Psychological Methods
2008, Vol. 13, No. 4, 337–353

Copyright 2008 by the American Psychological Association
1082-989X/08/$12.00 DOI: 10.1037/a0013269

A Double-Structure Structural Equation Model for Three-Mode Data

Jorge González, Paul De Boeck, and Francis Tuerlinckx
K. U. Leuven

Structural equation models are commonly used to analyze 2-mode data sets, in which a set of objects is measured on a set of variables. The underlying structure within the object mode is evaluated using latent variables, which are measured by indicators coming from the variable mode. Additionally, when the objects are measured under different conditions, 3-mode data arise, and with this, the simultaneous study of the correlational structure of 2 modes may be of interest. In this article the authors present a model with a simultaneous latent structure for 2 of the 3 modes of such a data set. They present an empirical illustration of the method using a 3-mode data set (person by situation by response) exploring the structure of anger and irritation across different interpersonal situations as well as across persons.

Keywords: structural equation model, three-mode data, individual differences, situational differences

Supplemental materials: http://dx.doi.org/10.1037/a0013269.supp

Structural equation models (SEMs) are popular tools in the social sciences. They are often used for investigating two-mode data arrays in which a set of objects (Mode 1) is measured on a set of variables (Mode 2). Examples of such two-mode data sets are abundant. As a specific example, which is elaborated upon in the remainder of the article, consider the case in which a group of persons is asked to answer a set of questions regarding emotional responses to a particular kind of situation (e.g., various negative affect responses to a frustrating situation). In this case, the object mode consists of a random sample of persons, and the emotional responses comprise the variable mode. The resulting data set is a two-dimensional array of persons by responses. An SEM can then be used to explain the covariation between the manifest variables (i.e., the emotional responses) by making use of latent variables (and the relationships between them). The latent variables model individual differences.

It is not uncommon, however, for a third mode to be present in the data. Following the previous example, consider P persons who have been asked to answer R questions regarding their emotions in S different situations. These kinds of data are represented in a three-dimensional array and are referred to as *three-mode data* (e.g., Cattell, 1946, 1952; Kroonenberg, 2005), in which persons and responses are extended with a third mode that is, in this case, the situation mode. In other cases it could be time, as in a longitudinal design, with moments in time instead of situations as the third mode (Cattell, 1952). Persons, responses, situations, and time are the four evident modes that can be used in a psychological study. In theory all four modes can be combined in one study, but in practice one is either interested in the impact of situations, such as contexts and experimental conditions, or in time, as in a longitudinal study.

For the data that are analyzed in this article, a set of 679 persons has been measured with respect to four emotional responses (frustration, tendency to act antagonistically, irritation, and anger) in a set of 11 situations. Thus, the situations are the third mode of the data. Figure 1 shows a schematic representation of the three-dimensional data array in which the observation y_{prs} represents the response of person p ($p = 1, \ldots, P$) to question r ($r = 1, \ldots, R$) in situation s ($s = 1, \ldots, S$).

In the context of the emotional response data set, it can be hypothesized that the propensity of individuals to one response will be positively correlated to the propensity to other responses because of general underlying traits related to negative and positive affect. On the other hand, it is possible that an emotional response may be less correlated, or even negatively correlated, to similar emotional responses when one looks at the responses from a situational

Jorge González, Paul De Boeck, and Francis Tuerlinckx, Department of Psychology, K. U. Leuven, Leuven, Belgium.

We have been partly supported by the Fund for Scientific Research—Flanders (FWO) Grant No. G-0148-04 and by the K. U. Leuven Research Council Grant No. GOA/2005/04. We want to thank Betsy Feldman for her invaluable help with the editing of the article, in particular the English.

Correspondence concerning this article should be addressed to Jorge González, who is now at the Centro de Medición MIDE UC, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Macul, Santiago, Chile. E-mail: jagonzal@uc.cl

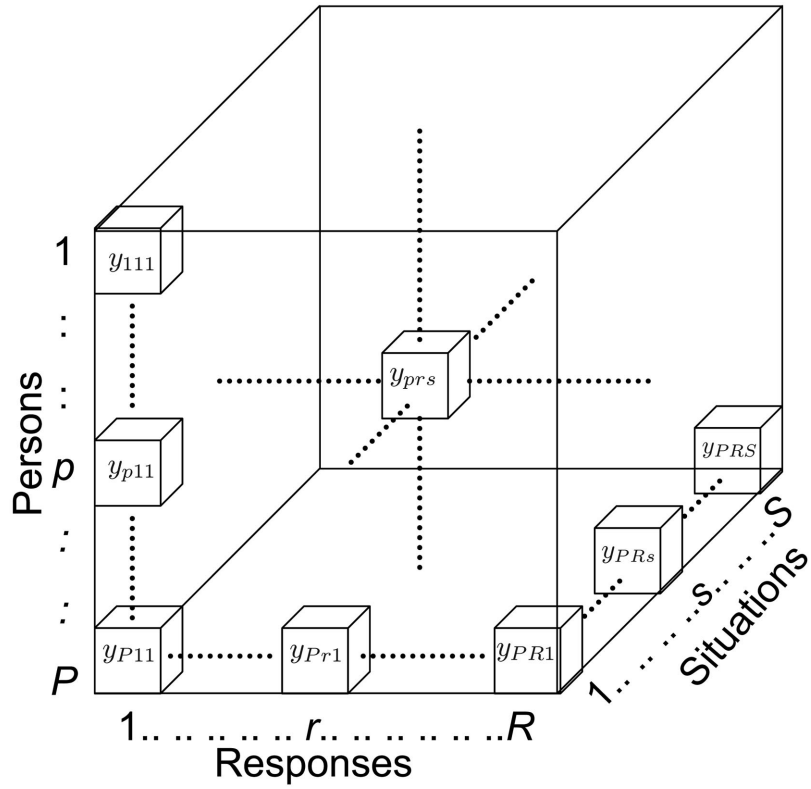

Figure 1. Three-mode data. The observation y_{prs} represents the response of person p to question r under situation s.

perspective because, for example, there is room for only one response at a time or because there are strong implicit or explicit situational rules. As a concrete example, the same person may experience both anger and guilt across situations, but situations that generally provoke anger may be different from those that provoke guilt (e.g., Vansteelandt, Van Mechelen, & Nezlek, 2005; Zelenski & Larsen, 2000). This would be reflected in a low correlation between anger and guilt across situations, but it would not be reflected in an analysis of correlated individual differences, such as in the common SEM approach.

Given that researchers are interested in the structures of both the individual differences and situational differences, an approach is needed that allows one to model the structures of two sides of the data array simultaneously. This requires that the psychometric approach and the experimental approach be combined in a single method, as suggested by Cronbach (1957) in his article on the two disciplines in scientific psychology. This is exactly what the *double-structure structural equation model* (2sSEM) aims to accomplish. If time had been chosen as the third mode, the 2sSEM would allow for a combination of psychometric and developmental approaches. The combined approach allows us to evaluate if between-persons phenomena can be generalized to the situational or time facet.

It is assumed that not only the persons but also the situations (or moments in time) are sampled from a population (often called a *universe*). In the next step, a joint model is specified for the individual differences and situational differences structure of the data. The notion of randomly sampled situations is also present in other contexts such as the random-effects analysis of variance (where the levels of a factor may be considered a random sample from a larger population) and in generalizability theory (e.g., Brennan, 2001; Cronbach, Rajaratnam, & Gleser, 1963), where, in the simplest case, the items are considered a random sample. Treating the situations as a random sample is closely linked with the idea of exchangeability among the situations (see also Lindley & Novick, 1981; Snijders, 2005). The model presented here is an extension of a model presented by De Boeck and Smits (2006), which is itself an extension of a basic item response theory model (i.e., the linear logistic test model). However, the model by De Boeck and Smits treats only one of the three modes as random; the other two are fixed.

Various models for three-mode data can be found in the psychometric literature. These models can be differentiated on the basis of three types of features: (a) the nature of the effects of modes (random vs. fixed), (b) the type of interactions among the modes implicit in the model, and (c) the

nature of dependence or independence among the responses within and across situations and within and across persons. These three features are discussed for all models under consideration. For a short overview of these models, and in line with the application that is described later, it is assumed that the three modes are persons, situations, and responses. Table 1 summarizes the differences among various possible models. In Table 1 we use Fienberg's (1980; see also Agresti, 2002) notation to express the interactions between variables. The use of the term *interaction* may seem unusual, but it makes sense when one considers that factors and dimensions arise when elements of one mode (e.g., persons) differ as a function of elements in a different mode (e.g., responses).

Two families of three-mode models can be discerned. The first family originated with Tucker (1966), who proposed a three-mode principal-component model, now known as the *Tucker3 model*. This type of model is the most general one, as far as interactions are concerned, because it includes three-way interactions between persons, situations, and responses. Each of the modes is analyzed in terms of its principal components, and in principle, all combinations (all products) of the three kinds of components are included. The *core matrix* of these models determines the degree to which each of the combinations plays a role. Because three-way interactions are considered, the lower order effects are implied, unless the data array is first processed to eliminate main effects (and possibly also pairwise interactions).

Depending on the specification of the core matrix, more restricted versions of Tucker3 models can be obtained (see Kroonenberg, 1983, for an annotated bibliography). For example, in the CANDECOMP/PARAFAC model (Carroll & Chang, 1970; Harshman, 1970), each of the three modes has the same number of factors. Within each mode, each factor is exclusively linked to one factor of each of the other two modes. As a result the core matrix takes a diagonal shape. In the original Tucker3 family, all parameters are fixed-effect parameters (see the column labeled Tucker3 (a) in Table 1). In a further development, the parameters referring to persons are considered to be random (e.g., Bentler & Lee, 1978, 1979; Bloxom, 1968; Lee & Fong, 1983; Oort, 1999), as indicated in the column with heading Tucker3 (b) in Table 1. When all parameters are fixed, as is commonly the case for the Tucker3 family, the issue of dependence and independence is irrelevant, as indicated in Table 1. However, if the persons are random, then the dependency structure is the same as in the SEM for multitrait–multimethod (MTMM) data. This is explained next.

A second family of models is the set of SEMs for MTMM data (e.g., M. W. Browne, 1984; Eid, 2000; Eid, Lischetzke, & Nussbeck, 2006; Eid, Lischetzke, Nussbeck, & Trierweiler, 2003). MTMM models are represented in the column labeled SEM–MTMM in Table 1. Let us assume that the traits correspond to responses and that the methods correspond to situations. In such models, two types of factors (latent traits) occur: response (trait) factors and situation (method) factors. The effects of trait and method factors on observed responses are additive. Response factors reflect the interactions between persons and responses and express individual differences based on those responses. Situation factors reflect the interactions between persons and situations and express how the individual differences depend on the situations. In other words, SEM–MTMM is focused on individual differences, so it is not surprising that the person mode is described in terms of random effects. The model utilizes a covariance structure across persons, which is decomposed into two kinds of covariation: the first resulting from individual differences related to responses and the second from individual differences related to situations. In SEM–MTMM, with observed variables representing situation–response pairs, responses are assumed to be independent across individuals but correlated within individuals.

In contrast to the previous two models (Tucker3 and SEM–MTMM), the 2sSEM models the interactions of the responses with persons *and* responses with situations (see Table 1). In the former case, the latent factors refer to

Table 1
Three Approaches for Modeling Three-Mode Data

Variable	Tucker3 (a)	Tucker3 (b)	SEM–MTMM	2sSEM
Interactions	[PSR]	[PSR]	[PS] [PR]	[PR] [SR]
Parameters				
Persons	fixed	random	random	random
Situations	fixed	fixed	fixed	random
Responses	fixed	fixed	fixed	fixed
Dependence structure				
Dependent	irrelevant	pairs (S, R)	pairs (S, R)	R
Independent	irrelevant	P	P	nonoverlapping pairs (P, S)

Note. Variables that interact are grouped together inside a set of brackets; variables that are independent of one another are placed in separate brackets. SEM–MTMM = structural equation model for multitrait–multimethod data; 2sSEM = double-structure structural equation model; P = persons; S = situations; R = responses.

individual differences, and in the latter case they refer to situational differences. Furthermore, the person mode and the situation mode are both treated as random, so two different covariance structures are considered: one over persons and a second over situations. Finally, the observations from different persons are not required to be independent, as is the case in the SEM–MTMM. Instead, person–situation pairs are assumed to be independent when they have neither the person nor the situation in common. See Equation 5 for the covariance structure of the 2sSEM. Because the SEM–MTMM is widely accepted and commonly used for modeling Person × Situation × Response data and because the 2sSEM we are proposing is relatively novel, we highlight the differences between the two models; see Table 1 for a summary.

1. The first distinction between the two models concerns individual differences and situational differences. All latent variables in the SEM–MTMM are individual-difference variables because the persons are treated as random. In fact, two kinds of factors are named in SEM–MTMM: trait factors and method factors. In the present context, situations are equivalent to methods. Trait factors refer to interactions between persons and responses. For example, anger and irritation are two of the responses in the example used in this article. Without any interaction between persons and responses, anger and irritation would have equal loadings. Generally speaking, if the individual trait levels do not differentially depend on the responses, then a single factor is required and all loadings will be equal. However, if the individual differences do depend on the responses, multiple factors with different loadings are likely. The presence of more than one factor indicates that there is interaction between persons and responses. In a similar way, situation (or method) factors refer to interactions between persons and situations (or methods). Hence, situational factors also refer to individual differences.

The 2sSEM also uses two types of latent variables. The first type models individual differences, and the second models situational differences. The individual difference factors come about as a result of the interaction between persons and responses (identical to the SEM–MTMM), whereas the situational difference factors are due to the interaction between situations and responses; they capture differences that depend on the situation (but do not vary across individuals). Suppose, for example, that a 2sSEM includes anger and irritation factors. This implies that some situations tend to generate anger, whereas other situations are more likely to cause irritation. The 2sSEM and the SEM–MTMM situational factors are, therefore, fundamentally different.

2. The second difference concerns situations being treated as random. In the 2sSEM, in contrast with the SEM–MTMM, the situation mode is also treated as random. This implies that there is a universe of exchangeable situations from which the researcher has randomly sampled a representative subset with the intent of generalizing across all situations. For example, in this study we consider all possible situations that may elicit anger or irritation and have sampled 11 exchangeable situations. In contrast, situations or *methods* in SEM–MTMM are fixed and nonexchangeable. There is an inherent interest in discriminating among individual responses across a few specific and distinct situations or methods. There could be reasons to be interested in particular situations. For example, in a study investigating situational differences in anger, one may concentrate on the home environment and the work environment. These situations are fixed and nonexchangeable, and the interest is in the covariance among the responses in these two major situations. Another implication of the assumption of random situation in 2sSEM is that the total number of situations is relatively large compared to the number of fixed methods in SEM–MTMM. It is clear that they are to be treated as fixed and that there is not a large number of them.

3. The third distinction concerns independence assumptions. Another important difference between SEM–MTMM and 2sSEM is the dependence structure they assume for the observations. In the SEM–MTMM, the observations from different persons are independent. However, observations within an individual (i.e., from different situation–response pairs) may be dependent. This is because situation–response pairs, which constitute the variables or indicators in an SEM–MTMM, may correlate due to common (response or situation) underlying latent variables. In the 2sSEM, observations within an individual are expected to show dependence (this holds for responses both in the same situation and in different situations). However, observations from different persons within the same situation may also be dependent upon each other, because the situation may affect all persons in the same way. It is only for distinct person–situation pairs (with no common elements) that independence is assumed.

The remainder of the article is organized as follows: First, we introduce the data set we analyzed and formulate the basic research questions. Second, we introduce the 2sSEM using the concepts underlying the regular SEM; Bayesian methods for model estimation, model selection, and model checking are discussed. To explain the model and to compare it with the SEM–MTMM, we use the normal linear version. Next, a three-mode data set (persons by situations by responses) about four emotions is used to evaluate the model, and finally, conclusions and a discussion are presented.

Description of the Data

A central question in emotion research pertains to the differences between irritation and anger (e.g., Averill, 1982; Frijda, Kuipers, & ter Schure, 1989; Van Coillie & Van Mechelen, 2006). Although the two emotions can be distin-

guished quite easily phenomenologically, it is still unclear in exactly what ways they differ. The componential theory in emotion research forms a fruitful framework for studying this type of problem (Scherer, Schorr, & Johnstone, 2001).

The basic hypothesis underlying the theory is that emotions rely on two components, a feeling component and an action-tendency component, and that the weight of each may differ depending on the emotion. Specifically, anger is hypothesized to be more action-oriented than irritation, whereas irritation is hypothesized to be more feeling-based. In addition, it is hypothesized that anger is primarily associated with an antagonistic action tendency and to a lesser degree with frustration, whereas irritation is thought to be primarily associated with the feeling of frustration.

The data in this article came from a sample of 679 Dutch-speaking students in their last year of high school (Kuppens, Van Mechelen, Smits, De Boeck, & Ceulemans, 2007). The students were presented with 11 situations, considered to be randomly drawn from a population of situations that might elicit negative emotional responses. They were asked to use a 4-point scale, ranging from 0 (*not at all*) to 3 (*very strong*), to indicate the degree to which they would experience frustration, a tendency to take antagonistic action, irritation, and anger. To simplify the model construction, the original data were dichotomized by recoding the values 0 and 1 as 0, and 2 and 3 as 1. Descriptions of the 11 situations are shown in Table 2.

The histograms in Figure 2 show the proportions of positive answers (or endorsements) for each of the four response types, aggregated over persons (Panel A) and situations (Panel B). Panel A shows that, on average, approximately 70% of the 11 situations elicited endorsements for anger, frustration, and antagonistic action. Endorsements for irritation were more common, and there was variability in all four responses across the 11 situations. Panel B shows the proportions of endorsements for each of the four responses and for the 679 persons. The variability across persons is visible in the plots. Many individuals showed negative affect in all of the situations. The skewed distribution on the manifest level might suggest a problem, but note that this is not necessarily a problem on the latent level. A normal distribution on the latent level may appear as a skewed distribution on the manifest level, depending on the precise values of the situation and response parameters. These histograms show that, for both situations and persons, there is sufficient variability to investigate the covariance structures.

Table 2
Summary Descriptions of the Situations Used in the Study

Situation
You are blamed for someone else's failures after a sports match
A fellow student loses your notes, causing you to fail the exam
A floppy disk holding an important school assignment is destroyed by your computer
Your sibling sneaks out when you both have to clean up the house
Being a work–study student yourself, an employee makes you do all his chores
A fellow student fails to return your notes the day before an exam
After working hard on an assignment, your teacher says it's still not better than your previous work
It's hard to study when the neighbors make a lot of noise and it's a hot day
A friend lets you down on an appointment to go out with his/her friend
You're at a party, and someone tells you that a friend outside has smashed your bike
You hear that a friend is spreading gossip about you

After aggregating over either the situation or the person mode, the resulting two-way data can be fitted with a standard SEM; however, this approach has serious disadvantages. After aggregating over situations, one can answer only questions with respect to individual differences and correlations across individuals. Of course, it is possible to then investigate the differences and the correlational structure of the situations by aggregating over the person mode, but the net result of these actions is two unrelated models. For example, when aggregating over situations, it is possible to investigate whether persons who tend to be angry also tend to be irritated. Or, when aggregating over persons, one might ask whether situations which tend to elicit anger also tend to elicit irritation. The relationships between emotional responses in these two models may be similar, but this is not required.

The basic research questions addressed with these data are as follows (see Figure 3 for a graphical representation):

1. What is the structure of the individual differences? If each emotional response (frustration, antagonistic action tendency, irritation, anger) is considered a separate latent variable over which individuals can differ, how are these latent variables related to one another? Is frustration the primary component of irritation (i.e., vs. antagonistic action tendency), and is the tendency to act antagonistically the primary component of anger (i.e., vs. frustration)? In terms of Figure 3, the hypothesis implied by the first research question is that $a > c$ and that $d > b$.

2. What is the structure on the situation side? Is it also the case for situations that frustration is the primary component of irritation and that antagonistic action tendency is the primary component of anger? The hypothesis is again that $a > c$ and that $d > b$ in Figure 3, but now the figure refers to situations rather than individuals.

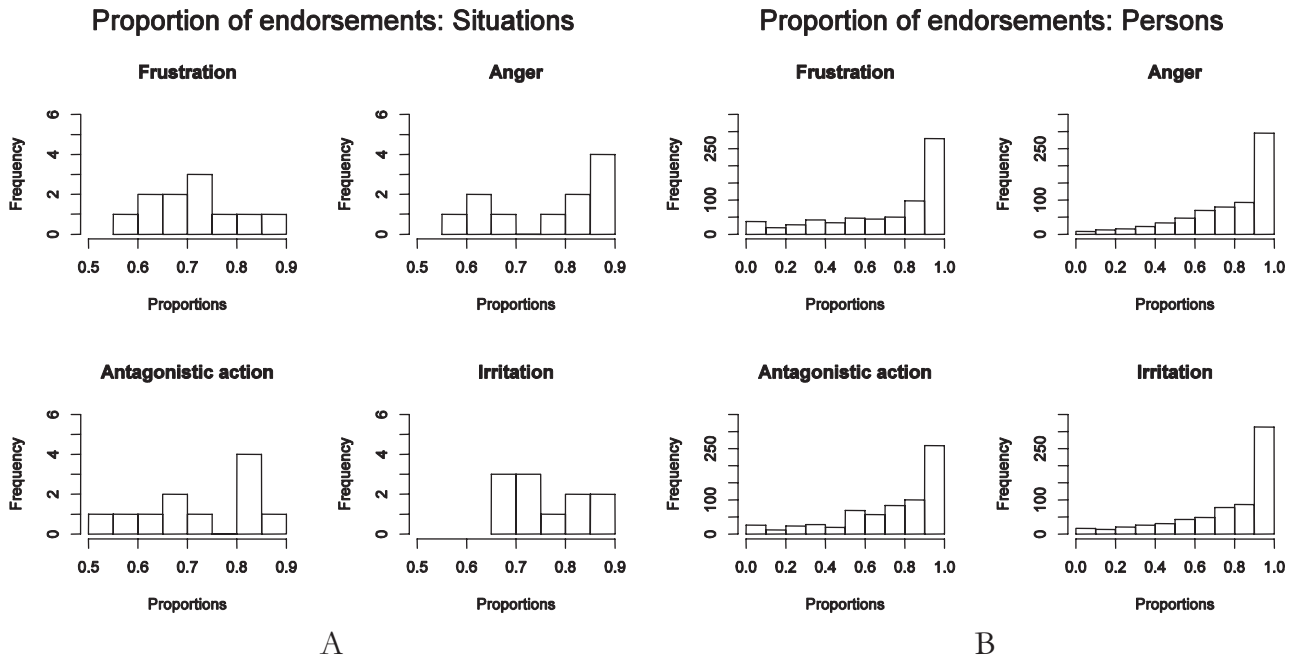

Figure 2. Proportion of endorsements for each of the four responses. A: Situation side (data aggregated over persons). B: Person side (data aggregated over situations).

3. If the structures of the person side and the situation side are similar, as hypothesized in 1 and 2, are the population values for effects that link the four latent variables the same for persons and situations? In Figure 3, this hypothesis states that the standardized values for *a*, *b*, *c*, and *d* are identical for the structure of individual differences and the structure of the situational differences. The hypothesis is phrased in terms of standardized values, in order for the estimates to be independent of the variances.

It is neither a statistical nor psychological necessity that the structures for the two modes be identical or even similar. For example, the structure may be one-dimensional for the persons, with all four emotional responses resulting from a single general factor, whereas situations may require multiple dimensions to explain the same four emotional responses.

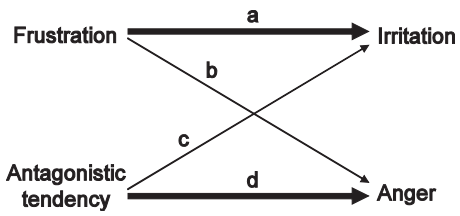

Figure 3. Graphical representation of the research questions. *a*, *b*, *c*, and *d* are effect parameters.

Double-Structure Structural Equation Model (2sSEM)

In this section, the 2sSEM that takes into account both individual and situational differences is described. Following Bollen (1989), an SEM encompasses two parts. The first part is the *measurement model*, which contains the equations representing the link between the observed and latent variables. The second part is the *structural model*, which contains the equations that represent the hypothesized relations between the latent variables. We explicitly define the measurement and structural model used for the study of both individual and situational differences simultaneously later.

For ease of presentation, we introduce the 2sSEM using continuous observed variables. However, because the data of our central application are binary, we later demonstrate how the 2sSEM can be extended in a straightforward manner to handle dichotomous data.

The 2sSEM

In a general formulation of the model, assume that there are Q latent person variables and T latent situation variables. The following notation is used: The observed random vari-

able y_{prs} represents the "magnitude" of person p's response r in situation s. In the emotion data set, the index r may take one of four values—1, 2, 3, or 4, which correspond to frustration, antagonistic action tendency, irritation, and anger, respectively. The symbol η_{pq}^{PER} refers to the score for person p on the qth latent trait ($q = 1, \ldots, Q$). The superscript PER is somewhat superfluous because the subscript p already indicates that the latent trait refers to the person side and not the situation side. However, for the factor loadings (λs, as in Equation 1) this difference will not be as clear so we will retain the superscript notation. Likewise, η_{st}^{SIT} refers to the score for situation s on the tth latent trait ($t = 1, \ldots, T$). The effect of responses is represented by τ_r. From the description of the effects it is clear whether an effect is fixed or random.

Measurement model. The measurement model we use is

$$y_{prs} = \sum_{q=1}^{Q} \lambda_{rq}^{\mathrm{PER}} \eta_{pq}^{\mathrm{PER}} + \sum_{t=1}^{T} \lambda_{rt}^{\mathrm{SIT}} \eta_{st}^{\mathrm{SIT}} + \tau_r + \varepsilon_{prs}$$
$$= \boldsymbol{\lambda}_r^{\mathrm{PER}\,\intercal} \boldsymbol{\eta}_p^{\mathrm{PER}} + \boldsymbol{\lambda}_r^{\mathrm{SIT}\,\intercal} \boldsymbol{\eta}_s^{\mathrm{SIT}} + \tau_r + \varepsilon_{prs}, \quad (1)$$

where $\lambda_{rq}^{\mathrm{PER}}$ is the loading of response r on the qth ($q = 1, \ldots, Q$) latent person variable, and $\lambda_{rt}^{\mathrm{SIT}}$ is the loading of response r on the tth ($t = 1, \ldots, T$) latent situation variable. τ_r is the fixed general effect for response r, and ε_{prs} is a random error term with mean zero. Both η_{pq}^{PER} and η_{st}^{SIT} are random variables, which are described when the structural model is presented.

The measurement model in Equation 1 can also be written in matrix form. First, all responses pertaining to the same person–situation combination are collected in the $R \times 1$ vector, $\mathbf{y}_{ps}$. The row vectors of loadings ($\boldsymbol{\lambda}_r^{\mathrm{PER}\,\intercal}$ and $\boldsymbol{\lambda}_r^{\mathrm{SIT}\,\intercal}$) are stacked in the loading matrices $\boldsymbol{\Lambda}^{\mathrm{PER}}$ and $\boldsymbol{\Lambda}^{\mathrm{SIT}}$, respectively, and the τ_r parameters are gathered in one $R \times 1$ vector $\boldsymbol{\tau}$. The model then becomes

$$\mathbf{y}_{ps} = \boldsymbol{\Lambda}^{\mathrm{PER}} \boldsymbol{\eta}_p^{\mathrm{PER}} + \boldsymbol{\Lambda}^{\mathrm{SIT}} \boldsymbol{\eta}_s^{\mathrm{SIT}} + \boldsymbol{\tau} + \boldsymbol{\varepsilon}_{ps}, \quad (2)$$

or

$$\begin{pmatrix} y_{p1s} \\ \vdots \\ y_{prs} \\ \vdots \\ y_{pRs} \end{pmatrix} = \begin{pmatrix} \lambda_{11}^{\mathrm{PER}} & \ldots & \lambda_{1Q}^{\mathrm{PER}} \\ \vdots & \ddots & \vdots \\ \lambda_{r1}^{\mathrm{PER}} & \ldots & \lambda_{rQ}^{\mathrm{PER}} \\ \vdots & \ddots & \vdots \\ \lambda_{R1}^{\mathrm{PER}} & \ldots & \lambda_{RQ}^{\mathrm{PER}} \end{pmatrix} \begin{pmatrix} \eta_{p1}^{\mathrm{PER}} \\ \vdots \\ \eta_{pq}^{\mathrm{PER}} \\ \vdots \\ \eta_{pQ}^{\mathrm{PER}} \end{pmatrix}$$
$$+ \begin{pmatrix} \lambda_{11}^{\mathrm{SIT}} & \ldots & \lambda_{1T}^{\mathrm{SIT}} \\ \vdots & \ddots & \vdots \\ \lambda_{r1}^{\mathrm{SIT}} & \ldots & \lambda_{rT}^{\mathrm{SIT}} \\ \vdots & \ddots & \vdots \\ \lambda_{R1}^{\mathrm{SIT}} & \ldots & \lambda_{RT}^{\mathrm{SIT}} \end{pmatrix} \begin{pmatrix} \eta_{s1}^{\mathrm{SIT}} \\ \vdots \\ \eta_{st}^{\mathrm{SIT}} \\ \vdots \\ \eta_{sT}^{\mathrm{SIT}} \end{pmatrix} + \begin{pmatrix} \tau_1 \\ \vdots \\ \tau_r \\ \vdots \\ \tau_R \end{pmatrix} + \begin{pmatrix} \varepsilon_{p1s} \\ \vdots \\ \varepsilon_{prs} \\ \vdots \\ \varepsilon_{pRs} \end{pmatrix}. \quad (3)$$

In a model for continuous and normally distributed data, one may also assume that

$$\boldsymbol{\varepsilon}_{ps} \sim N(\mathbf{0}, \boldsymbol{\Phi}),$$

where $\boldsymbol{\Phi}$ is an $R \times R$ variance–covariance matrix of the error terms. Usually $\boldsymbol{\Phi}$ is taken to be diagonal. Note that we define a different distribution for the measurement error terms in the case of binary data. In the most general case, there are as many latent variables as there are responses ($Q = R$ and $T = R$). However, it is perfectly possible to formulate a model in which $Q < R$ and $T < R$, and also $Q \neq T$. Given that, in the example, $Q = R$ and $T = R$, note that y_{pr1} to y_{prS} are multiple indicators for η_{pr}^{PER}, and similarly y_{1rs} to y_{Prs} are multiple indicators for η_{sr}^{SIT}.

(Double-structure) structural model. The two-fold structural model can now be defined as follows:

$$\boldsymbol{\eta}_p^{\mathrm{PER}} = \mathbf{B}^{\mathrm{PER}} \boldsymbol{\eta}_p^{\mathrm{PER}} + \boldsymbol{\zeta}_p^{\mathrm{PER}}$$
$$\boldsymbol{\eta}_s^{\mathrm{SIT}} = \mathbf{B}^{\mathrm{SIT}} \boldsymbol{\eta}_s^{\mathrm{SIT}} + \boldsymbol{\zeta}_s^{\mathrm{SIT}}, \quad (4)$$

where $\mathbf{B}^{\mathrm{PER}}$ and $\mathbf{B}^{\mathrm{SIT}}$ are $Q \times Q$ and $T \times T$ parameter matrices for regressions among the Q latent person variables and among the T latent situation variables, respectively, and $\boldsymbol{\zeta}_p^{\mathrm{PER}}$ and $\boldsymbol{\zeta}_s^{\mathrm{SIT}}$ are residual vectors. It is assumed that both ($\mathbf{I} - \mathbf{B}^{\mathrm{PER}}$) *and* ($\mathbf{I} - \mathbf{B}^{\mathrm{SIT}}$) are nonsingular. A value of zero in a $\mathbf{B}$ matrix means that there is no effect of one latent variable on another. In particular, the diagonals of both $\mathbf{B}^{\mathrm{PER}}$ and $\mathbf{B}^{\mathrm{SIT}}$ always contain only zeros, meaning that a variable is not explained by itself.

Note that in Equation 4, a structure is imposed on both the person latent variables and the situation latent variables. Both $\boldsymbol{\eta}_p^{\mathrm{PER}}$ and $\boldsymbol{\eta}_s^{\mathrm{SIT}}$ contain subvectors of latent explanatory (independent or exogenous) and latent response (dependent or endogenous) variables. The residual vectors $\boldsymbol{\zeta}_p^{\mathrm{PER}}$ and $\boldsymbol{\zeta}_s^{\mathrm{SIT}}$ follow normal distributions $N(\mathbf{0}, \boldsymbol{\Psi}^{\mathrm{PER}})$ and $N(\mathbf{0}, \boldsymbol{\Psi}^{\mathrm{SIT}})$, respectively, in which $\boldsymbol{\Psi}^{\mathrm{PER}}$ and $\boldsymbol{\Psi}^{\mathrm{SIT}}$ contain covariances within and between the sets of explanatory and explained variables. In the models that are considered here, these covariances are assumed to be zero when an explained variable is involved but not when both variables are explanatory.

The standard graphical way of representing the SEM is a path diagram (e.g., Bollen, 1989). This graphical representation of a system of simultaneous equations can be transferred to the 2sSEM as well. Directional influence between the variables is represented by single-headed arrows, and curved two-headed arrows represent correlations. In the application that follows, we make use of the graphical representation to introduce several models (see Figure 4). The observed variables and error terms are omitted from the path diagrams for clarity.

Having defined this model, it is possible to derive the model-implied covariance matrices between the response vectors $\mathbf{y}_{ps}$ and $\mathbf{y}_{p's'}$:

$$\begin{aligned}\mathrm{Cov}(\mathbf{y}_{ps}, \mathbf{y}_{p's'}) &= \boldsymbol{\Lambda}^{\mathrm{PER}} (\mathbf{I} - \mathbf{B}^{\mathrm{PER}})^{-1} \boldsymbol{\Psi}^{\mathrm{PER}} \\ &\quad \times (\mathbf{I} - \mathbf{B}^{\mathrm{PER}})^{\mathsf{T}-1} \boldsymbol{\Lambda}^{\mathrm{PER}\,\mathsf{T}} + \quad (5) \\ \boldsymbol{\Lambda}^{\mathrm{SIT}} (\mathbf{I} - \mathbf{B}^{\mathrm{SIT}})^{-1} \boldsymbol{\Psi}^{\mathrm{SIT}} (\mathbf{I} - \mathbf{B}^{\mathrm{SIT}})^{\mathsf{T}-1} \boldsymbol{\Lambda}^{\mathrm{SIT}\,\mathsf{T}} + \boldsymbol{\Phi} & \\ \text{if } p = p' \text{ and } s = s' & \\ = \boldsymbol{\Lambda}^{\mathrm{PER}} (\mathbf{I} - \mathbf{B}^{\mathrm{PER}})^{-1} \boldsymbol{\Psi}^{\mathrm{PER}} (\mathbf{I} - \mathbf{B}^{\mathrm{PER}})^{\mathsf{T}-1} \boldsymbol{\Lambda}^{\mathrm{PER}\,\mathsf{T}} & \\ \text{if } p = p' \text{ and } s \neq s' & \\ = \boldsymbol{\Lambda}^{\mathrm{SIT}} (\mathbf{I} - \mathbf{B}^{\mathrm{SIT}})^{-1} \boldsymbol{\Psi}^{\mathrm{SIT}} (\mathbf{I} - \mathbf{B}^{\mathrm{SIT}})^{\mathsf{T}-1} \boldsymbol{\Lambda}^{\mathrm{SIT}\,\mathsf{T}} & \\ \text{if } p \neq p' \text{ and } s = s' & \\ = 0 \text{ if } p \neq p' \text{ and } s \neq s'. &\end{aligned}$$

As can be seen, in the linear 2sSEM, the model-implied covariance is nonzero for responses in two person–situation pairs that have at least one common element.

The model-implied covariance for an SEM–MTMM is a little different. The covariance is zero when $p \neq p'$, and it is $\boldsymbol{\Delta} + \boldsymbol{\Theta} + \boldsymbol{\Xi}$ when $p = p'$. The symbols $\boldsymbol{\Delta}$, $\boldsymbol{\Theta}$, and $\boldsymbol{\Xi}$ refer to the trait covariance (covariance between individual differences in the traits), the situation covariance (covariance between individual differences associated with the situations), and the residual covariance matrices, respectively. All three refer to individual differences.

Finally, note that the implied covariance structures hold only under the normal linear model. In the remainder of the article we work with discrete (binary) data. For such data it is less trivial (but also less informative) to derive the implied covariance structure. However, under the 2sSEM it is still the case that responses of different persons in the same situation and responses of the same person in different situations are both permitted to be dependent (as well as responses by the same individual in the same situation).

The Model for Dichotomous Data

A latent threshold formulation is used to make the model suitable for analyzing binary data (details are given in Appendix A). Under this formulation, the measurement model in Equation 1 becomes a model for the conditional probability $\pi_{prs} = \Pr(y_{prs} = 1 | \boldsymbol{\eta}_p^{\mathrm{PER}}, \boldsymbol{\eta}_s^{\mathrm{SIT}})$ and is expressed as

$$\pi_{prs} = F(\boldsymbol{\lambda}_r^{\mathrm{PER}\,\mathsf{T}} \boldsymbol{\eta}_p^{\mathrm{PER}} + \boldsymbol{\lambda}_r^{\mathrm{SIT}\,\mathsf{T}} \boldsymbol{\eta}_s^{\mathrm{SIT}} + \tau_r). \quad (6)$$

We chose F to be $F(x) = \exp(x) / (1 + \exp(x))$, leading to a logistic model of the form

$$\begin{aligned}\log\left(\frac{\pi_{prs}}{1 - \pi_{prs}}\right) &= \mathrm{logit}(\pi_{prs}) \\ &= \boldsymbol{\lambda}_r^{\mathrm{PER}\,\mathsf{T}} \boldsymbol{\eta}_p^{\mathrm{PER}} + \boldsymbol{\lambda}_r^{\mathrm{SIT}\,\mathsf{T}} \boldsymbol{\eta}_s^{\mathrm{SIT}} + \tau_r. \quad (7)\end{aligned}$$

The scale of the person and situation latent variables is the same in terms of the effect on the logit scale, but the scale is not standardized with respect to variance. Other selections of F are possible as well (e.g., if F equals the standard normal cumulative distribution function, the result is a probit model). The approach taken here to accommodate binary variables (and elaborated in Appendix A) can be extended easily to, for instance, ordinal data, by using the generalized linear model framework (GLM; McCullagh & Nelder, 1989). For ordered-category data, one must make a choice between adjacent and cumulative logits. That is, either a given category is compared with the one immediately below it (adjacent logit), or the category of interest, together with all higher categories, is compared with all lower categories (cumulative logit). A simple way of doing this is to assume that the distance between the category thresholds is the same for all responses, so that the rating scale model (RSM; Andrich, 1978) can be used for either adjacent or cumulative logits. Skrondal and Rabe-Hesketh (2004) discussed how GLMs may be incorporated into SEMs. Such an approach greatly expands the area of application.

Statistical Inference

The 2sSEM includes random effects for two of the three modes in the data; therefore, it is a *crossed-random-effects model* (e.g., Janssen, Schepers, & Peres, 2004; Janssen, Tuerlinckx, Meulders, & De Boeck, 2000; Snijders & Bosker, 1999). When data are not normally distributed (such that the integral over the latent distribution has no analytic solution), the estimation of model parameters is not trivial (Tuerlinckx, Rijmen, Verbeke, & De Boeck, 2006). In this case, the likelihood contains an integral of a very high dimension because, unlike a traditional random-effects model, it cannot be factorized into separate contributions of the persons (because the responses of different persons are not independent). The dimension of the integral over the latent trait distributions is, at minimum, as high as the sum of the number of elements of the modes with random effects (in the central example in this article, that would mean at least 690 dimensions), but in most cases, the dimensionality will be larger.

For such cases, the well-known Gauss–Hermite quadrature method becomes unfeasible, because of the high dimensionality of the integrals in the likelihood function. Also Laplace approximations tend to break down because the dimensionality grows as a function of the sample size (e.g., Shun, 1997; Shun & McCullagh, 1995). Possible alternative solutions are Monte Carlo and Quasi-Monte Carlo methods for numerical integration (e.g., González, Tuerlinckx, De Boeck, & Cools, 2006) or Bayesian methodology, which was used in this article. In the following section, we outline the Bayesian method that was used to estimate the 2sSEM. Details of the estimation and software implementation are

given in Appendix B and the online supplemental material, respectively. Readers interested in more information about the Bayesian approach for SEMs are referred to Lee (2007).

Bayesian Estimation

Bayesian inference is based on the posterior distribution of all parameters of interest, given the observed data. Let $\boldsymbol{\vartheta}$ be the vector of parameters, as well as latent variables, and $\mathbf{y}$ the observed data. For the 2sSEM, $\boldsymbol{\vartheta}$ contains all of the parameters in the matrices $\mathbf{B}^{\text{PER}}$, $\mathbf{B}^{\text{SIT}}$, $\boldsymbol{\Lambda}^{\text{PER}}$, $\boldsymbol{\Lambda}^{\text{SIT}}$, $\boldsymbol{\Psi}^{\text{PER}}$, $\boldsymbol{\Psi}^{\text{SIT}}$, and $\boldsymbol{\tau}$ and the latent residuals $\boldsymbol{\zeta}_p^{\text{PER}}$, $\boldsymbol{\zeta}_s^{\text{SIT}}$. If $\mathbf{y} \mid \boldsymbol{\vartheta} \sim f(\mathbf{y} \mid \boldsymbol{\vartheta})$, then Bayes's theorem leads to the following relation:

$$f(\boldsymbol{\vartheta} \mid \mathbf{y}) \propto f(\mathbf{y} \mid \boldsymbol{\vartheta}) f(\boldsymbol{\vartheta}), \tag{8}$$

which shows that the posterior density is proportional to the likelihood function times the prior. A sample is obtained from the posterior distribution and the posterior mean, and posterior standard deviation and confidence intervals are used to summarize the distribution of each parameter of interest.

In order to obtain a sample from the posterior distribution, different iterative methods belonging to the class of Markov Chain Monte Carlo (MCMC) techniques (e.g., Gelman, Carlin, Stern, & Rubin, 2003) have been developed. After the prior distributions are specified for all the parameters in the model, these algorithms generate a Markov chain. Following an initial burn-in period (which often consists of several thousand iterations), the draws can be used as a sample from the posterior distribution in Equation 8 on the condition that convergence is reached (which can be approximately checked using the $\hat{R}$ diagnostic; Gelman & Rubin, 1992).

WinBUGS (Spiegelhalter, Thomas, Best, & Lunn, 2003) is a software program for Bayesian analysis of statistical models that uses MCMC techniques. After the user provides a likelihood and prior distribution, the program automatically draws a sample of all parameters from the posterior distribution. Once convergence has been reached, parameter estimates can be obtained and inferences made.

All models in this article were fitted with WinBUGS, which was called from R (R Development Core Team, 2006) by using the R2WinBUGS package (Sturtz, Ligges, & Gelman, 2005). Convergence was assessed using the CODA package (Plummer, Best, Cowles, & Vines, 2006), which implements standard convergence criteria (e.g., Cowles & Carlin, 1996). The main R/WinBUGS code that was used is available as supplemental material online.

Bayesian Model Selection and Model Checking

When different models for data are fitted, the researcher is confronted with a model selection issue: Which model fits the data best among the set of models under consideration? Because we are working in a Bayesian framework, we may calculate Bayes factors (Kass & Raftery, 1995) and select the model with the largest posterior probability given the data. However, it is quite computationally demanding to compute Bayes factors, and therefore we considered other solutions. The Akaike information criterion (AIC; Akaike, 1974) and Bayesian information criterion (BIC; Schwarz, 1978) provide alternative measures. However, these criteria are derived in the frequentist framework and, therefore, are less applicable in our situation (although it is possible to estimate posterior AIC and BIC distributions on the basis of the draws from the posterior and compare those across different models). Our preferred index for model selection is the deviance information criterion (DIC; Spiegelhalter, Best, Carlin, & Van der Linde, 2002), which is easier to compute than Bayes factors yet is a sufficiently theoretically sound measure for model selection in a Bayesian framework. The DIC is calculated as a compromise between model fit and the number of effective parameters of the model (in the same spirit as the more traditional model selection methods such as AIC or BIC) and is defined as

$$\text{DIC} = \overline{D(\boldsymbol{\vartheta})} + p_D, \tag{9}$$

where $\overline{D(\boldsymbol{\vartheta})}$ is the posterior mean of the deviance, used to measure model fit, and p_D is an estimate of the effective number of parameters. Lower values of the criterion indicate better fitting models (for more details, see Spiegelhalter et al., 2002).

After a model is selected, it is recommended that the researcher evaluate the model's global fit to the data. For this purpose we use samples from the posterior distribution via posterior predictive checks (PPC; e.g., Gelman et al., 2003). The main idea behind PPC is that "if the model fits, then replicated data generated under the model should look similar to observed data" (Gelman et al., 2003, p. 165). In addition, a test quantity or discrepancy measure, $T(\mathbf{y},\boldsymbol{\vartheta})$, which is calculated on the data and may depend on the model parameters, may be used to measure the discrepancy between specific features of the model and data. When the test quantity $T(\cdot,\cdot)$ does not depend on the model parameters, then it is called a test *statistic*, and it is denoted by $T(\mathbf{y})$. In this case, lack of fit can be assessed by the tail-area probability or p value of the test quantity. Note that the PPC procedure allows one to choose any test quantity to check a particular relevant aspect of the model. In practice, given a discrepancy measure, $T(\mathbf{y},\boldsymbol{\vartheta})$, the posterior predictive check p value is calculated as follows:

1. Draw a vector $\boldsymbol{\vartheta}^{(k)}$ from the posterior distribution.
2. Simulate a replicated data set $\mathbf{y}_{rep}^{(k)}$ from $f(\mathbf{y} \mid \boldsymbol{\vartheta}^{(k)})$.
3. Calculate $T(\mathbf{y},\boldsymbol{\vartheta}^{(k)})$ and $T(\mathbf{y}_{rep}^{(k)},\boldsymbol{\vartheta}^{(k)})$.
4. Repeat Steps 1 to 3 K times.

5. To obtain the posterior p value, count the proportion of replicated data sets $\mathbf{y}_{rep}^{(k)}$ for which $T(\mathbf{y}_{rep}^{(k)}, \boldsymbol{\vartheta}^{(k)}) \geq T(\mathbf{y}, \boldsymbol{\vartheta}^{(k)})$.

Note that when a test statistic $T(\mathbf{y})$ is used, Step 3 does not require a model estimation at each iteration k but only the calculation of $T(\mathbf{y})$ and $T(\mathbf{y}_{rep}^{(k)})$ for each replicated data set. Details of the implementation of PPC in WinBUGS are given in the R/WinBUGS supplemental material online.

Another useful tool for assessing model fit is the graphical display. In the case of discrepancy measures, one can plot $T(\mathbf{y}_{rep}^{(k)}, \boldsymbol{\vartheta}^{(k)})$ and $T(\mathbf{y}, \boldsymbol{\vartheta}^{(k)})$ against each other. In the case of test statistics one can compare $T(\mathbf{y})$ with $T(\mathbf{y}_{rep}^{(1)}), \ldots, T(\mathbf{y}_{rep}^{(k)})$ by localizing $T(\mathbf{y})$ in the frequency distribution of $T(\mathbf{y}_{rep})$. In the Application section of this article we use both a test quantity to check a particular aspect of our model and a global goodness-of-fit discrepancy measure.

Application

Model

Six models were fitted to the data.[1] For convenience of notation, we represent frustration, antagonistic action tendency, irritation, and anger with indices 1, 2, 3, and 4, respectively.

The six models have the same measurement model but differ structurally. In general, the equations can be written as

$$\text{logit}(\pi_{prs}) = \eta_{pr}^{\text{PER}} + \eta_{sr}^{\text{SIT}} + \tau_r \quad \text{Measurement model} \tag{10}$$

and

$$\boldsymbol{\eta}_p^{\text{PER}} = \mathbf{B}^{\text{PER}} \boldsymbol{\eta}_p^{\text{PER}} + \boldsymbol{\zeta}_p^{\text{PER}} \quad \text{Structural model} \tag{11}$$

$$\boldsymbol{\eta}_s^{\text{SIT}} = \mathbf{B}^{\text{SIT}} \boldsymbol{\eta}_s^{\text{SIT}} + \boldsymbol{\zeta}_s^{\text{SIT}},$$

with $p = 1, \ldots, 679$, $r = 1, \ldots, 4$, and $s = 1, \ldots, 11$. Given that the hypothesized latent variables are response specific (four latent variables, one for each emotion response), it seemed reasonable to adopt a strong hypothesis, stating that each latent variable plays a role in only one kind of emotion response. Under this hypothesis, the factor-loading matrices in the measurement model do not contain any free parameters, which is why $\lambda_{rq}^{\text{PER}}$ and $\lambda_{st}^{\text{SIT}}$ from Equation 1 are no longer needed in Equations 10 and 11. For every response, there is a single latent variable on the person side and a separate one on the situation side so that in Equation 1 $\lambda_{rq}^{\text{PER}} = 1$ when $r = q$, and 0 otherwise, and correspondingly, $\lambda_{rt}^{\text{SIT}} = 1$ when $r = t$, and 0 otherwise, with $Q = T = 4$.

The six structural models are most clear when shown by graphical representations. Figure 4 displays the path diagrams for three of the models (the remaining three models are derived from these). Model 1 (with parallel and crossed effects) is the most general of the models; the latent response variables, irritation and anger, are each determined by both of the latent explanatory variables, frustration and antagonistic tendency. Model 2 (with parallel effects) is the extreme form of the hypothesis mentioned earlier, stating that irritation is based only on frustration and that anger results solely from antagonistic action tendency. Model 3 (with only crossed effects) is the opposite of Model 2, with irritation based exclusively on action tendency and anger resulting only from frustration. The corresponding parameter matrices $\mathbf{B}^{\text{PER}}$ and $\mathbf{B}^{\text{SIT}}$ of the three structural models are shown in Table 3. The three remaining models are constructed by restricting the parameter matrices of the person and situation side to be equal ($\mathbf{B}^{\text{PER}} = \mathbf{B}^{\text{SIT}}$). These models are denoted as Models 4, 5, and 6, respectively.

Finally, as stated earlier, it is assumed that $\boldsymbol{\zeta}_p^{\text{PER}} \sim N(\mathbf{0}, \boldsymbol{\Psi}^{\text{PER}})$ and $\boldsymbol{\zeta}_s^{\text{SIT}} \sim N(\mathbf{0}, \boldsymbol{\Psi}^{\text{SIT}})$. Because we are interested in a complete explanation of the third and fourth latent variables, the covariances between ζ_{p3}^{PER}, ζ_{p4}^{PER}, ζ_{s3}^{SIT}, and ζ_{s4}^{SIT} are all set to zero. In addition, the variances of ζ_{p3}^{PER} and ζ_{p4}^{PER}, and of ζ_{s3}^{SIT} and ζ_{s4}^{SIT}, are constrained to be equal ($\psi_{\zeta_3}^{\text{PER}} = \psi_{\zeta_4}^{\text{PER}} = \psi_{\zeta}^{\text{PER}}$, and $\psi_{\zeta_3}^{\text{SIT}} = \psi_{\zeta_4}^{\text{SIT}} = \psi_{\zeta}^{\text{SIT}}$). A model allowing different variances was also estimated, but this did not result in a better fit. Table 4 shows the $\boldsymbol{\Psi}$ matrices for the fitted models. Note that because of the zero correlations among the residuals, the model is not a saturated SEM.

Estimation and Model Selection

A reparametrization of the model was used with hierarchical centering (e.g., W. J. Browne, 2004; Gelfand, Sahu, & Carlin, 1995) in order to avoid poor mixing within chains. Five chains were run starting from different randomly selected initial values for the parameters of interest in each of the fitted models. This approach helps the researcher monitor convergence and choose an appropriate burn-in period. Each chain was run with 5,000 iterations, and the first half of each chain was discarded as a burn-in stage. For the inferences and model checking we accepted every 5th draw of the remaining 2,500 draws of each of the five chains. Using nonconsecutive draws helps reduce autocorrelation and avoid dependence between subsequent draws. Thus, the results that are reported are based on a final sample of 2,500 iterations.

The $\hat{R}$ (Gelman & Rubin, 1992) was used to assess convergence. Values of $\hat{R}$ near 1.0 (say, below 1.1) are considered acceptable (Gelman et al., 2003, pp. 296–297). Also, graphical tools implemented in CODA were used to

[1] The data and the R-WinBUGS code used in this section are available in the supplemental online materials.

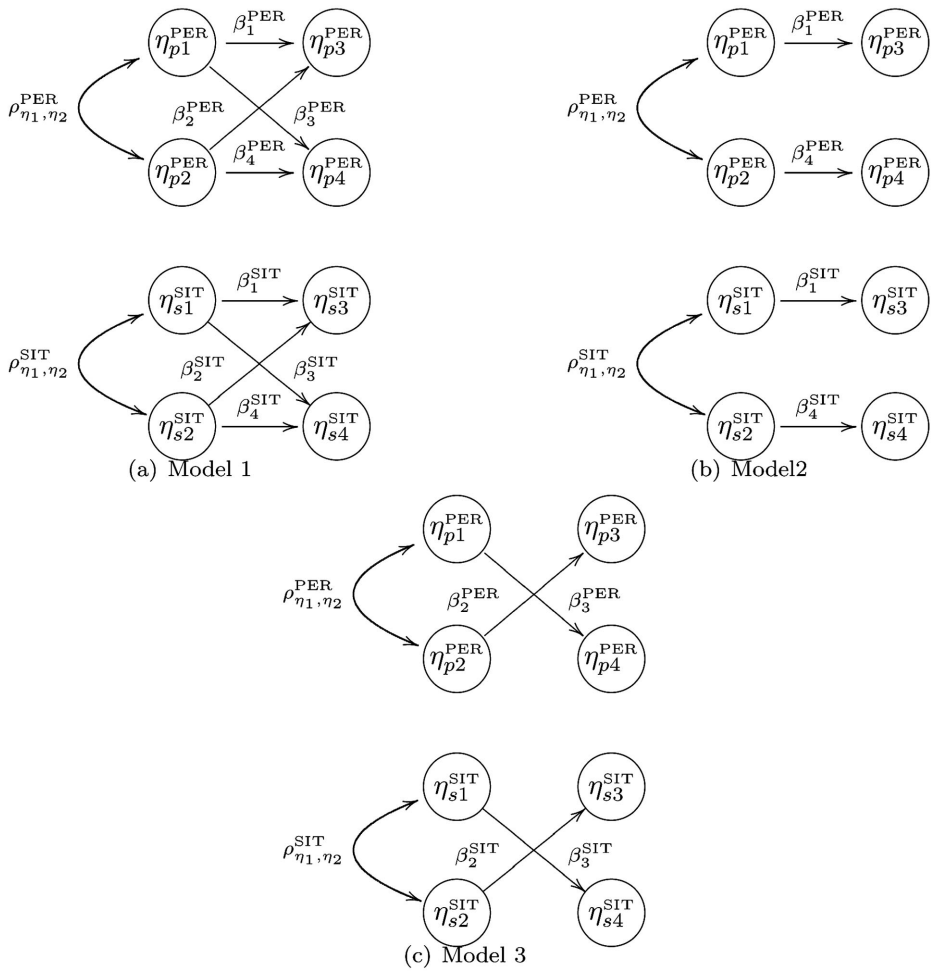

Figure 4. Three different models: (a) Model 1 (parallel and crossed effects), (b) Model 2 (parallel effects), and (c) Model 3 (crossed effects). η = latent variable; β = regression effect; ρ = correlation.

check the mixing and autocorrelation of the chains. The $\hat{R}$ statistic for all of the parameters was less than or equal to 1.03 for all fitted models, so we concluded that convergence had been established.

Table 5 shows the obtained DIC values for the fitted models. From these results, it is clear that Models 1 and 4 are the preferred models. The model with the lowest DIC was Model 4, in which the person and situation structure are restricted to be equal, both qualitatively and quantitatively. It must be noted, however, that between Models 1 and 4, the difference in DIC values seems to be relatively small.

Model Checking

Two types of tests were used to assess model adequacy. The first test was done to evaluate a particular aspect of our model; that is, given that our main goal was to evaluate the correlational structure of the four traits (frustration, antagonistic action tendency, irritation, and anger), the correlations between the sum scores over persons, $Corr(y_{+rs},y_{+r's})$, and over situations, $Corr(y_{pr+},y_{pr'+})$, for each construct ($r = 1, \ldots, 4$) were used as test statistics. The second test, an omnibus goodness-of-fit statistic, is the chi-square discrepancy measure (Gelman et al., 2003). In the context of this application, the Pearson chi-square measure is defined as

$$\chi^2(y,\boldsymbol{\vartheta}) = \sum_{p=1}^{P}\sum_{r=1}^{R}\sum_{s=1}^{S}\frac{(y_{prs} - E(y_{prs}|\boldsymbol{\vartheta}))^2}{Var(y_{prs}|\boldsymbol{\vartheta})}. \qquad (12)$$

The estimated p values for correlations between the y_{+rs} varied between .25 and .52 with a median of .36, and the estimated p values for the correlations over the y_{pr+} varied between .08 and .97 with a median of .84. A reasonable range for the p value is considered to be the interval between .05 and .95 (Gelman et al., 2003); on this basis, we concluded that the model fits the data. The more extreme p

Table 3
Parameter Matrices in the Structural Model for the Three 2sSEM Estimated Models

B matrices	Model 1	Model 2	Model 3
$\mathbf{B}^{\mathrm{PER}}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \beta_1^{\mathrm{PER}} & \beta_2^{\mathrm{PER}} & 0 & 0 \\ \beta_3^{\mathrm{PER}} & \beta_4^{\mathrm{PER}} & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \beta_1^{\mathrm{PER}} & 0 & 0 & 0 \\ 0 & \beta_4^{\mathrm{PER}} & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \beta_2^{\mathrm{PER}} & 0 & 0 \\ \beta_3^{\mathrm{PER}} & 0 & 0 & 0 \end{pmatrix}$
$\mathbf{B}^{\mathrm{SIT}}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \beta_1^{\mathrm{SIT}} & \beta_2^{\mathrm{SIT}} & 0 & 0 \\ \beta_3^{\mathrm{SIT}} & \beta_4^{\mathrm{SIT}} & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \beta_1^{\mathrm{SIT}} & 0 & 0 & 0 \\ 0 & \beta_4^{\mathrm{SIT}} & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \beta_2^{\mathrm{SIT}} & 0 & 0 \\ \beta_3^{\mathrm{SIT}} & 0 & 0 & 0 \end{pmatrix}$

Note. 2sSEM = double-structure structural equation model. βs = regression parameters; PER = person mode; SIT = situation mode.

values obtained when using correlations for y_{pr+} are a consequence of the smaller number of situations considered.

Figure 5 shows a scatter plot of the chi-square test quantity evaluated for observed and replicated data. The posterior p value (.70) reported in the figure is calculated as the proportion of points above the diagonal line (i.e., the proportion of replicated values that are larger than the observed ones). From Figure 5 it can be concluded that the model fits the data reasonably well in a global way as well.

Results

Table 6 shows the posterior means and standard deviations of the parameters in Model 4, and Figure 6 contains the graphical representation of the structural part of Model 4, including the parameter estimates. In latent person and latent situation structures, frustration and the antagonistic action tendency affected both irritation and anger. As expected, the feeling of frustration was more important for irritation than the antagonistic tendency, whereas the opposite was true in the case of anger. In the case of irritation, the difference was very small (.49 vs. .46), but for anger it was much larger (.54 vs. .30). The correlation between the explanatory latent person variables was moderately high and positive (.68), whereas the correlation between the explanatory latent situation variables was much lower but still positive (.25). The posterior standard deviation for correlation on the situation side was larger than the one on the person side due to the smaller sample of situations. The difference in correlations can be explained by the difference in variance; the situations seem much more homogeneous than the persons. This difference in correlations is also in agreement with earlier findings (Vansteelandt et al., 2005; Zelenski & Larsen, 2000).

The τ parameters must be interpreted as indicating the probability of showing each of the four responses when all latent variables are set to zero (on the logistic scale). For instance, for an average person in an average situation, the probabilities of endorsing responses 1 and 2 (see Table 6) equal

$$\pi_{p1s} = \frac{\exp(1.56)}{1 + \exp(1.56)} = 0.83, \quad \text{and}$$

$$\pi_{p2s} = \frac{\exp(1.51)}{1 + \exp(1.51)} = 0.82, \tag{13}$$

respectively. The two responses are approximately equally likely.

The primary hypotheses of the study were confirmed by the analyses. The questions mentioned in the second section of this article can now be answered in the following way: Regarding Question 1, on the basis of the $\mathbf{B}^{\mathrm{PER}}$ matrix, we conclude that irritation seems to be slightly more feeling based but the difference is negligible, whereas anger is

Table 4
Covariance Matrices in the Structural Model for the Three 2sSEM Estimated Models

Persons	Situations
$\mathbf{\Psi}^{\mathrm{PER}} = \begin{pmatrix} \Psi_{\zeta_1}^{\mathrm{PER}} & \Psi_{\zeta_1,\zeta_2}^{\mathrm{PER}} & 0 & 0 \\ \Psi_{\zeta_2,\zeta_1}^{\mathrm{PER}} & \Psi_{\zeta_2}^{\mathrm{PER}} & 0 & 0 \\ 0 & 0 & \Psi_{\zeta_3}^{\mathrm{PER}} & 0 \\ 0 & 0 & 0 & \Psi_{\zeta_4}^{\mathrm{PER}} \end{pmatrix}$	$\mathbf{\Psi}^{\mathrm{SIT}} = \begin{pmatrix} \Psi_{\zeta_1}^{\mathrm{SIT}} & \Psi_{\zeta_1,\zeta_2}^{\mathrm{SIT}} & 0 & 0 \\ \Psi_{\zeta_2,\zeta_1}^{\mathrm{SIT}} & \Psi_{\zeta_1}^{\mathrm{SIT}} & 0 & 0 \\ 0 & 0 & \Psi_{\zeta_3}^{\mathrm{SIT}} & 0 \\ 0 & 0 & 0 & \psi_{\zeta_4}^{\mathrm{SIT}} \end{pmatrix}$

Note. 2sSEM = double-structure structural equation model. Ψs = covariance parameters; PER = person mode; SIT = situation mode.

Table 5
Deviance Information Criterion (DIC) and p_D Values for the Six Fitted Models

Model	DIC	p_D
1	25,211.80	1,456.53
2	25,335.80	1,440.09
3	25,424.60	1,454.99
4	25,209.50	1,451.83
5	25,337.80	1,444.47
6	25,423.10	1,457.20

Note. Lower values for DIC are preferable. p_D = estimate of the effective number of parameters.

clearly more action based. Regarding Question 2, which deals with situational structure, in this particular case, our conclusions match those stated for the person side, because the best-fitted model was the one in which $\mathbf{B}^{\text{PER}}$ and $\mathbf{B}^{\text{SIT}}$ matrices were assumed to be equal. Regarding Question 3, our conclusion is that the equality of the structures' individual and situational differences was not contradicted by the data.

Discussion and Conclusion

An extension of the regular SEM, called a double-structure structural equation model (2sSEM), was presented and shown to be useful for studying underlying structures in two modes of a three-mode data set. A study of emotions, using a persons by situations by responses three-mode data set, was used to demonstrate how the model can be utilized to assess individual differences and situational differences simultaneously, without the need to use two separate, unrelated models. Using a common model for both structures can be advantageous in the sense that no information is lost.

Unlike other models for three-mode data, the 2sSEM treats two of the three modes as random. The 2sSEM may be also considered as a variant of a double-structure confirmatory factor analysis model (2sCFA); however, in the CFA model, correlations among latent variables can be assessed, but latent variables are never regressed on other latent variables. The 2sSEM adds a feature to the CFA model by imposing restrictions on how latent variables may affect each other. In the present application, there is only one latent (person, situation) variable for each response mode, and, as a result, the 2sSEM looks like a path model. Each element of the response mode (anger, irritation, etc.) is in fact associated with a Person × Situation matrix rather than just the single observed variable. In other applications, fewer latent variables may be needed, in which case the results would look more like the standard SEM.

A special feature of the data we considered is that they are binary. Thus, the model developed here is an extension of a nonlinear mixed model for binary data. However, the model can easily be adapted for continuous, ordered-category (rating-scale), or count data. In fact, the explanation of the model was fairly general and oriented toward continuous data. The use of binary data for our example illustrates the generality of the model. In addition, as a consequence of the way the 2sSEM is estimated here (via Bayesian methodology), missing data can be easily accommodated, as long as the missingness mechanism is ignorable (see Little & Rubin, 2002).

In this article, the focus is on a data set of a particular kind: a fully crossed three-dimensional data array with binary data on persons, situations, and emotional responses. An important asset of such data is that a combination of two approaches is possible, one based on individual differences and another based on situation differences. Often the study of emotions is either based on the structure of individual differences and correlations relating to these differences, or it is based on experimental methods, which only measure effects of situational differences. With three-mode data and the 2sSEM, it is possible to combine the two perspectives into a single analysis of individual differences and differences based on circumstances, situations, conditions, and so forth.

Note that the number of situations is rather low in the application. As a consequence, the standard errors of the parameter estimates for the situation structure are larger than for the person structure when the scale is comparable, as can be seen in the large values of the standard error of the correlation between ζ_1 and ζ_2 (see Table 6). However, the low number of 11 is still sufficiently high for the regression

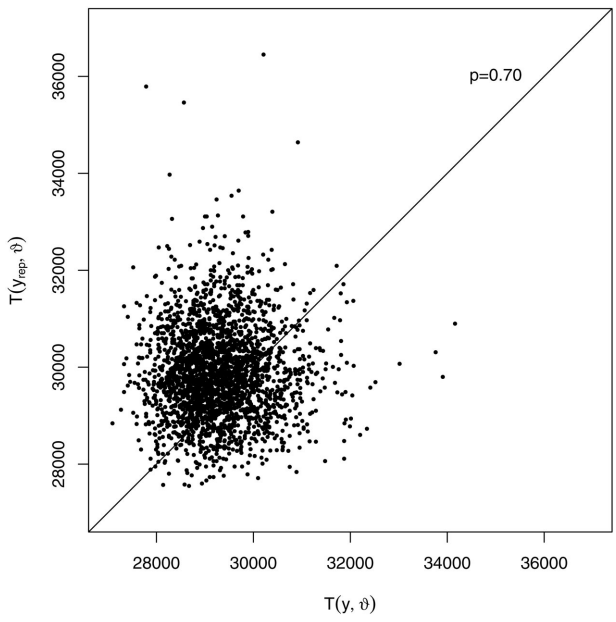

Figure 5. Chi-square values for the observed versus the replicated data based on 2,500 simulations from the posterior distribution of $(\mathbf{y}_{rep}, \vartheta)$.

Table 6
Results for Model 4

Parameter	Value	*SD*
Regression parameters in the **B** matrices		
$\beta_1^{PER} = \beta_1^{SIT}$	0.49	0.04
$\beta_2^{PER} = \beta_2^{SIT}$	0.46	0.05
$\beta_3^{PER} = \beta_3^{SIT}$	0.30	0.04
$\beta_4^{PER} = \beta_4^{SIT}$	0.54	0.05
Response effect parameters in τ		
τ_1	1.56	0.26
τ_2	1.51	0.27
τ_3	0.46	0.11
τ_4	0.34	0.12
Covariance parameters in the Ψ matrices		
$\Psi_{\zeta_1}^{PER}$	3.93	0.34
$\Psi_{\zeta_1,\zeta_2}^{PER}$	2.19	0.20
$\Psi_{\zeta_2}^{PER}$	2.70	0.24
$\Psi_{\zeta_1}^{SIT}$	0.66	0.36
$\Psi_{\zeta_1,\zeta_2}^{SIT}$	0.17	0.27
$\Psi_{\zeta_2}^{SIT}$	0.71	0.42
$\rho_{\zeta_1,\zeta_2}^{PER}$	0.68	0.03
$\rho_{\zeta_1,\zeta_2}^{SIT}$	0.25	0.28
$\Psi_{\zeta_3}^{PER} = \Psi_{\zeta_4}^{PER}$	0.51	0.06
$\Psi_{\zeta_3}^{SIT} = \Psi_{\zeta_4}^{SIT}$	0.10	0.04

parameters in the situation structure to be significantly different from zero in the case that no equality constraints are imposed. Model 1 is a model without such constraints; Model 4 is the equivalent with equality constraints. The number of situations will always be rather small, although perhaps larger than in the application, because this number is a multiplicative factor for the number of items.

Latent variable modeling in psychology is exclusively linked with individual differences. Borsboom, Mellenbergh, and Van Heerden (2002) attributed this "to a century of operating on silent uniformity-of-nature assumptions by focusing almost exclusively on between-subjects models" (p. 215), by which they meant that the latent variable model based on individual differences is also assumed to apply within persons. For example, in SEM approaches, often causal, between-persons effects are modeled and then generalized to within-persons changes. Looking within persons can mean looking at moments in time, as Borsboom et al. suggested, or across situations, as in the present example. If the relations of latent person variables also apply to latent situation variables, then we have tested the validity of generalizing from between persons to within persons. In other words, if the 2sSEM fits the data, and the structure of the person side is the same as the structure of the situation side, as was the case in the analyses presented here, then this type of generalization is supported. There are two other possibilities: First, the structure could be different for persons versus situations. The results of the 2sSEM can show whether this is the case, and if it is, the model would still be valid. Second, the latent situation variables might differ across persons. In this case, the 2sSEM approach would not be appropriate, because a basic assumption of the model is that the persons are all alike with respect to the situational latent variables. However, applying the 2sSEM and testing its goodness of fit provides a way to assess the validity of this assumption.

Other kinds of three-mode data can be imagined. For instance, suppose that a sample of persons has been asked to rate another sample of persons not on just one personality trait (as in one of the cases of Shrout & Fleiss, 1979), but on a set of trait terms. The items can be thought of as indicators of latent variables. The structure of individual differences between the raters is not necessarily equal to the structure of the individual differences between the persons who were rated, even though both structures are based on the same personality items. One could even argue that the equality of the two structures is suspicious, because that may indicate that the former is a mere reflection of beliefs from the part of the raters. A more challenging data set is one in which the first mode is a random set of persons; the second mode, a set of responses; and the third mode, points in time. In this case, an assumption of a random time mode requires some thought. For a time mode to be random, it must be assumed that the residuals at the different time points are independent for a given response but can be correlated between responses (for example, a larger residual for Response 1 at

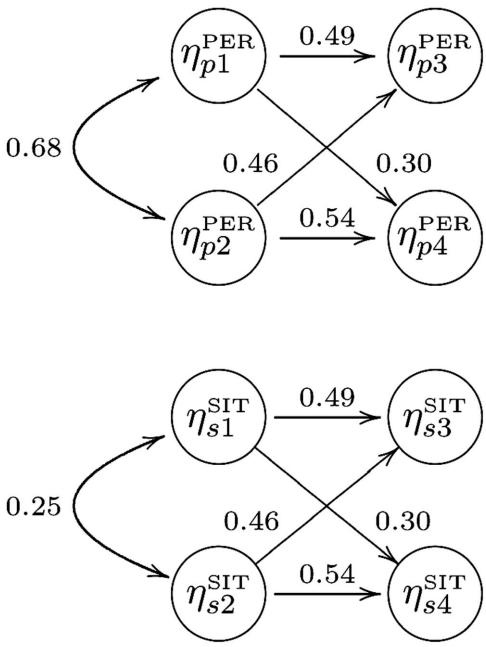

Figure 6. Results for Model 4 (parallel and crossed effects, equal for person and situation latent variables). η = latent variable.

time point *t* is more likely to co-occur with a larger residual for Response 2 at the same time point). Additionally, one may assume more refined structural relations between these within-time-point residuals.

The 2sSEM is a useful model for combining and comparing the structures of two modes in terms of a third mode within a data set concerning, for example, emotions, as in the application shown here, or personality, as in the application about which we speculated. However, a 2sSEM is not appropriate for all three-mode data. For example, only two types of interactions may be accounted for, as shown in Table 1. In our application, these were persons by responses and situations by responses. The roles of the three kinds of entities may be changed, but if these two pairwise interactions are modeled, then it is not possible to include the third interaction—that of persons and situations. In other words, for this model to be useful, the introduction of two types of latent variables, one for each of two of the three modes, must be meaningful.

References

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). Hoboken, NJ: Wiley.

Akaike, M. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716–723.

Andrich, D. (1978). A rating scale formulation for ordered response categories. *Psychometrika, 43,* 561–573.

Averill, J. (1982). *Anger and aggression: An essay on emotion.* New York: Springer.

Bentler, P., & Lee, S.-Y. (1978). Statistical aspects of a three-mode factor analysis model. *Psychometrika, 43,* 343–352.

Bentler, P., & Lee, S.-Y. (1979). A statistical development of three-mode factor analysis. *British Journal of Mathematical and Statistical Psychology, 32,* 87–104.

Bloxom, B. (1968). A note on invariance in three-mode factor analysis. *Psychometrika, 33,* 347–350.

Bollen, K. (1989). *Structural equations with latent variables.* New York: Wiley.

Borsboom, D., Mellenbergh, G., & Van Heerden, J. (2002). The theoretical status of latent variables. *Psychological Review, 110,* 203–219.

Brennan, R. L. (2001). *Generalizability theory.* New York: Springer-Verlag.

Browne, M. W. (1984). The decomposition of multitrait–multimethod matrices. *British Journal of Mathematical and Statistical Psychology, 37,* 1–21.

Browne, W. J. (2004). An illustration of the use of reparameterisation methods for improving MCMC efficiency in crossed random effects models. *Multilevel Modelling Newsletter, 16,* 13–25.

Carroll, J., & Chang, J. (1970). Analysis of individual differences in multidimensional scaling via an *N*-way generalization of Eckart–Young decomposition. *Psychometrika, 35,* 283–319.

Cattell, R. (1946). *Description and measurement of personality.* New York: World Book Cie.

Cattell, R. (1952). *Factor analysis: An introduction and manual for the psychologist and social scientist.* New York: Harper.

Cowles, M., & Carlin, B. (1996). Markov Chain Monte Carlo convergence diagnostics: A comparative study. *Journal of the American Statistical Association, 91,* 883–904.

Cronbach, L. (1957). The two disciplines of scientific psychology. *American Psychologist, 12,* 671–684.

Cronbach, L., Rajaratnam, N., & Gleser, G. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Mathematical and Statistical Psychology, 16,* l37–163.

De Boeck, P., & Smits, D. (2006). A double-structure structural equation model for the study of emotions and their components. In Q. Jing et al. (Eds.), *Progress in psychological science around the world: Vol. 1. Neural, cognitive and developmental issues* (pp. 349–365). Hove, United Kingdom: Psychology Press.

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

Eid, M., Lischetzke, T., & Nussbeck, F. (2006). Structural equation models for multitrait–multimethod data. In M. Eid & E. Diener (Eds.), *Handbook of psychological measurement: A multimethod perspective* (pp. 283–299). Washington, DC: American Psychological Association.

Eid, M., Lischetzke, T., Nussbeck, F., & Trierweiler, L. (2003). Separating trait effects from trait-specific method effects in multitrait–multimethod analysis: A multiple indicator CTC(M-1) model. *Psychological Methods, 8,* 38–60.

Fienberg, S. (1980). *The analysis of cross-classified, categorical data* (2nd ed.). Cambridge, MA: MIT Press.

Frijda, N., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology, 65,* 942–958.

Gelfand, A., Sahu, S., & Carlin, B. (1995). Efficient parameterizations for normal linear mixed models. *Biometrika, 82,* 479–488.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2003). *Bayesian data analysis* (2nd ed.). New York: Chapman and Hall.

Gelman, A., & Rubin, D. (1992). Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science, 7,* 457–511.

González, J., Tuerlinckx, F., De Boeck, P., & Cools, R. (2006). Numerical integration in logistic-normal models. *Computational Statistics and Data Analysis, 51,* 1535–1548.

Harshman, R. (1970). Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-mode factor analysis. *UCLA Working Papers in Phonetics, 16,* 1–84.

Janssen, R., Schepers, J., & Peres, D. (2004). Models with item and item group predictors. In P. De Boeck & M. Wilson (Eds.), *Explanatory item response models: A generalized linear and nonlinear approach* (pp. 189–212). New York: Springer.

Janssen, R., Tuerlinckx, F., Meulders, M., & De Boeck, P. (2000). A hierarchical IRT model for criterion-referenced measurement. *Journal of Educational and Behavioral Statistics, 25,* 285–306.

Kass, R., & Raftery, A. (1995). Bayes factors. *Journal of the American Statistical Association, 90,* 773–796.

Kroonenberg, P. (1983). Annotated bibliography of three-mode factor analysis. *British Journal of Mathematical and Statistical Psychology, 36,* 81–113.

Kroonenberg, P. (2005). Three-mode component and scaling methods. In B. Everitt & D. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (Vol. 4, pp. 2032–2044). New York: Wiley.

Kuppens, P., Van Mechelen, I., Smits, D., De Boeck, P., & Ceulemans, E. (2007). Individual differences in patterns of appraisal and anger experience. *Cognition & Emotion, 21,* 689–713.

Lee, S.-Y. (2007). *Structural equation modeling: A Bayesian approach.* Chichester, United Kingdom: Wiley.

Lee, S.-Y., & Fong, W.-K. (1983). A scale invariant model for three-mode factor analysis. *British Journal of Mathematical and Statistical Psychology, 36,* 217–223.

Lindley, D. V., & Novick, M. R. (1981). The role of exchangeability in inference. *Annals of Statistics, 9,* 45–58.

Little, R., & Rubin, D. (2002). *Statistical analysis with missing data* (2nd ed.). Chichester, United Kingdom: Wiley.

McCullagh, P., & Nelder, J. (1989). *Generalized linear models.* New York: Chapman & Hall.

Oort, F. (1999). Stochastic three-mode models for mean and covariance structures. *British Journal of Mathematical and Statistical Psychology, 52,* 243–272.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006, March). CODA: Convergence diagnosis and output analysis for MCMC. *R News, 6*(1), 7–11. Available from http://cran.r-project.org/doc/Rnews/Rnews_2006–1.pdf

R Development Core Team. (2006). *R: A language and environment for statistical computing* [Computer software manual]. Available from http://www.R-project.org

Scherer, K., Schorr, A., & Johnstone, T. (2001). *Appraisal processes in emotion: Theory, methods, research.* New York: Oxford University Press.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6,* 461–464.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 2,* 420–428.

Shun, Z. (1997). Another look at the salamander mating data: A modified Laplace approximation approach. *Journal of the American Statistical Association, 92,* 341–349.

Shun, Z., & McCullagh, P. (1995). Laplace approximation of high dimensional integrals. *Journal of the Royal Statistical Society, 57*(B), 749–760.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equations models.* Boca Raton, FL: Chapman and Hall/CRC.

Snijders, T. (2005). Fixed and random effects. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (Vol. 2, pp. 664–665). Chichester, United Kingdom: Wiley.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* London: Sage.

Spiegelhalter, D., Best, N., Carlin, B., & Van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, 64*(B), 583–639.

Spiegelhalter, D., Thomas, A., Best, N., & Lunn, D. (2003). *WinBUGS Version 1.4 user manual* [Computer software manual]. Retrieved from http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual14.pdf

Sturtz, S., Ligges, U., & Gelman, A. (2005). R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software, 12*(3), 1–16. Retrieved from http://www.jstatsoft.org/v12/i03/paper

Tucker, L. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika, 31,* 279–311.

Tuerlinckx, F., Rijmen, F., Verbeke, G., & De Boeck, P. (2006). Statistical inference in generalized linear mixed models: A review. *British Journal of Mathematical and Statistical Psychology, 59,* 225–255.

Van Coillie, H., & Van Mechelen, I. (2006). A taxonomy of anger-related behaviors in young adults. *Motivation and Emotion, 30,* 57–74.

Vansteelandt, K., Van Mechelen, I., & Nezlek, J. (2005). The co-occurrence of emotions in daily life: A multilevel approach. *Journal of Research in Personality, 39,* 325–335.

Zelenski, J., & Larsen, R. (2000). The distribution of basic emotions in everyday life: A state and trait perspective from experience sampling data. *Journal of Research in Personality, 34,* 178–197.

Appendix A

Latent Threshold Formulation of the Model for Analyzing Binary Data

The measurement part of the 2sSEM can handle both continuous and binary observed variables. To illustrate this, we relate the observed y_{prs} variables to a continuous latent response variable y^*_{prs}.

Consider the following linear regression model for continuous response variable y^*_{prs}:

$$y^*_{prs} = \nu_{prs} + \varepsilon_{prs}, \tag{14}$$

where $\nu_{prs} = \boldsymbol{\lambda}_r^{\mathrm{PER}\,\intercal}\boldsymbol{\eta}_p^{\mathrm{PER}} + \boldsymbol{\lambda}_r^{\mathrm{SIT}\,\intercal}\boldsymbol{\eta}_s^{\mathrm{SIT}} + \tau_r$ is the bilinear predictor used for the 2sSEM, and ε_{prs} is an error term of mean zero. This model can alternatively be defined as

$$E(y^*_{prs}|\nu_{prs}) = \nu_{prs}. \tag{15}$$

In the case of continuous responses, the latent response equals the observed response: $y_{prs} = y^*_{prs}$.

In the case of binary responses, the response variable is defined by the following threshold model:

$$y_{prs} = \begin{cases} 1, & \text{if } y^*_{prs} > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

It follows that

$$\begin{aligned} E(y_{prs}|\nu_{prs}) &= \Pr(y_{prs} = 1|\nu_{prs}) = \pi_{prs} \\ &= \Pr(y^*_{prs} > 0|\nu_{prs}) \\ &= \Pr(\nu_{prs} + \varepsilon_{prs} > 0) \\ &= \Pr(\varepsilon_{prs} > -\nu_{prs}) \\ &= F(\nu_{prs}), \end{aligned} \tag{17}$$

where the last equality is due to the symmetry of the distribution function F of ε_{prs}. Typically, F is chosen to be either the cumulative standard normal or logistic distribution. In this article, the logistic distribution was assumed such that

$$\begin{aligned} F^{-1}(\pi_{prs}) &= \log\left(\frac{\pi_{prs}}{1-\pi_{prs}}\right) = \mathrm{logit}(\pi_{prs}) = \nu_{prs} \\ &= \boldsymbol{\lambda}_r^{\mathrm{PER}\,\intercal}\boldsymbol{\eta}_p^{\mathrm{PER}} + \boldsymbol{\lambda}_r^{\mathrm{SIT}\,\intercal}\boldsymbol{\eta}_s^{\mathrm{SIT}} + \tau_r, \end{aligned} \tag{18}$$

where $F^{-1}(x) = \log\left(\frac{x}{1-x}\right) = \mathrm{logit}(x)$ is the link function.

Appendix B

Derivation of the Posterior Distribution

Assuming conditional independence, the likelihood function for binary data is written as

$$f(\mathbf{y}|\boldsymbol{\vartheta}) = \left(\prod_{\mathrm{p}=1}^{\mathrm{P}}\prod_{\mathrm{r}=1}^{\mathrm{R}}\prod_{\mathrm{s}=1}^{\mathrm{S}}\mathrm{F}(\nu_{\mathrm{prs}})^{y_{\mathrm{prs}}}[1-\mathrm{F}(\nu_{\mathrm{prs}})]^{1-y_{\mathrm{prs}}}\right), \tag{19}$$

where $F(\nu_{prs}) = \pi_{prs}$ (see Appendix A).

In order to have a sample of $\boldsymbol{\vartheta}$ we need to specify prior distributions for all the involved parameters. For the parameters in $\mathbf{B}^{\mathrm{PER}}$, $\mathbf{B}^{\mathrm{SIT}}$, and $\boldsymbol{\tau}$, normal diffuse priors with mean zero and variance 10^4 were chosen. For the upper left block of the covariance matrices $\boldsymbol{\Psi}^{\mathrm{PER}}$ and $\boldsymbol{\Psi}^{\mathrm{SIT}}$ (i.e., the variances and covariance of the latent explanatory variables), inverse-Wishart prior distributions with scale matrix $10^3 I$ and 2 degrees of freedom were chosen. For the variances of the latent residuals (i.e., the diagonal elements of the lower right block of the $\boldsymbol{\Psi}$ matrices), inverse-gamma distributions with both shape and scale parameters equal to 10^4 were chosen. Finally, the distribution for the latent person and situation residuals are given by the model assumptions and are multivariate normal distributions with zero vector mean and covariance matrices $\boldsymbol{\Psi}^{\mathrm{PER}}$ and $\boldsymbol{\Psi}^{\mathrm{SIT}}$, respectively.

Using the relation in Equation 8, the posterior distribution is then given by the product of the likelihood in Equation 19 and

$$\begin{aligned} f(\boldsymbol{\vartheta}) &= \prod_{p=1}^{P} N(\boldsymbol{\zeta}_p^{\mathrm{PER}};\mathbf{0},\boldsymbol{\Psi}^{\mathrm{PER}})\prod_{s=1}^{S} N(\boldsymbol{\zeta}_s^{\mathrm{SIT}};\mathbf{0},\boldsymbol{\Psi}^{\mathrm{SIT}}) \\ &\quad \times p(\mathbf{B}^{\mathrm{PER}})p(\mathbf{B}^{\mathrm{SIT}})p(\boldsymbol{\tau}) \times p(\boldsymbol{\Psi}^{\mathrm{PER}})p(\boldsymbol{\Psi}^{\mathrm{SIT}}), \end{aligned} \tag{20}$$

where $N(Z;\mu,\boldsymbol{\Sigma})$ means that the random vector $\mathbf{Z}$ is distributed as a multivariate normal distribution with vector mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Note that independence of the priors is assumed.

Received December 5, 2006
Revision received July 7, 2008
Accepted July 7, 2008 ■