British Journal of Mathematical and Statistical Psychology (2005), 58, 117–143
© 2005 The British Psychological Society



The British Psychological Society

www.bpsjournals.co.uk

Latent variable models for partially ordered responses and trajectory analysis of anger-related feelings

Michel Meulders[1]*, Edward H. Ip[2] and Paul De Boeck[1]

[1]Department of Psychology, University of Leuven, Belgium

[2]Biostatistics and Social Sciences and Health Policy, Wake Forest University School of Medicine, USA

A general framework is presented for the analysis of partially ordered set (poset) data. The work is motivated by the need to analyse poset data such as multi-componential responses in psychological measurement and partially accomplished cognitive tasks in educational measurement. It is shown how the generalized loglinear model can be used to represent poset data that form a lattice and how latent-variable models can be constructed by further specifying the canonical parameters of the loglinear representation. The approach generalizes a class of latent-variable models for completely ordered data. We apply the methods to analyse data on the frequency and intensity of anger-related feelings. Furthermore, we propose a trajectory analysis to gain insight into the response function of partially ordered emotional states.

1. Introduction

Partially ordered sets (posets) arise in theories of learning, cognition and developmental psychology and have a rather long history (e.g. Flavell, 1971; Piaget, 1950). Recently, with the increasing emphasis on educational tests as tools for cognitive diagnosis rather than for simply ranking students (Mislevy, 1996), and with increased attention being given to the delineation of granular components of complex psychological concepts (as in facet analysis: Costa & McCrae, 1995), to appraisal and action tendency analysis of emotion (Ortony & Turner, 1990), and to componential modelling of learning tasks (Hoskens & De Boeck, 2001), poset responses are becoming commonplace.

In educational measurement, the analysis of partially correct responses to test items could help to reveal information about a student's cognitive states in learning. This is an important issue as pedagogic interest in the cognitive states underlying student responses has recently increased, causing a growing demand for tests that can be used as

*Correspondence should be addressed to Michel Meulders, Department of Psychology, Tiensestraat 102, B-3000 Leuven, Belgium (e-mail: michel.meulders@psy.kuleuven.ac.be).

DOI:10.1348/000711005X38555

diagnostic tools for improving instruction (Mislevy, 1996). Consider, for instance, the following examples.

Different problem-solving strategies may be considered partially ordered because some strategies lead to superior performance whereas other strategies are equally successful. In particular, Siegler (1987) conducted a study in which students had to describe how they solved elementary addition problems (see also Wilson, 1992). The answers were classified into five categories (the following labels are for identifying the category in a Hasse diagram): (f) retrieval, in which the answer is retrieved from memory; (e) the min strategy, in which one counts up from the larger addend the number of times indicated by the smaller addend; (d) decomposition, in which the original problems are decomposed into several simpler problems; (b) the counting-all strategy, in which one counts from one the number of times indicated by the sum; and (a) guessing, in which one guesses or does not know the answer. Based on criteria of speed and accuracy, the strategies can be partially ordered as shown in Fig. 1a. For purposes of illustration, we add a sixth hypothetical strategy (c), which is more successful than (a) and less successful than (f) but cannot be ordered with respect to (d), (b), or (e) because it is qualitatively different.

As a second example, partially ordered responses may arise when scoring responses to open-ended questions on different aspects. For instance, the answers to a mathematical problem-solving task may be judged with respect to various features related to the correctness of the reasoning behind the answer and the correctness of computations that have to be carried out. The responses can therefore be scored in several respects, so that for each response a binary vector is obtained. These response vectors would in general not form a full order, but only a partial order. In the same way, open-ended responses to items that involve the subtraction of fractions could be scored with respect to the cognitive skills that are required to obtain a particular response such as transforming a whole number into a fraction, finding the common denominator of two fractions, and putting a fraction in its simplest form (see Tatsuoka, 2002). Again the response patterns formed by the binary scores on the different skills can be considered partially ordered.

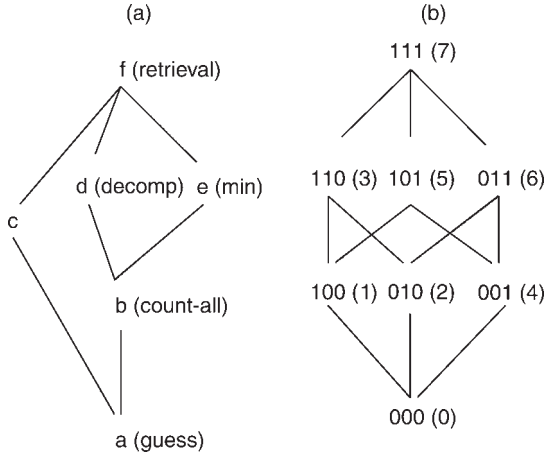


Figure 1. Hasse diagram for (a) partial order of strategies in solving elementary addition problems and (b) partially ordered responses based on a multi-component structure with three binary components.

As a third example, consider the componential delineation of the feelings of individuals in personality research. In a study of guilt feelings (Smits & De Boeck, 2003), it was hypothesized that three identifiable components form the main structure of a feeling of guilt – norm violation (appraisal of what one did as being a transgression of a moral norm), brooding (cognitive tendency to continually reflect upon the act), and tendency to repair (action tendency to undo the fault). Multiple situations under which feelings of guilt could be evoked were presented to the participants. An example of a situation of guilt is that one promises a friend to keep a secret but breaks the promise by telling the secret to someone else. Responses to a family of items that stem from each situation were collected. Each family included three items, each of which corresponded to a guilt component. To illustrate, suppose that the presence and absence of a guilt component are coded as 1 and 0, respectively. Accordingly, the possible responses are encoded into the row vector (000,100,010,110,001,101,011,111). Each combination of presence or absence is mapped to a category of response. Figure 1b shows the Hasse diagram of the order structure of the combinations. Accordingly, each situation solicits a poset with a maximal element (category 7) and a minimal element (category 0). The interest of inference is in the characteristics of each response category as well as in the interaction or dependency between components. Finally, Tatsuoka (2002) and Wilson (1992) also describe some other contexts in which the modelling of posets may be useful.

This paper proposes methods for analysing a general class of posets. The general loglinear methods that we use to analyse posets have been used in a previous paper by Ip, Wang, De Boeck, and Meulders (2004) for the analysis of multivariate polytomous responses, which can be represented as a Boolean lattice (e.g. lattice represented by Fig. 1b). The present paper shows that the general loglinear approach can be extended to model any form of poset as long as it can be represented as a (possibly non-Boolean) lattice. In this sense, this paper generalizes the methods of Ip *et al.* (2004) to the analysis of poset responses. Furthermore, the present paper describes the theoretical relation between the hypothesized partial order model and the observed trajectory of dominant responses along the latent scale. In addition, the paper proposes techniques for visualizing and for estimating the stability of the empirically observed trajectory in the context of a multi-componential latent variable analysis on anger-related feelings, with intensity and frequency as the components. The study of the trajectory allows researchers and practitioners to gain insight into the relationship between the partially ordered states and the latent construct that supposedly drives the responses.

The remainder of the paper is organized as follows. Section 2 highlights special statistical features in this analysis. Section 3 provides the representation of a general class of partially ordered responses using a generalized loglinear model for the case in which the posets follow a lattice structure. Section 4 describes how various latent variable models can be constructed by further specifying the canonical parameters of the loglinear representation. Section 5 deals with the derivation of tracelines for the proposed model. Section 6 describes parameter estimation for the models presented. Section 7 describes methods for model selection and model checking. Section 8 presents the results of two studies on the frequency and intensity of anger-related feelings. Finally, Section 9 contains a brief discussion.

2. Statistical features

The study described in this paper involves several special statistical features. First, the method for representing posets is closely related to that of multivariate categorical

variables. Our representation of posets can be seen as a generalization of a computationally efficient form of the loglinear model (Zhao & Prentice, 1990). We shall refer to the special representation of the loglinear model (Bishop, Fienberg, & Holland, 1975) as the generalized loglinear model (GLLM: Cox, 1972; Holland, 1990; Laird, 1991).

A second feature of this analysis is that it is presented in an item response theory (IRT) framework (Lord & Novick, 1968). Item response models can be considered as latent variable models that posit an individual-specific underlying trait that drives a respondent's multiple responses to a set of stimuli (often in the form of questions or items). The latent variable is modelled as a random variate on a continuous, ordered scale – or a random effect – which is ultimately integrated out to provide the manifest probability of the responses. While there exist well-developed IRT models in the literature for analysing responses that are either totally ordered or totally unordered, to the best of our knowledge measurement tools for partially ordered responses have not been fully developed. Except for a few isolated efforts (e.g. Tatsuoka, 2002; Wilson, 1992), the general lack of advanced techniques for structures other than ordered or nominal categories has been criticized as having a negative impact on the development of theory for psychology and educational measurement (Glaser, Lesgold, & Lajoie 1987).

Wilson (1992) developed the ordered partition model to analyse equivalent classes of nominal categories that are strictly ordered. The method presented in this paper generalizes Wilson's model in that it can be used to analyse any structure of partial order that can be represented as a lattice. Furthermore, our method can also be considered as an extension to several important IRT approaches that are used to analyse polytomous item responses. These IRT approaches include the nominal category response model (Bock, 1972), the graded response model (Samejima, 1969), and the partial credit model (Masters, 1982). The nominal category response model elaborates a formal model for choice between two alternatives (Luce, 1959) and is closely related to the multinomial-logit model (McFadden, 1974). The graded response and partial credit models, on the other hand, are commonly used to analyse polytomous items with a full order of response categories. The partial credit model is an adjacent-categories logit model (e.g. Agresti, 2002), while the graded response model is a cumulative-logit model (McCullagh, 1980), also called a proportional-odds model. A covariate term that is used in a regression setting in the adjacent- and cumulative-logit models is replaced by a latent variable in both the partial credit and graded response models. Statistically, the nominal, graded response and partial credit approaches can all be treated as random effect models in, respectively, multinomial-, adjacent- and cumulative-logit regressions. However, for posets, it is not clear how adjacency and cumulative sum should be defined and analysed because mathematically not all response categories are necessarily ordered. That is, any one response category may neither dominate nor be dominated by another category. In this paper, we use incidence algebra to provide a solution to this problem.

3. Matrix representation of posets

In the following paragraphs, we will first provide a general discussion of posets that can be represented as a lattice. These posets are further denoted as *general* posets. Next, we will discuss, as a special case, the formal representation of *multi-component* posets that are derived from responses to multiple component-items each of which has totally ordered categories.

3.1 General poset structure

Consider a finite set $G = \{g_1, \ldots, g_M\}$ and a relation $\prec$ on G such that g_n is said to dominate g_m if $g_m \prec g_n$ for $g_n, g_m \in G$. The relationship $\prec$ is reflexive and transitive but not symmetric. Assume that G constitutes a finite lattice that includes a maximal element or supremum $\hat{1}$ (i.e. $\forall g_m \in G: g_m \prec \hat{1}$) and a minimal element or infimum $\hat{0}$ (i.e. $\forall g_m \in G: g_m \succ \hat{0}$). For this lattice, we define a *dominance* matrix $A = (a_{mn})$ so that $a_{mn} = 1$ if $g_n \prec g_m$ and 0 otherwise. For the poset (a, b, c, d, e, f) of which the lattice structure is shown in Fig. 1a, $a = \hat{0}$ and $f = \hat{1}$. The corresponding dominance matrix A is given by:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

To represent the general poset structure, let us start from a categorical approach without any order information. Let Y indicate a polytomous random variable that can take values g_m, and let $P(Y = g_m) = \pi_m$. Furthermore, let $\mathbf{X} = (X_{g_1}, \ldots, X_{g_M})^{\mathrm{T}}$ with $X_{g_m} = I(Y = g_m)$ be a vector of indicator functions. The multinomial distribution of the response then has likelihood function $q(\mathbf{x}) \exp(\mathbf{x}^{\mathrm{T}} \log \boldsymbol{\pi})$, with $q(\mathbf{x})$ being a normalizing constant and with $\sum \pi_m = 1$. This multinomial distribution does not contain any information about the (partial) order of the categories. However, we may include this kind of information by applying the following lattice-based reparameterization to the kernel (see Ip *et al.*, 2004; Wang, 1986):

$$\mathbf{x}^{\mathrm{T}} \log \boldsymbol{\pi} = (\mathbf{x}^{\mathrm{T}} A)(A^{-1} \log \boldsymbol{\pi}) = s^{\mathrm{T}} \boldsymbol{\omega}, \tag{1}$$

where $\mathbf{s} = A^{\mathrm{T}} \boldsymbol{x}$, $\omega = A^{-1} \log \boldsymbol{\pi}$.

The elements of the vector $\mathbf{s}$ include information on the partial order of responses as they indicate for each element of the poset whether or not it dominates the chosen response. In particular, for the poset in Fig. 1a, $\mathbf{s}^{\mathrm{T}} = (x_a + x_b + x_c + x_d + x_e + x_f, x_b + x_d + x_e + x_f, x_c + x_f, x_d + x_f, x_e + x_f, x_f)$.

The elements of the vector ω depend on the contrast matrix A^{-1} and on the multinomial parameters π_m. For example, for the poset in Fig. 1a, the contrast matrix is given by

$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 1 & -1 & -1 & -1 & 1 \end{pmatrix}.$$

As a result of $A^{-1}\log\boldsymbol{\pi}$, we obtain canonical parameters that form contrasts between multinomial parameters. In particular, for the poset in Fig. 1a, $\boldsymbol{\omega} = (\log\pi_a, \log(\pi_b/\pi_a), \log(\pi_c/\pi_a), \log(\pi_d/\pi_b), \log(\pi_e/\pi_b), \log(\pi_a\pi_b\pi_f/\pi_c\pi_d\pi_e))$.

The assumption that the elements of a poset form a lattice is important because the invertibility of the dominance matrix A is guaranteed by a Möbius inversion under the framework of incidence algebra associated with the poset (Aigner, 1979, Chapter 4; Berge, 1971).

To identify the model in (1), the canonical parameter associated with the minimal element of the poset is put equal to zero. Accordingly, the log-likelihood equation for the reparameterized model is given by

$$\ell(\boldsymbol{\omega}) = s^{\mathrm{T}}\boldsymbol{\omega} - \kappa(\boldsymbol{\omega}), \tag{2}$$

where $\kappa(\boldsymbol{\omega})$ is the normalizing constant.

3.2 Multi-component poset structure

As a special case, we consider the representation of posets that are derived from responses to multiple component items with ordered categories. More specifically, let the componentwise responses be represented by a vector $\mathbf{Y} = (Y_1, \ldots, Y_I)$, where Y_i takes an ordered integer value y_i from 0 to $M_i - 1$. The patterns $\mathbf{y}$ can be considered the response categories of a poset. Let $\mathbf{X}$ be a vector of indicator functions the elements of which refer to the categories of the poset – that is, $X_{y_1,\ldots,y_I} = I(Y_1 = y_1, \ldots, Y_I = y_I)$ – and let $\boldsymbol{\pi}$ be a vector that comprises the probabilities of answering in a particular category of the poset – that is, $\pi_{y_1,\ldots,y_I} = P(X_{y_1,\ldots,y_I} = 1)$. The vectors $\mathbf{X}$ and $\boldsymbol{\pi}$ are arranged in lexicographical order so that the first index changes fastest. For instance, for the poset in Fig. 1b, $\mathbf{X}^{\mathrm{T}} = (X_{000}, X_{100}, X_{010}, X_{110}, X_{001}, X_{101}, X_{011}, X_{111})$. In general, the lexicographical order specifies indices $m = y_1 + \sum_{j=2}^{I}\left[(y_j - 1)\prod_{t=1}^{j-1}M_t\right] - 1$ that vary from 0 for the minimal element to $M_1 \times \ldots \times M_I - 1$ for the maximal element. To formally represent the poset, we may use the lattice-based reparameterization of the multinomial kernel in (1).

For posets with a multi-component structure, the dominance matrix A and the contrast matrix A^{-1} can be derived from the dominance matrices of individual components. In particular, for posets with I components, it can be shown by induction that the dominance matrix A can be expressed as the tensor product of I component-level dominance matrices. That is, $A = B_I \otimes B_{I-1} \cdots \otimes B_1$. For example, the dominance matrix in Fig. 1b can be computed using component matrices B_1, B_2, and B_3 that are defined for 0–1 chains of the binary components Y_1, Y_2, and Y_3, namely,

$$B_1 = B_2 = B_3 = B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

In particular, the overall dominance matrix for three-component poset is given by

$$B_3 \otimes B_2 \otimes B_1 = \begin{pmatrix} B & 0 & 0 & 0 \\ B & B & 0 & 0 \\ B & 0 & B & 0 \\ B & B & B & B \end{pmatrix}.$$

The contrast matrix A^{-1} can be computed using the inverted dominance matrices of individual components – that is, $A^{-1} = B_I^{-1} \otimes \cdots \otimes B_I^{-1}$.

Applying the transformation (1) to the multinomial kernel of a multi-component poset yields a vector **s** of indicator functions and an associated vector $\boldsymbol{\omega}$ of canonical parameters. The elements of **s** identify the elements of the poset that dominate the chosen response. For instance, for the poset in Fig. 1b the dominance matrix is presented in Table 1. The two last rows present (a different formulation of) the elements of **s** associated with the column elements of the dominance matrix. As $\mathbf{s} = A^{\mathrm{T}}\mathbf{x}$ the ith element of **s** is obtained as the sum of indicator functions in **x** that have a 1 in the ith column of A. For instance, for column '100' the value of s equals $\sum_{i=0}^{1}\sum_{j=0}^{1} x_{1ij} = x_{1++}$. Furthermore, defining component-specific indicator functions $Z_i(y_i) = I(Y_i > y_i)$ $(i = 1, \ldots, I;$ $y_i = 0, \ldots, M_i - 1)$, the elements in **s** can be reformulated as in the bottom row of Table 1 by using the fact that a row element only dominates a column element if all components of the row element dominate the corresponding component of the column element. For instance, for the column element '110', the indicator function x_{11+} can be reformulated as $I(Y_1 > 1, Y_2 > 1, Y_3 > 0) = Z_1(1)Z_2(1)Z_3(0) = Z_1(1)Z_2(1)$. In sum, we see that when evaluated at the ith element of **x**, **s** produces the ith row of the dominance matrix. For example, for the category of the poset corresponding to the response pattern '101', $\mathrm{s} = (1, 1, 0, 0, 1, 1, 0, 0)^{\mathrm{T}}$.

Table 1. Dominance matrix and vector of indicator functions in s for a multi-component poset with three binary components

Row	Column							
	000	100	010	110	001	101	011	111
000	1	0	0	0	0	0	0	0
100	1	1	0	0	0	0	0	0
010	1	0	1	0	0	0	0	0
110	1	1	1	1	0	0	0	0
001	1	0	0	0	1	0	0	0
101	1	1	0	0	1	1	0	0
011	1	0	1	0	1	0	1	0
111	1	1	1	1	1	1	1	1
	X_{+++}	X_{1++}	X_{+1+}	X_{11+}	X_{++1}	X_{1+1}	X_{+11}	X_{111}
	1	$Z_1(1)$	$Z_2(1)$	$Z_1(1)\,Z_2(1)$	$Z_3(1)$	$Z_1(1)Z_3(1)$	$Z_2(1)Z_3(1)$	$Z_1(1)\,Z_2(1)Z_3(1)$

The canonical parameters of the vector $\boldsymbol{\omega}$ are defined as contrasts of multinomial parameters π. For example, for the lattice structure in Fig. 1b, let $\boldsymbol{\omega} = A^{-1}\log \boldsymbol{\pi} = (\omega_0, \omega_{1(1)}, \omega_{2(1)}, \omega_{12(11)}, \omega_{3(1)}, \omega_{13(11)}, \omega_{23(11)}, \omega_{123(111)})^{\mathrm{T}}$. The components of the vector $\boldsymbol{\omega}$ can be grouped into four categories: a zeroth-order term $\omega_0 = \log \pi_{000}$, first-order conditional logits such as $\omega_{1(1)} = \log(\pi_{100}/\pi_{000})$, second-order log-odds ratios such as $\omega_{12(11)} = \log[(\pi_{110}\pi_{000})/(\pi_{100}\pi_{010})]$, and finally the third-order interaction term $\omega_{123(111)} = \log[(\pi_{111}\pi_{001})/(\pi_{101}\pi_{011})] - \log[(\pi_{110}\pi_{000})/(\pi_{100}\pi_{010})]$. Details are developed in Ip *et al.* (2004).

More generally, for a multi-component poset with I components, **s** and $\boldsymbol{\omega}$ are $(M_1 \times \ldots \times M_I)$-vectors that contain the following elements:

$$\mathbf{s} = (1, z_{1(1)}, \ldots, z_{I(m_i)}, z_{1(1)}z_{2(1)}, \ldots, z_{I-1(m_{I-1})}z_{I(m_I)}, \ldots, z_{1(m_1)}\ldots z_{I(m_I)})^{\mathrm{T}} \quad (3)$$

$$\boldsymbol{\omega} = (\omega_0, \omega_{1(1)}, \ldots, \omega_{I(m_I)}, \omega_{12(11)}, \ldots, \omega_{I-1I(m_{I-1}m_I)}, \ldots, \omega_{1\ldots I(m_1\ldots m_I)})^{\mathrm{T}}. \quad (4)$$

The order of the elements in $\mathbf{s}$ and $\boldsymbol{\omega}$ is determined by the order of the elements in $\mathbf{X}$, which is arbitrary. In (4) the order is such that subsequent canonical parameters refer to main effects, second-order interactions and so on. As indicated earlier, an alternative is to define a lexicograpical order for the elements in $\mathbf{X}$.

To identify the multi-component poset model, the canonical parameter ω_0 associated with the minimal element is constrained to be zero. Accordingly, the log-likelihood equation log $P(\mathbf{Y} = \mathbf{y})$ is similar as for general posets (i.e. (2)). Note that the log-likelihood equation that is reparameterized under the incidence algebra for multi-component responses is a generalization of the GLLM of Zhao and Prentice (1990). More specifically, the GLLM for I dichotomous 0–1 components Y_i can be expressed as

$$\log P(\mathrm{Y} = \mathrm{y}) = \sum_{i=1}^{I} y_i\omega_i + \sum_{i<j} y_i y_j \omega_{ij} + \ldots + y_1\ldots y_I\omega_{1\ldots I} - \kappa(\omega), \quad (5)$$

where $\omega^{\mathrm{T}} = (\omega_1, \ldots, \omega_{1\ldots I})$ is the set of loglinear interaction terms.

4. Latent variable models for posets

The canonical parameters may be further specified as a function of other parameters. For example, in IRT models a linear mixed formulation is given to these parameters with a random intercept as latent person variable. The partial credit model (PCM: Masters, 1982) is an IRT model for the analysis of totally ordered responses. To explain the PCM we may consider items Y_k (k = 1, . . . , K) which take ordered integer values m from 0 to $M_k - 1$.

Using the transformation in (1), we may derive that under the PCM the canonical parameters contrast adjacent categories – that is, $\omega_{mk} = \log[P(Y_k = m)/P(Y_k = m - 1)]$. Note that the first canonical parameter ω_0 is put equal to zero in order to identify the model. The PCM is obtained by specifying these logits of adjacent categories to be a linear function of a latent variable θ (Masters, 1982; Molenaar, 1983). Specifically,

$$\omega_{mk}(\theta) = \alpha\theta - \lambda_{mk} \quad (k = 1, \ldots, K;\ m = 1, \ldots, M_k - 1). \quad (6)$$

Note that the linear relationship for the logit of the probability that the response belongs to category m holds given that the response is either m or $m - 1$. The PCM can be extended in various ways by using other specifications for $\omega_{mk}(\theta)$. For instance, one may use item-specific slope parameters ($\omega_{mk}(\theta) = \alpha_k\theta - \lambda_{mk}$), as in the generalized partial credit model (Muraki, 1992). In case all variables Y_k have the same number of categories, one may, as in the rating scale model (Andrich, 1978), constrain parameters λ_{mk} to be a function of category- and item-specific parameters – that is, $\lambda_{mk} = \gamma_m + \xi_k$. Furthermore, one may specify multidimensional models (i.e. using $\sum_q \alpha_q\theta_q$ instead of θ), or one may include (person or item) covariates in order to model latent trait values or category parameters.

4.1 Latent variable models for general posets

To build latent variable models for general posets, one may use the same building blocks as for the PCM and specify canonical parameters to be a linear function of the latent trait (see (6)). This basic model, henceforth called the partial order model (POM), uses the incidence algebra described above for generalizing the concept of 'adjacency' in the PCM to posets. In the same way as described for the PCM, the POM can be extended by using other types of building blocks for the canonical parameters.

4.2 Latent variable models for multi-component posets

Extending the above models to multi-component posets is straightforward. Let Y_{ik} denote the ith component of poset k ($k = 1, \ldots, K$; $i = 1, \ldots, I$) that takes totally ordered values y_{ik} from 0 to $M_{ik} - 1$. One may construct latent variable models by specifying the parameters of $\boldsymbol{\omega}$ in (4) to be linear functions of θ. For example, for poset k we could specify first-order terms $\omega_{ik}(y_{ik}) = \alpha\theta - \lambda_{ik(y_{ik})}$, and second-order terms $\omega_{ijk}(y_{ik}y_{jk}) = \alpha\theta - \lambda_{ijk(y_{ik}y_{jk})}$.

In describing the extensions for multi-component posets it is useful to distinguish between modelling first-order terms and interaction terms. Interaction terms are specific for multi-component posets as they refer to response combinations of different items. For first-order terms, slope parameters α can be made poset-specific (i.e. $\omega_{ik}(y_{ik}) = \alpha_k\theta - \lambda_{ik(yik)}$), component-specific (i.e. $\omega_{ik}(y_{ik}) = \alpha_i\theta - \lambda_{ik(yik)}$), or specific for pairs of components and posets (i.e. $\omega_{ik}(y_{ik}) = \alpha_{ik}\theta - \lambda_{ik(yik)}$). For interaction terms, slope parameters can be poset-specific, but they cannot be component-specific. Partial order models that include specific slope parameters are henceforth called generalized partial order models (GPOM).

The basic building blocks as they are presented thus far, as in (4) and its variants, are all dimension-dependent because they are modelled as linear functions of the latent trait. A first useful restriction is to specify interaction terms to be constant across latent trait values (e.g. $\omega_{ijk}(y_{ik}y_{jk}) = -\lambda_{ijk}(y_{ijk(y_{ik}y_{jk})}$, and so on). For a distinction between dimension-dependent and dimension-independent interaction models, see Hoskens and De Boeck (1997). A second useful restriction is to put certain interaction terms equal to zero. If all interaction terms are made equal to zero, one obtains an independence model.

4.3 Nominal model

As the nominal model (Bock, 1972) does not involve any hypothesis about the order of the responses it may be considered a useful basis of comparison for models that hypothesize a specific (partial) order structure on the observed responses (Thissen & Steinberg, 1984). Although the nominal model does not hypothesize ordered responses, it can be formulated as a special case of the POM. Consider items $Y_k (k = 1, \ldots, K)$ that take nominal values m from 0 to $M_k - 1$. Using a dominance matrix in which each category dominates itself and the baseline category 0, we may derive that the canonical parameters contrast each category to the baseline, that is, $\omega_{mk} = \log[P(Y_k = m)/P(Y_k = 0)]$. The nominal model is now obtained by specifying the canonical parameters to be a linear function of the latent variables with category-specific slopes, that is,

$$\omega_{mk} = \alpha_{mk}\theta - \lambda_{mk} \quad (k = 1, \ldots, K; m = 1, \ldots, M_k - 1).$$

The first canonical parameter ω_{0k} is put equal to 0 in order to identify the model. Finally, we note that when analysing multi-component posets, a nominal model would be specified for the response patterns on the component items.

5. Tracelines and trajectories

When applying latent variable models, inference about the probability of answering in a particular response category, as a function of the latent variable, is often of primary interest since it shows the impact of individual differences. This information is provided by the so-called *traceline* $P(Y = c|\theta)$, which indicates the probability of responding in category c as a function of the latent trait. The traceline describes the change in the response probability with increasing value of the latent variable, for example a latent trait such as anger-proneness (see Section 8). We will discuss the derivation of the tracelines for general posets. The result for the special case of multi-component posets is straightforward. Let G_m denote the set of responses that are dominated by category g_m, that is, $G_m = \{g_n : g_n \prec g_m\}$, and let G denote the total set of responses.

Theorem 1. *The probability of responding in category g_m can be expressed as*

$$\pi_m = P(Y = g_m|\theta) = \frac{\exp\left\{\sum_{v\in G_m}\omega_v(\theta)\right\}}{\sum_{c\in G}\exp\left\{\sum_{v\in G_c}\omega_v(\theta)\right\}}. \tag{7}$$

To identify the model, we put the canonical parameter of the minimal element equal to zero. This implies that $\sum_{v\in G_{\hat{0}}}\omega_v(\theta) = 0$ *and, hence,* $\sum_{v\in G_m}\omega_v(\theta) \equiv \sum_{v\in G_m\setminus\hat{0}}\omega_v(\theta)$.

Proof. As $A_m A^{-1}$ is a vector of zeros, except for the mth element which equals one, it is easy to see that $\log \pi_m = (A_m A^{-1})\log \boldsymbol{\pi}$. Then using the fact that $\boldsymbol{\omega}(\theta) = A^{-1}\log \boldsymbol{\pi}$ and taking the normalization of the probabilities $\boldsymbol{\pi}$ into account we may derive that $\log \pi_m = A_m\omega(\theta) - \log\kappa(\theta)$. As $A_m\omega(\theta) = \sum_{v\in G_m}\omega_v(\theta)$ it follows that $\pi_m = \exp\{\sum_{v\in G_m}\omega_v(\theta)\}/\kappa(\theta)$ with $\kappa(\theta)$ the normalizing constant in the denominator of (7).

As a first example of the theorem, we consider the derivation of tracelines for the PCM. Let Y denote a random variable that takes totally ordered categories m from 0 to $M - 1$. Using (7), the traceline for category m is given by

$$P(Y = m|\theta) = \frac{\exp\{m\theta - \sum_{v=0}^{m}\lambda_v\}}{\sum_{c=0}^{M-1}\exp\{\sum_{v=0}^{c}(\theta - \lambda_v)\}},$$

with $\sum_{v=0}^{0}\lambda_v \equiv 0$ and $\sum_{v=0}^{c}(\theta - \lambda_v) \equiv \sum_{v=1}^{c}(\theta - \lambda_v)$.

As a second example, consider the traceline of category '6' in Fig. 1b. Using a generalization of (7) for multi-component posets, we may derive that

$$P(Y_1 = 0, \quad Y_2 = 1, \quad Y_3 = 1) = \frac{\exp\{\omega_{2(1)}(\theta) + \omega_{3(1)}(\theta) + \omega_{23(11)}(\theta)\}}{\kappa(\theta)},$$

with $\kappa(\theta)$ being a normalizing constant.

In addition to a thorough inspection of the tracelines of each category in the poset, one may also be interested in the combination of all tracelines, for instance in the categories that have the highest probability of being chosen across the latent scale. This pattern of dominant categories will henceforth be referred to as the *trajectory* of the dominant responses. Depending upon the context, the trajectory may provide useful information about the order of categories when θ varies across the latent scale. For instance, in Section 8, we will investigate the trajectory through combinations of

intensity and frequency of anger-related states when anger-proneness varies from low to high.

It is worthwhile to mention that the partial order structure which is hypothesized in the lattice restricts the order of the traceline peaks of the categories in the poset. More specifically, we may state the following theorem for a generalized partial order model with positive slopes:

Theorem 2. *In a lattice with minimal element* $\hat{0}$ *and maximal element* $\hat{1}$, if $g_m \prec g_n$ *then* π_m *will attain its peak at a smaller value of* θ *than* π_n.

Proof. *See Appendix A.*

From Theorem 2 it follows that the hypothesized partial order structure also puts clear restrictions on the empirical trajectory of categories that become dominant (i.e. have the highest probability) if one moves along the latent scale. More specifically, if $g_m \prec g_n$ then π_n cannot become dominant at a smaller value of θ than π_m. But note that, in general, neither g_n nor g_m need be dominant at some point of the scale. Finally, we note that the hypothesized partial order structure does not put any a priori restrictions on the traceline peaks of categories that are unordered in the poset. Rather, the order of these peaks is an empirical matter which corresponds, as in the nominal model (see Heinen, 1996), to the order of the (estimated) category-specific weights of the latent variable.

6. Parameter estimation

Assuming conditional independence of responses given the latent trait allows the likelihood equation of the canonical parameters to be expressed as a product of poset responses. Suppose Y_{pk} denotes the response of individual p to general poset k. The log-likelihood equation of the POM is given by

$$\ell(\boldsymbol{\omega}) = \log \prod_{p=1}^{P} \int \prod_{k=1}^{K} P(Y_{pk} = y_{pk}|\theta, \boldsymbol{\omega}) \mathrm{d}F(\theta), \tag{8}$$

where $\theta \sim F(\theta)$.

For the POM and its extensions discussed in Section 8, parameter estimates can be computed with the recently developed NLMIXED procedure of SAS (SAS Institute Inc., 1999). This procedure allows for marginal maximum likelihood (MML) estimates of the canonical parameters of the POM and can be programmed to provide empirical Bayes estimates of the person parameters. The MML approach of NLMIXED involves the maximization of the marginal likelihood formed from (8) using a normal distribution (e.g. N(0, 1)) for θ. Note that the dispersion of the latent scale is estimated by the slope parameter in (6). Various options are available in the NLMIXED procedure for approximating the integrated likelihood and for maximizing the approximation to the likelihood. In our data analysis, we used a non-adaptive Gaussian quadrature method to approximate the likelihood (Pinheiro & Bates, 1995), and a Newton-Raphson procedure for maximizing the approximate likelihood.

As an alternative, one may use a Markov chain Monte Carlo method such as the Metropolis algorithm (Metropolis, Rosenbluth, Rosenbluth, Teller, & Teller, 1953; Metropolis & Ulam, 1949) to compute a sample of the posterior distribution of the POM.

Note that, unlike the analysis with SAS, in this type of Bayesian analysis it may be necessary to specify proper priors for the category parameters (λ) and the slopes (α) as well (i.e. not only for θ) in order to ensure the propriety of the posterior.

Having available a sample of the posterior distribution has several advantages. First, it allows for the computation of 100 $(1 - \alpha)$% posterior intervals of the parameters that are valid in small samples because they do not rely on an asymptotic normal approximation to the posterior. Second, it allows for assessing the fit of the model with the technique of posterior predictive checks (Gelman, Carlin, Stern, & Rubin, 2004). Third, it allows for evaluating the uncertainty in any estimand of interest. In particular, in the context of partial order models, it may be of interest to evaluate the uncertainty in tracelines and corresponding trajectories for a particular model.

7. Model selection and model checking

The aim of *model selection* is to select one model from a set of competing models which best captures the process that generated the data. In order to achieve this goal, model selection criteria aim to select the model that optimally balances between parsimoniousness and goodness of fit (GOF) because only maximizing GOF will typically lead to the selection of models that are unnecessarily complex and that poorly generalize to other contexts. Well-known model selection criteria are Akaike's information criterion (AIC: Akaike, 1973) and the Bayesian information criterion (BIC: Schwarz, 1978). These criteria are defined as the sum of a badness-of-fit measure (minus twice the log-likelihood of the fitted model) and a complexity measure ($2k$ and $k \log(N)$ for AIC and BIC, respectively, where k is the number of free parameters and N being the sample size). For model selection purposes, there is no clear choice between these different information criteria (Hastie, Tibshirani, & Friedman, 2001). The BIC is asymptotically consistent (i.e. given a set of models including the correct one, the probability of selecting the correct model approaches 1 as $N \rightarrow \infty$), whereas asymptotically AIC will tend to select models that are too complex. On the other hand, with finite sample sizes BIC tends to select too simple models as it heavily penalizes model complexity.

As model selection criteria only concern the relative fit of models it is recommended to use *model checking* procedures in order to assess whether the best model fits specific aspects of the data that are of substantive interest and whether it fits the data in a global way. We will discuss four types of fit measures for multi-component posets: (1) a likelihood-ratio measure for evaluating the relative global GOF of the model; (2) a Pearson χ^2 measure for evaluating the absolute global GOF of the model; (3) a measure for evaluating whether (the strength of) the relation between component items and the latent trait (i.e. component-item slope) is captured by the model; and (4) a measure for evaluating whether the latent trait is able to capture the heterogeneity between persons. Extension to the case of general posets is straightforward.

7.1 Relative global goodness of fit

In order to evaluate the relative global GOF of the hypothesized partial order model, one may use a likelihood-ratio (LR) test in order to compare the model of interest to the nominal model. Using $\hat{\boldsymbol{\xi}}_{\mathrm{r}}$ and $\hat{\boldsymbol{\xi}}_{\mathrm{u}}$ to denote the maximum likelihood estimates of the

restricted POM and the unrestricted nominal model, the LR statistic is given by:

$$LR(y) = -2\log\frac{p(\hat{\boldsymbol{\xi}}_{\mathrm{r}}|y)}{p(\hat{\boldsymbol{\xi}}_{\mathrm{u}}|y)}. \tag{9}$$

Under the null hypothesis that the hypothesized partial order model holds, the LR statistic has asymptotically a χ^2 distribution with degrees of freedom equal to the number of parameters of the unrestricted model minus the number of parameters of the restricted model.

7.2 Absolute global goodness of fit

To construct absolute global GOF measures for a specific poset or for the entire collection of posets we form G homogeneous person groups on the basis of the percentiles of the distribution of the estimated person parameters. Let the vector $\boldsymbol{\xi}$ comprise all the model's parameters. A Pearson χ^2 GOF measure for poset k is given by

$$\chi^2_k(\xi, y_k) = \sum_{g=1}^{G}\sum_{l=1}^{L}\frac{[O_{glk}(y_k) - E_{glk}(\xi, y_k)]^2}{E_{glk}(\xi, y_k)}, \tag{10}$$

with O_{glk} and E_{glk} being the observed and expected number of persons in group g who respond in category l of poset k, respectively. The expected number of persons in group g who respond in category l of poset k is computed as the product of the number of persons in group g and the average probability of members in group g responding in category l of poset k. A GOF measure for the entire collection of posets is obtained by summing χ^2 the measures in (10) across posets, that is, $\chi^2_{tot}(\xi, \mathrm{y}) = \sum_k \chi^2_k(\xi, \mathrm{y}_k)$.

7.3 Component-item slope

For reasons of parsimony it is often interesting to restrict item-component slope parameters α_{ik} to be equal across item components, posets, or both item components and posets. In order to assess whether such restrictions hold one has to investigate whether possibly different item-component slopes that occur in the data are captured by the model. A specific GOF measure for evaluating whether the strength of the relation between a component item and the latent trait is captured by the model is the correlation between the component item and the sum of component items:

$$D_{ik}(y) = \mathrm{corr}_p\left(y_{pik}, \sum_i\sum_k y_{pik}\right).$$

7.4 Person heterogeneity

In building partial order models for posets it is important to investigate whether observations are conditionally independent given the latent trait(s) included in the model. If this is not the case, other latent traits or interactions between component items may have to be included in order to fully capture the heterogeneity between persons. A specific GOF measure for evaluating whether person heterogeneity is captured by the model is the correlation between pairs of component items:

$$H_{iki'k'}(\boldsymbol{y}) = \mathrm{corr}_p(y_{pik}, y_{pi'k'}). \tag{12}$$

If all correlations between pairs of component items are captured by the model we may conclude that the latent trait(s) and item-component interactions included in the model sufficiently explain the person heterogeneity observed in the data.

If a sample of the posterior distribution is available, it is straightforward to use posterior predictive checks (PPCs: Gelman *et al.*, 2004) to evaluate the significance of measures of global or specific fit. Gelman, Meng, and Stern (1996) define PPC *p*-values and describe related computational procedures for GOF measures $T(\mathbf{Y})$ that depend on the data only (e.g. (11), (12)) and for GOF measures $T(\boldsymbol{\xi}, \mathbf{Y})$ that depend on both the data and the parameters (e.g. (10)). These measures are labelled statistics and discrepancy measures, respectively. For statistics the PPC *p*-value can be computed by generating new data sets $\mathbf{Y}^{\text{rep}}$ (using the draws from the posterior distribution) and by computing the proportion of simulated data sets in which $T(\mathbf{y}^{\text{rep}}) \geq T(\boldsymbol{y})$. For discrepancy measures the PPC *p*-value is obtained by generating new data sets and by computing the proportion of generated data sets in which $T(\boldsymbol{\xi}, \mathbf{y}^{\text{rep}}) \geq \mathrm{T}(\boldsymbol{\xi}, \boldsymbol{y})$. Note that, for the computation of (10), for each draw of the posterior sample, persons are grouped on the basis the person parameter values.

8. Application

8.1 Study 1

We applied the approach described above to poset data on anger-related feelings. In particular, we studied two important aspects, namely, the frequency and the intensity (Diener, Larsen, Levine, & Emmons, 1985; Schimmack & Diener, 1997). In a first study, we collected responses from 420 first-year psychology students. The participants rated the frequency (0 = seldom, 1 = sometimes, 2 = often) and the intensity (0 = usually mild, 1 = sometimes mild and sometimes strong, 2 = usually strong) with which they experienced four anger-related feelings, namely 'anger', 'irritation', 'disgust' and 'rage' (Diener, Smith, & Fujita, 1995). In the data analysis, the categories 'sometimes' and 'often' of the frequency variable were joined (i.e. 0 = seldom, 1 = sometimes or often) because respondents rarely responded in the latter category, which led to unreliable parameter estimates. Thus, the data can be formally represented as four multi-component posets (one for each feeling) which consist of one trichotomous and one binary component.

Table 2 presents the canonical parameters of a 2×3 poset k and their definition in terms of local log-odds ratios. Specific latent variable models are built by parameterizing the canonical parameters as linear functions of the latent variable θ. In this application, θ can be interpreted as anger-proneness. In particular, we distinguish between four models by manipulating the type of interaction and the slope parameters in each of two ways. First, either interaction terms are put equal to 0, which leads to local *independence* (IND) between components within a poset, or they are set equal to a constant, which leads to *constant interaction* (CI) between components (i.e. independent of the latent trait). Second, either one global slope parameter α is assumed, as in the POM, or slopes α_{ik} are specified for each component i within a poset k, as in a GPOM.

Important questions that can be answered by analysing the data with the proposed models are:

(1) What is the *trajectory* of dominant responses across the latent scale for a particular feeling? In other words, which responses become dominant as one moves along the θ continuum? A related question concerns the stability of the trajectory if the uncertainty of the model's item parameters is taken into account.

Table 2. Canonical parameters for the partial order model (POM) and for the generalized partial order model (GPOM) assuming independence (IND) or constant interaction (CI) between components of a 2×3 poset k

		IND-POM	CI-POM	IND-GPOM	CI-GPOM
Parameter	$\omega_0 = \log \pi_{00}$	0	0	0	0
	$\omega_{1(1)} = \log\left(\frac{\pi_{10}}{\pi_{00}}\right)$	$\alpha\theta - \lambda_{1k(1)}$	$\alpha\theta - \lambda_{1k(1)}$	$\alpha_{1k}\theta - \lambda_{1k(1)}$	$\alpha_{1k}\theta - \lambda_{1k(1)}$
	$\omega_{2(1)} = \log\left(\frac{\pi_{01}}{\pi_{00}}\right)$	$\alpha\theta - \lambda_{2k(1)}$	$\alpha\theta - \lambda_{2k(1)}$	$\alpha_{2k}\theta - \lambda_{2k(1)}$	$\alpha_{2k}\theta - \lambda_{2k(1)}$
	$\omega_{2(2)} = \log\left(\frac{\pi_{02}}{\pi_{01}}\right)$	$\alpha\theta - \lambda_{2k(2)}$	$\alpha\theta - \lambda_{2k(2)}$	$\alpha_{2k}\theta - \lambda_{2k(2)}$	$\alpha_{2k}\theta - \lambda_{2k(2)}$
	$\omega_{12(11)} = \log\left(\frac{\pi_{00}\pi_{11}}{\pi_{10}\pi_{01}}\right)$	0	$-\lambda_{12k(11)}$	0	$-\lambda_{12k(11)}$
	$\omega_{12(12)} = \log\left(\frac{\pi_{01}\pi_{12}}{\pi_{02}\pi_{11}}\right)$	0	$-\lambda_{12k(12)}$	0	$-\lambda_{12k(12)}$
Measure	Npar* [2]	13	21	20	28
	Deviance	4942	4889	4916	4842
	AIC	4968	4931	4956	4898
	BIC	5020	5016	5037	5011

(2) How well do the frequency and intensity of the various feelings measure the underlying trait? To put it another way, to what extent do frequency and intensity levels rise if the latent trait increases?
(3) Are there any interactions between the components within a poset? These interactions are important because they indicate the dependencies between frequency and intensity categories beyond those based on the latent trait.

In the following paragraphs we discuss each of these three issues in turn.

8.1.1 Trajectories of dominant responses

The models are estimated with the SAS NLMIXED procedure. The SAS code used for estimating the CI-GPOM is explained in Appendix B. Table 2 summarizes the number of parameters involved in each model and the values of model selection criteria (AIC and BIC) that are provided by the NLMIXED procedure. Both AIC and BIC indicate that the GPOM with constant interactions between components (CI-GPOM) has the best balance between complexity and fit.

To further investigate the fit of the models we evaluate the four proposed fit measures of global and specific GOF. In order to evaluate relative global GOF, a nominal model was specified for the six poset categories that result from combining all frequency and intensity levels. This model has 40 parameters (i.e. 5 slope parameters and 5 category parameters for each of 4 posets). The AIC, BIC and the deviance of the nominal model equal 4904, 5065, and 4824, respectively. Comparison of the CI-GPOM and the nominal model with the LR test in (9) indicates that the restricted CI-GPOM cannot be rejected ($LR = 18$, df $= 12, p = .12$). In the same way, AIC and BIC for the CI-GPOM (4898 and 5011, respectively) are lower than for the nominal model.

To evaluate global and specific GOF measures in (10), (11) and (12) we computed a sample of the posterior distribution for each model and computed PPC p-values. More specifically, posterior samples were computed with the Metropolis algorithm (see Gelman *et al.*, 2004). For each model, two chains of 10,000 iterations were run, and

from the second halves of the chains evenly spaced draws were gathered to construct a sample of 2,000 iterations. Each parameter converged according to the diagnostic proposed by Gelman and Rubin (1992). Unlike the analysis with SAS, the Bayesian analysis also assumed normal priors for the item parameters, that is, $\lambda \sim \mathrm{N}(\mu_\lambda, \sigma_\lambda^2)$ and $\alpha \sim \mathrm{N}(\mu_\alpha, \sigma_\alpha^2)$, in order to guarantee the propriety of the posterior distribution. For purposes of comparison, the hyperparameters of the prior distributions in the Bayesian analysis were put equal to the means and variances of the parameter estimates that were obtained with SAS.

Table 3 shows PPC *p*-values for absolute global GOF measures and for the specific measures related to component-item discrimination. The results related to person heterogeneity are summarized by listing the proportion of correlations between component items that lie within their simulated 95% posterior interval.

Table 3. PPC *p*-values for global GOF measures (χ^2) and for specific measures of component-item discrimination (*D*). As a measure of person heterogeneity the last row lists the proportion of correlations between component items that lie within their simulated 95% posterior interval

Measure	IND-POM	CI-POM	IND-GPOM	CI-GPOM
χ^2_{anger}	.06	.44	.04	.42
χ^2_{irrit}	.40	.43	.44	.49
χ^2_{disg}	.08	.32	.02	.32
χ^2_{rage}	.01	.49	.01	.44
χ^2_{tot}	.004	.33	.002	.33
$D_{\text{freq, anger}}$	.01	.00	.13	.15
$D_{\text{int, anger}}$	.43	.43	.29	.24
$D_{\text{freq, irrit}}$	.99	.98	.76	.55
$D_{\text{int, irrit}}$	.00	.01	.13	.18
$D_{\text{freq, disg}}$	.50	.44	.28	.37
$D_{\text{int, disg}}$	.99	.99	.73	.65
$D_{\text{freq, disg}}$	.12	.13	.14	.10
$D_{\text{int, rage}}$	.03	.04	.15	.17
Person heterogeneity	.82	.82	.93	.96

As can be seen in Table 3, models without interactions between frequency and intensity components are rejected on the basis of absolute global GOF tests, whereas models that include such interactions yield an acceptable fit to the data. Apparently, the Pearson χ^2 measures are not sensitive to misspecification of the slope parameters as both models with restricted and unrestricted slope parameters (CI-POM and CI-GPOM) yield about the same results. Note that the results of the GOF tests are based on four equally sized person groups. GOF tests based on eight groups yield the same conclusions.

As indicated by the results of the specific GOF measure for component-item slopes, models that contain only a global slope parameter (IND-POM and CI-POM) over- or underestimate several observed correlations between component items and the sum score of component itemsms whereas the generalized partial order models (IND-GPOM and CI-GPOM) do capture all correlations.

Finally, the proportion of the ($8 \times 7/2 = 28$) correlations between all pairs of component items that lie within their 95% simulated posterior interval indicates that

models with restricted slope parameters (IND-POM and CI-POM) yield 18% of significant residual correlations, whereas the generalized partial order models without and with interactions (IND-GPOM and CI-GPOM, respectively) yield only 7% and 4% of significant residual correlations. We may conclude that the CI-GPOM captures observed person heterogeneity sufficiently well as we expect 5% of significant residuals by chance.

In sum, the CI-GPOM, which was also selected on the basis of information criteria, may be considered a suitable model for the data as it passes all the GOF tests. Note that a bivariate latent trait model with a separate latent trait for frequency and intensity items would also be a meaningful model in view of the way the data are collected. However, such a model does not fit the data, significantly better than the CI-GPOM (i.e. AIC is higher for the bivariate model (4920) than for the CI-GPOM (4898)). Moreover, comparing the trajectories of dominant responses for different emotions, which is one of the main purposes of the present study, would be more complicated in a bivariate latent trait analysis. In the next paragraphs we discuss the results of the CI-GPOM in more detail.

Figure 2 shows, for each feeling, the tracelines and the trajectories of the responses for the CI-GPOM. For example, in the top left-hand panel, the trajectory for anger is 00 → 01 → 11 → 12. This is also visualized in the box diagram of Fig. 3. Note that the trajectories are computed for the interval $[-4, 4]$, which covers the distribution of the latent variable θ. This implies that the trajectory consists only of categories that are

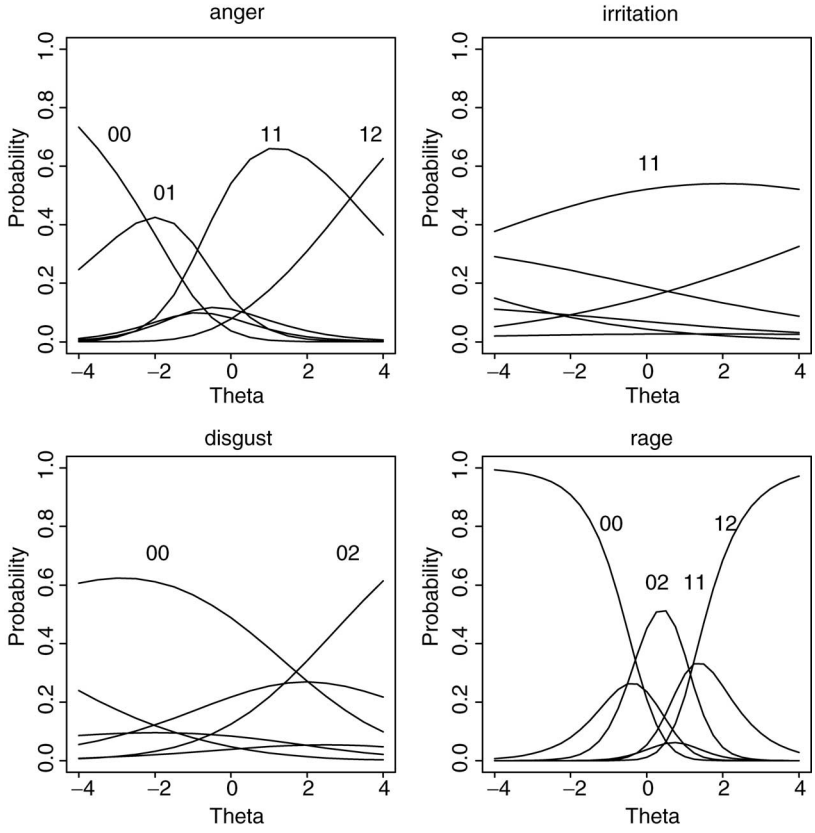


Figure 2. Tracelines and trajectories of CI-GPOM applied to feelings of Study 1.

dominant in this interval. For example, for irritation, the trajectory consists only of the category '11'.

In order to give a valid description of the differences between the trajectories of different feelings in the figure it is recommended to assess the variety of the trajectories for each feeling that is due to the uncertainty in the item parameters. This can be done by comparing the trajectories for each of the 2,000 draws of item parameters. Table 4 summarizes, for each feeling, the proportion of trajectories that occurred in more than 10% of the simulations.

From Fig. 2 and Table 4, it is clear that each feeling is characterized by a specific trajectory. For increasing values of θ, feelings of *anger* are subsequently experienced with low frequency and low intensity (00), low frequency and middle intensity (01), high frequency and middle intensity (11), and high frequency and high intensity (12). As indicated in Table 4, the high intensity level is not reached for 22% of the simulated trajectories. Feelings of *irritation* occur frequently and may have varying intensity levels depending upon θ. In particular, the trajectories 11, 10 → 11 and 10 → 11 → 12 occur in 63%, 13%, and 15% of the simulations, respectively. *Disgust* feelings occur seldom and may have varying intensity levels depending upon θ. More specifically, the trajectories 00 → 02 and 00 → 01 → 02 occur in 57% and 35% of the simulations, respectively. Finally, with increasing θ, feelings of *rage* are subsequently experienced with low frequency and low intensity (00), low frequency and high intensity (02), and high frequency and high intensity (12) in 93% of the simulations.

Finally, we note that the results of the Bayesian analysis and the analysis with SAS coincide rather well, as the displayed trajectories in Fig. 2 (NLMIXED results) also have the highest probability of occurring in Table 4 (Bayesian results). An exception is the trajectory for 'rage' in Fig. 2, which also has '11' as the dominant category for a very small part of the latent scale. This type of trajectory only rarely occurs in the Bayesian analysis (i.e. in less than 7% of the simulated cases).

8.1.2 Relation between observed variables and latent trait

The slope parameters of the CI-GPOM indicate the strength of the relationship between the latent trait and the observed variables. As shown in Table 5, the latent trait is best measured by the frequency of experienced anger (1.47) and by the frequency and intensity of experienced rage (3.20 and 1.28, respectively). However, for 'rage' the slope parameters could not be estimated very reliably (standard errors are 1.85 and 0.81, respectively). Frequency and intensity of experienced irritation are only weakly related to θ (0.20 and 0.19, respectively). Hence, the latter variables are unreliable measures of

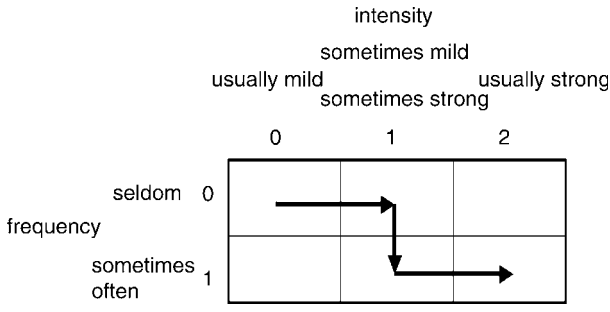


Figure 3. Box diagram of the trajectory for anger in Study 1.

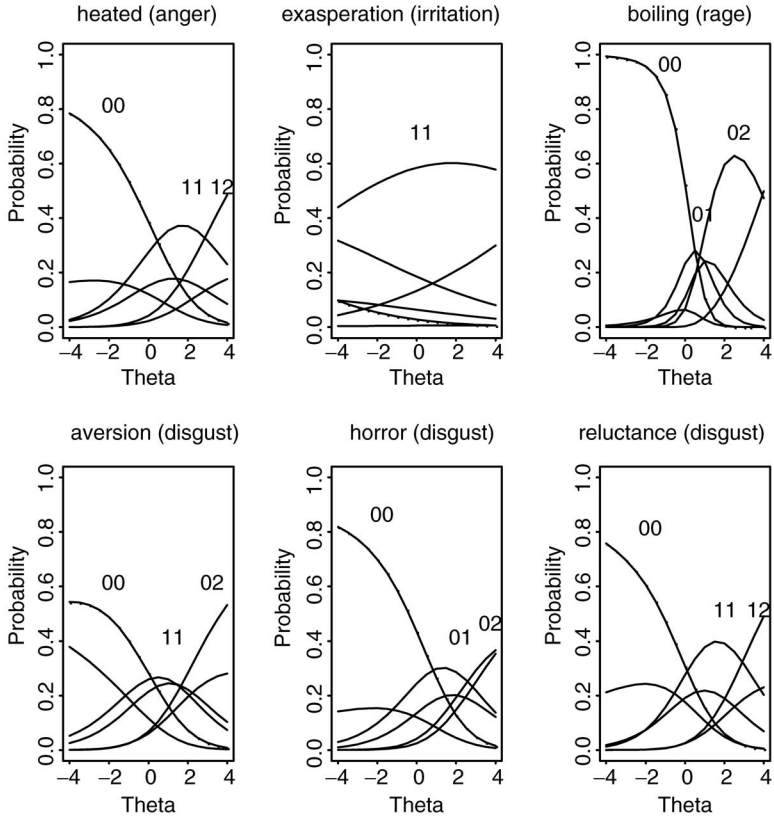


Figure 4. Tracelines and trajectories of CI-GPOM applied to replication feelings of Study 2. The original feelings are given in parentheses.

Table 4. Observed proportion of trajectories for the CI-GPOM that occur in more than 10% of the simulations

		Proportion	
Feeling	Trajectory	Study 1	Study 2
Anger	00 → 01 → 11	.22	
	00 → 01 → 11 → 12	.72	.99
Irritation	11	.63	.46
	10 → 11	.13	.20
	10 → 11 → 12	.15	.28
Disgust	00 → 02	.57	.35
	00 → 01 → 02	.35	.56
Rage	00 → 02 → 12	.93	.40
	00 → 01 → 02 → 12		.44

the latent trait. The low slopes for irritation may indicate the homogeneity of the feeling with respect to frequency and intensity in the group under investigation.

Finally, it is interesting to note that the frequency of experienced disgust is negatively related to the latent trait, whereas the intensity of experienced disgust is positively

Table 5. Slope parameters of CI-GPOM

		Study 1		Study 2	
Variable	Feeling	$\hat{\alpha}$	se($\hat{\alpha}$)	$\hat{\alpha}$	se($\hat{\alpha}$)
Frequency	Anger	1.47*	0.51	1.75*	0.65
	Irritation	0.20	0.19	0.15	0.19
	Disgust	−0.36*	0.18	−0.09	0.18
	Rage	3.20	1.85	1.52*	0.58
Intensity	Anger	0.62*	0.22	1.41*	0.50
	Irritation	0.19	0.11	0.21*	0.11
	Disgust	0.40*	0.11	0.37*	0.11
	Rage	1.28	0.81	0.99*	0.35

*$p < .05$.

related to the latent trait. We conjecture that disgust is an avoidance feeling (while anger is an approach feeling) and that frequency reflects a style while intensity is related to how negative the experience is. In this case, the negative slope of disgust would indicate an avoidance tendency, one that suppresses anger, while the moderately positive slope for intensity reflects the negative appreciation.

8.1.3 Interactions between poset components

Table 6 presents parameter estimates and standard errors for interaction parameters of the CI-GPOM. After controlling for the latent trait, 'anger' and 'rage' show a significant negative correlation ($p < .05$) between frequency and intensity from the middle of the intensity scale – that is, $(\pi_{01}\pi_{12})/(\pi_{02}\pi_{11}) < 1$. Hence, for rage and anger, from a given point on, intensity goes against frequency. In contrast, disgust shows a positive correlation between frequency and intensity in the first part of the intensity scale, that is, $(\pi_{00}\pi_{11})/(\pi_{01}\pi_{10}) > 1$. When disgust is not very strong, intensity seems to be associated with frequency.

More generally, we may observe in Table 6 that, for all feelings, the correlation between frequency and intensity for the first part of the intensity scale is larger than the

Table 6. Interaction parameters of CI-GPOM

		Study 1		Study 2	
Feeling	Parameter	$\hat{\lambda}$	se ($\hat{\lambda}$)	λ	Se ($\hat{\lambda}$)
Anger	$-\lambda_{k12(11)}$	0.19	0.41	−1.13	0.71
	$-\lambda_{k12(12)}$	−1.33*	0.42	−2.15*	0.73
Irritation	$-\lambda_{k12(11)}$	0.55	0.33	0.18	0.34
	$-\lambda_{k12(12)}$	−0.27	0.38	−0.02	0.40
Disgust	$-\lambda_{k12(11)}$	1.38*	0.31	1.07*	0.32
	$-\lambda_{k12(12)}$	−0.21	0.36	−0.10	0.32
Rage	$-\lambda_{k12(11)}$	0.25	1.17	0.65	0.58
	$-\lambda_{k12(12)}$	−2.22*	1.09	−1.18*	0.50

*$p < .05$.

correlation between frequency and intensity for the second part of the scale and that correlations for the first part of the scale are positive, whereas for the second part of the scale they are negative (though not always significantly so). In other words, it seems that feelings with low to middle intensity tend to occur more frequently, whereas feelings with middle to high intensity tend to occur more seldom. A possible explanation is that very intense anger-related feelings with high frequency are not really bearable for the person in question, and that they are often perceived by others as socially unacceptable.

8.2 Study 2

A second study was conducted in order to investigate whether the most important findings of Study 1 could be replicated and generalized to other (similar) feelings. Study 2 included 10 feelings. For replication of Study 1, it included the four feelings of the first study (i.e. 'anger', 'irritation', 'disgust', and 'rage'). For generalization to other feelings, it included four additional feelings similar to those of the first study-namely, 'heated' (for 'anger'), 'exasperation' (for 'irritation'), 'boiling' (for 'rage') and 'aversion' (for 'disgust'), as well as two additional terms for 'disgust' – 'horror' and 'reluctance'. The reason for adding two 'disgust' replicates is that we wanted to test the negative slope for the frequency variable. The 10 feelings were rated by 376 first-year psychology students with respect to the frequency (0 = seldom, 1 = sometimes, 2 = often) and the intensity (0 = usually mild, 1 = sometimes mild and sometimes strong, 2 = usually strong) with which they were experienced. As in the first study, a dichotomized frequency variable (0 = seldom, 1 = sometimes or often) was used for the analysis.

8.2.1 Replication of Study 1

As a literal replication of Study 1, the four feelings of the first study were analysed with the CI-GPOM. As can be seen in Table 4, replication of Study 1 is very succesful because, with a few exceptions, the two studies yielded the same types of trajectories for each feeling. Furthermore, as summarized in Table 5, the strengths of the relationships between the latent trait and the observed variables were very similar to those in Study 1. In particular, the finding that the frequency and the intensity of experienced 'disgust' are, respectively, negatively (though not significantly) and positively related to θ was confirmed. Finally, as summarized in Table 6, interactions between components within a poset were similar for both studies. In particular, we found again that the residual correlations between frequency, and intensity (i.e. after controlling for θ) are always higher for the first part of the intensity scale than for the second part. of the scale and that the former tend to be positive while the latter tend to be negative.

8.2.2 Generalization to similar feelings

To investigate whether the findings of the first study can be generalized to similar feelings, we applied the CI-GPOM to the 10 feelings that were included in Study 2. This analysis yielded especially high slope parameters for frequency and intensity of 'disgust' and its replicates. Hence, the fact that relatively more 'disgust' replicates are included changes the interpretation of the latent trait. Therefore, in order to make a valid comparison with the first study, we decided to perform three separate analyses with CI-GPOM, each of which contained eight feelings: the four original feelings, and four

replication terms including a different one for 'disgust' ('aversion', 'horror', 'reluctance'), depending on the analysis. Because these three analyses differed only in the particular replicate that was used for 'disgust', they yielded very similar results.

In general, trajectories of replicates behave in the same way as those of the feelings of Study 1. Figure 4 shows trajectories for 'heated', 'exasperation', 'boiling', and for the replicates of 'disgust'. A Bayesian analysis of the stability of trajectories confirmed the main conclusions that were drawn in the first study. Inspection of the slope parameters indicates that the relationships between the frequency and intensity variables and the latent trait are rather similar for replication feelings and original feelings. For 'disgust' and its replicates, the slope of the frequency variable is not negative but close to zero, whereas the slope of the intensity variable is positive. Thus, the conjecture of Study 1 that 'disgust' is an avoidance feeling that goes against 'anger' cannot be generalized to similar feelings. A more plausible interpretation is that this category of feelings is independent of anger-proneness. Finally, for all feelings we found again that residual correlations between frequency and intensity are larger for the first part than for the second part of the intensity scale.

9. Discussion

This paper presents a general framework for the analysis of posets that form a lattice. Formally, the approach presented involves a lattice-based reparameterization of the multinomial kernel of multivariate polytomous responses into an extended form of the generalized loglinear model. Latent variable item-response models can be constructed by parameterizing the canonical parameters of the generalized loglinear model. In particular, it is shown that the partial credit model for completely ordered data can be generalized to the case of poset data by specifying each canonical parameter of the generalized log-linear model as a linear function of the latent trait with equal slopes across items and item-specific category parameters. Moreover, this basic model can be easily extended or restricted by using alternative parameterizations for the canonical parameters. As the partial credit model can be considered an adjacent-categories logit model, an interesting topic for future research would be the extension of a cumulative-logit model for totally ordered data to the case of partially ordered data.

In addition to general posets, we also considered as a special case the analysis of posets that are derived from responses to multiple-component items. An important restriction of our approach is that it is only applicable when responses to each of the component items are observed. However, the aim of componential analysis is often to determine unobserved components underlying a set of observed item responses. Thus, an interesting topic for future research would be to extend our approach so that it is applicable to posets derived from unobserved components items. Tatsuoka (2002) developed a model to classify persons as masters/non-masters for a number of cognitive operations, based on observed responses of the persons to test items and based on knowledge about the cognitive operations that are involved in a particular item. Hence, in this model the (unobserved) patterns of mastery/non-mastery for a set of cognitive operations form a poset with a lattice representation as in Fig. 1b. A related model in which the cognitive operations involved in a set of items are extracted from the data rather than derived from cognitive analysis was developed by Maris (1999) (see also Meulders, De Boeck, & Van Mechelen, 2003).

Tracelines, and particularly trajectories of dominant categories, are proposed as a useful graphical device to summarize the results of our approach and to serve as a basis

for interpretation. The trajectory can be used to derive an order for the unordered categories of a poset through the location of the maxima of the tracelines along the latent scale. However, a particular category may be absent from the trajectory because it is never dominant along the latent scale. Tracelines and trajectories can easily be extended to multidimensional models. For instance, in a two-dimensional model involving variables θ_1 and θ_2, each category is characterized by a response surface. The trajectory can be defined as the areas of the plane (θ_1, θ_2) in which particular categories have the highest probability of being chosen. When reduced to marginal unidimensional trajectories, in principle, each axis in the multidimensional-space can have its own trajectory.

Acknowledgements

The research reported in this paper was supported by GOA/2000/02 awarded to Paul De Boeck and Iven Van Mechelen.

References

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

Aigner, M. (1979). *Combinatorial theory*. Berlin: Springer-Verlag.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Second international symposium on information theory* (pp. 271–281). Budapest: Akadémiai Kiadó.

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

Berge, C. (1971). *Principles of combinatorics*. New York: Academic Press.

Bishop, Y., Fienberg, S. E., & Holland, P. (1975). *Discrete multivariate analysis*. Cambridge, MA: MIT Press.

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, *37*, 29–51.

Costa, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment*, *64*, 21–50.

Cox, D. R. (1972). The analysis of multivariate binary data. *Applied Statistics*, *21*, 113–120.

Diener, E., Larsen, R. J., Levine, S., & Emmons, R. A. (1985). Intensity and frequency: Dimensions underlying positive and negative affect. *Journal of Personality and Social Psychology*, *48*, 1253–1265.

Diener, E., Smith, H., & Fujita, F. (1995). The personality structure of affect. *Journal of Personality and Social Psychology*, *69*, 130–141.

Flavell, J. H. (1971). Stage-related properties of cognitive development. *Cognitive Psychology*, *2*, 421–453.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). London: Chapman & Hall/CRC.

Gelman, A., Meng, X. M., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, *4*, 733–807.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*, 457–472.

Glaser, R., Lesgold, A., & Lajoie, S. (1987). Toward a cognitive theory for the measurement of achievement. In R. Rorminp, J. Glover, J. C. Conoley & J. Witt (Eds.), *The influence of cognitive psychology on testing and measurement: The Buros-Nebraska symposium on measurement and testing* (Vol. 3, pp. 41–58). Hillsdale, NJ: Erlbaum.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning: Data mining, inference and prediction*. New York: Springer-Verlag.

Heinen, T. (1996). *Latent class and discrete latent trait models: Similarities and differences*. Thousand Oaks, CA: Sage.

Holland, P. (1990). The Dutch identity: A new tool for the study of item response models. *Psychometrika*, *55*, 5–18.

Hoskens, M., & De Boeck, P. (1997). A parametric model for local dependence among test items. *Psychological Methods*, *2*, 261–277.

Hoskens, M., & De Boeck, P. (2001). Multidimensional componential item response models for polytomous items. *Applied Psychological Measurement*, *25*, 19–37.

Ip, E. H., Wang, Y. J., De Boeck, P., & Meulders, M. (2004). Locally dependent latent trait models for polytomous responses. *Psychometrika*, *69*, 191–216.

Laird, N. M. (1991). Topics in likelihood-based methods for longitudinal data analysis. *Statistica, Sinica*, *1*, 33–50.

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

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika*, *64*, 187–212.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*, 149–174.

McCullagh, P. (1980). Regression models for ordinal data (with discussion). *Journal of the Royal Statistical Society, B*, *42*, 109–142.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). New York: Academic Press.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, *21*, 1087–1092.

Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. *Journal of the American Statistical Association*, *44*, 335–341.

Meulders, M., De Boeck, P., & Van Mechelen, I. (2003). A taxonomy of latent structure assumptions for probability matrix decomposition models. *Psychometrika*, *68*, 61–77.

Mislevy, R. J. (1996). Test theory reconceived. *Journal of Educational Measurement*, *33*, 379–416.

Molenaar, I. W. (1983). *Item steps (Heyman Bulletins 83-630-EX)*. Groningen: Heymans Bulletins Psychologische Instituten, R.U. Groningen.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, *16*, 159–176.

Ortony, A., & Turner, T. (1990). What's basic about emotions? *Psychological Review*, *97*, 315–331.

Piaget, J. (1950). *The psychology of intelligence*. New York: Harcourt Brace Jovanovich.

Pinheiro, J. C., & Bates, D. M. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*, *4*, 12–35.

Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores* (Psychometrika, Monograph Supplement No. 17). Richmond, VA: Psychometric Society.

SAS Institute Inc. (1999). SAS OnlineDoc (Version 8) [software manual on CD-ROM]. Cary, NC: SAS Institute Inc.

Schimmack, U., & Diener, E. (1997). Affect intensity: Separating intensity and frequency in repeatedly measured affect. *Journal of Personality and Social Psychology*, *73*, 1313–1329.

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

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, *116*, 250–264.

Smits, D. J. M., & De Boeck, P. (2003). An application of the model with internal restrictions on item difficulties on guilt-feelings. *Multivariate Behavioral Research*, *38*, 161–188.

Tatsuoka, C. (2002). Data analytic methods for latent partially ordered classification models. *Applied Statistics*, *51*, 337–350.

Thissen, D., & Steinberg, L. (1984). A response model for multiple choice items. *Psychometrika*, *49*, 501–519.

Wang, Y. J. (1986). Ordered-dependent parameterization of multinomial distributions. *Scandinavian Journal of Statistics*, *13*, 199–205.

Wilson, M. (1992). The ordered partition model: An extension of the partial credit model. *Applied Psychological Measurement*, *16*, 309–325.

Zhao, L. P., & Prentice, R. L. (1990). Correlated binary regression using a generalized quadratic model. *Biometrics*, *77*, 642–648.

Received 27 June 2003; revised version received 27 August 2004

Appendix A: Relation between hypothesized partial order structure and order of peaks in a generalized partial order model with positive slopes

Consider a poset with minimal element $\hat{0}$ and maximal element $\hat{1}$ and let $g_m \prec g_n$. Specifying the canonical parameters in (7) as linear functions of θ, it is easy to see that, for g_m and g_n, the derivative with respect to θ of the logarithm of the traceline can be expressed as

$$\frac{\partial \log(\pi_m)}{\partial \theta} = c_m - \frac{\partial \log(\kappa(\theta))}{\partial \theta} \tag{13}$$

and

$$\frac{\partial \log(\pi_n)}{\partial \theta} = c_n - \frac{\partial \log(\kappa(\theta))}{\partial \theta}, \tag{14}$$

with $\kappa(\theta)$ being the normalizing constant in the denominator of (7),and with c_m and c_n being the weights of the latent variable in the numerator of the tracelines for g_m and g_n. Note that, in a generalized partial order model with positive slopes, $g_m \pi g_n$ implies that $c_m < c_n$. Consider for instance the poset in Fig. 1b. Specifying first-order terms $\omega_i(y_i) = \alpha_i\theta - \lambda_i(y_i)$ and dimension-independent interaction terms, it is easy to derive that $c_{011} = \alpha_2 + \alpha_3$,$c_{010} = \alpha_2$, $c_{001} = \alpha_3$ and $c_{000} = 0$ and hence $c_{011} > c_{010}$, $c_{011} > c_{001}$, $c_{011} > c_{000}$ if all slopes are positive.

Suppose that $\hat{\theta}_n$ is the value of θ for which the traceline π_n attains its maximal value (i.e.$\partial \log(\pi_n)/\partial\theta = 0$). Evaluating (14) at $\hat{\theta}_n$ yields

$$\left.\frac{\partial \log(\kappa(\theta))}{\partial \theta}\right|_{\hat{\theta}_n} = c_n$$

Hence evaluating (13) at $\hat{\theta}_n$ yields

$$\left.\frac{\partial \log(\pi_m)}{\partial \theta}\right|_{\hat{\theta}_n} = c_m - c_n, \tag{15}$$

which is negative because $c_m < c_n$ when $g_m \prec g_n$. As π_m and π_n are single-peaked (see Heinen, 1996), a negative value in (15) implies that π_m attains its peak before π_n.

Appendix B: Sample SAS code for a generalized partial order model assuming constant interactions between frequency and intensity measures of two emotions

```
1. DATA cigpom;
2. INFILE 'c:\data.txt';
```

```
 3. INPUT person Y I1 I2 CA1 CA2
          CL11_1 CL21_1 CL21_2 CL121_11 CL121_12
          CL12_1 CL22_1 CL22_2 CL122_11 CL122_12;
 4. RUN;
 5. PROC SORT data = cigpom;
 6. BY person;
 7. RUN;
 8. PROC  NLMIXED  data = cigpom  noad  technique = newrap
    qpoints = 20;
 9. PARMS L11_1 = 0 L21_1 = 0 L21_2 = 0 L121_11 = 0 L121_12 = 0
          L12_1 = 0 L22_1 = 0 L22_2 = 0 L122_11 = 0 L122_12 = 0
          A11 = 1 A21 = 1
          A12 = 1 A22 = 1;
          num1 = (A11*CA1 + A21*CA2)*theta
                  + L11_1*CL11_1 + L21_1*CL21_1 + L21_2*CL21_2
                  + L121_11*CL121_11 + L121_12*CL121_12;
10. denom1 = 1 + exp(A21*theta-L21_1)
             + exp(2*A21*theta-L21_1-L21_2)
             + exp(A11*theta-L11_1)
             + exp((A11 + A21)*theta-L11_1-L21_1-L121_11)
             + exp((A11 + 2*A21)*theta-L11_1-L21_1
              -L21_2-L121_11-L121_12);
11. num2 = (A12*CA1 + A22*CA2)*theta
           + L12_1*CL12_1 + L22_1*CL22_1 + L22_2*CL22_2
           + L122_11*CL122_11 + L122_12*CL122_12;
12. denom2 = 1 + exp(A22*theta-L22_1)
             + exp(2*A22*theta-L22_1-L22_2)
             + exp(A12*theta-L12_1)
             + exp((A12 + A22)*theta-L12_1-L22_1-L122_11)
             + exp((A12 + 2*A22)*theta-L12_1-L22_1
              -L22_2-L122_11-L122_12);
14. loglik = I1*(num1-log(denom1)) + I2*(num2-log(denom2));
15. MODEL Y ~ general(loglik);
16. RANDOM theta ~ normal(0,1) subject = person;
17. RUN;
```

Explanatory notes for SAS codes

In lines 1–4 the data set is read from the file 'data.txt'. This data set has the following structure:

Person	Y	I1	I2	CA1	CA2	CL11_1	. . .	CL122_12
1	1	1	0	0	0			
1	2	0	1	0	1			
2	3	1	0	0	2			
2	4	0	1	1	0			
. . .								
420	5	1	0	1	1			
420	6	0	1	1	2			

In line 3 the 16 variables that are listed in the columns of the data set are specified: (1) 'person' is the person ID, (2) 'Y' is the response variable taking values 1 to 6 if frequency and intensity components equal (1,1), (1,2), (1,3), (2,1), (2,2), (2,3), respectively, (3) 'I1' and 'I2' are indicators of whether an observation pertains to emotion 1 or 2 (1 = yes, 0 = no), (4) 'CA1' and 'CA2' variables weight the discrimination parameters of the frequency and the intensity component (A1k and A2k, respectively) for each category of the emotion *k*, (5) 10 'CL' design variables to code the lambda parameters in Table 1 (e.g. CL22_1 codes the parameter $\lambda_{22(1)}$; CL122_12 codes the parameter $\lambda_{122(12)}$).

The weights of the discrimination parameters and the design variables are defined as follows:

Frequency	Intensity	Y	CA1	CA2	CL1k_1	CL2k_1	CL2k_2	CL12k_11	CL12k_12
0	0	1	0	0	0	0	0	0	0
0	1	2	0	1	0	−1	0	0	0
0	2	3	0	2	0	−1	−1	0	0
1	0	4	1	0	−1	0	0	0	0
1	1	5	1	1	−1	−1	0	−1	0
1	2	6	1	2	−1	−1	−1	−1	−1

As an illustration, consider the response function of category 5 for emotion 1 derived from the design variables:

$$\Pr(Y_1 = 5) = \frac{\exp\left[(\alpha_{11} + \alpha_{12})\theta - \lambda_{11(1)} - \lambda_{21(1)} - \lambda_{121(11)}\right]}{\kappa(\theta)},$$

where the numerator corresponds to the 5th row of the design matrix and the denominator is a normalizing constant which equals the sum of the numerators for all categories.

In lines 5–7 the observations are grouped by person. This preprocessing of the data is recommended for increasing the efficiency of the computations involved in the analysis.

In lines 8–17 the model is fitted with the SAS NLMIXED procedure. In line 8 the options are specified. In line 9 initial values are assigned to the model parameters. In lines 10–15 the general form of the log-likelihood is computed. In line 16 'theta' is specified to be a random person effect which is normally distributed with mean 0 and standard deviation 1.

Remark: the run time of one analysis on a data set with 4 emotions and 420 subjects (see data set in study 1) took about 6 minutes on a Pentium III 800mHz machine with 128 MB RAM.

Copyright of British Journal of Mathematical & Statistical Psychology is the property of British Psychological Society and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.