PSYCHOMETRIKA—VOL. 69, NO. 3, 401–420
SEPTEMBER 2004

A CONJUNCTIVE PARALLELOGRAM MODEL FOR PICK ANY/*N* DATA

IWIN LEENEN AND IVEN VAN MECHELEN

KATHOLIEKE UNIVERSITEIT LEUVEN

This paper proposes a multidimensional generalization of Coombs' (1964) parallelogram model for "pick any/*n*" data, which result from each of a number of subjects having selected a number of objects (s)he likes most from a prespecified set of *n* objects. In the model, persons and objects are represented in a low dimensional space defined by a set of ordinal variables with a prespecified number of categories; objects are represented as points and persons as intervals on each dimension. A conjunctive combination rule is assumed implying that a person selects an object if and only if the object is within the subject's interval on each dimension. An algorithm for fitting the model to a data set is presented and evaluated in a simulation study. The model is illustrated with data on preferences regarding holiday trips.

Key words: Choice data analysis, binary data, parallelogram analysis, noncompensatory combination rule.

A natural and widely used method to obtain information on an individual's preferences with respect to a set of n choice objects involves presenting him with the full set of objects and having him select the subset of objects he likes most. When this method is used to have each of h persons select from the same set of n objects, it gives rise to "pick any/n" data (Coombs, 1964), which can be represented by an $h \times n$ binary matrix $\mathbf{D}$, with $d_{ij} = 1$ if person i selects object j and 0 otherwise. This paper presents a new model for such data, which can be considered a generalization of Coombs' (1964) parallelogram model. The latter model assumes that a single underlying characteristic of the objects accounts for the choices made by the persons. (For instances of early parallelogram modeling, see Leik & Matthews, 1968; Coombs & Smith, 1973.) In the new model, though, it is assumed that persons judge the objects along several characteristics; these judgements are subsequently combined in a conjunctive way: A person selects all objects he considers acceptable with respect to *each* characteristic. In the domain of decision making, several studies have shown that people prefer such a conjunctive strategy over compensatory and other noncompensatory strategies when combining several characteristics of choice objects, especially in early stages of the decision process, if the number of choice alternatives is large and if there is a lot of information on each object available (see Beach, 1990; Ford, Schmitt, Schechtman, Hults, & Doherty, 1989; Ogilvie & Schmitt, 1979; Westenberg & Koele, 1992). An underlying reason for this may be that in using a conjunctive choice rule, unlike a compensatory one, commensurability problems (i.e., problems with respect to trading off values on different characteristics against one another, which requires relating the characteristics to a common dimension) are avoided (Abelson & Levi, 1985).

The new model will be called a *c*onjunctive *pa*rallelogram model and will be abbreviated as the CPA-model. In Section 1, we will introduce the model starting from the unidimensional

The research reported in this paper was partially supported by the Research Council of K.U.Leuven (PDM/99/037).

The authors gratefully acknowledge the contribution of Veerle De Wael in running the study used in the application section and of Paul De Boeck, Luc Delbeke, and an anonymous referee for their helpful comments on a previous draft of the manuscript.

All correspondence concerning this paper is to be addressed to Iven Van Mechelen, Department of Psychology, K.U.Leuven, Tiensestraat 102, B-3000 Leuven, Belgium; email: iven.vanmechelen@psy.kuleuven.ac.be.

0033-3123/2004-3/1998-0698-A $00.75/0

parallelogram model and a generalization of it. In Section 2, an algorithm is described to fit a CPA-model to a data set. Section 3 reports the results of a simulation study to evaluate to algorithm's performance. In Section 4, the CPA-model is illustrated with an application on holiday choice data. Section 5, finally, discusses the relation of the new model to several other models for choice data.

1. Model

1.1. The Unidimensional Parallelogram Model

Parallelogram analysis of pick any/*n* data approximates a binary $h \times n$ (person by object) matrix **D** by a binary $h \times n$ matrix **M**, which can be represented by a unidimensional parallelogram model (Coombs, 1964). In a unidimensional parallelogram model, a joint scale for persons and objects is assumed to underly the data with each object j having a position q_j and each person i being characterized by his ideal point c_i on the scale and his *latitude of acceptance* ϵ_i (a term adopted from Hovland, Harvey, & Sherif, 1957; see Coombs, 1964, p. 298). A person is predicted to select a given object iff the object's position is within the acceptance interval around the person's ideal point. Formally, for any person $i (1 \leq i \leq h)$ and any object $j (1 \leq j \leq n)$, it holds that:

$$m_{ij} = 1 \quad \text{iff} \quad |c_i - q_j| \leq \epsilon_i. \tag{1}$$

As an example, one may think of a person's preference with respect to bath temperature: Different bath temperatures have a position on a common temperature scale and the person's ideal point coincides with the (hypothetical) temperature he consideres ideal or prefers most; a temperature different from the ideal point still may be judged comfortable, though, provided that the difference to the ideal point is "acceptable."

A reparametrization of the model presented above reads as follows: Given an $h \times n$ (person by object) binary matrix **M**, each person i can be assigned an interval with lower bound a_i and upper bound b_i, and each object j a value q_j such that for any person $i (1 \leq i \leq h)$ and object $j (1 \leq j \leq n)$:

$$m_{ij} = 1 \quad \text{iff} \quad a_i \leq q_j \leq b_i. \tag{2}$$

Note that the model allows for a_i exceeding b_i, which causes the right-hand side of (2) to be false for any object j with respect to person i and, as a result, the model structurally includes the possibility of persons that do not select any objects. It is clear that in the latter case the specific values for a_i and b_i on the latent dimension are arbitrary. Incidentally, any person that selects 0, 1, or n objects does not convey any information on the ordering of the objects (Van Schuur, 1993b). Note further that a reparametrization of the parallelogram model in terms of intervals has been proposed before by DeSarbo and Hoffman (1986), Andrich and Luo (1993) and Van Schuur (1993a); in the models of the latter two authors, though, intervals are assigned to the objects rather than to the persons (and points to the persons rather than to the objects).

Although Equation (2) formally is a mere reparametrization of (1), both model variants are conceptually distinct. Most importantly, Equation (1) links up closely with an underlying symmetric preference function that peaks at the subject's ideal point, whereas Equation (2) links up with a broader class of preference functions that may not even include the concept of a subject's ideal point. Regarding the latter, under the assumption of a single-peaked preference function, parametrization (2) leaves open the possibility of a subject's ideal point to be anywhere within the acceptance interval, which implies that the acceptance interval does not need to be symmetric around the ideal point.

TABLE 1.
Hypothetical two-way two-mode binary matrix representing pick any/n data

	Objects									Objects							
	a	b	c	d	e	f	g	h		d	e	a	g	h	b	c	f
person 1	0	0	0	1	1	0	0	0	person 1	1	1	0	0	0	0	0	0
person 2	1	1	1	1	1	1	1	1	person 4	1	1	1	1	0	0	0	0
person 3	1	0	0	0	0	0	1	1	person 6	1	1	1	1	1	0	0	0
person 4	1	0	0	1	1	0	1	0	person 2	1	1	1	1	1	1	1	1
person 5	0	1	1	0	0	1	0	0	person 3	0	0	1	1	1	0	0	0
person 6	1	0	0	1	1	0	1	1	person 7	0	0	1	1	1	1	1	1
person 7	1	1	1	0	0	1	1	1	person 5	0	0	0	0	0	1	1	1

Note. In the matrix at the right side the rows and columns have been permuted such that the one-entries form a "parallelogram."

If a matrix **M** can be represented by a parallelogram model then its columns can be permuted such that the ones at each row are consecutive, as illustrated by the hypothetical matrix in Table 1. In this table, the rows have been permuted as well such that the ones in the matrix form a "parallelogram" (with an irregular right side). One may note that the name "parallelogram relation" stems from the study of pick k/n data, which result from persons having selected a fixed number k from the set of n objects.

The concept of Coombs' (1964) parallelogram relation has been studied independently under different names in several domains. In mathematics, a matrix representing a parallelogram relation is known as satisfying the consecutive ones property for rows (Fulkerson & Gross, 1965); a theorem by Tucker (1972) characterizes such matrices and Booth and Lueker (1976) presented an algorithm to test matrices for the consecutive ones property in a time linear in the size of the input. Apart from mathematics, matrices with properties similar to the consecutive ones property have been used in archeology within the context of seriation methods (Kendall, 1969; Goldmann, 1971; see also Arabie & Hubert, 1992) and in computer science for information retrieval (Ghosh, 1972). Moreover, in the psychological literature, the notion of a parallelogram relation is at the basis of Gati and Tversky's (1982) notion of a qualitative dimension (as opposed to a quantitative dimension).

1.2. An Extension of the Unidimensional Model

The concept of a unidimensional parallelogram model can be slightly extended by including a prespecified integer k (≥ 2) that indicates the maximum number of different values allowed for the parameters a_i, b_i and q_j in (2). In principle, this would imply prespecifying a set of constants $\{c_1, c_2, \ldots, c_k\}$ and restricting $a_i, b_i, q_j \in \{c_1, c_2, \ldots, c_k\}$. However, as only ordinal information is relevant in (2), it can be assumed, without loss of generality, that a_i, b_i and q_j take values from the set $\{0, 1, \ldots, k-1\}$. The value k will further be called the position-rank, or p-rank, as it corresponds to the number of different positions allowed on the underlying scale for individuals and objects; the unidimensional parallelogram model with p-rank k will further be denoted a k-CPA-model. Clearly, the model reduces to the original parallelogram model if k equals (or exceeds) the number of objects n. Table 2 presents a 4-CPA-model for the hypothetical matrix in Table 1.

A restriction on the number of positions on the underlying dimension may be advisable from two considerations: Firstly, from a theoretical point of view, if the number of positions is small compared to the number of persons or objects, the model yields classifications for both persons

TABLE 2.
A 4-CPA-model for the matrix in Table 1

	Persons' intervals		Objects' positions
Person 1	[0, 0]	object a	1
Person 2	[0, 3]	object b	3
Person 3	[1, 2]	object c	3
Person 4	[0, 1]	object d	0
Person 5	[3, 3]	object e	0
Person 6	[0, 2]	object f	3
Person 7	[1, 3]	object g	1
		object h	2

and objects; generally, the smaller the value for k, the more the parallelogram model is turned into a classification model. Secondly, from a data-analytical point of view, k-CPA analyses with small k can be expected to yield more reliable results when applied to data sets perturbed with random error (e.g., when cell values in the observed data matrix are the result of independent changes of cell values in some underlying true model matrix according to some random Bernoulli type of process).

1.3. The Conjunctive Parallelogram Model

The unidimensional model implies that a single dimension can account for person by object choice data. Often, it may be more realistic to assume that differences in choice behavior result from persons considering several characteristics of the objects differently. In the conjunctive parallelogram model, positions and intervals are assigned to objects and persons, respectively, on each of r dimensions, and a person is predicted to select an object iff the object is within the person's interval on each dimension. The number of dimensions r is called the dimension-rank, or d-rank, of the model, and a conjunctive parallelogram model of d-rank r and p-rank k is denoted a k^r-CPA-model. Formally, an $h \times n$ binary matrix $\mathbf{M}$ can be represented by a k^r-CPA-model iff matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{Q}$ with entries in $\{0, \ldots, k-1\}$ exist such that for any $i(1 \leq i \leq h)$ and $j(1 \leq j \leq n)$ it holds that:

$$m_{ij} = 1 \quad \text{iff} \quad \forall v(1 \leq v \leq r) : a_{iv} \leq q_{jv} \leq b_{iv}. \tag{3}$$

As an example, the binary 8×10 matrix in Table 3 presents the preferences (like/dislike) of each of 8 persons with respect to each of 10 liquors; in Table 4, a 3^2-CPA-model for this matrix is given. The preferences of the subjects can be easily read from the model: For example, Person 5 likes Amaretto and Porto as these liquors are within the person's acceptance interval on the first dimension (interval [0, 0]) and on the second dimension (interval [0, 1]). In this example, both dimensions can be given a substantive interpretation: The first dimension is related to the sweetness/bitterness of the liquor, while the second dimension is correlated with the percentage of alcohol. People only like the spirits for which both the level of sweetness/bitterness and the amount of alcohol is acceptable. The model further yields a classification of persons and objects: For example, Persons 4 and 7 constitute a class as they share the same acceptance intervals; with respect to the spirits, Vodka and Whiskey constitute a class like Cognac and Rum do as in both cases the liquors have identical positions on each dimension.

It is trivial to prove that a positive integer r can be found so that a k^r-CPA-model ($k \geq 2$) exists for any binary matrix $\mathbf{M}$: For any arbitrary person i and object j, a parallelogram dimension can be constructed where all persons select all objects except for person i who does not

TABLE 3.
Hypothetical preference matrix of 8 persons for 10 liquors

	Cognac	Rum	Whiskey	Martini	Vodka	Gin	Porto	Sherry	Dry white wine	Amaretto
Person 1	0	0	0	1	0	1	1	1	1	1
Person 2	1	1	1	1	1	1	0	0	0	0
Person 3	1	1	0	1	0	0	1	1	0	1
Person 4	1	1	0	1	0	0	0	1	0	0
Person 5	0	0	0	0	0	0	1	0	0	1
Person 6	0	0	0	0	0	0	1	1	0	0
Person 7	1	1	0	1	0	0	0	1	0	0
Person 8	1	1	1	1	1	1	1	1	1	1

TABLE 4.
A 3^2-CPA-model for the preference matrix in Table 3

	Persons' intervals			Objects' positions	
	I	II		I	II
Person 1	[0, 2]	[0, 1]	Cognac	1	2
Person 2	[1, 2]	[1, 2]	Rum	1	2
Person 3	[0, 1]	[0, 2]	Whiskey	2	2
Person 4	[1, 1]	[0, 2]	Martini	1	1
Person 5	[0, 0]	[0, 1]	Vodka	2	2
Person 6	[0, 1]	[0, 0]	Gin	2	1
Person 7	[1, 1]	[0, 2]	Porto	0	0
Person 8	[0, 2]	[0, 2]	Sherry	1	0
			Dry white wine	2	0
			Amaretto	0	1

select object j; such dimensions can be conjunctively combined to represent any arbitrary matrix **M**. However, as for any fixed d-rank r, matrices **M** can be found that cannot be represented by a k^r-CPA-model for any positive integer k, the converse does not hold. (Note that if the converse were true, any matrix **M** could be represented by a unidimensional parallelogram model, which would make multidimensional generalizations needless.)

Parallelogram models generally are not unique: For example, each dimension can be reflected and dimensions can be permuted without affecting the model's predicted values. Moreover, apart from this trivial nonuniqueness, structurally different k^r-CPA-models may exist for a given binary matrix **M**. For special cases, however, it can be proven that the model is unique upon a reflection and/or permutation of the dimensions. This is, for example, the case if at each of the k^r positions of a k^r-CPA-model an object is located and if for any pair of neighboring objects j and j', a person exists that selects j and j' only. (Two objects j and j' are considered neighbors if at one dimension v', $q_{jv'} = q_{j'v'} \pm 1$, and at all other dimensions v ($v \neq v'$), $q_{jv} = q_{j'v}$.) It is clear that under these conditions any object must have the same set of neighbors in any alternative solution. It can further be derived that if one of the 2^r objects with r neighbors is fixed in a corner position and subsequently its r neighbors are fixed in the adjacent positions, the position of all other elements immediately follows. Note that fixing an object in the corner position comes down to fixing the direction of the dimensions, while fixing its r neighbors corresponds to choosing a permutation of the dimensions.

Recently, Gelman, Leenen, Van Mechelen, De Boeck, and Poblome (2001) have proposed a Bayesian approach to deal with nonuniqueness problems in deterministic models. Their procedure can be considered to develop a Bayesian variant of the conjunctive parallelogram model presented in this paper as well.

1.4. Order-Preservance

Any binary matrix **M** induces an implication relation on both the set of row elements and the set of column elements: For each row element, the set of column elements it is associated with in **M** (i.e., the set of columns with a 1 in the row vector $\mathbf{m}_i$) can be considered, and subsequently relations like "for any column j, *if* row A is associated with j *then* row B is associated with j as well" can be derived. These implication relations often are of substantive interest. In an ability context, for example, with **M** representing person by item success/failure data, the if-then relations among the items have been extensively studied in the domain of knowledge spaces (see Falmagne, Koppen, Vilano, Doignon, & Johannesen, 1990). Other studies that illustrate the substantive importance of if-then type relations include work of Gara and Rosenberg (1979) in the domain of social psychology, of Vansteelandt and Van Mechelen (1998) in the personality domain, and of Ganter and Wille (1996) in the domain of concept analysis. Also in the study of choices, implication relations, particularly among the persons, may be informative. In a marketing context, for example, it may be useful to know that the products selected by a particular set of consumers includes the set of products chosen by some other type of consumers.

Models that represent the implication relations (on the row elements or the column elements) are generally called order-preserving (with respect to the rows and the columns, respectively). For the case of the parallelogram model, it can be proven that any CPA-model can be restricted to be order-preserving with respect to the persons (rows) (Leenen, Van Mechelen, & De Boeck, 1999).

The order-preservance restriction can be formalized as follows: A quasi-order (i.e., a relation that is both reflexive and transitive; see Barbut & Monjardet, 1970), $\preceq_R$, is defined on the rows of the $h \times n$ matrix **M** as:

$$i \preceq_R i' \quad \text{iff} \quad \forall j : m_{ij} \leq m_{i'j},$$

for rows i and $i'(1 \leq i, i' \leq h)$. (Note that $\preceq_R$ is generally defined without reference to any model.) In the conjunctive parallelogram model, order-preservance with respect to the rows (Leenen et al., 1999) implies for any pair of persons i, i':

$$i \preceq_R i' \quad \text{iff} \quad [\forall j(1 \leq j \leq n), \forall v(1 \leq v \leq r) : a_{iv} \leq q_{jv} \leq b_{iv} \Longrightarrow a_{i'v} \leq q_{jv} \leq b_{i'v}]. \quad (4)$$

Equation (4) states that on any dimension all objects that are in the interval of person i are in the interval of person i' as well if i precedes i' in the quasi-order on the persons. In view of a meaningful substantive interpretation, though, the parallelogram model is slightly restricted further such that the acceptance interval of person i' includes the acceptance interval of person i on each dimension:

$$i \preceq_R i' \quad \text{iff} \quad [\forall v(1 \leq v \leq r) : a_{i'v} \leq a_{iv} \text{ and } b_{iv} \leq b_{i'v}]. \quad (5)$$

Obviously, any k^r-CPA-model can be modified to a k^r-CPA-model that additionally satisfies (5) by redefining the boundaries a_{iv} and b_{iv} of the persons' intervals as follows:

$$a_{iv} \leftarrow \min_{i'|i' \preceq_R i} a_{i'v}$$

$$b_{iv} \leftarrow \max_{i'|i' \preceq_R i} b_{i'v}.$$

Equation (5) implies that the sets of objects selected by any two persons are nested iff the persons' intervals are nested for each dimension. For the model in Table 4, for example, the acceptance intervals of Person 6 are part of the intervals of Person 3, the latter in their turn being part of the acceptance intervals of Person 8.

Similarly, one might require the model to be order-preserving with respect to the objects. This implies for any pair of objects $j, j'(1 \leq j, j' \leq n)$:

$$j \preceq_C j' \quad \text{iff} \quad [\forall v(1 \leq v \leq r), \forall i(1 \leq i \leq h) : a_{iv} \leq q_{jv} \leq b_{iv} \Longrightarrow a_{iv} \leq q_{j'v} \leq b_{iv}], \quad (6)$$

with $\preceq_C$ a quasi-order on the columns being defined as:

$$j \preceq_C j' \quad \text{iff} \quad \forall i : m_{ij} \leq m_{ij'},$$

for columns j and j'.

However, as proven in Appendix A, for any integers k (≥ 2) and r (≥ 1) a matrix **M** exists that can be represented by a k^r-CPA-model and that cannot be represented by a k^r-CPA-model that additionally satisfies property (6). As a result, conjunctive parallelogram models can generally be restricted to represent the quasi-order on the persons only. Note that, if intervals were assigned to the objects and points to the persons (like, e.g., in Van Schuur, 1993a), a representation of the object quasi-order would be possible. Note further that, since the set of CPA-models for a matrix **M** that are order-preserving generally is a proper subset of the full set of CPA-models for **M**, requiring order-preservance generally reduces the lack of identifiability of the model.

1.5. Graphical Representation

A conjunctive parallelogram model can be given a graphical representation which depicts the association relation (3) and which may additionally depict the quasi-order on the persons if restriction (5) is added to the model. A k^r-CPA-model generally represents persons and objects in a joint r-dimensional space, with each object corresponding to a point and each person to a hypercube. Since in the model only intra-dimensional distances are relevant (i.e., person and object parameters are compared per dimension only), there is no restriction for this r-dimensional space to be Euclidean.

Figure 1 shows the graphical representation for the 3^2-CPA-model in Table 4. As the model has d-rank 2, the graph can be drawn in the plane. Each object is represented by a point and each person by a rectangle defined by the Cartesian product of his acceptance intervals on both dimensions. For reasons of clarity, only a subset of the persons is represented. The selection made by any person corresponds with the set of liquors that is within his rectangle. For example, Person 3, who likes any nonbitter liquor, selects Cognac, Rum, Martini, Sherry, Amaretto, and Porto. Also the quasi-order on the persons can be read from the graph by the inclusion relation on the person's rectangles. For example, the liquors selected by Person 5 are a subset of the liquors selected by Person 3 and, hence, Person's 3 rectangle includes Person's 5 rectangle.

2. Algorithm

Given a binary $h \times n$ matrix **D** and prespecified values r and k for the d-rank and p-rank, respectively, the algorithm for conjunctive parallelogram analysis aims at finding a binary $h \times n$ matrix **M**, which can be represented by a conjunctive k^r-CPA-model, such that the loss function

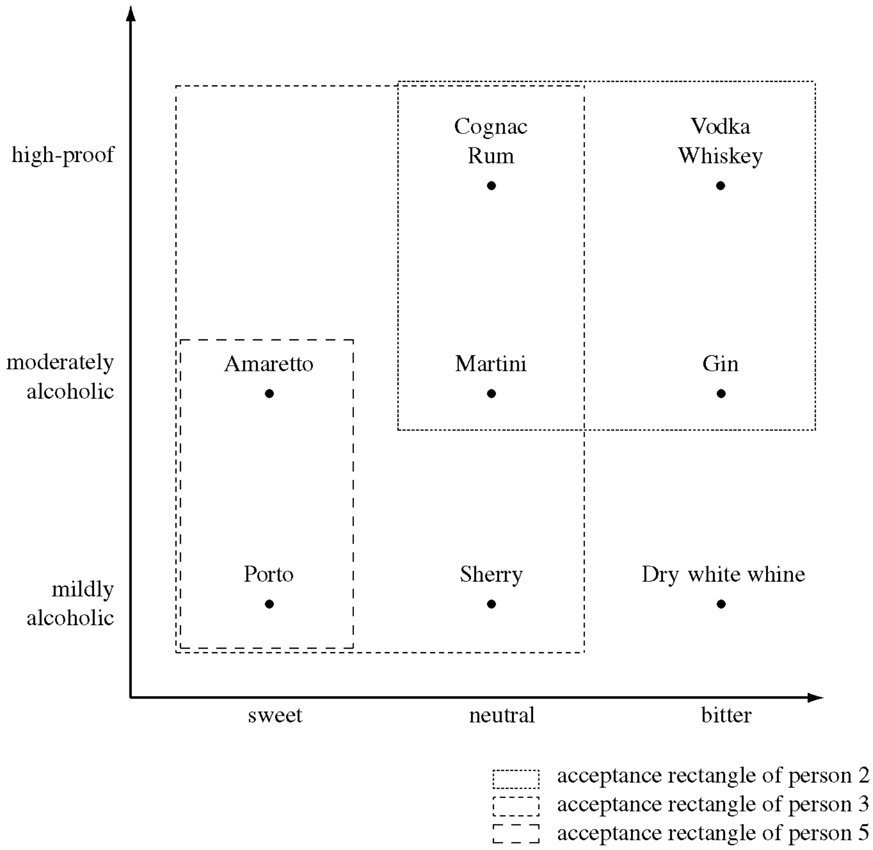

FIGURE 1.
Graphical representation of the 3^2-CPA-model in Table 4.

$$f(\mathbf{M}) = \sum_{i=1}^{h} \sum_{j=1}^{n} (d_{ij} - m_{ij})^2 \tag{7}$$

has minimal value. As **D** and **M** are binary, Equation (7) can be considered both a least squares and a least absolute deviations loss function (Carroll & Chaturvedi, 1995).

The algorithm is an alternating least squares procedure, which starts by constructing initial estimates for the positions of the objects in the $n \times r$ matrix $\mathbf{Q}^{(0)}$. This initial configuration can be obtained either (a) (pseudo-)randomly, or (b) rationally by a built-in heuristic in the algorithm, or (c) from the user. A user-provided initial configuration may be particularly useful in a confirmatory analysis, whereas random and rational heuristics are the obvious means in exploratory analyses. Otherwise, in order to avoid local minimum problems as much as possible, it is advisable to run the algorithm using both rational and (several different) random configurations as a starting point.

Once an initial configuration has been obtained, the procedure iterates between successive conditional estimations of (a) the $h \times r$ matrices $\mathbf{A}^{(w)}$ and $\mathbf{B}^{(w)}$, conditionally upon $\mathbf{Q}^{(w-1)}$, and (b) the matrix $\mathbf{Q}^{(w)}$, conditionally upon $\mathbf{A}^{(w)}$ and $\mathbf{B}^{(w)}$ ($w = 1, 2, \ldots$), with the entries in $\mathbf{A}^{(w)}$, $\mathbf{B}^{(w)}$ and $\mathbf{Q}^{(w)}$ being elements from the set $\{0, \ldots, k-1\}$. The alternating procedure continues until no further improvement in the loss function (7) is observed.

An analysis of the loss function shows that this function is separable (Carroll & Chaturvedi, 1995): From the combination rule (3), it is clear that the position $(q_{j1}, \ldots, q_{jr})$ of object j in the r-dimensional space only affects the jth column, $\mathbf{m}_{.j}$, of $\mathbf{M}$, while, regarding $\mathbf{Q}$, $\mathbf{m}_{.j}$ only depends on $(q_{j1}, \ldots, q_{jr})$. Similarly, a person's i set of intervals, $([a_{i1}, b_{i1}], \ldots, [a_{ir}, b_{ir}])$, only affects $\mathbf{m}_{i.}$; regarding $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{m}_{i.}$ only depends on $([a_{i1}, b_{i1}], \ldots, [a_{ir}, b_{ir}])$. As a result, the contribution of object j (resp. person i) to the loss function can be separated from the contribution of other objects (resp. persons) and an optimal estimate of the objects' positions $\mathbf{Q}$ conditionally upon $\mathbf{A}$ and $\mathbf{B}$ may be found by a sequential (conditionally optimal) estimation of the position $(q_{j1}, \ldots, q_{jr})$ of each single object j. Similarly, a conditionally optimal estimate of $\mathbf{A}$ and $\mathbf{B}$ (simultaneously) can be obtained by sequentially estimating the set of r intervals, $([a_{i1}, b_{i1}], \ldots, [a_{ir}, b_{ir}])$ for each person i. Appendix B describes a branch-and-bound procedure to find the optimal set of intervals for a person i conditionally upon $\mathbf{Q}$. The procedure used for optimally estimating an object's position is similar, though simpler as only half as many parameters are to be estimated.

The algorithm has been implemented in a computer program, which can be obtained from the authors by simple request.

3. Simulation Study

The simulation study in the present section has a two-fold aim: First, the algorithm described in the previous section is evaluated with respect to goodness of fit and goodness of recovery. Second, a heuristic for selecting a model among several parallelogram models of different d-ranks and/or p-ranks for a given data set is evaluated. In Subsection 3.1, the design of the study is dealt with and the results are presented in Subsection 3.2 (goodness of fit), Subsection 3.3 (goodness of recovery) and Subsection 3.4 (rank selection).

3.1. Design and Procedure

Three different types of binary $h \times n$ matrices are to be distinguished in the simulation study: the matrices $\mathbf{T}$ and $\mathbf{M}$, both of which can be represented by a k^r-CPA-model, and the data matrix $\mathbf{D}$; $\mathbf{T}$ is the true matrix underlying $\mathbf{D}$, and $\mathbf{M}$ is the model matrix which results from applying the parallelogram algorithm to $\mathbf{D}$.

The levels of the following independent variables were orthogonally crossed in a complete five-factorial design:

(a) the number of persons, h, in $\mathbf{T}$, $\mathbf{D}$, and $\mathbf{M}$, at 3 levels: 60, 120, 240;
(b) the number of objects, n, in $\mathbf{T}$, $\mathbf{D}$, and $\mathbf{M}$, at 3 levels: 40, 90, 150;
(c) the d-rank, r, of the conjunctive parallelogram model for $\mathbf{T}$ and $\mathbf{M}$ at 3 levels: 1, 2, 3;
(d) the p-rank, k, of the conjunctive parallelogram model for $\mathbf{T}$ and $\mathbf{M}$ at 4 levels: 2, 4, 6, 10;
(e) the error level, ε, which is the proportion of cells d_{ij} that is different from t_{ij}, at 4 levels: .00, .05, .10, .20.

The number of persons and objects, the p-rank and the error level will be considered random effects, whereas the d-rank will be considered fixed (Kirk, 1982, p. 64).

For each combination of number of persons h, number of objects n, d-rank r, p-rank k, and error level ε, 20 triplets of an $h \times r$ matrix $\mathbf{A}$, an $h \times r$ matrix $\mathbf{B}$, and an $n \times r$ matrix $\mathbf{Q}$ were generated with:

$$Q_{jv} \overset{\text{i.i.d.}}{\sim} \mathrm{U}\{0, \ldots, k-1\} \qquad (1 \leq j \leq n; 1 \leq v \leq r)$$

$$A_{iv}|(W_{iv} = w_{iv}) \overset{\text{i.i.d.}}{\sim} \text{U}\{0, \ldots, k - w_{iv}\} \qquad (1 \le i \le h; 1 \le v \le r)$$
$$B_{iv} = A_{iv} + W_{iv} - 1 \qquad (1 \le i \le h; 1 \le v \le r)$$

where

$$W_{iv} \overset{\text{i.i.d.}}{\sim} \text{Bin}(k, \pi) \qquad (1 \le i \le h; 1 \le v \le r)$$

and $\pi = \sqrt[r]{.5}$. W_{iv} may be interpreted as the width of person's i acceptance interval (i.e., the number of positions covered by person i) on dimension v. Matrices **T** were calculated by combining **A**, **B** and **Q** by the parallelogram combination rule:

$$t_{ij} = 1 \quad \text{iff} \quad \forall v (1 \le v \le r) : a_{iv} \le q_{jv} \le b_{iv} \quad (1 \le i \le h; 1 \le j \le n).$$

Note that by the chosen value for π, the expected proportion of one-entries in **T** equals .50. Next, a data matrix **D** was constructed from each **T** by altering the value of a randomly selected set of entries in **T**, consisting of a proportion ε of the total number of entries. Finally, the parallelogram algorithm with the initial configuration obtained by the rational heuristic was applied to find for each matrix **D** a matrix **M**, which can be represented by k^r-CPA-model. As a result of this procedure, 20×3 (numbers of persons) $\times$ 3 (numbers of objects) $\times$ 3 (d-ranks) $\times$ 4 (p-ranks) $\times$ 4 (error levels) $= 8{,}640$ triplets (**T**, **D**, **M**) were obtained.

3.2. Goodness of Fit

For each triplet, the proportion of discrepancies between **D** and **M**, which is a badness-of-fit (BOF) statistic, is used to evaluate the fit of the models yielded by the algorithm:

$$\text{BOF} = \frac{\sum_{i=1}^{h} \sum_{j=1}^{n} (d_{ij} - m_{ij})^2}{h \times n}. \tag{8}$$

An analysis of variance with BOF as the dependent variable and the five factors in the design as independent variables yields for the main effect of error level an intraclass correlation $\hat{\rho}_I$ (Haggard, 1958; Kirk, 1982) of .95, which means that almost all variance in badness of fit is accounted for by error level. Mean badness-of-fit values across the 2,160 observations within each error level are: .008, .057, .106, and .200 for ε equal to .00, .05, .10, and .20, respectively. The latter implies that the matrices **M** yielded by the algorithm are about as close to the data as the true model **T** is. In an analysis with the difference between BOF and ε as the dependent variable, the effect of error level is much lower, though still important ($\hat{\rho}_I = .06$) (see Table 5). The difference between BOF and ε being close to 0 for any of the error levels shows that the algorithm does not merely tail after the data. The main effect of error level may be attributed to BOF (slightly) coming closer to ε with increasing ε. Two more important main effects on BOF$-\varepsilon$ are to be considered: BOF $-\, \varepsilon$ increases with increasing p-rank ($\hat{\rho}_I = .12$) and increasing d-rank ($\hat{\omega}^2 = .09$):[1] The more complex the model that is fitted, the larger the difference between BOF and ε. Moreover, the interaction effect of d-rank by p-rank was found to be significant ($\hat{\rho}_I = .07$) due to d-rank having no effect at the lowest p-rank 2 and, vice versa, p-rank hardly having an effect for d-rank equal to 1. The interaction between p-rank and error level ($\hat{\rho}_I = .06$) can be similarly explained; for, BOF and ε are about equal at the lowest p-rank and at the highest error level. The other effects in the analysis are not important: Because almost any of the effects in the

[1]For fixed effects, ω^2 is used; for random effects, ρ_I is used (Kirk, 1982). Both statistics can be interpreted as proportions of variance accounted for (Hays, 1994).

TABLE 5.
Mean differences between badness of fit and ε at levels of d-rank $\times$ p-rank, and error level $\times$ p-rank

	p-rank				
	2	4	6	10	Overall
d-rank					
1	−.000	−.000	.000	.002	.000
2	−.000	.008	.011	.012	.008
3	−.000	.009	.012	.014	.009
Error level					
.00	.000	.007	.011	.015	.008
.05	−.000	.006	.010	.014	.007
.10	−.000	.006	.008	.010	.006
.20	−.001	.002	.001	−.002	.000
Overall	−.000	.005	.008	.009	.006

model reaches statistical significance due to the large number of observations (and the resulting high power), only effects accounting for at least 5% of the variance of the dependent variable (i.e., $\hat{\rho}_I$, $\hat{\omega}^2 \geq .05$) are discussed in the analyses of variance in this and the next subsection.

3.3. *Goodness of Recovery*

For each triplet (**T**, **D**, **M**), the proportion of discrepancies between **T** and **M** was calculated as a measure of the badness of recovery (BOR):

$$\text{BOR} = \frac{\sum_{i=1}^{h}\sum_{j=1}^{n}(t_{ij} - m_{ij})^2}{h \times n}.$$

The mean BOR across the 8,640 observations equals .022, which implies that the model yielded by the algorithm differs on average 2.2% from the underlying truth. An analysis of variance with BOR as the dependent variable reveals important main effects of error level ($\hat{\rho}_I = .16$), p-rank ($\hat{\rho}_I = .12$) and d-rank ($\hat{\omega}^2 = .07$). Badness of recovery clearly increases with higher error levels, higher p-ranks and higher d-ranks (see Table 6). However, the effect of error level strongly depends on the d-rank (with $\hat{\rho}_I = .11$ for the interaction effect), the p-rank ($\hat{\rho}_I = .09$), and the number of objects ($\hat{\rho}_I = .09$): At a low d-rank or p-rank, or for a high number of objects, error level hardly affects the recovery, whereas with increasing d-rank and p-rank and decreasing number of objects the effect of error level considerably increases. The smaller effect size of the error level by number of persons interaction ($\hat{\rho}_I = .02$) might be explained by the fact that adding a person generally yields less extra information as compared to adding an item (as persons have intervals instead of positions on the latent dimensions, which implies more uncertainty). Also the interaction between p-rank and d-rank should be considered ($\hat{\rho}_I = .06$), mainly because the combination of a high d-rank and a high p-rank gives worse results than expected on the basis of the individual main effects. The other effects in the analysis accounted for less than 5% of the variance of BOR.

An alternative and probably more informative statistic to evaluate goodness of recovery is the relative recovery gain (RRG) defined as:

$$\text{RRG} = \frac{\varepsilon - \text{BOR}}{\varepsilon}.$$

TABLE 6.
Mean badness of recovery at levels of d-rank × error level, p-rank × error level, and number of objects × error level

		Error level				
		.00	.05	.10	.20	Overall
d-rank						
	1	.000	.002	.003	.011	.004
	2	.012	.013	.019	.056	.025
	3	.013	.019	.031	.090	.038
p-rank						
	2	.000	.000	.001	.006	.002
	4	.007	.009	.013	.043	.018
	6	.011	.015	.022	.066	.029
	10	.015	.023	.035	.094	.042
Number of objects						
	40	.010	.016	.029	.086	.035
	90	.008	.010	.014	.043	.019
	150	.008	.008	.011	.027	.014
Overall		.008	.012	.018	.052	.022

for $\varepsilon > 0$ (RRG being undefined for $\varepsilon = 0$). The relative recovery gain equals 1 in case of perfect recovery and 0 if the model **M** yielded by the algorithm is as far from the true model **T** as the data **D** are; as a consequence, RRG may be interpreted as the progress with respect to **T** that is achieved by applying the algorithm, or, loosely speaking, as the proportion of the erroneous cells in **D** that have been corrected in **M**. The mean relative recovery gain amounts to .777 (across the 6,480 observations for which $\varepsilon > 0$). An analysis of variance with RRG as the dependent variable shows important main effects of p-rank ($\hat{\rho}_I = .25$), d-rank ($\hat{\omega}^2 = .14$) and the number of objects ($\hat{\rho}_I = .09$): The higher the p-rank or the d-rank, or the lower the number of persons, the lower the relative recovery gain. The p-rank by d-rank and p-rank by number of objects interactions additionally account for respectively 12% and 6% of the variance: As appears from Table 7, the combination of a high p-rank with a high d-rank or a low number of objects considerably deteriorates the relative recovery gain.

3.4. Rank Selection

In real-life applications, the d-rank r and the p-rank k of the k^r-CPA-model underlying a given data set are most often unknown (just like the number of factors in factor analysis or the number of clusters in cluster analysis). In that case, the user may typically fit parallelogram models with varying p-rank and d-rank and select one of them using some heuristic like, for example, the elbow criterion. In this section, we will present an alternative heuristic, a pseudo-binomial test (similar to the pseudo F-tests proposed by Hartigan (1975, pp. 89–91) to determine the number of clusters in k-means cluster analysis) and we will subsequently evaluate it in an extension of the simulation study discussed in Subsection 3.1.

We generally assume that for a given $h \times n$ data matrix **D** analyses have been carried out with d-ranks varying from 1 to R and p-ranks varying from 2 to K, yielding k^r-CPA-models ($r = 1, \ldots, R$; $k = 2, \ldots, K$) and associated matrices $\mathbf{M}_{r,k}$, and that for each $\mathbf{M}_{r,k}$ the value on the loss function (7), $f(\mathbf{M}_{r,k})$, and the badness-of-fit value defined in (8), $\text{BOF}_{r,k}$, have been calculated. The pseudo-binomial rule first selects, for each d-rank r, the smallest p-rank

TABLE 7.
Mean relative recovery gain at levels of d-rank × p-rank and number of objects × p-rank

	p-rank 2	4	6	10	Overall
d-rank					
1	.999	.990	.963	.876	.957
2	.991	.812	.679	.545	.757
3	.972	.680	.510	.305	.617
Number of objects					
40	.970	.748	.542	.357	.654
90	.995	.860	.771	.628	.813
150	.998	.875	.838	.741	.863
Overall	.988	.828	.717	.575	.777

k for which $f(\mathbf{M}_{r,k+1})$ exceeds the first percentile of the Binomial distribution $\text{Bin}(hn, \text{BOF}_{r,k})$. Let $k[r]$ denote the selected p-rank for d-rank r. Next the smallest r is selected for which $f(\mathbf{M}_{r+1,k[r+1]})$ exceeds the first percentile of the Binomial distribution $\text{Bin}(hn, \text{BOF}_{r,k^*})$, where $k^* = \max(k[r], k[r+1])$. This number r and the associated $k[r]$ are the d-rank and p-rank, respectively, that are finally retained. This rule is similar to the selection heuristic proposed by Leenen, Van Mechelen, and De Boeck (2001) and Leenen and Van Mechelen (2001).

In order to evaluate the proposed heuristic, we ran additional analyses on the data sets of the simulation study discussed in the previous subsections by applying the algorithm on data arrays in d-ranks and p-ranks that are different from the d-rank and p-rank of the underlying true matrix **T**. For clarity, we will denote the d-rank and the p-rank of the true matrix **T** using Greek letters, ρ and κ respectively, and the d-rank and the p-rank of the matrix **M** yielded by the analysis by Roman letters, r and k respectively. For computational reasons, we analyzed only a subset of the generated data arrays described in Subsection 3.1, that is, the data sets with true p-rank κ lower than or equal to 6, and, furthermore only the first 10 data sets in each cell of the design were analyzed. Analyses were done on the remaining 3,120 matrices **D** in d-ranks r varying from 1 to 4 and p-ranks k varying from 2 to 7. Subsequently, we applied the pseudo-binomial rule to select one model from the resulting set of $4(7-1) = 24$ models for each data matrix **D**.

The pseudo-binomial rule was found to select in 54.1% of the cases both the true d-rank ρ and the true p-rank κ. The true d-rank in itself is selected in 67.4% of the cases, and the true p-rank in 60.4%. The goodness of fit and the goodness of recovery of the k^r-CPA model selected by the pseudo-binomial rule are very close to the corresponding values for the κ^ρ-CPA model, with both the overall mean BOF-statistics and BOR-statistics differing less than .005.

Finally, we examined the effect of overestimating the d-rank or the p-rank on the goodness of recovery. To this end, we calculated for each data set the difference in badness of recovery between the κ^ρ-CPA model and both the $\kappa^{\rho+1}$-CPA model and the $(\kappa+1)^\rho$-CPA model. On the average, overestimating the p-rank or the d-rank was not found to have an effect on the badness of recovery, both calculated differences being close to 0.

4. Illustrative Application: Preferences for Holiday Trips

The conjunctive parallelogram model may be a useful model for choice data when the choice objects are believed to differ from each other with respect to several underlying dimensions. In this section, we illustrate the new model with an analysis of data on preferences for holiday trips, which are believed to be judged with respect to several characteristics. Other possible relevant applications include the study of preferences for political parties (as most parties can be considered to take a position on several underlying opinion dimensions) and the study of the applicability of emotions to different situations or persons as, according to several emotion theories, emotions can be decomposed in terms of a limited set of more basic dimensions (see, e.g., Smits, De Boeck, Kuppens, & Van Mechelen, 2002).

For the present application, descriptions of 45 widely varying holiday trips (from several travel brochures) were presented to 98 students to have them select those trips on which they would like to have more information at the travel agency. Because the students' choices were expected to be affected by their exam results, the subjects were presented the task twice, that is, under both the assumption of having resits and the assumption of not having resits. (In Belgium, students who succeed in their first examination period have three-month summer holidays, whereas students who have resits have vacation for about two or three weeks only.)

Both resulting 98×45 binary matrices were analyzed by the conjunctive parallelogram algorithm. As the dimensional structure for the objects turned out to be very similar regardless of having resits or not, we present here the results of the parallelogram analyses on the concatenated 196×45 matrix. Note that an analysis of such concatenated data implies that the structure of the holiday trips is fixed across resit and no-resit data but that the persons are allowed to differ across the two conditions. We fitted k^r-CPA-models with d-ranks r varying from 1 to 4 and p-ranks k varying from 2 to 7; the proportion of discrepancies (i.e., the badness of fit) of the resulting models is presented in Table 8. Applying the pseudo-binomial rule to these results leads to the selection of the 4^2-CPA-model; hence, this model is retained for further discussion.

Figure 2 shows the graphical representation of the selected model. The figure displays the number of holiday trips at each of the 4^2 positions in the model; also, the four acceptance rectangles with the highest person frequencies are displayed. Based on the trips' descriptions, trip classes can be given a substantive interpretation: For example, all 9 trips at position $\mathbf{q} = (2, 3)$ are city trips (in total, 10 city trips were included in the study); the three trips at positions $\mathbf{q} = (3, 1)$ are group hiking tours, whereas both trips at the neighbouring position $\mathbf{q} = (3, 2)$ are individual hiking tours. Furthermore, both dimensions can be given a substantive interpretation. The first dimension is strongly related to the relative price, that is, the price per week, the Spearman rank correlation between position on the first dimension and relative price being equal to .62. The second dimension is related to the level of organization, with at one side trips characterized by travelling in group, arranged meals, and so on, and at the other side trips in which the participants have to organize each day themselves. The latter interpretation was validated by relating

TABLE 8.
Proportion of discrepancies for the models for the holiday preference data as a function of the d-rank and the p-rank

	d-rank			
	1	2	3	4
p-rank				
2	.206	.180	.157	.143
3	.190	.154	.133	.123
4	.178	.141	.124	.110
5	.169	.134	.117	.104
6	.163	.130	.113	.101
7	.162	.129	.110	.097

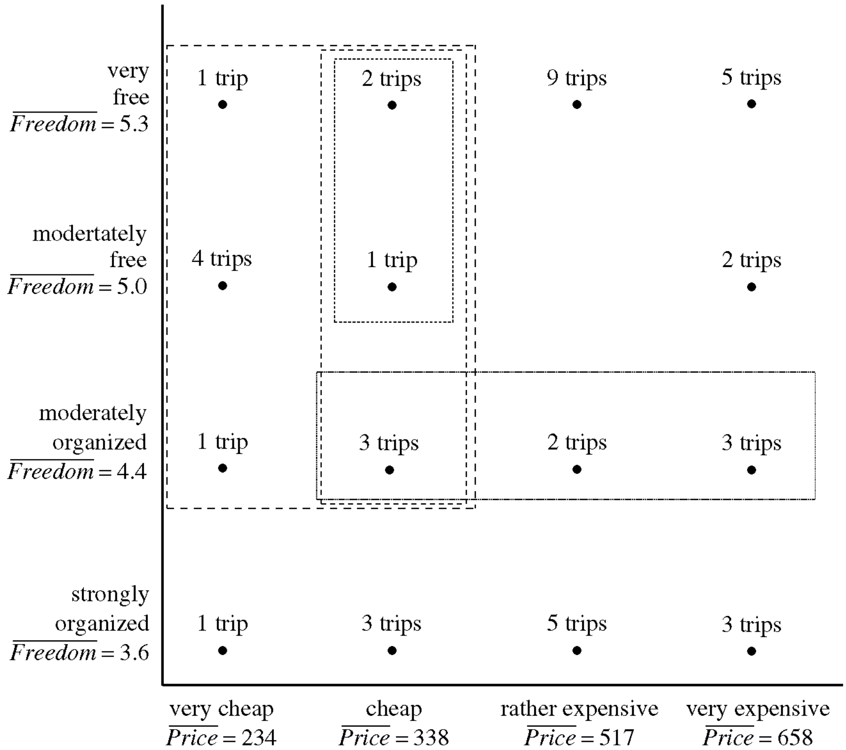

FIGURE 2.
Graphical representation of the 4^2-CPA-model for the holiday trip preference data. *Note*. At each position of the first dimension, the mean price (in Euro) of the trips is given; at the second dimension, the mean judged freedom across judges and trips on each position is given. Additionally, the four most frequent acceptance rectangles are drawn.

the dimension to external ratings of the 45 trips on a 7-point scale (ranging from "a complete lack of freedom" versus "a lot of freedom"), with the Spearman rank correlation equal to .49. With respect to the persons, subjects characterized by the intervals [0, 1] and [1, 3] on the first and second dimension, respectively, are interested in (very) cheap trips that are not strongly organized, whereas subjects with intervals [1, 1] and [2, 3] are more choosy and only select trips with at least a moderate degree of freedom that are cheap, but not very cheap, the latter possibly being considered "suspiciously" cheap. Finally, it is interesting to note that the mean width of a person's interval on the level-of-organization dimension is significantly smaller for the "resit"-condition as compared to the no resit condition ($t(97) = 2.10$, $p < .05$), which implies that if persons have to resit exams, they are in general more choosy with respect to the level of organization. An explanation of why students with resits are more choosy may be that due to the limited amount of time they have available for their holiday trip, they prefer to play it safe: They may not want to risk their trip—and, hence, their complete holiday period—to be disappointing, whereas students without resits may still have plenty of time to recover from a disappointing trip.

This illustrative example shows that the conjunctive parallelogram model may yield substantive insights into the psychological mechanisms underlying decision making and choice. First, a meaningful clustering of the holiday trips into homogeneous classes as well as two well interpretable dimensions were induced from the data. The solution being well interpretable with respect to both the clustering and the dimensions can further be considered an indication of the psychological validity of the model assumed, and provides a strong support for the hybrid nature of that model, which simultaneously is a classification and a dimensional model. Also, the conjunctive nature of the association rule in the CPA-model was supported: At the person side, most students were found to conjunctively combine the conditions on both dimensions (see, e.g., the four most frequent acceptance rectangles in Figure 2), although the model allows students to have an acceptance interval that fully covers one dimension and, hence, to take into account a single dimension only.

5. Relation to Other Models

As a unidimensional, deterministic model for choice data, Coombs' parallelogram model (and the CPA-model of d-rank 1) closely relates to unidimensional unfolding models (Coombs, 1964). The latter models are typically used in case of rank-order data (i.e., data on persons who have rank-ordered the n objects) rather than pick any/n data. In unidimensional unfolding a joint scale for persons and items is assumed to underly the data with a person's preference order corresponding to the order of the distances between the respective objects and the person's position. Furthermore, several probabilistic extensions of the unidimensional parallelogram (and unfolding) model have been proposed (Andrich, 1988, 1989, 1995; Andrich & Luo, 1993; Hoijtink, 1990, 1991; Hoijtink & Molenaar, 1992; Roberts & Laughlin, 1996; Verhelst & Verstralen, 1993). Luo (1998) presents a general model subsuming unidimensional and multidimensional variants of these models. Typically, in probabilistic parallelogram models the probability of a person selecting an object is assumed to be some prespecified function of the distance between the object's and the person's position on an underlying scale. (Note that, as a consequence of the model matrix being a probability matrix, order-preservance in terms of quasi-orders defined on binary matrices is lost in these models.)

A key ingredient of both the unidimensional and multidimensional parallelogram models as proposed in this paper is the option to restrict the number of positions on the dimensions underlying the data. This option was motivated by the desire to (a) turn the parallelogram models into classification models and (b) improve the quality (particularly, the reliability and the robustness) of the output of the data analysis associated with the models. Interestingly, both desires have their counterpart within the framework of some related, probabilistic models for pick any/n data. In particular, regarding (a), several types of latent class models have been proposed, in which persons are classified into a (usually small) number of homogeneous subgroups (Böckenholt and Böckenholt, 1991; Croon, 1993; Formann, 1988, 1993). Moreover, regarding (b), Böckenholt and Böckenholt (1991) indicated that in stochastic models parameter estimation problems may show up if the number of positions on underlying dimensions as included in those models is not constrained. In this respect, one may also note that for estimation of dominance models for two-way binary data, such as the conjunctive Rasch-model, a discretization of the underlying dimensions into a prespecified number of "nodes" has been proposed (van Leeuwe and Roskam, 1991). For the estimation of probabilistic parallelogram models, a similar approach might be considered.

As a multidimensional deterministic choice model, the conjunctive parallelogram model relates to multidimensional unfolding (Coombs, 1964), which assumes that the (Euclidean) distances of the objects' positions to a person's ideal point in a multidimensional space reflects the person's preference order with respect to the objects. As a multidimensional generalization of Coombs' unidimensional parallelogram model, the new conjunctive parallelogram model is also related to Feger's (1994) parallelogram model. Unlike the conjunctive model (but similar to multidimensional unfolding models), Feger uses a compensatory combination rule, though, with the unidimensional notion of acceptance intervals being multidimensionally generalized into acceptance (hyper)spheres in a metric space. However, Feger's model when used with a city-block metric (Arabie, 1991; Borg and Groenen, 1997, pp. 277–281) turns out to be equivalent to a special case of the conjunctive parallelogram model, namely with the restriction that the width of a person's intervals is equal across dimensions.

As a deterministic decomposition model for binary relations, the new model is a member of a generic class of decomposition models for n-ary relations described by Leenen et al. (1999). The latter paper includes general propositions on, among others, existence and order-preservance of decomposition models as well as generic tools to find the minimal d-rank r such that for a given binary matrix $\mathbf{M}$ and p-rank k, a k^r-CPA-model exists. The parallelogram model is also related to another member of this generic class: the conjunctive ORDCLAS model (Leenen et al., 2001). The latter model also assumes that a number of dimensions underly the data and, in particular, when applied within a choice context, assumes that persons have minimal requirements with respect to each relevant dimension and that a person selects an object iff the object meets the minimal requirement for that person on each dimension. The CPA-model may be considered a generalization of the conjunctive ORDCLAS model as the CPA combination rule (3) with the lower bounds of the persons' intervals restricted to 0 (i.e., $a_{iv} = 0$, $\forall i, v$: $1 \leq i \leq n$, $(1 \leq v \leq r)$ is equivalent to the ORDCLAS combination rule. The difference between the parallelogram model and the ORDCLAS model conceptually relates to Gati and Tversky's (1982) distinction between qualitative and quantitative dimensions. This distinction implies that an object j can be characterized by a set of elementary features F_j; on a quantitative dimension, the features sets are ordered by inclusion, that is, for any pair j, j' either $F_j \subseteq F_{j'}$ or $F_j \supseteq F_{j'}$, such that objects higher on the dimension include all features of lower objects. A qualitative dimension, on the other hand, merely requires $F_j \cap F_{j''} \subseteq F_{j'}$ if j' is between j and j'' on the dimension. The latter clarifies why the conjunctive parallelogram model, contrary to an ORDCLAS model, cannot be guaranteed to be order-preserving with respect to the objects.

Appendix A

Theorem 1. For any pair of integers k (≥ 2) and r (≥ 1) a binary $h \times n$ matrix $\mathbf{M}$ exists for which:

(a) an $h \times r$ matrix $\mathbf{A}$, an $h \times r$ matrix $\mathbf{B}$ and an $n \times r$ matrix $\mathbf{Q}$ exist with entries

$$a_{iv}, b_{iv}, q_{jv} \in \{0, \ldots, k-1\}(1 \leq i \leq h, 1 \leq j \leq n, 1 \leq v \leq r)$$

such that (3) holds;

(b) no matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{Q}$ exist that satisfy both (3) and (6).

Proof. Define the $h \times n$ matrix $\mathbf{M}$ with $h = k^r - 1$, $n = k^r$, and entries $m_{ij} = 1$ if $i = j$ and $m_{ij} = 0$, otherwise. For any given pair of objects j and j', a person i can be found such that $m_{ij} \neq m_{ij'}$. From (3) follows:

$$\forall j, j'(1 \leq j, j' \leq n) : q_{jv} \neq q_{j'v} \text{ for some } v. \tag{A1}$$

The latter results in a unique pattern $(q_{j1}, \ldots, q_{jr})$ for any j, and because $n = k^r$, an object exists for any pattern $(q_{j1}, \ldots, q_{jr})$ with $q_{jd} \in \{0, \ldots, k-1\}$. From (3) and the definition of $\mathbf{M}$, it is further derived that $a_{iv} < b_{iv}$ for some i and v implies $m_{ij} = 1$ for any object j which satisfies $q_{jv'} = q_{iv'}$ for $v' \neq v$ and $q_{jv} \in \{a_{iv}, b_{iv}\}$. As the latter is true for at least two objects (because all patterns $(q_{j1}, \ldots, q_{jr})$ occur) and as $m_{ij} = 1$ iff $i = j$, it follows that $a_{iv} = q_{iv} = b_{iv}$ for any i and v. Clearly, matrices $\mathbf{Q}$, $\mathbf{A}$ and $\mathbf{B}$ that satisfy the above restrictions can be constructed, which proves part (a).

By the imposed restrictions on $\mathbf{Q}$, $\mathbf{A}$ and $\mathbf{B}$ and by $n \preceq_C j$ for any j, (6) implies

$$\forall j(1 \leq j \leq n), \forall v(1 \leq v \leq r), \forall i(1 \leq i \leq h) : [q_{nv} = a_{iv} \Longrightarrow q_{jv} = a_{iv} = q_{nv}]$$

Since for any v a person i exists with $q_{nv} = a_{iv}$, it follows from the latter equation that $q_{jv} = q_{nv}$ for all v and j, which contradicts (A1) and consequently proves part (b). □

Appendix B

Given an array $\mathbf{d} = (d_1, \ldots, d_n)$ with $d_j \in \{0, 1\}(1 \leq j \leq n)$ and an $n \times r$ matrix $\mathbf{Q}$ with $q_{jv} \in \{0, \ldots, k-1\}(1 \leq j \leq n, 1 \leq v \leq r)$, the following procedure returns arrays $\mathbf{a} = (a_1, \ldots, a_r)$ and $\mathbf{b} = (b_1, \ldots, b_r)$, such that

$$f(\mathbf{a}, \mathbf{b}) = \sum_{j=1}^{n} (d_j - m_j)^2$$

is minimal, where $m_j = 1$ if $\forall v(1 \leq v \leq r)$: $a_v \leq q_{jv} \leq b_v$, and $m_j = 0$ otherwise. The procedure is a branch-and-bound procedure and relies on the following bounding function:

$$f_{\text{b}}(\mathbf{a}, \mathbf{b}) = \#\{j | (1 \leq j \leq n) \text{ and } d_j = 1 \text{ and } m_j = 0\},$$

where $\#S$ denotes the cardinality of set S. Clearly, $f(\mathbf{a}, \mathbf{b}) \geq f_{\text{b}}(\mathbf{a}, \mathbf{b})$ for any $(\mathbf{a}, \mathbf{b})$. Given some $\mathbf{a}', \mathbf{b}'$ for which $a_v \leq a'_v \leq b'_v \leq b_v$ for all $v(1 \leq v \leq r)$, it holds that $f_{\text{b}}(\mathbf{a},' \mathbf{b}') \geq f_{\text{b}}(\mathbf{a}, \mathbf{b})$ and, consequently, $f_{\text{b}}(\mathbf{a}, \mathbf{b}) \geq L$ implies $f(\mathbf{a},' \mathbf{b}') \geq L$ given some lower bound L. The procedure essentially starts from the maximal interval $[a_v, b_v]$ on each dimension and, subsequently, the intervals are narrowed until $f_{\text{b}}(\mathbf{a}, \mathbf{b})$ yields a value larger than the current minimum.

In the procedure below, $\mathbf{a}$ and $\mathbf{b}$ hold the current optimal values and, hence, $f(\mathbf{a}, \mathbf{b})$ holds the current minimum. The n-dimensional arrays $\mathbf{a}'$ and $\mathbf{b}'$, and the integers d and l, are to be considered local.

```
procedure EstimateIntervals (input variables: n, r, k, d, Q; output variables: a, b);
begin
    for all v(1 ≤ v ≤ r) do {a_v ← 1; b_v ← 0};
    for all v(1 ≤ v ≤ r) do {a'_v ← 0; b'_v ← k − 1};
    repeat
        if f(a,'b') < f(a, b)) then {a ← a'; b ← b'};
        v ← r;
        if f_b(a,'b') ≥ f(a, b) then
            {while (v ≥ 1) and (a'_v = 0) and (b'_v = k − 1) do v ← v − 1;
            if (v ≥ 1) and (b'_v < k − 1)
                then {a'_v ← 1; b'_v ← b_v + 1; continue};
                else v ← v − 1};
        while (v ≥ 1) and (a'_v = k − 1) do v ← v − 1;
        if v ≥ 1 then
            {if a'_v = b'_v
                then {a'_v ← 1; b'_v ← b'_v + 1};
                else {if a'_v = 0 then b'_v ← b'_v − 1 else a'_v ← a'_v + 1}};
        if v ≥ 1 then
            {for all l (v < l ≤ r) do {a'_l ← 0; b'_l ← k − 1}};
    until v <= 0;
end;
```

References

Abelson, R. P., & Levi, A. (1985). Decision making and decision theory. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology: Vol. 1. Theory and method* (3rd ed., pp. 231–309). New York: Random House.
Andrich, D. (1988). The application of an unfolding model of the PIRT type to the measurement of attitude. *Applied Psychological Measurement, 12*, 33–51.
Andrich, D. (1989). A probabilistic IRT model for unfolding preference data. *Applied Psychological Measurement, 13*, 193–216.
Andrich, D. (1995). Hyperbolic cosine latent trait models for unfolding direct responses and pairwise preferences. *Applied Psychological Measurement, 19*, 269–290.
Andrich, D., & Luo, G. (1993). A hyperbolic cosine latent trait model for unfolding dichotomous single-stimulus responses. *Applied Psychological Measurement, 17*, 253–276.
Arabie, P. (1991). Was Euclid an unnecessarily sophisticated psychologist? *Psychometrika, 56*, 567–587.
Arabie, P., & Hubert, L. (1992). Combinatorial data analysis. *Annual Review of Psychology, 43*, 169–203.
Barbut, M., & Monjardet, B. (1970). *Ordre et classification: Algèbre et combinatoire* [Order and classification: Algebra and combinatorics]. Paris: Hachette.
Beach, L. R. (1990). Image theory: Decision making in personal and organizational contexts. Chichester, England: Wiley.
Böckenholt, U., & Böckenholt, I. (1991). Constrained latent class analysis: Simultaneous classification and scaling of discrete choice data. *Psychometrika, 56*, 699–716.
Booth, K. S., & Lueker, G. S. (1976). Testing for the consecutive ones property, interval graphs, and graph planarity using *PQ*-tree algorithms. *Journal of Computer and System Sciences, 13*, 335–379.
Borg, I., & Groenen, P. (1997). *Modern multidimensional scaling: Theory and applications*. New York: Springer–Verlag.
Carroll, J. D., & Chaturvedi, A. (1995). A general approach to clustering and multidimensional scaling of two-way, three-way, or higher-way data. In R. D. Luce, M. D'Zmura, D. Hoffman, G. J. Iverson & A. K. Romney (Eds.), *Geometric representations of perceptual phenomena* (pp. 295–318). Mahwah, NJ: Erlbaum.
Coombs, C. H. (1964). *A theory of data*. New York: Wiley.
Coombs, C. H., & Smith, J. E. (1973). On the detection of structure in attitudes and developmental processes. *Psychological Review, 80*, 344–358.
Croon, M. (1993). Ordinal latent class analysis for single-peaked items. *Kwantitatieve Methoden, 42*, 127–142.
DeSarbo, W. S., & Hoffman, D. L. (1986). Simple and weighted unfolding treshold models for the spatial representation of binary choice data. *Applied Psychological Measurement, 10*, 247–264.
Falmagne, J.-C., Koppen, M., Vilano, M., Doignon, J.-P., & Johannesen, L. (1990). Introduction to knowledge spaces: How to build, test and search them. *Psychological Review, 97*, 201–224.
Feger, H. (1994). *Structure analysis of co-occurence data*. Aachen, Germany: Shaker.
Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M., & Doherty, M. L. (1989). Process tracing methods: contributions, problems, and neglected research questions. *Organizational Behavioral and Human Decision Processes, 43*, 75–117.
Formann, A. K. (1988). Latent class models for nonmonotone dichotomous items. *Psychometrika, 53*, 45–62.
Formann, A. K. (1993). Latent class models for monotone and nonmonotone dichotomous items. *Kwantitatieve Methoden, 42*, 143–160.
Fulkerson, D. R., & Gross, O. A. (1965). Incidence matrices and interval graphs. *Pacific Journal of Mathematics, 15*, 835–855.
Ganter, B., & Wille, R. (1996). *Formale begriffsanalyse: Mathematische grundlagen* [Formal concept analysis: Mathematical foundations]. Berlin, Germany: Springer–Verlag.
Gara, M. A., & Rosenberg, S. (1979). The identification of persons as supersets and subsets in free-response personality descriptions. *Journal of Personality and Social Psychology, 37*, 2161–2170.
Gati, I., & Tversky, A. (1982). Representations of qualitative and quantitative dimensions. *Journal of Experimental Psychology, 8*, 325–340.
Gelman, A., Leenen, I., Van Mechelen, I., De Boeck, P., & Poblome, J. (2001). *Bridges between deterministic and probabilistic classification models*. Submitted for publication.
Ghosh, S. P. (1972). File organization: the consecutive retrieval property. *Communication of the Association for Computing Machinery, 9*, 802–808.
Goldmann, K. (1971). Some archaeological criteria for chronological seriation. In F. R. Hodson, D. G. Kendall & P. Tăutu (Eds.), *Mathematics in the Archaeological and Historical Sciences: Proceedings of the Anglo-Romanian Conference* (pp. 202–208). Edinburgh, Scotland: University Press.
Haggard, E. A. (1958). *Intraclass correlation and the analysis of variance*. New York: Dryden.
Hartigan, J. A. (1975). *Clustering algorithms*. New York: Wiley.
Hays, W. L. (1994). *Statistics* (Fifth ed.). New York: Harcourt Brace.
Hoijtink, H. (1990). A latent trait model for choice data. *Psychometrika, 55*, 641–656.
Hoijtink, H. (1991). The measurement of latent traits by proximity items. *Applied Psychological Measurement, 15*, 153–169.
Hoijtink, H., & Molenaar, I. W. (1992). Testing for DIF in a model with single peaked item characteristic curves: The PARELLA model. *Psychometrika, 57*, 383–397.
Hovland, C. I., Harvey, O. J., & Sherif, M. (1957). Assimilation and contrast effects in reactions to communication and attitude change. *Journal of Abnormal Social Psychology, 55*, 244–252.

Kendall, D. G. (1969). Incidence matrices, interval graphs and seriation in archaeology. *Pacific Journal of Mathematics*, *28*, 565–570.

Kirk, R. E. (1982). *Experimental design: Procedures for the behavioral sciences* (2nd ed.). Belmont, CA: Brooks/Cole.

Leenen, I., & Van Mechelen, I. (2001). An evaluation of two algorithms for hierarchical classes analysis. *Journal of Classification*, *18*, 57–80.

Leenen, I., Van Mechelen, I., & De Boeck, P. (1999). A generic disjunctive/conjunctive decomposition model for n-ary relations. *Journal of Mathematical Psychology*, *43*, 102–122.

Leenen, I., Van Mechelen, I., & De Boeck, P. (2001). Models for ordinal hierarchical classes analysis. *Psychometrika*, *66*, 389–404.

Leik, R. K., & Matthews, M. (1968). A scale for developmental processes. *American Sociological Review*, *33*, 62–75.

Luo, G. (1998). A general formulation of multidimensional unfolding models involving the latitude of acceptance. In A. Rizzi, M. Vichi & H.-H. Bock (Eds.), *Advances in data science and classification* (pp. 483–488). Berlin, Germany: Springer–Verlag.

Ogilvie, J. R., & Schmitt, N. (1979). Situational influences on linear and nonlinear use of information. *Organizational Behavior and Human Performance*, *4*, 337–352.

Roberts, J. S., & Laughlin, J. E. (1996). A unidimensional item response model for unfolding responses from a graded disagree-agree scale. *Applied Psychological Measurement*, *20*, 231–255.

Smits, D. J. M., De Boeck, P., Kuppens, P., & Van Mechelen, I. (2002). The structure of negative emotion scales: Generalization over contexts and comprehensiveness. *European Journal of Personality*, *16*, 127–141.

Tucker, A. C. (1972). A structure theorem for the consecutive 1's property. *Journal of Combinatorial Theory*, *12*, 153–162.

van Leeuwe, J. F. J., & Roskam, E. E. (1991). The conjunctive item response model: A probabilistic extension of the Coombs and Kao model. *Methodika*, *5*, 14–32.

Van Schuur, W. H. (1993a). MUDFOLD. *Kwantitatieve Methoden*, *42*, 39–54.

Van Schuur, W. H. (1993b). Unfolding models for pick any/n data: a summary and discussion. *Kwantitatieve Methoden*, *42*, 161–176.

Vansteelandt, K., & Van Mechelen, I. (1998). Individual differences in situation-behavior profiles: A triple typology model. *Journal of Personality and Social Psychology*, *75*, 751–765.

Verhelst, N. D., & Verstralen, H. H. F. M. (1993). A stochastic unfolding model derived from the partial credit model. *Kwantitatieve Methoden*, *42*, 73–92.

Westenberg, M. R. M., & Koele, P. (1992). Response modes, decision processes and decision outcomes. *Acta Psychologica*, *80*, 169–184.

Manuscript received 4 JAN 1999
Final version received 2 DEC 2001