(16,12)

! Set free the variances of the latent trait variables
FREE PS(9,9) PS(10,10)

!Fixing the scales of the Traits
VALUE 1.0 BE(1,9) BE(5,10)
```

```
! Set free the effects of the Traits
FREE  BE(2,9) BE(3,9) BE(4,9)
FREE  BE(6,10) BE(7,10) BE(8,10)

ST 1.0 BE(2,9) BE(3,9) BE(4,9)
ST 1.0 BE(6,10) BE(7,10) BE(8,10)
ST 0.7 PS(9,10)

! Set free the covariances between the latent traits
FREE  PS(9,10)

! Set free the covariances between the method factors
FREE  PS(11,12)

! Set free the variances of the common latent state residuals
FREE PS(13,13) PS(14,14) PS(15,15) PS(16,16)
ST 0.5 PS(13,13) PS(14,14) PS(15,15) PS(16,16)

! Fix the loadings of the common latent state residuals
VALUE  1.0 BE(1,13) BE(2,14) BE(3,15) BE(4,16)
FREE       BE(5,13) BE(6,14) BE(7,15) BE(8,16)
START -1.0 BE(5,13) BE(6,14) BE(7,15) BE(8,16)

! Set free the variances of the latent state residuals
!FREE PS(1,1)  PS(2,2)  PS(3,3)  PS(4,4)
!FREE PS(5,5)  PS(6,6)  PS(7,7)  PS(8,8)

! Set free the covariances between the latent state residuals
!FREE PS(1,5)  PS(2,6)  PS(3,7)  PS(4,8)

EQUAL   TE(1,1)   TE(3,3)    TE(5,5)    TE(7,7)
EQUAL   TE(2,2)   TE(4,4)    TE(6,6)    TE(8,8)
EQUAL   TE(9,9)   TE(11,11)  TE(13,13)  TE(15,15)
EQUAL   TE(10,10) TE(12,12)  TE(14,14)  TE(16,16)

PD
OU WP SI=LISOUT.MAT AD=OFF MI SE SC
```

TI Model C described in the paper

```
DA  NI = 16  NO = 470  MA=PM

LA
Good1 Cont1 Good2 cont2 Good3 Cont3 Good4 Cont4
Bad1 Discon1 Bad2 Discon2 Bad3 Discon3 Bad4 Discon4

PM  FI = set1.pm
AC  FI = set1.acm

MO NY=16  NE=16  LY=FU,FI  PS=SY,FI  TE=DI,FR  BE=FU,FI

LE
ETA1  ETA2  ETA3  ETA4
ETA*1 ETA*2 ETA*3 ETA*4
KSI   KSI*
MetF1  MetF*1
Zeta1 Zeta2 Zeta3 Zeta4

! Fixing the scales of the latent state variables
VALUE  1.0  LY(1,1)  LY(3,2)  LY(5,3)  LY(7,4)
VALUE  1.0  LY(9,5)  LY(11,6)  LY(13,7) LY(15,8)
```

```
! Set free the other loadings on the latent state variables
FREE LY(2,1)  LY(4,2)  LY(6,3)  LY(8,4)
FREE LY(10,5)  LY(12,6)  LY(14,7) LY(16,8)

! Set free the variances of the method factors
FREE PS(11,11) PS(12,12)

! Fixing the scales of the method factors
VALUE  1.0  LY(2,11)  LY(10,12)

! Set free the loadings of the method factors
FREE LY(4,11) LY(6,11) LY(8,11)
FREE LY(12,12) LY(14,12) LY(16,12)
START 1.0 LY(4,11) LY(6,11) LY(8,11)
START 1.0 LY(12,12) LY(14,12) LY(16,12)

! Set free the variances of the latent trait variables
FREE PS(9,9) PS(10,10)

! Fixing the scales of the Traits
VALUE 1.0 BE(1,9) BE(5,10)

! Set free the effects of the Traits
FREE  BE(2,9) BE(3,9) BE(4,9)
FREE  BE(6,10) BE(7,10) BE(8,10)

ST 1.0 BE(2,9) BE(3,9) BE(4,9)
ST 1.0 BE(6,10) BE(7,10) BE(8,10)
ST 0.7 PS(9,10)

! Set free the covariances between the latent traits
FREE  PS(9,10)

! Set free  the covariances between the method factors
FREE  PS(11,12)

! Set free the variances of the common latent state residuals
FREE PS(13,13) PS(14,14) PS(15,15) PS(16,16)
ST 0.5 PS(13,13) PS(14,14) PS(15,15) PS(16,16)

!Fix the loadings of the common latent state residuals
VALUE  1.0 BE(1,13) BE(2,14) BE(3,15) BE(4,16)
VALUE -1.0 BE(5,13)
FREE                BE(6,14) BE(7,15) BE(8,16)
START -1.0          BE(6,14) BE(7,15) BE(8,16)

! Set free the variances of the latent state residuals
! for the first occasion of measurement
FREE PS(1,1) PS(5,5)

EQUAL   TE(1,1)   TE(3,3)    TE(5,5)    TE(7,7)
EQUAL   TE(2,2)   TE(4,4)    TE(6,6)    TE(8,8)
EQUAL   TE(9,9)   TE(11,11)  TE(13,13)  TE(15,15)
EQUAL   TE(10,10) TE(12,12)  TE(14,14)  TE(16,16)

PD
OU WP SI=LISOUT.MAT AD=OFF MI SE SC
```