Tucker3 hierarchical classes analysis — **Source link**

Eva Ceulemans, Iven Van Mechelen, Iwein Leenen

Institutions: Katholieke Universiteit Leuven

Related papers:

- Hierarchical classes: Model and data analysis

- Indclas: a three-way hierarchical classes model

- The Conjunctive Model of Hierarchical Classes.

- An Evaluation of Two Algorithms for Hierarchical Classes Analysis

- Towards a standardized notation and terminology in multiway analysis

TUCKER3 HIERARCHICAL CLASSES ANALYSIS

EVA CEULEMANS, IVEN VAN MECHELEN, AND IWIN LEENEN

KATHOLIEKE UNIVERSITEIT LEUVEN

This paper presents a new model for binary three-way three-mode data, called Tucker3 hierarchical classes model (Tucker3-HICLAS). This new model generalizes Leenen, Van Mechelen, De Boeck, and Rosenberg's (1999) individual differences hierarchical classes model (INDCLAS). Like the INDCLAS model, the Tucker3-HICLAS model includes a hierarchical classification of the elements of each mode, and a linking structure among the three hierarchies. Unlike INDCLAS, Tucker3-HICLAS (a) does not restrict the hierarchical classifications of the three modes to have the same rank, and (b) allows for more complex linking structures among the three hierarchies. An algorithm to fit the Tucker3-HICLAS model is described and evaluated in an extensive simulation study. An application of the model to hostility data is discussed.

Key words: three-way three-mode data, binary data, hierarchical classes, multiway data analysis, clustering.

1. Introduction

Binary N-way N-mode data are data that define a N-ary relation among N different sets. Such data are often found in psychology. One can, for example, think of the "solves" relation between a set of children and a set of items, which implies binary two-way two-mode data. Another example is the "elicits" relation among a set of persons, a set of situations and a set of responses, which implies binary three-way three-mode data.

The family of hierarchical classes models (De Boeck & Rosenberg, 1988; Van Mechelen, De Boeck, & Rosenberg, 1995; Leenen, Van Mechelen, De Boeck, & Rosenberg, 1999, Leenen, Van Mechelen, & De Boeck, 2001) is a collection of structural models for binary N-way N-mode data. Being related to the family of principal component models for real-valued N-way N-mode data (for an overview, see Kroonenberg, 1983), all hierarchical classes models reduce each mode to a number of binary variables, called bundles. All hierarchical classes models further include the following three basic features: (a) a classification of the elements of each mode, (b) a hierarchical organization of the classification per mode, in terms of if-then type relations, and (c) a linking structure among the hierarchical classifications, which accounts for the N-ary relation among the modes.

In this paper we will focus on the case of binary three-way three-mode data. We will refer to the elements of the first mode as objects, to the elements of the second mode as attributes, and to the elements of the third mode as sources. Within the family of hierarchical classes models, the individual differences hierarchical classes model (INDCLAS; Leenen et al., 1999) has been developed for this type of data. As a hierarchical classes model, INDCLAS consists of a hierarchical classification of the elements of each of the three modes, the three hierarchies being connected by a linking structure. INDCLAS is a restrictive model, however: First, the three hierarchies are constrained to have the same number of base classes or rank, since the three modes are reduced to the same number of bundles; second, the linking structure among the three hierarchies is restricted to a one-to-one correspondence among the respective bundles. Obviously, the IND-

CLAS model constitutes the hierarchical classes counterpart of the PARAFAC/CANDECOMP model (Carroll & Chang, 1970; Harshman, 1970) for real-valued three-way three-mode data.

In this paper, we present a new hierarchical classes model for binary three-way three-mode data, which generalizes INDCLAS in the same way the Tucker3 model (Tucker, 1966) generalizes PARAFAC/CANDECOMP; therefore, this new model is named Tucker3-HICLAS. Like INDCLAS, Tucker3-HICLAS consists of three linked hierarchies. Unlike INDCLAS, however, Tucker3-HICLAS allows the number of bundles to differ across the three modes; hence, the three resulting hierarchies are not restricted to have the same number of base classes. Furthermore, Tucker3-HICLAS allows for more complex linking structures among the three hierarchies. As a consequence, Tucker3-HICLAS can, in principle, yield simpler classifications for one or more of the modes. Note that, although we can distinguish between a disjunctive and a conjunctive variant of the INDCLAS and Tucker3-HICLAS models, for ease of explanation, we will only consider the disjunctive type in this paper.

In section 2, starting from the INDCLAS model, we introduce the new Tucker3-HICLAS model. Section 3 describes the aim of and an algorithm for Tucker3-HICLAS data analysis. In section 4 the results of a simulation study to evaluate the algorithm's performance are reported. In section 5, the Tucker3-HICLAS model is illustrated with an application on hostility data. Section 6 contains some concluding remarks.

2. Model

2.1. The INDCLAS Model

An INDCLAS model approximates an I (objects) $\times J$ (attributes) $\times K$ (sources) binary data array $\underline{\mathbf{D}}$ by an $I \times J \times K$ binary model array $\underline{\mathbf{M}}$ which can be decomposed into an $I \times R$ binary matrix $\mathbf{A}$, a $J \times R$ binary matrix $\mathbf{B}$ and a $K \times R$ binary matrix $\mathbf{C}$, where R denotes the rank of the model. The R columns of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ define R (possibly overlapping) clusters, called bundles, of objects, attributes and sources, respectively; consequently, $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are called an object, attribute and source bundle matrix.

As any hierarchical classes model, an INDCLAS model represents three types of structural relations in $\underline{\mathbf{M}}$: equivalence, hierarchy and association. The hypothetical array $\underline{\mathbf{M}}$ in Table 1 will be used as a guiding example to illustrate the three types of structural relations; Table 2 presents the bundle matrices of an INDCLAS model for $\underline{\mathbf{M}}$.

Equivalence. Three equivalence relations are defined: one on each mode of $\underline{\mathbf{M}}$. In the case of the objects, object i is equivalent with object i' in $\underline{\mathbf{M}}$ iff $m_{ijk} = m_{i'jk}$ for all $j = 1..J, k = 1..K$. The equivalence relations among the attributes and sources are defined similarly. Equivalent

TABLE 1.
Hypothetical INDCLAS model array

Source A						Source B						Source C					
	Attributes						Attributes						Attributes				
Objects	a	b	c	d	e	Objects	a	b	c	d	e	Objects	a	b	c	d	e
1	0	0	1	1	1	1	0	0	1	1	1	1	0	0	1	1	1
2	0	0	0	0	0	2	0	0	0	0	0	2	0	0	0	0	0
3	0	0	1	1	1	3	1	0	1	1	1	3	1	0	1	1	1
4	0	1	0	1	0	4	0	0	0	0	0	4	0	1	0	1	0
5	0	0	1	1	1	5	1	0	1	1	1	5	1	0	1	1	1
6	0	1	1	1	1	6	0	0	1	1	1	6	0	1	1	1	1
7	0	0	0	0	0	7	1	0	0	1	0	7	1	0	0	1	0

TABLE 2.
INDCLAS Model for $\underline{\mathbf{M}}$ in Table 1

Objects	Object Bundles			Attributes	Attribute Bundles			Sources	Source Bundles		
	OB_1	OB_2	OB_3		AB_1	AB_2	AB_3		SB_1	SB_2	SB_3
1	0	1	0	a	1	0	0	A	0	1	1
2	0	0	0	b	0	0	1	B	1	1	0
3	1	1	0	c	0	1	0	C	1	1	1
4	0	0	1	d	1	1	1				
5	1	1	0	e	0	1	0				
6	0	1	1								
7	1	0	0								

objects (resp. attributes, sources) constitute an object (resp. attribute, source) class, with those classes implying a partition of the objects (resp. attributes, sources). An INDCLAS model for $\underline{\mathbf{M}}$ represents those equivalence relations in that equivalent elements have identical bundle patterns (i.e., the set of bundles to which an element belongs). For example, in Table 1, Objects 3 and 5 are equivalent; hence, Object 3 and 5 have identical bundle patterns in the INDCLAS model of Table 2.

Hierarchy. Hierarchical relations are defined among the elements of each mode of $\underline{\mathbf{M}}$. With respect to the objects, object i is hierarchically below object i' in $\underline{\mathbf{M}}$ iff $m_{ijk} \leq m_{i'jk}$ for all $j = 1..J, k = 1..K$. The hierarchical relations among the attributes and sources are defined similarly. Note that the definition of hierarchical relations between the elements of a mode implies the definition of hierarchical relations between the classes of that mode. In the INDCLAS model for $\underline{\mathbf{M}}$, hierarchical relations are represented as follows: an element (class) x is hierarchically below an element (class) y iff the bundle pattern of x is a subset of the bundle pattern of y. For example, in Table 1, Object 4 is hierarchically below Object 6; consequently, in Table 2, the bundle pattern of Object 4 is a subset of the bundle pattern of Object 6. The relations of equivalence and hierarchy are called *set-theoretical* relations.

Association. The association relation is the ternary relation among the objects, attributes and sources as defined by the 1-entries of the array $\underline{\mathbf{M}}$. The INDCLAS model represents the association relation by the following association rule:

$$m_{ijk} = \bigoplus_{r=1}^{R} a_{ir} b_{jr} c_{kr} \qquad (\forall i = 1..I, \forall j = 1..J, \forall k = 1..K) \qquad (1)$$

where $\bigoplus$ denotes the Boolean sum. In other words, an entry m_{ijk} equals one iff there is a bundle to which object i, attribute j and source k belong. For example, in Table 2, Object 1 is associated with Attribute c according to Source A, because all three elements belong to the second bundle, that is, OB_2, AB_2 and SB_2. Note that (1) defines a linking structure among the hierarchical classifications of the three modes that implies a one-to-one correspondence among the respective object, attribute and source bundles.

Leenen et al. (1999) proposed an overall graphical representation of the INDCLAS model that accounts for the three types of structural relations in the model. Figure 1 shows an overall graphical representation of the INDCLAS model of Table 2. To obtain the graphical representation, one first draws the hierarchical classification of objects and attributes in the upper and lower half of the representation, respectively. Equivalent elements are enclosed by boxes, which represent the equivalence classes, and the hierarchical relations between the classes are represented by lines between the respective boxes. Note that the attribute hierarchy is represented upside down.
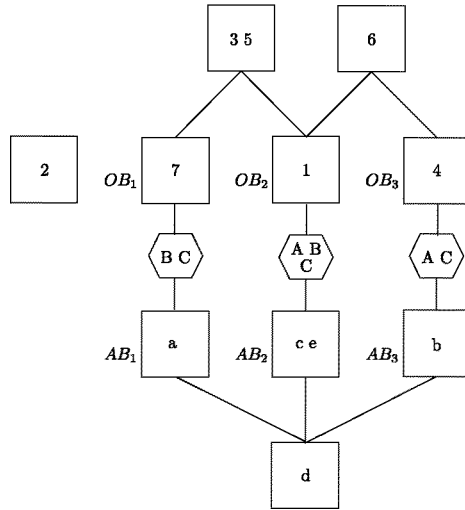
FIGURE 1.
Overall graphical representation of the INDCLAS model in Table 2.

Subsequently, the object and attribute hierarchies are linked by drawing paths between the base classes (i.e., classes of elements that belong to one bundle only) of the corresponding object and attribute bundles; each path includes a hexagon that contains the (classes of) sources that belong to the corresponding source bundle.

Whereas the set-theoretical relations between objects and attributes are immediately visible in the overall graphical representation, the derivation of the set-theoretical relations among the sources is less straightforward. Therefore, one might prefer to draw a separate representation of the source hierarchy; Figure 2 represents the source hierarchy of the INDCLAS model of Table 2. From the overall graphical representation, the association relation among objects, attributes and sources can be read as follows: An object i is associated with attribute j according to source k iff object i and attribute j are connected with each other by a downward path that goes via source k. For example, Object 7 is associated with Attribute d according to Sources B and C, since there
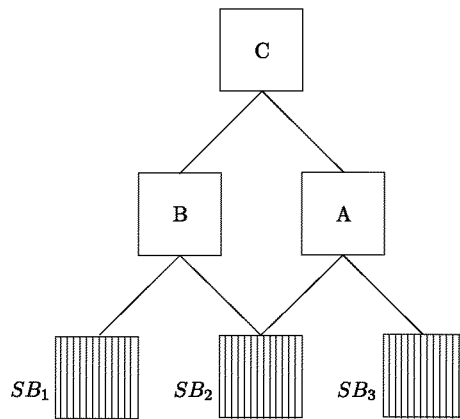

FIGURE 2.
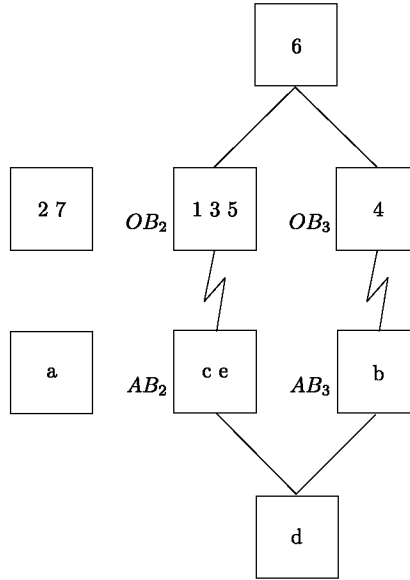Source hierarchy of the INDCLAS model in Table 2.

FIGURE 3.
Private structure of Source A in the INDCLAS model in Table 2.

exists a path from 7 to d that includes B and C. For Source A, however, there exists no such path; hence, 7 is not associated with d according to A.

A private structure can be derived for each source. The private structure of Source x is the HICLAS model (i.e., the original hierarchical classes model for binary two-way two-mode data, De Boeck & Rosenberg, 1988) that is obtained by only taking into account the object and attribute bundle(s) that correspond to the source bundle(s) Source x belongs to. For example, the private structure of Source A is graphically represented in Figure 3.

From the definition and the graphical representation of the INDCLAS model, it follows that INDCLAS is quite restrictive. First, the hierarchical classifications of the three modes are constrained to have the same rank, in that the number of bundles must be equal for each mode. Second, the linking structure among the three hierarchies, as defined by (1), is restricted to a one-to-one correspondence among the respective object, attribute and source bundles.

Finally, INDCLAS is closely related to the PARAFAC/CANDECOMP model for real-valued three-way three-mode data (Carroll & Chang, 1970; Harshman, 1970). More specifically, a PARAFAC/CANDECOMP model with binary components differs from INDCLAS in two respects only. First, INDCLAS involves a Boolean decomposition, whereas PARAFAC/CANDE-COMP is based on an algebraic decomposition: If one replaces $\oplus$ in (1) by a regular sum, one obtains the PARAFAC/CANDECOMP combination rule. Second, the INDCLAS model restricts the bundle matrices to represent the set-theoretical relations that exist in the model array.

2.2. A Generalized Three-Way Three-Mode Hierarchical Classes Model: Tucker3-HICLAS

A Tucker3-HICLAS model implies a decomposition of an $I \times J \times K$ binary model array $\mathbf{\underline{M}}$ into an $I \times R$ binary object bundle matrix $\mathbf{A}$, a $J \times S$ binary attribute bundle matrix $\mathbf{B}$, a $K \times T$ binary source bundle matrix $\mathbf{C}$ and a $R \times S \times T$ binary array $\mathbf{\underline{G}}$, with (R, S, T) being the rank of the model. $\mathbf{\underline{G}}$ defines a ternary association relation among the object, attribute and source bundles; this array will be called the core array.

TABLE 3.
Hypothetical Tucker3-HICLAS model array

Source A						Source B						Source C					
	Attributes						Attributes						Attributes				
Objects	a	b	c	d	e	Objects	a	b	c	d	e	Objects	a	b	c	d	e
1	1	0	0	1	1	1	0	1	1	1	0	1	1	1	1	1	1
2	0	0	0	0	0	2	0	0	0	0	0	2	0	0	0	0	0
3	0	0	0	0	0	3	1	0	0	1	1	3	1	0	0	1	1
4	0	1	1	1	0	4	0	0	0	0	0	4	0	1	1	1	0
5	1	0	0	1	1	5	1	1	1	1	1	5	1	1	1	1	1
6	1	1	1	1	1	6	0	1	1	1	0	6	1	1	1	1	1
7	0	0	0	0	0	7	1	0	0	1	1	7	1	0	0	1	1

TABLE 4.
Tucker3-HICLAS Model for $\underline{\mathbf{M}}$ in Table 3

Bundle matrices									Core array			
Object Bundles				Attribute Bundles			Source Bundles				Source Bundles	
Objects	OB_1	OB_2	OB_3	Attributes	AB_1	AB_2	Sources	SB_1	SB_2	Object Bundles	Attribute Bundles	SB_1	SB_2
1	0	1	0	a	1	0	A	0	1	OB_1	AB_1	1	0
2	0	0	0	b	0	1	B	1	0	OB_1	AB_2	0	0
3	1	0	0	c	0	1	C	1	1	OB_2	AB_1	0	1
4	0	0	1	d	1	1				OB_2	AB_2	1	0
5	1	1	0	e	1	0				OB_3	AB_1	0	0
6	0	1	1							OB_3	AB_2	0	1
7	1	0	0										

The hypothetical array $\underline{\mathbf{M}}$ in Table 3 serves as the guiding example in this subsection. Table 4 presents a Tucker3-HICLAS model for $\underline{\mathbf{M}}$. Since the definition and representation of the set-theoretical relations of equivalence and hierarchy are identical for the INDCLAS and Tucker3-HICLAS models, only the association relation will be discussed here in detail.

In the Tucker3-HICLAS model, the *association* relation (as defined in sec. 2.1) is represented by the following rule:

$$ m_{ijk} = \bigoplus_{r=1}^{R} \bigoplus_{s=1}^{S} \bigoplus_{t=1}^{T} a_{ir} b_{js} c_{kt} g_{rst} \qquad (\forall i = 1..I, \forall j = 1..J, \forall k = 1..K). \qquad (2) $$

This rule means that m_{ijk} equals 1 iff there exist object, attribute and source bundles to which object i, attribute j and source k, respectively, belong, and that are associated in the core array $\underline{\mathbf{G}}$. For example, from the model in Table 4, it can be derived that Object 4 is associated with Attribute b according to Source A, because OB_3, AB_2 and SB_2, to which the three elements belong respectively, are associated in $\underline{\mathbf{G}}$. Note that $\underline{\mathbf{G}}$ defines the linking structure among the hierarchical classifications of the three modes. Since replacing $\bigoplus$ in (2) by a regular sum yields the Tucker3 model formula, it is clear that Tucker3-HICLAS is the hierarchical classes counterpart of the Tucker3 model (Tucker, 1966).

The Tucker3-HICLAS model can be given a graphical representation that closely resembles the graphical representation of the INDCLAS model. Figures 4 and 5 contain the overall graph-
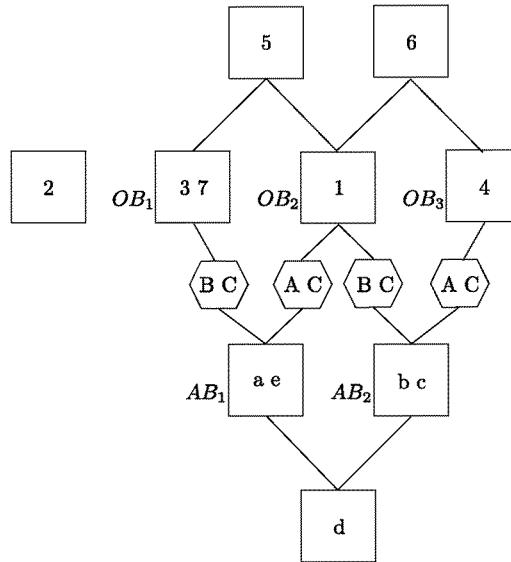
FIGURE 4.
Overall graphical representation of the Tucker3-HICLAS model in Table 4.
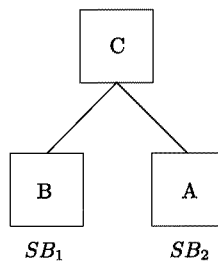

FIGURE 5.
Source hierarchy of the Tucker3-HICLAS model in Table 4.

ical representation and the source hierarchy of the Tucker3-HICLAS model of Table 4. The key difference between the graphical representation of INDCLAS and Tucker3-HICLAS concerns the way the object and attribute hierarchies are linked: In the Tucker3-HICLAS model representation, a path is drawn between each base class of the object hierarchy and each base class of the attribute hierarchy that are associated in $\underline{G}$, which further includes a hexagon that contains the (classes of) sources that belong to the source bundle for which the respective association holds. For example, in $\underline{G}$ OB_3 and AB_2 are associated according to SB_2; therefore, a path is drawn between OB_3 and AB_2, which includes a hexagon containing all sources that belong to SB_2 (i.e., sources A and C). Finally, a private structure can be derived for each source. The private structure of Source x is the HICLAS model that is obtained by taking into account only the object and attribute bundles that are associated according to the source bundle(s) Source x belongs to. The private structure of Source A is graphically represented in Figure 6.

The model definition and graphical representation of Tucker3-HICLAS clearly show that the latter model is less restrictive than INDCLAS. First, the number of bundles is allowed to differ across the three modes. Second, the linking structure among the three hierarchies is no longer restricted to take the form of a one-to-one correspondence; more specifically, every linking structure that can be represented by the core array $\underline{G}$ is allowed for. Consequently, Tucker3-
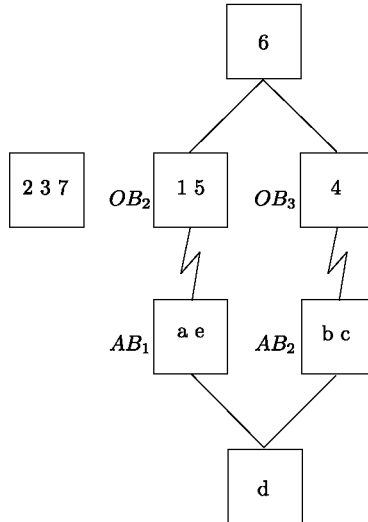
FIGURE 6.
Private structure of Source A in the Tucker3-HICLAS model in Table 4.

HICLAS can yield more parsimonious hierarchical classifications for one or more of the modes, by moving part of the complexity to the core array. In this respect, one may note that Tucker3-HICLAS reduces to INDCLAS iff R, S and T are equal and $\underline{\mathbf{G}}$ is a "unit superdiagonal" array (i.e., $R = S = T$, $g_{rst} = 1$ iff $r = s = t$ and $g_{rst} = 0$ otherwise; Kiers, 2000) and that every INDCLAS model can be rewritten as a Tucker3-HICLAS model by adding a $R \times R \times R$ "unit superdiagonal" core array $\underline{\mathbf{G}}$.

3. Data Analysis

3.1. Aim

The aim of a Tucker3-HICLAS analysis in rank (R, S, T) of a binary $I \times J \times K$ data array $\underline{\mathbf{D}}$ is to look for a binary $I \times J \times K$ model array $\underline{\mathbf{M}}$ which on the one hand has a minimal value on the loss function

$$f(\underline{\mathbf{M}}) = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} (d_{ijk} - m_{ijk})^2 \qquad (3)$$

and on the other hand can be represented by a Tucker3-HICLAS model of rank (R, S, T). As $\underline{\mathbf{D}}$ and $\underline{\mathbf{M}}$ are binary, (3) can be considered both a least squares and a least absolute deviations loss function.

In practice, the true rank of the Tucker3-HICLAS model underlying a given data set is unknown. Therefore, one will usually fit models of increasing ranks to a data set. Concerning the latter, some solutions (R, S, T) can however be omitted, as we have proven (see Appendix) that if a Tucker3-HICLAS model of rank (R, S, T) exists with $R > ST$, then an equivalent Tucker3-HICLAS model of rank (R', S, T) exists with $R' = ST$ (a similar result exists for the Tucker3 model, see Wansbeek & Verhees, 1989). Having obtained models of different ranks for a data set, a final solution may be picked out by means of rank selection heuristics. In section 4.4, we will propose and evaluate two such heuristics that are related to the DIFFIT-method proposed by Timmerman and Kiers (2000) to select the number of components in Tucker3 analysis (Tucker, 1966). More specifically, both the proposed heuristics as well as the DIFFIT-method aim at find-

ing a model with a good balance between fit to the data on the one hand and total number of bundles/components involved on the other hand. Note that the final choice for a specific rank should also be based on the interpretability of the corresponding model.

3.2. Algorithm

Given an $I \times J \times K$ data array $\underline{\mathbf{D}}$ and a rank (R, S, T), the Tucker3-HICLAS algorithm consists of two main routines: In the first routine, the algorithm searches for bundle matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and a core array $\underline{\mathbf{G}}$ that combine by (2) to an $I \times J \times K$ model array $\underline{\mathbf{M}}$ for which (3) is minimal. Since the first routine only restricts the bundle matrices to represent the association relation in $\underline{\mathbf{M}}$, the second routine transforms $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ so as to make them represent the set-theoretical relations correctly.

The first routine starts from an initial configuration for the three bundle matrices. Two initial configurations are obtained by performing (a) INDCLAS analyses on the $I \times J \times K$ data array $\underline{\mathbf{D}}$ in ranks R, S and T; and (b) HICLAS analyses (De Boeck & Rosenberg, 1988) on the matricized $I \times JK$, $J \times IK$ and $K \times IJ$ data array in ranks R, S and T, respectively (note that an initial configuration for both the HICLAS and INDCLAS algorithm are obtained via a built-in heuristic). Both the INDCLAS and the HICLAS analyses yield an object bundle matrix of rank R, an attribute bundle matrix of rank S and a source bundle matrix of rank T.

The first routine proceeds with an alternating least-squares procedure. Assuming an initial configuration $\mathbf{A}^{(0)}$, $\mathbf{B}^{(0)}$, $\mathbf{C}^{(0)}$ for the bundle matrices, the procedure calculates, conditionally upon $\mathbf{A}^{(0)}$, $\mathbf{B}^{(0)}$ and $\mathbf{C}^{(0)}$, the optimal core array $\underline{\mathbf{G}}^{(0)}$ that minimizes (3). In the next steps, $\mathbf{A}^{(w)}$ is re-estimated conditionally upon $\mathbf{B}^{(w-1)}$, $\mathbf{C}^{(w-1)}$ and $\underline{\mathbf{G}}^{(w-1)}$; $\mathbf{B}^{(w)}$ conditionally upon $\mathbf{A}^{(w)}$, $\mathbf{C}^{(w-1)}$ and $\underline{\mathbf{G}}^{(w-1)}$; $\mathbf{C}^{(w)}$ conditionally upon $\mathbf{A}^{(w)}$, $\mathbf{B}^{(w)}$ and $\underline{\mathbf{G}}^{(w-1)}$; and $\underline{\mathbf{G}}^{(w)}$ conditionally upon $\mathbf{A}^{(w)}$, $\mathbf{B}^{(w)}$ and $\mathbf{C}^{(w)}$ ($w = 1, 2, \ldots$). This alternating procedure continues until no updating of a bundle matrix or core array further improves the loss function (3) and, hence, the estimation procedure has converged. Since the bundle matrices can be estimated in six different orders, the alternating estimation of the bundle matrices and core array is repeated six times. This way, six solutions are obtained using the HICLAS bundle matrices as initial configuration and six solutions using the INDCLAS bundle matrices. Finally, out of the 12 resulting solutions, the solution which minimizes (3) is retained. Note that we estimate 12 solutions to avoid retaining a local minimum.

With respect to the estimation of a bundle matrix conditionally upon the two other bundle matrices and the core array, an analysis of the loss function (3) shows that it satisfies a separability property (Chaturvedi & Carroll, 1994). This property implies that the contribution of the bundle pattern of row i, $a_{i.}$, to the loss function (3) can be separated from the contribution of the bundle patterns of the other rows:

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(d_{ijk} - \bigoplus_{r=1}^{R} a_{ir} \bigoplus_{s=1}^{S} \bigoplus_{t=1}^{T} b_{js} c_{kt} g_{rst} \right)^2$$

$$= \sum_{j=1}^{J} \sum_{k=1}^{K} \left(d_{1jk} - \bigoplus_{r=1}^{R} a_{1r} \bigoplus_{s=1}^{S} \bigoplus_{t=1}^{T} b_{js} c_{kt} g_{rst} \right)^2$$

$$+ \cdots + \sum_{j=1}^{J} \sum_{k=1}^{K} \left(d_{Ijk} - \bigoplus_{r=1}^{R} a_{Ir} \bigoplus_{s=1}^{S} \bigoplus_{t=1}^{T} b_{js} c_{kt} g_{rst} \right)^2. \tag{4}$$

Consequently, an optimal estimate of $\mathbf{A}$ (conditional upon $\mathbf{B}$, $\mathbf{C}$ and $\underline{\mathbf{G}}$) can be found by successively optimizing the bundle patterns $a_{i.}$ ($i = 1..I$) in each of the I terms of (4). Note that this row-wise estimation of $\mathbf{A}$ is much faster than a matrix-wise estimation, since row-wise updating

implies I evaluations of 2^R possible bundle patterns, whereas matrix-wise updating implies the evaluation of 2^{IR} possible bundle matrices. While the separability property also holds for the conditional estimations of **B** and **C**, a further analysis of the loss function shows that the core array $\underline{\mathbf{G}}$ has to be updated array-wise.

The conditionally optimal estimates of the bundle patterns and the core array are obtained by means of Boolean regression (Leenen & Van Mechelen, 1998). For example, to estimate the bundle pattern $a_{i.}$ in the i-th term of (4) the values of the criterion variable are the data-entries d_{ijk} $(j = 1..J; k = 1..K)$ and the values of the r-th $(r = 1..R)$ predictor variable are the sums

$$\bigoplus_{s=1}^{S} \bigoplus_{t=1}^{T} b_{js} c_{kt} g_{rst} \quad (j = 1..J; k = 1..K).$$

In the estimation of the core array $\underline{\mathbf{G}}$ RST predictor variables are involved; the values of the criterion variable are the data-entries d_{ijk} $(i = 1..I; j = 1..J; k = 1..K)$ and the values of the predictor variable (r, s, t) are the products $a_{ir} b_{js} c_{kt}(i = 1..I; j = 1..J; k = 1..K)$.

In the second routine, the set-theoretical relations are added to the bundle matrices obtained at the end of the first routine. This is accomplished by applying a closure operation (Barbut & Monjardet, 1970; Birkhoff, 1940) to **A**, **B** and **C**. For example, to make **A** correctly represent the set-theoretical relations among the objects, each zero-entry a_{ir} is changed to one iff for all $j = 1 \ldots J, k = 1 \ldots K$ holds that

$$\bigoplus_{s=1}^{S} \bigoplus_{t=1}^{T} b_{js} c_{kt} g_{rst} \leq m_{ijk}$$

(implying that the modification of **A** does not alter $\underline{\mathbf{M}}$). It may be noted that this closure operation is a sufficient, though not necessary condition for set-theoretical consistency.

Observe that the described algorithm has been implemented in Delphi 5 and is freely available from the first author.

4. Simulation Study

In this section we present a simulation study with a twofold aim: First, the Tucker3-HICLAS algorithm is evaluated with respect to sensitivity for local minima, speed of convergence, goodness-of-fit and goodness of recovery. Second, two heuristics for rank selection are evaluated.

In section 4.1, the design of the simulation study is outlined. Next, the results are presented in sections 4.2 (local minima, speed of convergence and goodness-of-fit), 4.3 (goodness of recovery) and 4.4 (rank selection).

4.1. Design and Procedure

Three different types of binary $I \times J \times K$ arrays must be distinguished in this simulation study: a true array $\underline{\mathbf{T}}$, which can be represented by a Tucker3-HICLAS model of rank r (r being a triplet of three positive integers); a data array $\underline{\mathbf{D}}$, which is $\underline{\mathbf{T}}$ perturbed with error; and the model array $\underline{\mathbf{M}}$ yielded by the algorithm, which can be represented by a Tucker3-HICLAS model of rank r' (r' being a triplet of three positive integers).

Three parameters were systematically varied in a complete trifactorial design:

1. the *Size*, $I \times J \times K$, of $\underline{\mathbf{T}}$, $\underline{\mathbf{D}}$ and $\underline{\mathbf{M}}$, at 3 levels: $15 \times 15 \times 15$, $30 \times 20 \times 10$, $30 \times 20 \times 60$;
2. the *True rank*, r, of the Tucker3-HICLAS model for $\underline{\mathbf{T}}$, at 4 levels: (2,2,2), (2,3,3), (4,3,2), (4,4,4);

3. the *Error level*, ε, which is the proportion of cells d_{ijk} differing from t_{ijk}, at 5 levels: .00, .05, .10, .20, .30.

All parameters will be considered random effects.

For each combination of Size $I \times J \times K$, True rank r, and Error level ε, 20 true arrays $\underline{\mathbf{T}}$ were constructed, subject to the constraint that $\underline{\mathbf{T}}$ cannot be represented almost perfectly by a Tucker3-HICLAS model of lower rank; the latter constraint (which is similar to the procedure used by Timmerman and Kiers (2000) in their study on rank selection methods for Tucker3 analysis) is included in view of the evaluation of rank selection heuristics. More in particular, each $\underline{\mathbf{T}}$ was constructed as follows: Bundle matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ were generated with entries that were independent realizations of a Bernoulli variable with a probability parameter chosen such that the expected proportion of ones in $\underline{\mathbf{T}}$ equals 0.5. Like Timmerman and Kiers, we chose the core arrays such that a bundle in a particular mode is associated with few bundles for the other modes, in order to mimic simple underlying processes. More specifically, $\underline{\mathbf{G}}_{(2,2,2)}$ and $\underline{\mathbf{G}}_{(4,3,2)}$ were copied from Timmerman and Kiers's study:

$$\underline{\mathbf{G}}_{(2,2,2)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

$$\underline{\mathbf{G}}_{(4,3,2)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

and $\underline{\mathbf{G}}_{(2,3,3)}$ and $\underline{\mathbf{G}}_{(4,4,4)}$ were chosen as follows:

$$\underline{\mathbf{G}}_{(2,3,3)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

$$\underline{\mathbf{G}}_{(4,4,4)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Subsequently, the true array $\underline{\mathbf{T}}$, which resulted from combining $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\underline{\mathbf{G}}$ by the association rule (2), was selected if the Tucker3-HICLAS solution $\underline{\mathbf{M}}$ of true rank contained at least 2% more correct entries than all solutions of lower rank. Next, a data array $\underline{\mathbf{D}}$ was constructed from each true array $\underline{\mathbf{T}}$ by randomly altering the value of a proportion ε of the entries in $\underline{\mathbf{T}}$. Finally, all data arrays $\underline{\mathbf{D}}$ were analyzed with the Tucker3-HICLAS algorithm in ranks (1,1,1) through (5,5,5), yielding 74 model arrays $\underline{\mathbf{M}}$ for each data array $\underline{\mathbf{D}}$ (due to the result in the Appendix). As a result of this procedure, 20 $\times$ 3 (size) $\times$ 4 (true rank) $\times$ 5 (error level) $\times$ 74 (analysis rank) = 88,800 different triplets ($\underline{\mathbf{T}}$,$\underline{\mathbf{D}}$,$\underline{\mathbf{M}}$) were obtained.

4.2. Local Minima, Speed of Convergence and Goodness-of-Fit

In this section, the Tucker3-HICLAS algorithm is evaluated with respect to how well and how fast it succeeds in minimizing the loss function, that is, with respect to (i) goodness-of-fit and (ii) speed of convergence.

i. For the study of goodness-of-fit, the following badness-of-fit (BOF) statistic was used:

$$BOF = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} (d_{ijk} - m_{ijk})^2}{I \times J \times K}. \tag{5}$$

Since the Error level ε constitutes an upper bound for the BOF of the global minimum in case of analyses in the True rank (i.e., $r' = r$), only the triplets (**T**,**D**,**M**) with analyses in True rank will be considered. There are 1200 such analyses (5 Error levels $\times$ 240 analyses per Error level), each implying 12 runs of the algorithm. Of these 1200 analyses, 499 or 41.6% ended in a solution with a BOF value smaller than or equal to ε. If we split up the results according to the different Error levels, the percentages amount to 37.1, 60.8, 55.8, 34.2 and 20.0 for ε equal to .00, .05, .10, .20, and .30, respectively. To investigate further the issue of local minima, we examined how many out of the 12 runs per analysis ended in the retained solution: On average, this was the case for 5.88 of the 12 runs per analysis. The latter results imply that it is not unusual for the algorithm to end in a local minimum; however, all subsequent simulation results will show that the obtained solutions are reasonably close to the underlying truth.

An analysis of variance with BOF as the dependent variable yielded an intraclass correlation $\hat{\rho}_I$ (Haggard, 1958; Kirk, 1982) of .99 for the main effect of Error level. In other words, almost all variance in BOF is accounted for by Error level. Mean BOF-values across the 240 observations within each Error level are: .011, .054, .104, .211 and .312 for ε equal to .00, .05, .10, .20, and .30, respectively. These results imply that the retained solutions are about as close to the data as the truth is. An analysis of variance with the difference between BOF and ε as the dependent variable reveals a main effect of Error level ($\hat{\rho}_I = .07$): This main effect implies that the higher ε, the harder it is for the algorithm to find a model that is as close to the data as the truth is (see Table 5). The main effect of Error level is further qualified by a True rank $\times$ Error level interaction ($\hat{\rho}_I = .20$), indicating that the effect of Error level increases with higher True rank. Other effects are not discussed: In this and the following analyses of variance only effects accounting for at least 5% of the variance of the dependent variable will be considered (i.e., $\hat{\rho}_I \geq .05$). Note the unexpectedly high ($BOF - \varepsilon$)-value for errorfree data in Table 5. In the following sections, it will be shown that all simulation results are better for slightly error perturbed data than for errorfree data. This finding is in line with other hierarchical classes simulation studies (e.g., Leenen, Van Mechelen, De Boeck, & Rosenberg, 1999); so far, its cause has not been clarified.

ii. With respect to the speed of convergence, no improvement in the loss function (3) was observed after, on average, 4.91 iterations. An analysis of variance with as dependent variable the number of iterations before convergence, shows main effects of Error level ($\hat{\rho}_I = .06$)

TABLE 5.
Mean differences between Badness of Fit and ε and mean number of iterations before convergence at Levels of True rank $\times$ Error

		Error Level					
	True rank	.00	.05	.10	.20	.30	Overall
$\varepsilon - BOF$	(2,2,2)	.012	.005	.003	.002	.002	.005
	(2,3,3)	.018	.004	.003	.011	.014	.010
	(4,3,2)	.011	.004	.004	.012	.013	.009
	(4,4,4)	.004	.002	.008	.019	.020	.011
	Overall	.011	.004	.004	.011	.012	.009
Nr. of iterations	(2,2,2)	1.13	1.00	1.37	2.72	6.22	2.49
	(2,3,3)	1.95	1.12	1.80	6.22	12.72	4.76
	(4,3,2)	2.25	1.52	1.67	6.75	12.27	4.89
	(4,4,4)	2.03	1.55	3.55	12.98	17.42	7.51
	Overall	1.84	1.30	2.10	7.17	12.16	4.91

and True rank ($\hat{\rho}_I = .42$): apart from errorfree data, the number of iterations increases with higher ε and with higher True rank (see Table 5). Additionally, the analysis of variance yielded an Error level × True rank interaction ($\hat{\rho}_I = .10$): the higher the True rank, the bigger the effect of Error level.

4.3. Goodness of Recovery

Goodness of recovery will be evaluated with respect to (i) the equivalence relations, (ii) the hierarchical relations, and (iii) the association relation. Again only the 1200 triplets (**T,D,M**) with analyses in True rank (i.e., $r' = r$) are considered.

i. With respect to the recovery of the equivalence relations, the corrected Rand index (Hubert & Arabie, 1985) is used to compare the partition of the set of objects (resp. attributes, sources) in the Tucker3-HICLAS model for **T** with the partition of the set of objects (resp. attributes, sources) in the Tucker3-HICLAS model for **M**. This index equals 1 if the two partitions are identical and 0 if the two partitions do not correspond more than expected by chance. The results being very similar for the objects, attributes and sources, a combined corrected Rand index (c-CRI) was calculated by taking the average corrected Rand index for the object, the attribute and the source equivalence relation, weighted by the number of objects, attributes and sources, respectively.

The mean c-CRI across the 1200 observations equals .837, implying a high correspondence between the true and reconstructed equivalence relations. An analysis of variance with c-CRI as the dependent variable yields a main effect of Error level ($\hat{\rho}_I = .33$): except for errorfree data, it holds that the higher the Error level, the lower the c-CRI. The main effect of Error level is qualified by an Error level × True rank interaction ($\hat{\rho}_I = .18$), the latter resulting from the Error level effect being more important for higher True ranks (see Table 6).

ii. To assess the recovery of the hierarchical relations, we first define the object hierarchy matrix associated with **M** (resp. **T**) as the $I \times I$ binary matrix $\mathbf{U}^{(M)}$ (resp. $\mathbf{U}^{(T)}$), with $u_{ii'}^{(M)} = 1$

TABLE 6.
Mean Combined Corrected Rand Index, Mean Badness of Hierarchy Recovery and Mean Badness of Recovery at Levels of True Rank × Error

		Error Level					
	True rank	.00	.05	.10	.20	.30	Overall
c-CRI	(2,2,2)	.707	.793	.810	.835	.720	.773
	(2,3,3)	.901	.970	.979	.891	.667	.882
	(4,3,2)	.915	.962	.952	.821	.632	.856
	(4,4,4)	.986	.995	.959	.821	.428	.838
	Overall	.877	.930	.925	.842	.612	.837
c-$BOHR$	(2,2,2)	.108	.077	.069	.062	.103	.084
	(2,3,3)	.040	.012	.006	.034	.118	.042
	(4,3,2)	.029	.013	.015	.061	.130	.050
	(4,4,4)	.003	.003	.012	.056	.183	.051
	Overall	.045	.026	.025	.053	.134	.057
BOR	(2,2,2)	.012	.006	.004	.004	.012	.008
	(2,3,3)	.018	.005	.004	.020	.056	.021
	(4,3,2)	.011	.004	.005	.024	.061	.021
	(4,4,4)	.004	.002	.011	.041	.120	.036
	Overall	.011	.004	.006	.022	.062	.021

(resp. $u_{ii'}^{(\mathrm{T})}=1$) iff object i is hierarchically below object i' in $\underline{\mathbf{M}}$ (resp. $\underline{\mathbf{T}}$). Subsequently, the proportion of discrepancies between $\mathbf{U}^{(\mathrm{T})}$ and $\mathbf{U}^{(\mathrm{M})}$ was calculated yielding a badness-of-hierarchy-recovery (*BOHR*) statistic for the objects:

$$BOHR = \frac{\sum_{i=1}^{I} \sum_{i'=1}^{I} (u_{ii'}^{(\mathrm{T})} - u_{ii'}^{(\mathrm{M})})^2}{I^2}.$$

Similarly, a *BOHR* statistic for the attributes and sources was defined. A weighted average of the three statistics (weighted by the number of elements in each mode), denoted *c-BOHR*, was calculated and used in the following analyses.

The mean value on *c-BOHR* across the 1200 observations equals .057, implying that, on average, 94.3% of the true hierarchy is recovered. An analysis of variance with *c-BOHR* as the dependent variable shows a main effect of Error level ($\hat{\rho}_I = .30$): apart from errorfree data, the recovery deteriorates with increasing Error level (see Table 6). Additionally, the Error level × True rank ($\hat{\rho}_I = .13$) interaction has to be considered: the higher the True rank, the bigger the effect of Error level.

iii. The badness of recovery (*BOR*) of the association relation was assessed by the proportion of discrepancies between $\underline{\mathbf{T}}$ and $\underline{\mathbf{M}}$:

$$BOR = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} (t_{ijk} - m_{ijk})^2}{I \times J \times K}.$$

The mean *BOR* across the 1200 observations equals .021, which means that the model yielded by the algorithm differs on average 2.1% from the underlying truth. An analysis of variance with *BOR* as the dependent variable yields a main effect of Error level ($\hat{\rho}_I = .35$): except for errorfree data, badness of recovery clearly increases with higher Error levels (see Table 6). Furthermore, the True rank × Error level interaction ($\hat{\rho}_I = .29$) has to be taken in account, indicating that the effect of Error level increases with higher True rank.

Goodness of recovery of the association relation was also measured by the relative-recovery-gain (*RRG*) statistic, defined as:

$$RRG = \frac{\varepsilon - BOR}{\varepsilon},$$

RRG being undefined for $\varepsilon = 0$. This statistic equals 1 in case of perfect recovery and 0 if the model is as far from the truth as the data are. The mean *RRG* across the 960 observations for which $\varepsilon > 0$ amounts to .883. The latter means that $\underline{\mathbf{M}}$ contains, on average, 88.3% less erroneous entries than $\underline{\mathbf{D}}$.

4.4. Rank Selection

In this subsection two heuristics for rank selection are presented and evaluated in terms of the match between the selected model and the true model. Both heuristics aim at finding a model with a good balance between fit to the data on the one hand and total number of bundles/components involved on the other hand. The procedure is as follows: First, Tucker3-HICLAS model arrays $\underline{\mathbf{M}}$ in ranks (1,1,1) through (r_1, r_2, r_3) are fitted to an $I \times J \times K$ data array $\underline{\mathbf{D}}$. For each model, the sum s of the number of object, attribute and source bundles, and the *BOF*-value (5) is calculated. Next, per sum of bundles s only the best fitting model is retained. Finally, a variant of the scree test (Cattell, 1966) is applied to the resulting models. In the following, we will present two such variants. Table 7, which displays the *BOF* values of the retained models for a particular $\underline{\mathbf{D}}$ in our simulation study, serves as a guiding example; the associated *BOF* by sum of bundles s plot is shown in Figure 7.

TABLE 7.
Badness-of-fit values for a data set of True rank (4,3,2), Size $30 \times 20 \times 10$ and Error level .10 and statistics on which rank selection rules are based as a function of the sum of bundles s

s	Rank	BOF_s	Rule A statistic	Rule B statistic
3	(1,1,1)	.293	/	/
5	(2,2,1)	.237	.019	.0142
6	(2,2,2)	.228	−.061	.0073
7	(3,2,2)	.158	.056	.0264
8	(4,2,2)	.144	−.03	.0234
9	(4,3,2)	.100	.044	.032
10	(4,3,3)	.100	0	.0274
11	(4,3,4)	.100	0	.0239
12	(4,3,5)	.100	−.001	.0211
13	(5,4,4)	.099	.001	.0194
14	(5,4,5)	.099	0	.0176
15	(5,5,5)	.099	/	/

Rule A, which is the most similar to the scree test used in PCA, selects the model with sum of bundles s that maximizes

$$\frac{BOF_{s^{\text{prev}}} - BOF_s}{s - s^{\text{prev}}} - \frac{BOF_s - BOF_{s^{\text{next}}}}{s^{\text{next}} - s},$$

with s^{prev} and s^{next} denoting the left and right neighbor of s in the BOF plot. For the example in Table 7, this rule yields $s = 7$. In order to decrease the impact of local fluctuations in the BOF plot, one may also consider to compare the BOF for a particular s with the BOF for the smallest and biggest s obtained. Formally, Rule B selects the model with sum of bundles s which maximizes

$$\frac{BOF_{s^{\text{min}}} - BOF_s}{s - s^{\text{min}}} - \frac{BOF_s - BOF_{s^{\text{max}}}}{s^{\text{max}} - s},$$

with s^{min} denoting the smallest sum of bundles and s^{max} the biggest sum of bundles in the BOF plot. In the example in Table 7, this rule yields $s = 9$. Note that the result of Rule B is more
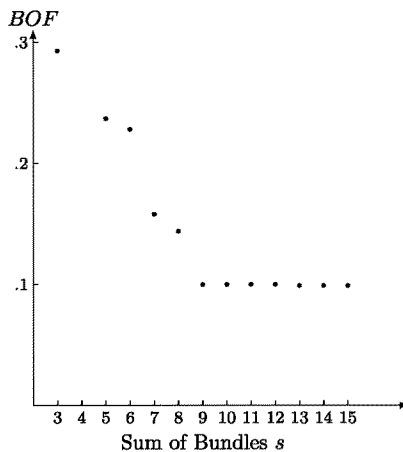

FIGURE 7.
The BOF by s plot for the example used in Table 7.

TABLE 8.
Percentage of correct estimation of the True rank

		Error					
	True Rank	.00	.05	.10	.20	.30	Overall
Rule A	(2,2,2)	86.7	95.0	98.3	95.0	86.7	92.3
	(2,3,3)	46.7	81.7	85.0	46.7	21.7	56.4
	(4,3,2)	60.0	70.0	73.3	40.0	26.7	54.0
	(4,4,4)	15.0	6.7	5.0	6.7	1.7	7.0
	Overall	52.1	63.3	65.4	47.1	34.2	52.4
Rule B	(2,2,2)	86.7	91.7	95.0	95.0	83.3	90.3
	(2,3,3)	40.0	80.0	86.7	41.7	18.3	53.3
	(4,3,2)	56.7	75.0	78.3	36.7	25.0	54.3
	(4,4,4)	50.0	56.7	30.0	15.0	1.7	30.7
	Overall	58.3	75.8	72.5	47.1	32.1	57.2

influenced by the choice of s^{min} and s^{max} than the result of Rule A; it is recommended to choose s^{min} as small as possible (i.e., $s^{min} = 3$) and $s^{max} \geq s^{exp} + 3$, where s^{exp} indicates the expected maximal value of s.

In the simulation study, each $\underline{D}$ was analyzed in ranks (1,1,1) through (5,5,5). Retaining only the best fitting solution per value of s therefore resulted in twelve solutions per data set ($s = 3, 5, 6 \ldots 15$), on which Rule A and Rule B were applied. On average, Rule B is superior to Rule A with 57.2% and 52.4% correct estimations, respectively (see Table 8). Rule A and Rule B were also compared with respect to the goodness of recovery of the selected models: The models selected by Rule B generally have a considerably higher goodness of recovery than the models selected by Rule A. For example, the mean RRG of the models selected by Rule A and Rule B equal .74 and .87, respectively.

5. Illustrative Application

In this section we present an illustrative application of the Tucker3-HICLAS model on data from a study in the personality domain. This study concerns individual differences in self-reported hostile behavior in frustrating situations (Vansteelandt & Van Mechelen, 1998). The data were situation by behavior by person data, which were originally analyzed by means of the INDCLAS algorithm. More specifically, in the study 54 persons (sources) had to indicate with yes or no whether they displayed 15 hostile behaviors (attributes) in 23 frustrating situations (objects). This resulted in a binary 23 × 15 × 54 data array $\underline{D}$ with $d_{ijk} = 1$ if person k displayed behavior j in situation i and 0 otherwise.

Tucker3-HICLAS models in ranks (1,1,1) through (6,6,6) were fitted to $\underline{D}$. Applying Rules A and B resulted in the selection of the (2,4,4) and (3,5,5) models, respectively. However, as both these models consist of 9 behavior classes for 15 behaviors only and, hence, are rather difficult to interpret, we prefer to discuss a simpler model here. Since the (2,3,2) model has about the same BOF as the INDCLAS model of rank 3 described by Vansteelandt and Van Mechelen, namely .27, and, therefore, allows for a fair comparison of the INDCLAS and Tucker3-HICLAS models for $\underline{D}$, we retain the (2,3,2) model. Figure 8 shows the overall graphical representation of the (2,3,2) model and Figure 9 represents the person hierarchy, with the number of persons in each class indicated between parentheses; note that the situations, responses and persons with a zero bundle pattern have not been included in the representation and that the situations have been indicated by the key words presented in Table 9. For clarity's sake, the overall graphical representation of Vansteelandt and Van Mechelen's (1998) INDCLAS model of rank 3 is provided in Figure 10.
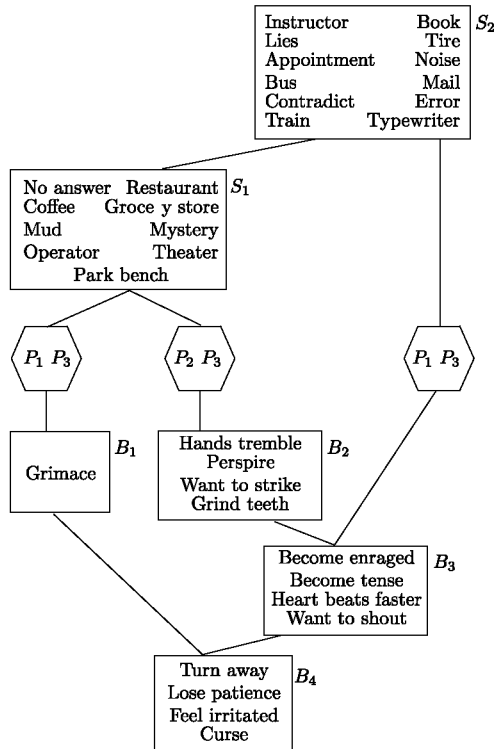
FIGURE 8.
Overall graphical representation of the Tucker3-HICLAS model for the hostility data.

As was the case for the INDCLAS model (see Figure 10), the hierarchy of situation classes takes the form of a total order, which may be conceived as a quantitative dimension (Gati & Tversky, 1982). In line with Vansteelandt and Van Mechelen (1998), this dimension can be interpreted as a frustration dimension with class S_1 containing mildly frustrating situations (e.g., someone pushes ahead of you in a ticket line) and class S_2 strongly frustrating situations (e.g., incessant noise when you are studying). The latter interpretation of the situation hierarchy was further validated by calculating the Spearman rank correlation between the position of the S_1 and S_2 situations on the quantitative dimension and external expert judgements on those situations: The strongly frustrating situations are characterized by a longer duration of the frustration ($r = .85, p < .0001$), more severe consequences ($r = .75, p < .0001$), a more external causal locus ($r = .78, p < .0001$) and a more ego-threatening character ($r = .45, p = .043$) than the mildly frustrating situations.
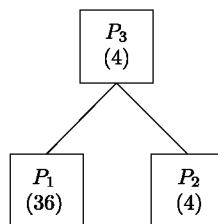

FIGURE 9.
Person hierarchy of the Tucker3-HICLAS model for the hostility data.

TABLE 9.
Key words for the 21 frustrating situations in the graphical representations of the INDCLAS and Tucker3-HICLAS models for the hostility data

Key word	Full situation description
instructor	Your instructor unfairly accuses you of cheating on an examination
book	Someone has lost an important book of yours
lies	You have just found out that someone has told lies about you
tire	You are driving to a party and suddenly your car has a flat tire
appointment	You arrange to meet someone and he (she) doesn't show up
noise	You are trying to study and there is incessant noise
bus	You are waiting at the bus stop and the bus fails to stop for you
mail	Someone has opened your personal mail
contradict	Someone persistently contradicts you when you know you are right
error	Someone makes an error and blames it on you
train	You miss your train because the clerk has given you faulty information
typewriter	You are typing a term paper and your typewriter breaks
no answer	You are talking to someone and he (she) does not answer you
restaurant	You are in a restaurant and have been waiting a long time to be served
coffee	You are carrying a cup of coffee to the table and someone bumps into you
grocery store	The grocery store closes just as you are about to enter
mud	Someone has splashed mud over your new clothing
mystery	You are reading a mystery and find that the last page of the book is missing
operator	You use your last 10 c to call a friend and the operator disconnects you
theater	Someone pushes ahead of you in a theater ticket line
park bench	You accidently bang your shins against a park bench

With the exception of response class B_1, the Tucker3-HICLAS behavior hierarchy also constitutes a quantitative dimension. Since the INDCLAS behavior hierarchy consists of six classes (see Figure 10), the latter means that the Tucker3-HICLAS behavior hierarchy, although involving the same number of bundles, is more parsimonious. The Spearman rank correlations between the position of the B_2, B_3 and B_4 behaviors on the quantitative dimension and external expert judgements on those behaviors, indicate that this dimension can be interpreted as a physiological dimension ($r = -.62$, $p = .03$): Class B_2 contains behaviors that are most physiologic in nature (e.g., perspire), class B_3 behaviors that are moderately physiologic (e.g., become enraged) and class B_4 behaviors that are least physiologic (e.g., turn away).

Making use of the INDCLAS association relation, Vansteelandt and Van Mechelen (1998) interpreted the person structure by characterizing each person class in terms of a set of if (situation class) then (behavior class) rules (Wright & Mischel, 1987). A similar interpretation can be given to the Tucker3-HICLAS person structure. For example, from Figure 8 it can be derived that person class P_1 is characterized by the following set of if-then rules: (a) *If* a person of this class finds himself in a strongly frustrating situation (S_2) *then* he or she will grimace (B_1), become enraged (B_3) and turn away (B_4); and (b) *If* a person of this class is in a frustrating situation (S_1 or S_2) *then* he or she will grimace (B_1) and turn away (B_4). Note that the full set of if (situation class) then (behavior class) rules that characterize a person class can be easily read from the private structure of the respective person class. For example, the above mentioned rules can be read from Figure 11, which represents the private structure of person class P_1. Finally, observe that the Tucker3-HICLAS model yields only three person classes to be interpreted, whereas the INDCLAS model yields seven of them (see Figure 10).

One may conclude that the Tucker3-HICLAS model gives an equally well fitting but more parsimonious description of Vansteelandt and Van Mechelen's (1998) data than the INDCLAS model and, hence, can be preferred.
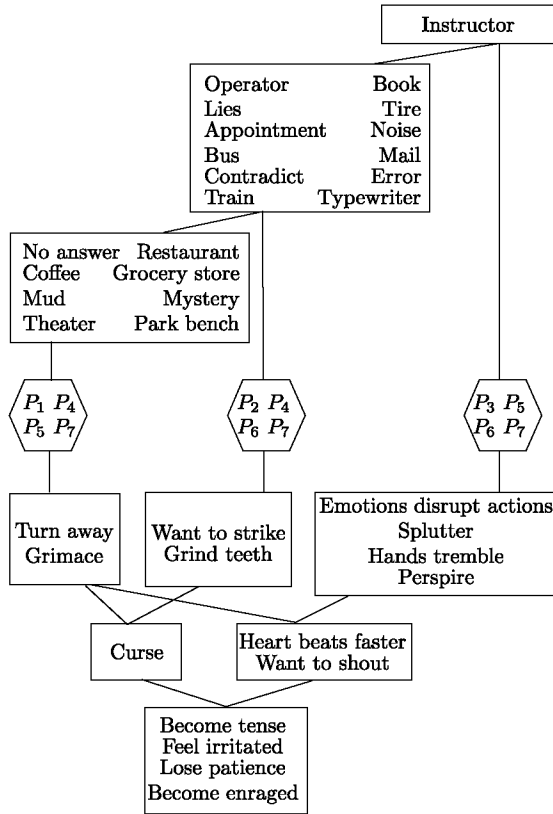
FIGURE 10.
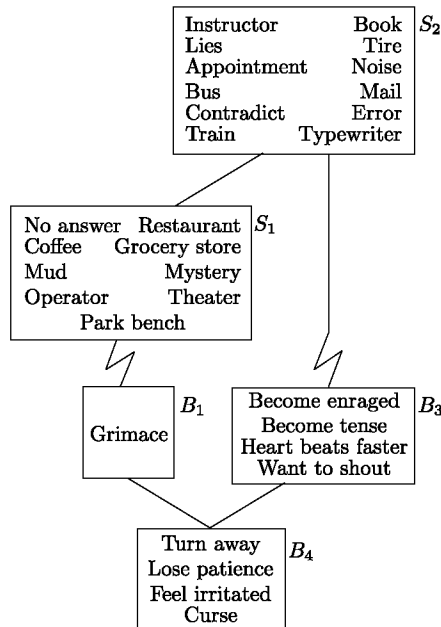Overall graphical representation of the INDCLAS model for the hostility data.

FIGURE 11.
rivate structure of person class P_1 in the Tucker3-HICLAS model for the hostility data.

6. Concluding Remarks

In this paper we proposed a novel hierarchical classes model for binary three-way three-mode data, the Tucker3-HICLAS model. The new model has been shown to be less restrictive than INDCLAS, in that it does not restrict the hierarchical classifications of the three modes to have the same rank and allows for more complex linking structures than a one-to-one correspondence among the three hierarchies. As a consequence, Tucker3-HICLAS can result in more parsimonious hierarchical classifications for one or more of the modes, since part of the complexity can be moved to the linking structure. Moreover, the Tucker3-HICLAS relaxations of the INDCLAS restrictions also make sense from a substantive point of view. For example, in the personality psychology domain there is no a priori reason for the person, situation and behavior structure to have the same rank, let alone for a one-to-one correspondence between the respective person, situation and behavior bundles.

The Tucker3-HICLAS model naturally follows from the relations within and between the family of hierarchical classes models on the one hand and the family of principal component models on the other hand. A further exploration of these relations may result in the development of yet other interesting three-way three-mode hierarchical classes models. For example, one may consider a hierarchical classes version of the Tucker2 model (Kroonenberg, 1983; Kroonenberg & de Leeuw, 1980), which reduces only two of the three modes of a data set to a few components. Such a model would allow for the maximal modeling of the differences in the not-reduced mode (Ceulemans & Van Mechelen, in press).

A. Appendix

Theorem. If a Tucker3-HICLAS decomposition of an $I \times J \times K$ array $\underline{\mathbf{M}}$ in rank (R, S, T) exists with $R > ST$, then also a Tucker3-HICLAS decomposition in rank (R^*, S, T) exists with $R^* = ST$.

Proof. Assume arrays $\mathbf{A}$ $(I \times R)$, $\mathbf{B}$ $(J \times S)$, $\mathbf{C}$ $(K \times T)$ and $\underline{\mathbf{G}}$ $(R \times S \times T)$ exist such that

$$m_{ijk} = \bigoplus_{r=1}^{R} \bigoplus_{s=1}^{S} \bigoplus_{t=1}^{T} a_{ir} b_{js} c_{kt} g_{rst} \qquad (A6)$$

holds for any i, j, k. We define an $I \times ST$ matrix $\tilde{\mathbf{A}}$ with entries

$$\tilde{a}_{i((s-1)T+t)} = \bigoplus_{r=1}^{R} a_{ir} g_{rst} \quad (s = 1..S; t = 1..T),$$

and an $ST \times S \times T$ array $\tilde{\underline{\mathbf{G}}}$ with entries $\tilde{g}_{r^*st} = 1$ if $r^* = (s-1)T + t$ and $\tilde{g}_{r^*st} = 0$ otherwise $(r^* = 1..ST; s = 1..S; t = 1..T)$. Then (6) can be rewritten as

$$m_{ijk} = \bigoplus_{s=1}^{S} \bigoplus_{t=1}^{T} b_{js} c_{kt} \bigoplus_{r=1}^{R} a_{ir} g_{rst} = \bigoplus_{s=1}^{S} \bigoplus_{t=1}^{T} b_{js} c_{kt} a_{i((s-1)T+t)} = \bigoplus_{s=1}^{S} \bigoplus_{t=1}^{T} b_{js} c_{kt} \bigoplus_{r^*=1}^{ST} a_{ir^*} \tilde{g}_{r^*st},$$

which proves that a Tucker3-HICLAS decomposition of $\underline{\mathbf{M}}$ in rank (ST, S, T) exists. $\square$

References

Barbut, M., & Monjardet, B. (1970). *Ordre et classification: Algèbre et combinatoire* [Order and classification: Algebra and combinatorics]. Paris: Hachette.

Birkhoff, G. (1940). *Lattice theory.* Providence, RI: American Mathematical Society.

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

Cattell, R.B. (1966). The meaning and strategic use of factor analysis. In R.B. Cattell (Ed.), *Handbook of multivariate experimental psychology* (pp. 174–243). Chicago, IL: Rand McNally.

Ceulemans, E., & Van Mechelen, I. (in press). *Tucker2 hierarchical classes analysis. Psychometrika.*

Chaturvedi, A., & Carroll, J.D. (1994). An alternating combinatorial optimization approach to fitting the INDCLUS and generalized INDCLUS models. *Journal of Classification, 11*, 155–170.

De Boeck, P., & Rosenberg, S. (1988). Hierarchical classes: Model and data analysis. *Psychometrika, 53*, 361–381.

Gati, I., & Tversky, A. (1982). Representations of qualitative and quantitative dimensions. *Journal of Experimental Psychology, 8*, 325–340.

Haggard, E.A. (1958). *Intraclass correlation and the analysis of variance.* New York, NY: Dryden.

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

Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification, 2*, 193–218.

Kiers, H.A.L. (2000). Towards a standardized notation and terminology in multiway analysis. *Journal of Chemometrics, 14*, 105–122.

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

Kroonenberg, P. M. (1983). *Three-mode principal component analysis: Theory and applications.* Leiden: DSWO.

Kroonenberg, P.M., & de Leeuw, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika, 45*, 69–97.

Leenen, I., & Van Mechelen, I. (1998). A branch-and-bound algorithm for Boolean regression. In I. Balderjahn, R. Mathar, & M. Schader (Eds.), *Data highways and information flooding, a challenge for classification and data analysis* (pp. 164–171). Berlin, Germany: Springer–Verlag.

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

Leenen, I., Van Mechelen, I., De Boeck, P., & Rosenberg, S. (1999). INDCLAS: A three-way hierarchical classes model. *Psychometrika, 64*, 9–24.

Timmerman, M.E., & Kiers, H.A.L. (2000). Three-mode principal components analysis: Choosing the numbers of components and sensitivity to local optima. *British Journal of Mathematical and Statistical Psychology, 53*, 1–16.

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

Van Mechelen, I., De Boeck, P., & Rosenberg, S. (1995). The conjunctive model of hierarchical classes. *Psychometrika, 60*, 505–521.

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.

Wansbeek, T., & Verhees, J. (1989). Models for multidimensional matrices in econometrics and psychometrics. In R. Coppi & S. Bolasco (Eds.), *Multiway data analysis* (pp. 543–552). Amsterdam, the Netherlands: Elsevier.

Wright, J.C., & Mischel, W. (1987). A conditional approach to dispositional constructs: The local predictability of social behavior. *Journal of Personality and Social Psychology, 55*, 454–469.