

Europäische Hochschulschriften



Sabine Krolak-Schwerdt

Modelle der  
dreimodalen Faktorenanalyse



Peter Lang Frankfurt am Main · Bern · New York · Paris



Sabine Krolak-Schwerdt

Modelle der  
dreimodalen Faktorenanalyse



**PETER LANG**  
Frankfurt am Main · Bern · New York · Paris

CIP-Titelaufnahme der Deutschen Bibliothek

Krolak-Schwerdt, Sabine:

Modelle der dreimodalen Faktorenanalyse / Sabine Krolak-Schwerdt. - Frankfurt am Main ; Bern ; New York ; Paris : Lang, 1991

(Europäische Hochschulschriften : Reihe 6, Psychologie ; Bd. 329)

Zugl.: Saarbrücken, Univ., Diss., 1990

ISBN 3-631-43445-6

NE: Europäische Hochschulschriften / 06

Der Druck der vorliegenden Arbeit  
erfolgt mit finanzieller Unterstützung  
der Universität des Saarlandes.

D 291

ISSN 0531-7347

ISBN 3-631-43445-6

© Verlag Peter Lang GmbH, Frankfurt am Main 1991

Alle Rechte vorbehalten.

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Printed in Germany 1 2 3 5 6 7

## Modelle der dreimodalen Faktorenanalyse

# Europäische Hochschulschriften

Publications Universitaires Européennes  
European University Studies

**Reihe VI**

**Psychologie**

Série VI · Series VI

Psychologie

Psychology

**Bd./Vol. 329**



**PETER LANG**

Frankfurt am Main · Bern · New York · Paris

Für Anna und Wilhelm



## Danksagungen

An dieser Stelle möchte ich Herrn Professor Dr. P. Orlik für die Überlassung des Themas und die kontinuierliche, freundliche Unterstützung meiner Bemühungen danken. Herrn Professor Dr. W. Tack danke ich für wertvolle Hinweise und die kritische Durchsicht des Manuskripts. Herrn Dr. A. Kohler danke ich für die hilfreiche Unterstützung in der Entwicklung der Pascal-Programme. Mein Dank gilt ferner meinem Mann, Dr. Dr. H. Schwerdt, für redaktionelle Hinweise.



Gliederung		Seite
1.	Einleitung	11
2.	Von der Covariation Chart zur dreimodalen Faktorenanalyse: einige geschichtliche Marksteine	23
THEORETISCHER TEIL		
3.	Modelle der dreimodalen Faktorenanalyse: Grundgleichungen, Algorithmen und Interpretationen	41
3.1.	Das Tuckersche Modell	47
3.1.1.	Die Grundgleichungen des Modells	47
3.1.2.	Der Tuckersche Algorithmus	50
3.1.3.	Der Algorithmus von Kroonenberg und De Leeuw	52
3.1.4.	Multidimensionale Skalierung individueller Differenzen nach Tucker als Sonderfall der dreimodalen Faktorenanalyse	56
3.1.5.	Die viermodale Erweiterung des Tuckerschen Modells	58
3.1.6.	Interpretation der Kern- und Faktormatrizen des dreimodalen Modells	59
3.2.	Das CANDECOMP/PARAFAC-Modell	63
3.2.1.	Die Grundgleichungen des Modells	63
3.2.2.	Der ALS-Algorithmus zum CANDECOMP/PARAFAC-Modell	67
3.2.3.	Multidimensionale Skalierung individueller Differenzen als Sonderfall von CANDECOMP/PARAFAC	70
3.2.4.	Die n-modale Erweiterung von CANDECOMP/PARAFAC	72
3.2.5.	Interpretation des dreimodalen CANDECOMP/PARAFAC-Modells	73
3.3.	Das SUMMAX-Modell	76
3.3.1.	Die Grundgleichungen des Modells	76
3.3.2.	Der SUMMAX-Algorithmus	80
3.3.3.	Multidimensionale Skalierung individueller Differenzen als Sonderfall von SUMMAX	83
3.3.4.	Die n-modale Erweiterung des SUMMAX-Modells	84
3.3.5.	Interpretation des dreimodalen SUMMAX-Modells	85
4.	Formale Eigenschaften und Spezifika der Modelle der dreimodalen Faktorenanalyse: eine Klassifikation	89
4.1.	Eigenschaften der trilinearen Modelle	91
4.2.	Eigenschaften der quadrilinearen Modelle	98
5.	Zusammenhänge und theoretische Bezüge zwischen den Modellen der dreimodalen Faktorenanalyse	103
5.1.	Der Bezug des SUMMAX-Modells zu CANDECOMP/PARAFAC	104
5.1.1.	Beweis des Zusammenhangs zwischen SUMMAX und CANDECOMP/PARAFAC	105

5.1.2.	Implikationen des Bezuges des SUMMAX-Modells zu CANDECOMP/PARAFAC	111
5.2.	Der Bezug von SUMMAX zu dem Tuckerschen Modell	112
5.2.1.	Beweis des Zusammenhangs zwischen SUMMAX und dem Tuckerschen Modell	115
5.2.2.	Implikationen des Bezuges zwischen SUMMAX und dem Tuckerschen Modell	119
5.3.	Die Rekonstruktion der dreimodalen Modelle aus SUMMAX	121
EMPIRISCHER TEIL		
6.	Anwendungen der Modelle der dreimodalen Faktorenanalyse: Vorbemerkungen	125
6.1.	Untersuchte Datensätze	127
6.2.	Algorithmen und deren Abbruchkriterien	128
7.	Reanalyse eines Modellexperimentes von Orlik zur Psychophysik des Polaritätsprofils	135
7.1.	Normierung des Datenkörpers	138
7.2.	Die CANDECOMP/PARAFAC-Lösung	138
7.2.1.	Verwendete Auswahlkriterien	139
7.2.2.	Interpretation der zweidimensionalen Konfiguration	142
7.2.3.	Überprüfung der faktorenanalytischen Befunde auf der Datenebene	144
7.3.	Die Lösung nach dem Tuckerschen Modell	144
7.3.1.	Kriterien zur Bestimmung der Anzahl substantieller Faktoren	144
7.3.2.	Interpretation der Lösung nach dem Tuckerschen Modell	147
7.4.	Die Lösung nach dem SUMMAX-Modell	153
7.4.1.	Kriterien zur Auswahl einer Lösung	153
7.4.2.	Interpretation der unrotierten SUMMAX-Lösung aus zwei Faktortripeln	153
7.4.3.	Interpretation der SUMMAX-Lösung aus neun Faktortripeln	158
7.5.	Die dreimodalen Konfigurationen im Vergleich	163
8.	Dimensionen der Farbwahrnehmung: ein Vergleich der MDS-Modelle zur Strukturanalyse individueller Differenzen	165
8.1.	Normierung des Datenkörpers	169
8.2.	Die INDSCAL-Lösung	169
8.2.1.	Verwendete Auswahlkriterien	169
8.2.2.	Interpretation der zweidimensionalen Konfiguration	170
8.3.	Die Lösung nach dem Tuckerschen Modell	174
8.4.	Die Lösung nach dem SUMMAX-Modell	179
8.4.1.	Kriterien zur Auswahl einer Lösung	179
8.4.2.	Interpretation der trilinearen SUMMAX-Lösung aus zwei Faktortripeln	181
8.4.3.	Interpretation der rotierten SUMMAX-Lösung	181
8.5.	Schlußbemerkung	186

9.	Varianten der Sortiertechnik: eine Methodenstudie von Rosenberg und Kim	187
9.1	Normierung des Datenkörpers	191
9.2	Die INDSCAL-Lösung	191
9.2.1.	Interpretation der dreidimensionalen Lösung	194
9.3	Die Lösung nach dem Tuckerschen Modell	195
9.3.1.	Interpretation der unrotierten Lösung	195
9.3.2.	Interpretation der rotierten Lösung	201
9.4	Die Lösung nach dem SUMMAX-Modell	202
9.4.1.	Interpretation der unrotierten Konfigura- tion aus drei Faktortripeln	203
9.4.2.	Interpretation der rotierten Konfiguration	206
9.5	Schlußbemerkung	210
10.	Osgood und Luria: ein Fall von multipler Persönlichkeit - ein empirischer Vergleich der dreimodalen Datenmodelle	213
10.1.	Die CANDECOMP/PARAFAC-Lösung	214
10.1.1.	Verwendete Auswahlkriterien	214
10.1.2.	Interpretation der CANDECOMP/PARAFAC- Lösung	219
10.1.3.	Überprüfung der faktorenanalytischen Be- funde auf der Datenebene	221
10.2.	Die ALS-Lösung nach Kroonenberg	222
10.2.1.	Interpretation der Lösung	222
10.2.2.	Überprüfung der faktorenanalytischen Be- funde auf der Datenebene	224
10.3.	Die Lösung nach dem Tuckerschen Modell	225
10.3.1.	Interpretation der Tuckerschen Lösung	225
10.3.2.	Überprüfung der faktorenanalytischen Be- funde auf der Datenebene	228
10.4.	Die Lösung nach dem SUMMAX-Modell	229
10.4.1.	Kriterien zur Auswahl einer Lösung	229
10.4.2.	Interpretation der Lösung	232
10.4.3.	Überprüfung der faktorenanalytischen Be- funde auf der Datenebene	233
10.4.4.	Interpretation der unrotierten SUMMAX- Lösung aus drei Faktortripeln	233
10.5	Die faktorenanalytischen Konfigurationen im Vergleich	236
11.	Diskussion	239
11.1.	Implikationen der theoretischen Bezüge zwischen den Modellen im Vergleich zu den Diagonalisierungsansätzen	240
11.2.	Algorithmus- und Prozessor-bedingte Ein- schränkungen der empirischen Befunde	249
11.3.	Alternative non-metrische Algorithmen	251
11.4.	Befunde aus den empirischen Datenanalysen	252
12.	Zusammenfassung	259
	Literatur	263



1.

Einleitung

Faktorenanalyse ist ein lineares mehrdimensionales Modell, das die Analyse quantifizierbarer Merkmale eines Merkmalsbereiches, gemessen oder beobachtet an einer Stichprobe von Individuen, gestattet. Liegen Beobachtungen oder Messungen zu vielen Merkmalsvariablen vor, so wird mittels der Faktorenanalyse untersucht, ob die Merkmale ohne entscheidenden Informationsverlust auf eine geringe Anzahl von Grundvariablen zurückgeführt werden können. Faktoren, die Linearkombinationen der beobachteten Merkmale sind, wurden in der Faktorenanalyse als Grundvariablen vorgeschlagen (Rao, 1964). Dabei wird von der Annahme ausgegangen, daß die beobachteten, voneinander abhängigen Merkmale aus Grundvariablen erklärbar sind, die selbst nicht direkt meßbar sind, aber den Abhängigkeiten zwischen den beobachteten Merkmalen zugrunde liegen (Weber, 1974; Pawlik, 1968). Dem Merkmalsgeflecht wird in der Faktorenanalyse eine möglichst einfach strukturierte Ordnung aus einer geringen Anzahl an Faktoren unterlegt; nach dieser Funktion wird die Faktorenanalyse als datenreduzierendes Verfahren charakterisiert. Hierbei kommt das Ökonomieprinzip als Grundidee der Faktorenanalyse zum Ausdruck; eine sparsame und befriedigende Darstellung des untersuchten Gegenstandsbereiches liegt vor, wenn die Meßwerte in den Merkmalsvariablen mit einer kleinen Zahl gedachter, hypothetischer Grundvariablen angenähert werden können (Holzinger & Harman, 1941; Pawlik, 1968).

Darüber hinaus ist die zweite Funktion der Faktorenanalyse, ein Ordnungssystem zu generieren, welches eine sinnvolle begriffliche Interpretation gestattet. In der Interpretation der Faktoren sucht man Hypothesen über ihre Bedeutung in dem untersuchten Merkmalsbereich zu gewinnen. Neben der Datenreduktion ermöglicht die Faktorenanalyse die explorative Analyse eines noch wenig vorstrukturierten Gegenstandsbereiches. Sie führt zu Hypothesen über das Geflecht der Merkmale, ohne daß man zuvor eine bestimmte Struktur kennen oder annehmen muß (Überla, 1968). Daher wird sie als heuristisches Verfahren bezeichnet.

Faktoren sind geometrisch Koordinatenachsen, welche ein Bezugssystem für die Darstellung der Merkmalsvariablen lie-

fern. In der Faktorenanalyse wird untersucht, wieviele Koordinatenachsen oder Dimensionen zur Erfassung der Merkmale benötigt werden. Dies führt zu der dritten Funktion der Faktorenanalyse: sie ist ein Verfahren zur Überprüfung der Dimensionalität der beobachteten Merkmale (Bortz, 1979).

Nach ihrer historischen Entwicklung und ihrem gegenwärtigen Stand ist die Faktorenanalyse ein eingeführtes Verfahren zur Aufklärung der internen Struktur multivariater Zusammenhänge (Seber, 1984; Jolliffe, 1986). Als multivariates Verfahren bezieht sie sich auf einen komplexen Merkmalsbereich, der nicht durch eine einzige Merkmalsvariable, sondern immer nur durch eine Gruppe von  $m$  Variablen erfaßt wird. Der multivariate Aspekt kommt in der Vielfalt der Variablen zum Ausdruck (Weber, 1974); eine Faktorenanalyse ist nicht möglich, wenn Beobachtungen oder Meßwerte anhand einer einzigen Variablen erhoben wurden (dies kennzeichnet die univariate Betrachtungsweise; Pawlik, 1968). Ferner wird in der Charakterisierung der Faktorenanalyse als Modell zur Analyse quantitativer Merkmale davon ausgegangen, daß die Beobachtungen nicht nur an einem Individuum, sondern an einer Stichprobe von  $n$  Individuen registriert wurden. Es werden also zwei Mengen von "Elementen" oder "Entitäten" (Cattell, 1966), in diesem Fall Individuen und Merkmalsvariablen, berücksichtigt. Dabei wird angenommen, daß die Erhebung aller Beobachtungen jeweils unter einer Untersuchungsbedingung erfolgt. Wird die Erhebung einer Reihe von Variablen an einer Stichprobe von Individuen erweitert von einer Untersuchungsbedingung auf mehrere Bedingungen, so wird hierdurch eine weitere Menge von Entitäten hinzugefügt. Die dreimodale Faktorenanalyse gestattet die Analyse von Meßwerten oder Beobachtungen, die der Kombination von drei Mengen von Entitäten zugeordnet sind. Der Terminus "Modus" wurde von Tucker (1964a, 1966) als Bezeichnung für eine Menge von Entitäten eingeführt. D.h. ein Modus ist definiert als eine Menge von Entitäten (Carroll & Arabie, 1980), und der Begriff "dreimodale Faktorenanalyse" charakterisiert die Ausgangsdaten, die mittels faktorenanalytischer Verfahren zu

analysieren sind, und das Ziel der Analyse, das in der Konstruktion eines interpretierbaren, dimensionalen Ordnungssystems für die Beobachtungen besteht.

Allgemein ist eine dreimodale Anordnung definiert als das kartesische Produkt von drei unterschiedlichen Modi. Ein Element dieser Anordnung ist ein bestimmter Wert des kartesischen Produktes, d.h. eine Kombination bestimmter Entitäten der Modi wie z.B.  $(V_i, P_j, S_k)$ , wobei  $V_i$  die  $i$ -te Variable,  $P_j$  das  $j$ -te Individuum und  $S_k$  die  $k$ -te Untersuchungsbedingung bezeichnet. Jedes beobachtete Datum  $x_{ijk}$  ist einer Kombination  $(V_i, P_j, S_k)$  zugeordnet; aus der Zuweisung numerischer Werte  $x_{ijk}$  zu den Elementen einer dreimodalen Anordnung entsteht eine dreimodale Datenmatrix  $X = (x_{ijk})$ . Unter dem Begriff "dreimodale Faktorenanalyse" sind daher lineare mehrdimensionale Modelle zur Analyse dreimodaler Datenmatrizen zu verstehen. In der Mehrzahl der dreimodalen Verfahren wird vorausgesetzt, daß die zu analysierenden Daten mindestens Intervallskalengüte besitzen und vollständig sind, indem für jede Kombination  $(V_i, P_j, S_k)$  ein Meßwert vorliegt.

Die seit der Jahrhundertwende entwickelte Faktorenanalyse ermöglichte zunächst die Analyse zweimodaler Datenmatrizen, aufgebaut aus der Kombination von Individuen und Variablen. Das erste Modell zur Analyse dreimodaler Datenmatrizen wurde 1964 von Tucker vorgestellt. Auf die geschichtliche Entwicklung der dreimodalen Faktorenanalyse und anwendungsbezogene Beispiele wird im zweiten Abschnitt der vorliegenden Arbeit eingegangen. In der klassischen, zweimodalen Faktorenanalyse war es möglich, ein allgemeines Faktorenmodell anzugeben, welches in wenigen Grundgleichungen das Modell der Zusammensetzung der Merkmale aus nicht direkt beobachteten Grundvariablen definierte. In der dreimodalen Faktorenanalyse gibt es bisher kein übergeordnetes, allgemeines Faktorenmodell, sondern es liegen mindestens zwei unterscheidbare Modellklassen vor. Man geht davon aus, daß jeder Meßwert  $x_{ijk}$  als eine Funktion mehrerer hypothetischer Größen, den Faktoren,

aufzufassen ist. Die Beziehung zwischen den Beobachtungen oder Meßwerten und den Faktoren wird durch ein trilineares Modell der Form

$$\begin{aligned} x_{ijk} &= a_{i1}b_{j1}c_{k1} + \dots + a_{ir}b_{jr}c_{kr} + \dots + a_{is}b_{js}c_{ks} \\ &= \sum_r a_{ir}b_{jr}c_{kr} \end{aligned}$$

oder durch ein quadrilineares Modell mit der Grundgleichung

$$x_{ijk} = \sum_m \sum_p \sum_q a_{im} b_{jp} c_{kq} g_{mpq}$$

dargestellt. In beiden Ansätzen werden die Beobachtungen als Linearkombinationen von Faktoren beschrieben, wobei die Verfahren zur Bestimmung der Faktoren auf Methoden der linearen Algebra zurückgehen. Im folgenden werden die begrifflichen Konzepte, die in der dreimodalen Faktorenanalyse zu berücksichtigen sind, hauptsächlich anhand des trilinearen Modells beschrieben.

Die Koeffizienten  $a_{ir}$ ,  $b_{jr}$  und  $c_{kr}$  heißen die Faktorladungen der Entitäten  $V_i$ ,  $P_j$ ,  $S_k$  der drei Modi. Ein Faktor  $a_r$  (bzw.  $b_r$  und  $c_r$ ) ist darstellbar als ein Vektor, dessen Komponenten die Faktorladungen  $a_{ir}$  für alle Entitäten  $V_i$  (bzw.  $b_{jr}$ ,  $c_{kr}$  für alle Entitäten  $P_j$  und  $S_k$ ) sind. Die Ladung  $a_{ir}$  gibt die Bedeutung des  $r$ -ten Faktors  $a_r$  für die Entität  $V_i$  an, und analog werden die Koeffizienten  $b_{jr}$  und  $c_{kr}$  als Indikatoren der Bedeutsamkeit des  $r$ -ten Faktors  $b_r$  und des Faktors  $c_r$  für die Entitäten  $P_j$  und  $S_k$  aufgefaßt. Je höher die Ladung  $a_{ir}$  ist, desto höher ist der Einfluß von Faktor  $a_r$  auf das Merkmal  $V_i$ . Umgekehrt ist der Faktor  $a_r$  ohne Einfluß auf  $V_i$ , wenn die Faktorladung  $a_{ir}$  nahe null ist. Vor der Analyse bekannt sind lediglich die Beobachtungen  $x_{ijk}$ . Aufgabe der Faktorenanalyse ist es, durch ein geeignetes Kalkül die Faktorladungen zu bestimmen. Hierzu wurden zwei wesentliche Faktorextraktionsverfahren entwickelt, die dreimodale Generalisierung der Hauptkomponentenmethode sowie der Centroidmethode. Beide Verfahren sind innerhalb der trilinearen Modellklasse einsetzbar. Die zu bestimmenden Faktoren

$a_r$ ,  $b_r$  und  $c_r$ ,  $r = 1, \dots, s$ , sind algebraisch definiert, sie können geometrisch als Koordinatenachsen gedeutet werden, und sie werden begrifflich interpretiert.

Jeder Faktor  $a_r$ ,  $b_r$  und  $c_r$  ist algebraisch als eine Linearkombination der Ausgangsdaten definiert, als eine Zusammenfassung der Beobachtungen über jeweils zwei Modi. Die Art der Zusammenfassung wird festgelegt durch die speziellen Annahmen der einzelnen dreimodalen Verfahren. Zur Veranschaulichung sei vereinfachend angenommen, daß die Koeffizienten  $b_{jr}$  und  $c_{kr}$  bereits bekannt sind. Dann können aus  $b_{jr}$  und  $c_{kr}$  Gewichte  $b_{jr}^*$  und  $c_{kr}^*$  bestimmt werden, aus deren Multiplikation mit den Ausgangsdaten der Faktor

$$a_r = \sum_j \sum_k x_{ijk} b_{jr}^* c_{kr}^*$$

resultiert. In analoger Weise ist jeder Faktor  $b_r$  ( $c_r$ ) aus der Bekanntheit der Koeffizienten  $a_{ir}$  und  $c_{kr}$  ( $a_{ir}$  und  $b_{jr}$ ) als Linearkombination der Ausgangsdaten konstruierbar. Geometrisch werden die Faktoren als Koordinatenachsen interpretiert, die einen  $s$ -dimensionalen Raum aufspannen. Die Faktoren  $a_r$  werden als Bezugsachsen aufgefaßt, die eine dimensionale Repräsentation der Entitäten  $V_i$  gestatten (analog  $b_r$  und  $c_r$ ). Die Koordinaten der Entitäten  $V_i$  auf den Koordinatenachsen sind durch ihre Faktorladungen  $a_{ir}$  gegeben. Auf diese Weise erhält man einen Variablenraum, einen Personenraum und einen Raum für die Untersuchungsbedingungen. Dabei sind jeweils die Bezugsachsen  $a_r$ ,  $b_r$  und  $c_r$  einander zugeordnet, sie bilden ein Tripel von Dimensionen oder Faktoren. In der Faktorenanalyse wird überprüft, welche Anzahl an Dimensionen zur näherungsweise Reproduktion der Ausgangsdaten benötigt werden. Daher wurde die Faktorenanalyse eingangs als ein Verfahren zur Überprüfung der Dimensionalität der Beobachtungen charakterisiert. Mittels der Faktorenanalyse wird der Struktur der Beobachtungen ein Referenzsystem von Bezugsachsen unterlegt (Cattell, 1952b; Guilford, 1954). Generell werden nur wenige der  $s$  Faktoren einen hohen Beitrag zur gesamten Varianz der Ausgangsdaten liefern. Werden  $t$  Faktoren mit  $t < s$  identifiziert, die

einen hohen Anteil der Varianz erfassen, und tragen die übrigen  $s - t$  Faktoren nur gering zur Aufklärung der Varianz in den Daten bei, so werden nur die  $t$  "varianzstarken" Faktoren berücksichtigt (Weber, 1974). In der geometrischen Interpretation bedeutet dies, daß von einem höherdimensionalen Raum auf einen niederdimensionalen Raum mit relativ wenigen Bezugsachsen übergegangen wird. Diese Vorgehensweise entspricht dem Ökonomieprinzip als Grundidee der Faktorenanalyse, indem den Beobachtungen eine möglichst einfach strukturierte Ordnung aus einer geringen Anzahl an Faktoren unterlegt wird.

Über die algebraische Betrachtungsweise hinaus wird die Transformation

$$a_r = \sum_j \sum_k x_{ijk} b_{jr}^* c_{kr}^* \quad , \quad r = 1, \dots, t$$

als Abbildung der alten Variablen oder allgemein Entitäten in neue hypothetische Variablen  $a_r$  gedeutet. In der begrifflichen Interpretation der Faktoren sucht man Hypothesen über ihre psychologische Bedeutung zu gewinnen (Pawlik, 1968). Daher ist die Faktorenanalyse, wie bereits erwähnt, ein heuristischer Ansatz. Inhaltlich sind Faktoren das, was die Variablen gemeinsam haben, welche hoch auf den Faktoren laden (Revenstorf, 1980). Zur Interpretation eines Faktors werden die Variablen ausgelesen, die auf dem Faktor betragsmäßig hohe Ladungen aufweisen. Der in Betracht gezogene Faktor faßt dann die ausgelesenen Variablen zusammen. Man sucht nun diejenige Eigenschaft begrifflich zu fassen, die diesen Variablen inhaltlich gemeinsam ist, und die Interpretation des Faktors besteht in der Zuordnung der gewählten Eigenschaft zu dem Faktor (Pawlik, 1968). In der dreimodalen Faktorenanalyse werden die Faktoren für alle drei Modi in dieser Weise begrifflich gedeutet. Beispielsweise ist ein Faktor, der hohe Ladungen für erfolgsorientierte und leistungsmotivierte Personen aufweist, als "Karriere-Faktor" oder "Karriere-Typ" interpretierbar. Ein Faktor, auf dem etwa Items wie "nervös", "unkonzentriert" und "gehemmt"

eines Persönlichkeitsfragebogens laden, kann als "emotionale Labilität" gedeutet werden. Oder Situationen des beruflichen Umfeldes wie "Einstellungsgespräch", "Diskussion mit dem Vorgesetzten", "Teamarbeit" würden durch gemeinsame hohe Ladungen einen Faktor "interpersonale Situationen im Beruf" konstituieren. Tucker (1966) bezeichnet Faktoren als "derivational" oder "idealized types" und differenziert zwischen dem Faktorkonzept für Variablen, Personen und Situationen. In dem Beispiel würde der "Karriere-Typ" als idealtypische Person, der Emotionsfaktor als latente Variable und der Faktor "interpersonale Situationen" als prototypische Situation aufgefaßt (vgl. Kroonenberg, 1983b). Durch die Interpretation der Faktoren ermöglicht die dreimodale Faktorenanalyse Hypothesen über das Zusammenwirken der Entitäten. Die Koppelung der Faktoren der verschiedenen Modi erfolgt durch ihre Anordnung in jeweils einem Tripel  $a_r$ ,  $b_r$ ,  $c_r$ , wobei inhaltliche Bezüge zwischen den einem Tripel zugeordneten Faktoren hergestellt werden und die Faktorladungen eines ausgewählten Modus als Indikatoren der Bedeutsamkeit einer Faktorkombination  $a_r$ ,  $b_r$ ,  $c_r$  interpretiert werden. Hieraus resultieren Hypothesen derart, daß das Erleben oder Verhalten der idealtypischen Person  $a_r$  durch die latente Variable  $b_r$  in der prototypischen Situation  $c_r$  beeinflußt wird. Das quadrilineare Grundmodell unterscheidet sich im wesentlichen von dem trilinearen Ansatz dadurch, daß die zusammengehörigen Faktoren  $a_m$ ,  $b_p$  und  $c_q$  aus den drei Modi nicht jeweils in einem Tripel angeordnet sind, sondern durch die Koeffizienten  $g_{mpq}$  verknüpft werden. D.h. die Koeffizienten  $g_{mpq}$ , die als Kernrelationen bezeichnet werden, enthalten Informationen zu dem Ausmaß des Zusammenhangs zwischen  $a_m$ ,  $b_p$  und  $c_q$ . Die Matrix  $G = (g_{mpq})$ , deren Zellen die Koeffizienten  $g_{mpq}$  enthalten, wird Kernmatrix genannt. Schließlich sehen Faktorenanalytiker wie Cattell und Thurstone in den Faktoren invariante Grundfunktionen des Verhaltens und allgemeine, voneinander abgrenzbare Grundtypen von Individuen und Situationen, wenn aus weiteren Untersuchungen eine Replikation der faktorenanalytischen Befunde resultiert, die auch für andere als die untersuchte

Personen-, Situationen- und Variablenstichprobe gelten.

Zu den Modellen der dreimodalen Faktorenanalyse gehören der PARAFAC-Ansatz von Harshman (1976), das von Cattell (1966) entwickelte "Interaktivmodell", das CANDECOMP-Modell von Carroll und Chang (1970), das von Orlik (1980) entwickelte SUMMAX-Modell, die dreimodale Hauptkomponentenanalyse von Tucker (1966) sowie zwei von Bloxom (1968) und Bentler und Lee (1978) vorgeschlagene Varianten des Tuckerschen Ansatzes (vgl. Snyder, Law & Hattie, 1984). Im anglo-amerikanischen Schrifttum fanden vor allem CANDECOMP, PARAFAC und das Tuckersche Modell als Grundkonzeptionen der dreimodalen Datenanalyse Beachtung (vgl. Kruskal, 1984). In empirischen Untersuchungen wurden bisher hauptsächlich CANDECOMP, PARAFAC, der Tuckersche Ansatz und das SUMMAX-Verfahren eingesetzt.

Diese Modelle sind in einer Reihe von Überblicksarbeiten hinsichtlich ihrer spezifischen Annahmen und der zugeordneten Verfahren zur Faktorenextraktion diskutiert worden (vgl. Law & Snyder, 1979; Carroll & Arabie, 1980; Kroonenberg, 1983b; Snyder, Law & Hattie, 1984). Demgegenüber wurden die Zusammenhänge zwischen den genannten Modellen noch nicht umfassend untersucht. Bekannt ist lediglich die generelle Äquivalenz von PARAFAC und dem Interaktivmodell von Cattell zu dem dreimodalen CANDECOMP-Ansatz. Formal gehören PARAFAC und CANDECOMP der trilinearen Modellklasse an wie auch das SUMMAX-Modell, das durch Einführung einer Kernmatrix zu einem quadrilinearen Ansatz erweitert wurde. Dabei ist bisher nicht untersucht worden, ob Bezüge zwischen CANDECOMP, PARAFAC und dem trilinearen SUMMAX-Verfahren bestehen und inwieweit diese Zusammenhänge formalisierbar sind. Ferner stehen bisher die trilinearen Modelle und der Tuckersche Ansatz als Repräsentant der quadrilinearen Modellklasse insofern unverbunden nebeneinander, als keine allgemeingültigen Zusammenhänge zwischen den beiden Klassen theoretisch aufgezeigt worden sind.

Darüber hinaus hat der Einsatz der verschiedenen Verfahren

in empirischen Datenanalysen eine kontroverse Diskussion um ihren Stellenwert ausgelöst. In Abhängigkeit vom untersuchten Gegenstandsbereich und der Fragestellung wurden teilweise die trilinearen Ansätze, teilweise das Tuckersche Modell favorisiert. Aus bisherigen Erfahrungen wurde die Schlußfolgerung gezogen, daß der Tuckersche Ansatz zu differenzierteren Aussagen führt als die trilinearen Verfahren, aber mit erheblichen Schwierigkeiten in der Interpretation der Kernrelationen behaftet ist (vgl. Rowe, 1979; MacCallum, 1974; Rösler, 1979). Von anderen Autoren (vgl. Carroll & Wish, 1974; Carroll & Pruzansky, 1984) wurden die trilinearen Ansätze als vorzuziehende Verfahren diskutiert, weil sie auf eine komplexe Verflechtung der Faktoren in Form der Kernrelationen verzichten und daher zu einer besser interpretierbaren Strukturierung der Ausgangsdaten führen.

Aus den bisher offenen Fragestellungen zu den formalen Zusammenhängen der Modelle und ihren inhaltlich-interpretativen Bezügen entstand eine zweiteilige Aufgabenstellung für die vorliegende Arbeit. Innerhalb der trilinearen Modellklasse stehen SUMMAX und CANDECOMP/PARAFAC unverbunden nebeneinander und mögliche Bezüge wurden bisher nicht diskutiert, obwohl diese Verfahren von übereinstimmenden Modellgleichungen ausgehen. Aufgrund von Ergebnissen aus dem Bereich der multilinearen Algebra (vgl. Kruskal, 1977) ist aber bekannt, daß auch Verfahren, die unter einem grundlegend verschiedenen Rationale entwickelt wurden, zu äquivalenten Lösungen führen können und daher direkt ineinander überführbar sind. Liegt eine wechselseitige Äquivalenz zwischen den Verfahren vor, so konstituieren die trilinearen Ansätze eine konzeptuell und formal einheitliche Modellklasse, deren Repräsentanten zu Konfigurationen führen, die aus der Bekanntheit einer einzigen Lösung rekonstruierbar sind. Ein solcher Bezug würde den Einsatz nur eines Verfahrens erfordern, aus dem die Lösungen der übrigen Ansätze ohne erneute Datenanalysen direkt hergestellt werden können. Daher soll zunächst untersucht werden, ob theoretisch-alge-

braische Bezüge zwischen CANDECOMP, PARAFAC und dem SUMMAX-Verfahren bestehen und inwieweit diese Ansätze formal ineinander überführt werden können.

Obgleich im Rahmen empirischer Datenanalysen vereinzelt versucht wurde, das Tuckersche Modell auf eine trilineare Form zurückzuführen, wurden bisher CANDECOMP/PARAFAC und der Tuckersche Ansatz als Konzeptionen mit grundlegend verschiedenen Annahmen über die Komplexität der zu analysierenden Daten behandelt (vgl. Rösler, 1979). Nach dieser Auffassung sind die Konfigurationen aus den trilinearen und quadrilinearen Verfahren strukturell nicht vergleichbar, und vom Anwender ist eine Entscheidung zugunsten eines Modells zu treffen. Diese prinzipielle Inkompatibilität ist hauptsächlich durch empirische Befunde begründet worden, während modelltheoretische Untersuchungen der Bezüge zwischen beiden Modellklassen nur in speziellen Sonderfällen erfolgreich waren. Wenn dagegen die trilinearen Konfigurationen auf zugeordnete quadrilineare Grundstrukturen generell abgebildet werden können, so wird hierdurch impliziert, daß die Ansätze im wesentlichen gleiche Informationen enthalten und diese nur in unterschiedlicher Weise darstellen. Auf der Grundlage der strukturellen Verknüpfungen der trilinearen Verfahren ist daher zu untersuchen, ob die trilinearen Ansätze unter allgemeinen Bedingungen in das Tuckersche Modell überführt werden können und damit die faktorenanalytischen Konfigurationen nach dem Tuckerschen Modell formal aus den Lösungen trilinearere Verfahren rekonstruierbar sind.

Im Rahmen empirischer Datenanalysen wurde bisher nur in sehr wenigen Untersuchungen ein direkter Methodenvergleich vorgenommen. In der Mehrzahl wurde ein einziges Verfahren zur Analyse eines Datensatzes eingesetzt, dessen Brauchbarkeit anhand der Plausibilität der Ergebnisse ohne Analyse der durch andere Verfahren gewonnenen Strukturaussagen beurteilt wurde. Um einen systematischen Vergleich der Verfahren und eine Untersuchung ihrer Unterschiede oder Gemeinsamkeiten in der Erzeugung eines dimensional Ordnungssystems und daraus ableitbarer Hypothesen zu ermöglichen, wird eine

Modelluntersuchung durchgeführt, in welcher die Daten aus ausgewählten Methodenexperimenten jeweils durch die verschiedenen Verfahren analysiert werden.

Entsprechend dieser Aufgabenstellung ist die vorliegende Arbeit in einen theoretischen und einen empirischen Teil gegliedert. Nach einem historischen Abriß zur Entwicklung des Gebietes der dreimodalen Faktorenanalyse in Abschnitt 2 werden im theoretischen Teil zunächst die Grundgleichungen der betrachteten Modelle, Algorithmen zur Faktorenextraktion und mögliche begriffliche Interpretationen in Abschnitt 3 dargestellt. Die Klassifikation der Verfahren in eine trilineare und eine quadrilineare Modellklasse gestattet eine Abgrenzung unterschiedlicher algebraischer Eigenschaften der Ansätze; die Diskussion der modellspezifischen Eigenschaften ist Gegenstand von Abschnitt 4. In Abschnitt 5 werden generelle algebraische Bezüge zwischen CANDECOMP, PARAFAC und SUMMAX sowie die formale Transformation dieser Verfahren in quadrilineare Modelle abgeleitet. Im empirischen Teil wird nach einer Darstellung der mit den verschiedenen Verfahren verknüpften Abbruchkriterien der Faktorenextraktion in Abschnitt 6 ein Methodenvergleich anhand der Daten aus vier Untersuchungen vorgenommen, deren zugrunde liegender Versuchsaufbau jeweils eine Vorhersage der relevanten faktoranalytischen Befunde und eine Bewertung wesentlicher Hypothesen aus der Anwendung der verschiedenen Verfahren auf dem Hintergrund bekannter Strukturmerkmale der Daten gestattet.

## 2.

Von der Covariation Chart zur drei-  
modalen Faktorenanalyse: einige ge-  
schichtliche Marksteine

Traditionell wird in faktorenanalytischen Untersuchungen von einer Reihe von Variablen ausgegangen, die einen bestimmten Merkmalsbereich (z.B. Gedächtnisleistungen oder Persönlichkeitsmerkmale) erfassen und an jedem Individuum einer Stichprobe gemessen oder beobachtet wurden (Weber, 1974). Klassische Untersuchungen stammen aus dem Gebiet der Intelligenzforschung und nahmen ihren Anfang 1904 mit dem englischen Psychologen C. Spearman, der sich mit Korrelationen zwischen Aufgaben befaßte, die der Messung der Intelligenz dienten. In der Folgezeit waren die Korrelationen zwischen den Aufgaben bzw. Merkmalen die Ausgangsdaten, deren Struktur aufgeklärt werden sollte. Seit der 1909 veröffentlichten Arbeit von Burt und dem Erscheinen der "multiple factor analysis" von Thurstone im Jahre 1931 wurde von der Modellvorstellung ausgegangen (vgl. Pawlik, 1968), daß jedes der gemessenen Merkmale als eine gewichtete Summe von mehreren fundamentalen Einflußgrößen aufzufassen sei, wobei diese Einflußgrößen hypothetischer Natur sind und Faktoren genannt werden. Jedes einzelne Datum aus den erhobenen Meßwerten oder Beobachtungen ist der Kombination einer bestimmten Person und eines bestimmten Merkmals zugeordnet und enthält Informationen über den Zusammenhang zwischen den "Entitäten" beider Mengen, die von Tucker (1964) als "Modi" bezeichnet wurden. Einen Überblick über die geschichtliche Entwicklung der klassischen faktorenanalytischen Ansätze gibt Harman (1960).

Durch Erweiterung der Messung auf beispielsweise unterschiedliche Untersuchungsbedingungen wird eine weitere Menge von "Entitäten" hinzugefügt, aus der ursprünglich zweimodalen Datenmatrix, aufgebaut aus der Kombination der Probanden und Merkmale, entsteht ein dreimodaler Datenquader aus Personen, Merkmalen und Untersuchungsbedingungen. Tucker (1964) formulierte das erste Modell zu einer faktorenanalytischen Behandlung dreimodaler Fragestellungen, die eine gleichzeitige Darstellung der Struktur in allen Modi gestattet. Dabei ist die Diskussion um einen angemessenen Zugang zu komplexeren Datenstrukturen, die auf der Verknüpfung von mehr als

zwei Modi beruhen, wesentlich älter, und ihr Beginn wird nach Cattell (1966) und Snyder, Law und Hattie (1984) auf das Jahr 1946 datiert, in dem Cattells Buch "Description and measurement of personality" erschien. Durch die Anordnung der Personen, Tests und Untersuchungsbedingungen (vgl. auch Cattell, 1950, 1952a) entlang dreier Achsen der "Covariation Chart" wurden über die herkömmlichen R- und P-Techniken hinaus zusätzliche Möglichkeiten der Datenanalyse aufgezeigt. In der Weiterführung der Überlegungen von Spearman wurden bis zu diesem Zeitpunkt hauptsächlich Korrelationen zwischen Merkmalen aufgrund von untersuchten Probanden (R-Technik) und Korrelationen zwischen Probanden aufgrund von Merkmalen (P-Technik) für eine faktorenanalytische Behandlung zugrunde gelegt. Cattell diskutierte auch die traditionellen Verfahren zur Analyse zweimodaler Datenkörper; jedoch, wie die "Covariation Chart" zeigte, eröffnet ein solcher Datenquader die Berechnung vier weiterer Korrelationsmaße, die Grundlage der sogenannten Q-, O-, S- und T-Techniken sind (vgl. Abbildung 1). Zunächst war damit die Konzeption eines dreimodalen Designs in den Mittelpunkt gerückt, und es war bekannt, daß die sechs unterschiedlichen Techniken in drei Paaren - R-P, Q-S und O-T - zusammengefaßt werden können, die jeweils systematische Bezüge zueinander aufweisen. Mögliche Zusammenhänge zwischen diesen drei Paaren sollten zwar untersucht werden, aber waren noch unbekannt: "These alone yield factors in statistically independent, unbridgeable measurement systems, and are the pillars between which bridges of paired scientific inference and relation can be built" (Cattell, 1952a, S. 506). Der faktorenanalytische Zugang zu dreimodalen Daten beschränkte sich weiterhin während dieser und der folgenden Dekade auf die klassischen ein- bzw. zweimodalen Methoden, wobei einer der Modi des zugrunde liegenden Datenkörpers durch Mittelung über die zugehörigen Elemente eliminiert wurde. Dennoch wurde die dreimodale Betrachtungsweise in diesem Zeitraum vermehrt diskutiert, etwa von Burt (1955) und Guttman (1958). So betont Burt (1955), daß sich die elementarste Form der Beobachtung ("the most elementary observation", S.

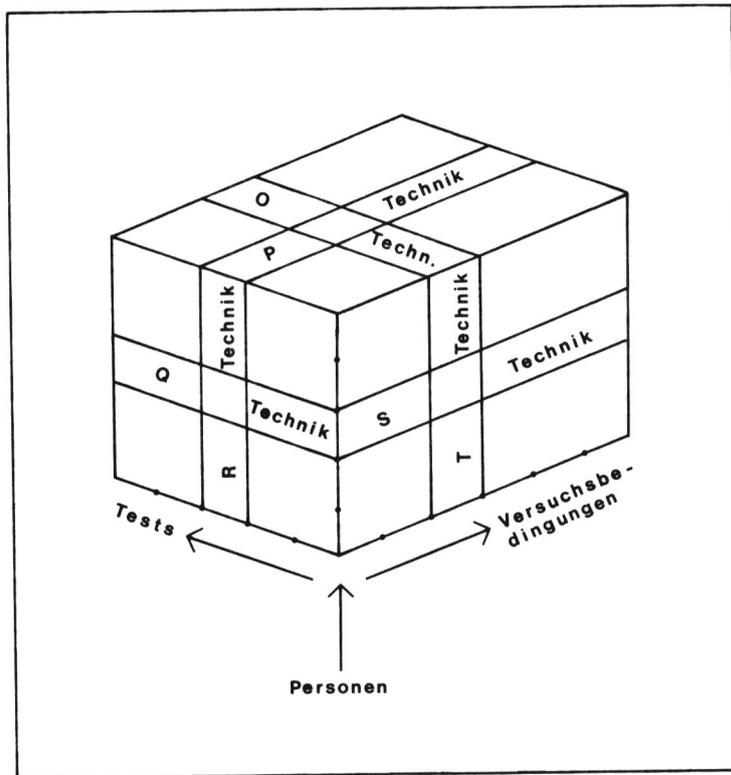


Abb. 1: Die Covariation Chart von Cattell (1946)

104) aus dem Tripel Person, Merkmal und Versuchsbedingung konstituiert und daher aus drei Dimensionen gebildet wird. Ferner klassifiziert Burt die Elemente, die an der gesamten Variation der Daten beteiligt sind, in einer Weise, die mit der Covariation Chart völlig identisch ist (vgl. auch Mahmoud, 1955).

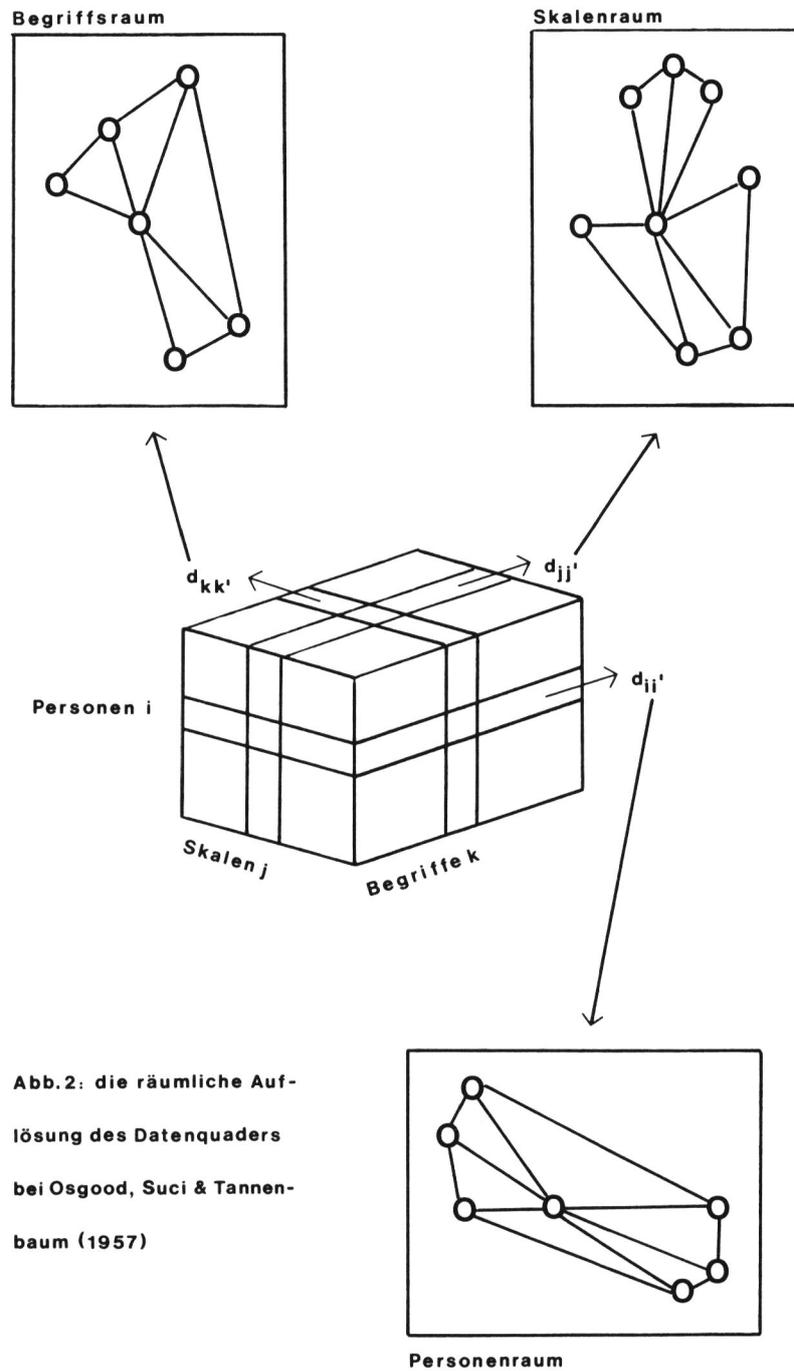
Im gleichen Jahr - 1955 - findet ein internationales Kolloquium zur Faktorenanalyse in Paris statt. Guttman faßt 1958 die aus seiner Perspektive wesentlichen Beiträge des Kolloquiums zu der aktuellen Intelligenzforschung zusammen und stellt einen dreimodalen Zugang zur Struktur der Intelligenz in den Vordergrund, der eng mit der Guilfordschen Theorie verknüpft ist. Guttman stellt zunächst den Zusammenhang zwischen den wesentlichen Elementen der Guilfordschen Theorie, den Denkopoperationen, Denkinhalten und Denkergebnissen (Hofstätter, 1977) in Form eines kartesischen Produktes dieser Elemente dar und postuliert zur "Erklärung" (Guttman, 1958, S. 509) der Relationen zwischen Denkopoperationen, -inhalten und -ergebnissen eine Kombination aus Faktorenanalyse im Sinne von Spearman und Varianzanalyse. Zur Entwicklung weiterer methodischer Zugänge schlägt Guttman die Erweiterung seiner Radex-Theorie vor, später als Facettentheorie bekannt geworden (Borg, 1986), aber er konstatiert gleichzeitig, daß in Zukunft die Analyse solch komplexer Datenstrukturen eine algebraische Form mit sich bringen könnte, die von der bekannten weit abweicht.

Über die genannten Arbeiten aus dem Bereich der Intelligenzforschung hinaus gehören Osgood, Suci und Tannenbaum (1957) zu den Wegbereitern der dreimodalen Datenanalyse. In ihrer Monographie "The measurement of meaning" stellen sie ein neues Verfahren zur Erfassung konnotativer Wortbedeutungen vor, das Semantische Differential. Dieses Instrument besteht aus einer Reihe bipolarer Ratingskalen, die inhaltlich die drei nach Osgood, Suci und Tannenbaum wesentlichen Dimensionen repräsentieren, welche dem "semantischen Raum" zugrunde liegen, und mit "evaluation", "activity" und "potency" bezeichnet werden. Entsprechend der Theorie von Osgood,

Suci und Tannenbaum bieten die bipolaren Ratingskalen einen Zugang zu dem "semantischen Raum", und sie werden im Raum der drei Dimensionen dargestellt. In der Regel werden den Versuchspersonen verschiedene Begriffe vorgegeben, die auf den Ratingskalen einzustufen sind, so daß aus der Kombination Personen x Begriffe x Skalen ein Datenquader entsteht. Im Mittelpunkt der Analyse steht die Frage nach den semantischen Ähnlichkeitsrelationen zwischen den Begriffen. Es wird angenommen, daß die konnotative Bedeutung jedes der einzustufenden Begriffe durch seine Lokalisation auf den Dimensionen "evaluation", "activity" und "potency" im semantischen Raum erfaßt wird, und diese Lokalisation erfolgt aus den Einstufungen des Begriffs auf den Ratingskalen. Unterschiede in der Bedeutung von zwei Begriffen sind dann eine Funktion ihrer unterschiedlichen Lokalisationen im gleichen Raum. Darüberhinaus sind auch Unterschiede in den Urteilsystemen der Personen von Interesse, die durch Divergenzen in den individuellen semantischen Räumen bedingt sind.

In Ermangelung einer Methode, durch welche Relationen zwischen den Elementen der drei Modi simultan analysiert werden können, wählten Osgood, Suci und Tannenbaum ein Verfahren zur Datenauswertung, in welchem für eine geometrische Repräsentation für jeden Modus getrennt zunächst die beiden übrigen Modi durch Summation bzw. Mittelwertberechnung ausgeblendet werden und in einem zweiten Schritt aus den so reduzierten Daten Distanzen bestimmt werden. Diese Auswertungen werden entweder auf der Basis der Rohdaten oder einer vorhergehenden Faktorenanalyse der Skalen vorgenommen, wobei Osgood, Suci und Tannenbaum auch verschiedene Alternativen zu dieser Vorgehensweise diskutieren. Abbildung 2 zeigt in schematischer Form die Auflösung des Datenquaders bei Osgood, Suci und Tannenbaum.

Abelson (1958) diskutiert indessen einen diskriminanzanalytischen Zugang zu Datenstrukturen von der Art, die Osgood, Suci und Tannenbaum betrachten. Auch in diesem Ansatz werden Linearkombinationen aufgrund der Personen und Skalen gebildet, und so stellt Abelson (1960) fest, daß seine Methode in



**Abb. 2: die räumliche Auf-  
lösung des Datenquaders  
bei Osgood, Suci & Tannen-  
baum (1957)**

wesentlichen Aspekten dem Vorgehen von Osgood, Suci und Tannenbaum äquivalent ist. Bis 1960 wird damit eine Reihe von Modellen vorgestellt, die der Komplexität von dreimodalen Daten gerecht werden sollen, aber im Grunde die Priorität eines Modus implizieren. Hierzu bemerkt Abelson (1960, S. 170): "The traditional thinking about the factor problem with three-way tables has been that information must be skillfully lost by picking one 'way', forming correlations or covariances over another 'way' and ignoring the effect of the third 'way' ."

Mit Tuckers Abhandlung (1964a) und der Dissertation von Levin (1963) wird ein Wendepunkt in der Auseinandersetzung um die Analyse dreimodaler Daten eingeleitet. Mit diesen Arbeiten wird das erste dreimodale faktorenanalytische Modell unter der Bezeichnung "three-mode factor analysis" eingeführt. Das Modell ist ein Beispiel für eine theoretische Konzeption, die in der Mathematik als multilineare Algebra bekannt ist, aber bis 1976 gibt es kaum Querbezüge zu der Mathematik. Während die Vorläufer von 1946 bis 1960 latente Strukturen aus einem Modus in den Vordergrund stellten, führt das Tuckersche Modell zu einer dimensional Repräsentation für jedem Modus, wobei der wechselseitige Zusammenhang zwischen den drei faktoriell reduzierten Modi der sogenannten "Kernmatrix" zugeordnet ist (vgl. Tucker, 1966; Levin, 1965). Das Modell geht beispielsweise bei einem Design wie in der Covariation Chart davon aus, daß die Tests aus idealisierten Variablen (Testfaktoren) bestehen, die für idealisierte Personen (Personenfaktoren) in idealisierten Versuchsbedingungen (Situations- oder Versuchsbedingungsfaktoren) unterschiedlich bedeutsam sind (Bartussek, 1973, S. 170). Die variierende Bedeutsamkeit bzw. Interaktion ergibt sich aus der Kernmatrix, die selbst eine Art kondensierte Form des zugrunde liegenden Datenkörpers darstellt (Snyder, Law & Hattie, 1984). Bloxom (1968) und Snyder (1969) formulierten zwei Varianten des Tuckersche Modells, welche neben gemeinsamen Faktoren, die jeweils mindestens zwei beträchtlich von null verschiedene Faktorladungen aufweisen, auch

spezifische Faktoren enthalten, die nur eine erheblich von null verschiedene Ladung aufweisen und eine einzige Variable repräsentieren. Der Ansatz von Bloxom wurde von Bentler und Lee (1978, 1979) weiterentwickelt und mündete in den statistischen "Covariance Structure"-Modellen.

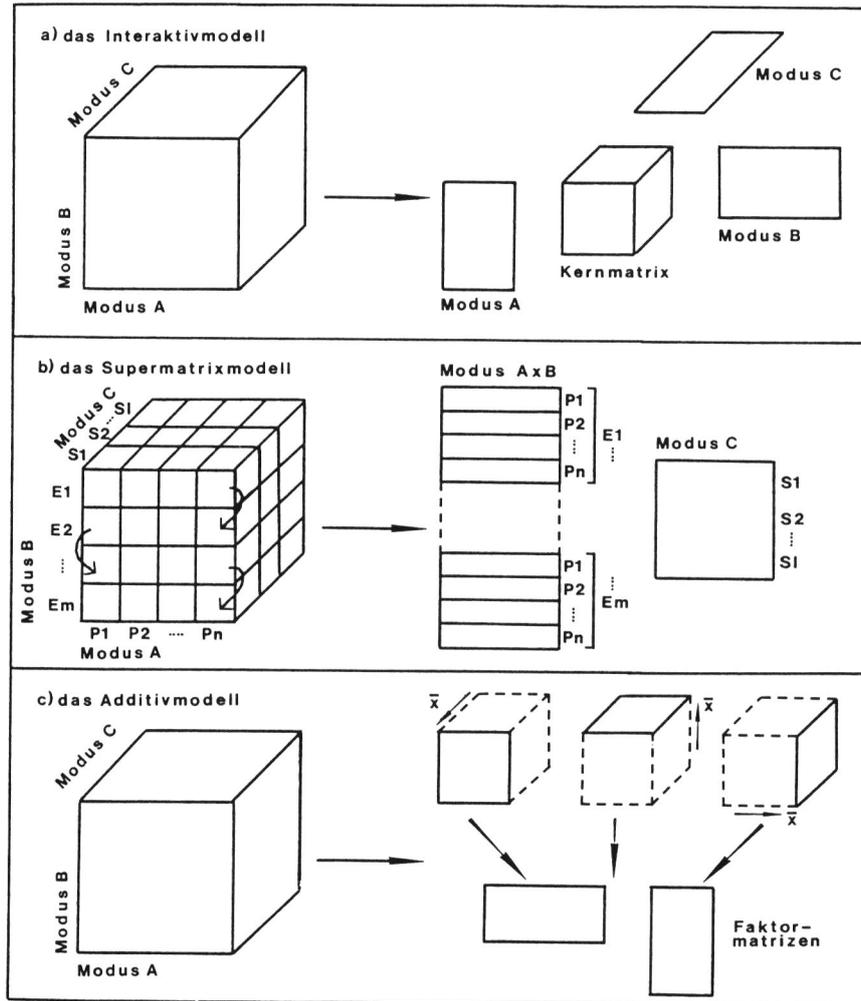
Carroll und Chang publizierten 1970 einen weiteren Markstein zur faktorenanalytischen Dekomposition dreimodaler Datensätze - das CANDECOMP-Modell. Während die Verknüpfung der Faktoren aus den verschiedenen Modi im Tuckerschen Modell in der Kernmatrix vorgenommen wird, enthält CANDECOMP ein andersgeartetes Rationale. Hier wird auf eine komplexe Verflechtung der Faktormatrizen verzichtet; stattdessen werden die Faktorladungen in einem der drei Modi direkt als Indikatoren der Bedeutsamkeit der Faktortripel interpretiert. Dabei war das Modell von Carroll und Chang weniger durch Fragestellungen im Rahmen der Faktorenanalyse motiviert, sondern als Reaktion auf neuere Ansätze der multidimensionalen Skalierung, die eine Analyse individueller Differenzen in der Beurteilung der Ähnlichkeit einer gegebenen Objektmenge enthielten. Carroll und Chang beziehen sich insbesondere auf ein von Tucker und Messick (1963) entwickeltes Modell, das "Points-of-View"-Modell. Die Datenbasis besteht hier aus den Ähnlichkeitsurteilen mehrerer Versuchspersonen über alle Paare aus einer Reihe von Objekten. Um solche Gruppen von Personen zu bestimmen, die in relativ übereinstimmender Weise urteilten, ermittelten Tucker und Messick zunächst Cluster von Versuchspersonen, wobei ein Repräsentant aus jedem Cluster einem "point of view" einer idealisierten Person entspricht. Für jede idealisierte Person wird dann ein separater Objektraum konstruiert. Dieser Ansatz führt aber kaum zu Aussagen derart, worin sich die Urteilsysteme einzelner realer Personen unterscheiden bzw. welche Gemeinsamkeiten sie haben (Carroll & Chang, 1970, S. 284). Daher konzipieren Carroll und Chang ihr INDSCAL-Modell spezifisch zur multidimensionalen Skalierung individueller Ähnlichkeitsurteile, welches explizit eine Analyse der Unterschiede zwischen den Urteilssystemen der Versuchspersonen

gestattet. INDSCAL dient damit der Analyse von Daten, die zwar in drei Richtungen variieren, aber nur zwei Modi umfassen - Versuchspersonen und beurteilte Objekte. Die Koppelung zu der dreimodalen Faktorenanalyse stellt dabei das bereits genannte CANDECOMP-Modell dar, welches aus der Verallgemeinerung von INDSCAL entstand. INDSCAL wurde bereits 1969 von Horan vorweggenommen, aber Carroll und Chang's Novum bestand in der zusätzlichen Entwicklung eines Algorithmus und in der Generalisierung auf faktorenanalytische Fragestellungen.

Der Zeitabschnitt zwischen 1966 und 1971 stellt eine Art Brennpunkt des Interesses für dreimodale Datenanalysen dar, denn neben CANDECOMP erscheinen weitere Modellvorschläge zu diesem Themenbereich. Drei wichtige Autoren sind hier R.A. Harshman, K. Jöreskog und wiederum R.B. Cattell. Harshman stellt 1970 sein PARAFAC-Modell vor, abgeleitet von Cattell's "principle of proportional profiles" (vgl. Harshman, 1976; Harshman & Lundy, 1984). PARAFAC wurde unabhängig von dem dreimodalen CANDECOMP-Modell und unter einem anderen Rationale entwickelt, aber die beiden Ansätze sind algebraisch völlig äquivalent (Snyder, Law & Hattie, 1984; Harshman & Berenbaum, 1981). So spricht u.a. Kruskal (1984) von dem CANDECOMP/PARAFAC-Modell.

Cattell unterscheidet 1966 zwischen drei verschiedenen Modellklassen zur Analyse dreimodaler Datensätze, die er als Additivmodelle, Interaktivmodelle und Supermatrixmodelle bezeichnet (vgl. Abbildung 3). In den Supermatrixmodellen wird die Aufteilung von zwei Quellen der Variation - etwa der Untersuchungsbedingungen und der Variablen oder Tests - aufgegeben, und die zugehörigen Elemente der zwei Modi werden miteinander kombiniert, so daß aus dem dreimodalen Datenquader eine Anordnung entlang zweier Dimensionen entsteht - Versuchspersonen x Bedingungs- und Testkombinationen. Zu der Klasse der Interaktivmodelle gehören die Modelle von Tucker und Carroll und Chang, und in den Additivmodellen wird die klassische zweimodale Faktorenanalyse auf die Ausschnitte aus dem Datenquader angewendet, die jeweils zwei Modi umfassen, während der Mittelwert über den dritten Modus

Abb.3: Die Modellklassen von Cattell (1966) zur Analyse dreimodaler Datenquader



gebildet wird (Snyder, Law & Hattie, 1984; Lohmöller, 1979). Unter Bezugnahme auf die Covariation Chart entwickelt Cattell in der Folgezeit eigene Interaktiv- und Additivmodelle (Cattell, 1966, 1980a, 1980b; Cattell, Blaine & Brennan, 1984), wobei sein Interaktivmodell äquivalent zu CANDECOMP ist (Snyder, Law & Hattie, 1984). Schließlich betrachtet Jöreskog die Klasse der Additivmodelle und entwickelt Algorithmen für Modellgleichungen, deren Parameter unterschiedlich strengen Klassen der Restriktion unterliegen (Jöreskog, 1971; Lohmöller, 1979).

Wie bereits angedeutet wurde, können die Modelle, die seit der Tuckerschen Konzeption bis in die 70er Jahre entwickelt wurden, dem Gebiet der multilinearen Algebra zugeordnet werden, aber es existieren kaum Querbezüge zu der Mathematik (Kroonenberg, 1983a) bis zu den 1976 und 1977 erschienenen Abhandlungen von Kruskal. Auf dem Hintergrund des CANDECOMP- und des PARAFAC-Modells setzt sich Kruskal theoretisch mit den Eigenschaften der dreimodalen Datenmodelle und den zu analysierenden Datenquadranten auseinander und zeigt auf, daß sich die neuen Ansätze in wesentlichen Aspekten von den traditionellen ein- oder zweimodalen Modellen unterscheiden. So entfällt bei CANDECOMP/PARAFAC beispielsweise das Problem der angemessenen Faktorenrotation, das im Rahmen der Diskussion um die klassischen Modelle einen großen Raum eingenommen hatte (vgl. Horst, 1965; Harman, 1960), und es werden mehr Faktoren zur Reproduktion des Datenkörpers benötigt, als an Personen, Tests und Versuchsbedingungen in eine Untersuchung einbezogen waren (Kruskal, 1976, 1977, 1981, 1983). In Zusammenhang mit dieser hohen Anzahl an Faktoren und den neuartigen Computeralgorithmen von Carroll und Chang stand die Notwendigkeit, für den Anwender andere Strategien zur Auswahl einer angemessenen Lösung bereitzustellen. Unter dieser Perspektive hat insbesondere Harshman (1984) neue Kriterien vorgestellt.

In der Zwischenzeit sind vor allem CANDECOMP, INDSCAL und das Tuckersche Modell in einer Reihe von Untersuchungen zur

Datenanalyse eingesetzt worden und hinsichtlich der Qualität und Interpretierbarkeit der Ergebnisse und ihrer Anwendbarkeit von verschiedenen Seiten kritisiert worden. So bemerkt Bartussek (1973), daß die für das Tuckersche Modell zentrale Kernmatrix Schwierigkeiten aufwirft, wenn sie inhaltlich interpretiert wird. Ferner erfüllen die Lösungen nach dem Tuckerschen Modell auch nicht das Kriterium der kleinsten Abweichungsquadrate (Kroonenberg, 1983b). Zudem sind die Modelle für eine Reihe von Daten, wie sie in den Sozialwissenschaften betrachtet werden, nicht anwendbar, denn es wird vorausgesetzt, daß die Daten mindestens Intervallskalenniveau besitzen (Sands & Young, 1980). Und schließlich ist Carroll und Chang's Rückschluß von der festen Achsenorientierung auf die direkte Bedeutsamkeit der Faktoren in Zweifel gezogen worden (vgl. Ahrens, 1974; Lohmöller, 1979; Orlik, 1980). Neben diesen Kritiken, die sich jeweils auf spezifische Aspekte eines Modells beziehen, sei die ironische Kommentierung von Torgerson (1986, S. 60) zu der gesamten Modellklasse angefügt: " I seriously wonder whether the data we ordinarily obtain in psychological investigations are worthy of the elegance and power of these marvelous constructions. It sometimes seems like exploring the jungle with a Mercedes."

Auf dem Hintergrund dieser Kritikpunkte wurden 1980 drei Modelle vorgestellt, die zur Vermeidung der genannten Nachteile konzipiert wurden. Sands und Young (1980) greifen das Problem des geforderten Skalenniveaus auf und schlagen einen neuen Algorithmus, den ALSCOMP-Algorithmus, vor. ALSCOMP ist nicht auf intervall- oder verhältnisskalierte Daten eingeschränkt, sondern kann auch auf nominal- oder ordinalskaliertes Datenmaterial angewendet werden.

Kroonenberg und De Leeuw (1980) entwickeln ebenso einen Algorithmus, der jedoch metrische Daten voraussetzt und spezifisch der Parameterschätzung des Tuckerschen Modells dient. Dieser Ansatz enthält Vorteile gegenüber den von Tucker vorgeschlagenen Verfahren, da er auch dann zu einer optimalen Approximation des Datenquaders im Sinne der Methode der kleinsten Quadrate führt, wenn eine Konfiguration aus

möglichst wenigen Faktoren von Interesse ist. In einer Reihe von Anwendungen zeigt Kroonenberg (1983b) die Güte der Lösungen hinsichtlich der Interpretierbarkeit und des faktoriell erklärten Varianzanteils auf.

Während in diesen Arbeiten im Grunde neue Verfahren zur Datenauswertung unter Rückgriff auf bekannte Modelle vorgestellt wurden, diskutiert Orlik ein neues Modell zur faktorenanalytischen Dekomposition von dreimodalen Datenkörpern - das SUMMAX-Modell. SUMMAX entstand auf dem Hintergrund der Interpretationsprobleme, die mit der Tuckerschen Kernmatrix verbunden sind, und der Kritik an dem CANDECOMP-Modell. Das Grundprinzip der Faktorenextraktion ist eng verwandt mit dem Rationale des Centroidmodells von Thurstone (vgl. Orlik, 1973, 1976, 1980) und unterscheidet sich daher von der Tuckerschen Konzeption und von CANDECOMP, die beide eine Generalisierung der klassischen Hauptachsentransformation enthalten. Ferner umfaßt SUMMAX zwei Zugänge zur Analyse dreimodaler Daten:

wird von der Möglichkeit einer Faktorenrotation abgesehen, so werden die Faktoren der drei Modi direkt miteinander verknüpft, und das Modell weist Bezüge in den formalen Grundgleichungen und Interpretationen zu CANDECOMP auf. Zur Ermöglichung anschließender Faktorrotationen wird das Modell erweitert um eine Kernmatrix, wodurch SUMMAX zwar eine Reihe von formalen Zusammenhängen zu dem Tuckerschen Ansatz enthält, sich aber hinsichtlich des Aufbaus und der Interpretation der Kernmatrix wesentlich von Tuckers Modell unterscheidet.

Unter Bezugnahme auf das CANDECOMP-Modell greifen Carroll, Pruzansky und Kruskal (1980) Ansätze zur Interpretation der Faktoren auf und entwickeln einen Zugang, der zusätzlich erhobene Außenkriterien integriert. In diesem Modell, genannt CANDELINC, können beispielsweise Informationen zum Schulabschluß, zum ausgeübten Beruf oder ähnliches als Außenkriterien herangezogen werden, um die inhaltliche Bedeutung der Konfiguration für Personen zu klären, die an einer Untersuchung teilgenommen haben. Hierbei werden etwa

die Parameter für die Faktoren des Personenmodus so eingeschränkt, daß die resultierende Konfiguration in möglichst hohem Maße das erhobene Außenkriterium wiedergibt. Restriktionen der Modellparameter können im Prinzip in allen Modi vorgenommen werden.

CANDELINC wie auch bereits das CANDECOMP-Modell sind im Gegensatz zu dem Tuckerschen Modell nicht spezifisch zur Analyse dreimodaler Datensätze konzipiert worden, sondern umfassen "n-modale" Verallgemeinerungen. Ebenso gestatten die Grundgleichungen des SUMMAX-Modells eine direkte Erweiterung zur Analyse n-modaler Datensätze. D.h. es können Datenkörper betrachtet werden, die aus der Kombination von mehr als drei Modi aufgebaut sind, wie etwa in einer Untersuchung zur konnotativen Bedeutung einer Reihe von Begriffen, die von einer Anzahl von Versuchspersonen auf mehreren Ratingskalen einzustufen sind, wobei die Untersuchung zu verschiedenen Zeitpunkten wiederholt wird. Aus der Kombination Begriffe x Personen x Skalen x Zeitpunkte entsteht ein Datenkörper, der in vier Richtungen variiert. Eine entsprechende Generalisierung des Tuckerschen Modells wurde 1981 von Lastovicka für den viermodalen Fall und 1986 von Kapteyn, Neudecker und Wansbeek für den n-modalen Fall diskutiert. Allerdings sind mit Ausnahme von Lastovicka, Carroll und Chang kaum entsprechende Algorithmen entwickelt und implementiert worden. Offenbar sind die n-modalen Ansätze auch nicht von hoher praktischer Relevanz, da einerseits die Interpretation der Ergebnisse sehr komplex ist und andererseits kaum Untersuchungen vorliegen, die mehr als drei unterschiedliche Varianzquellen umfassen (Kroonenberg, 1983a). Während momentan die Entwicklung n-modaler Analysetechniken eher von theoretischen als von anwendungsorientierten Kriterien geleitet erscheint, akzentuiert die faktoranalytische Diskussion seit 1946 den Zugang zu dreimodalen Datenstrukturen in der Theorie und Anwendung. Dies wird auch deutlich an der Vielzahl der empirischen Datenanalysen mittels der dreimodalen Modelle. Im folgenden seien exemplarisch nur einige Studien erwähnt, um die Relevanz der drei-

modalen Methoden in verschiedenen Bereichen der psychologischen Forschung aufzuzeigen.

Im Rahmen der Organisationspsychologie erstellte Frederikson (1972) eine Taxonomie von verschiedenen Situationen, mit denen Manager in ihrem beruflichen Umfeld konfrontiert werden, und untersuchte die Beziehung der Situationen zu verschiedenen Leistungsaspekten. Cornelius, Hakel und Sackett (1979) diskutierten Kriterien der Personalauswahl aufgrund einer Klassifikation von Berufen und den mit unterschiedlichen Berufen verbundenen Graden der Verantwortung. Gegenstand der Studie von Zenisek (1980) ist die Messung der Arbeitszufriedenheit und deren Bezug zu Fehlzeiten am Arbeitsplatz bei Arbeitern, Büroangestellten und Managern.

In sozialpsychologischen Studien wurde u.a. der Einfluß der ethnischen Abstammung von Beobachtern auf die Wahrnehmung von Personen unterschiedlicher ethnischer Herkunft untersucht (Imada & London, 1979), wobei die individuellen Wahrnehmungen durch eine Reihe von Ratingskalen erhoben wurden. Kjerulff und Wiggins (1976) befaßten sich mit Charakteristika belastender Studiensituationen und möglichen Bewältigungsstrategien bei Hochschulstudenten.

Eine der ersten Anwendungen von Tuckers Modell stammt aus dem Gebiet der Differentiellen Psychologie mit Levins Untersuchung (1963, 1965) zur Beziehung der Versuchssituation und der Beantwortung des "S-R Inventory of Anxiousness" von Ender, Hunt und Rosenstein, die verschiedene "Typen" von Probanden hinsichtlich der Ängstlichkeit in den betrachteten Situationen aufzeigte.

Lilly (1965) ging mit seiner Studie zur Struktur und Veränderung des semantischen Raums bei Schulkindern verschiedener Alterstufen einer entwicklungspsychologischen Fragestellung nach, und Vavra (1974) verwendete eine dreimodale Faktorenanalyse zur Analyse der Effektivität verschiedener Marketing-Strategien in der Produktwerbung bei unterscheidbaren Gruppen von Konsumenten.

Und schließlich wurden im Rahmen der Psychophysik Aspekte der Sprachwahrnehmung (Soli & Arabie, 1979; Kruskal, 1984) sowie der Farbwahrnehmung (Carroll & Wish, 1974; Orlik, 1980, 1981) mittels dreimodaler Designs untersucht. Eine Übersicht über weitere Arbeiten seit 1964 gibt die umfangreiche Bibliographie von Kroonenberg (1983a, 1983b). In der Mehrzahl der seit 1964 publizierten Studien wurde das Tuckersche Modell zur Datenanalyse verwendet sowie das INDSCAL-Modell, wenn eine strukturelle Repräsentation von Ähnlichkeitsurteilen im Vordergrund stand. In einer Reihe von Untersuchungen kamen darüberhinaus CANDECOMP, PARAFAC und SUMMAX zur Anwendung, während für die übrigen Modelle, die in dem vorliegenden Abschnitt angesprochen wurden, kaum Erfahrungen in konkreten Anwendungen vorliegen. Dabei stehen allerdings die Modelle, die vermehrt Beachtung gefunden haben, größtenteils unverbunden nebeneinander sowohl hinsichtlich ihrer theoretisch-algebraischen Verknüpfungen als auch hinsichtlich inhaltlich-interpretativer Bezüge. Die folgenden Abschnitte sind der Untersuchung dieser beiden, bisher offenen Fragestellungen gewidmet.



### 3.

Modelle der dreimodalen Faktorenanalyse: Grundgleichungen, Algorithmen und Interpretationen

Gegenstand des vorliegenden Abschnitts ist eine Darstellung der Grundgleichungen der Modelle, die in den folgenden Kapiteln hinsichtlich ihrer algebraischen und inhaltlich-interpretativen Bezüge analysiert werden: das Tuckersche Modell, CANDECOMP, PARAFAC und SUMMAX. Neben den modelltheoretischen Schwerpunkten werden die zugeordneten Algorithmen zur Faktorextraktion betrachtet, und es werden die modellspezifischen Interpretationsmöglichkeiten der Faktor- und Kernmatrizen diskutiert. Jedes der angeführten Modelle besitzt eine verallgemeinerte Darstellung zur Analyse n-modaler Datenkörper und eine spezifische Form, die dem Gebiet der multidimensionalen Skalierung zugeordnet werden kann, und der Analyse individueller Differenzen in der Beurteilung von Reizähnlichkeiten dient. Zur vollständigen Beschreibung der Modelle werden daher jeweils die korrespondierenden Generalisierungen bzw. Sonderfälle aufgeführt.

Vor der Diskussion der Modelle sind eine Reihe von terminologischen Spezifika, die eng mit der Komplexität der zugrunde liegenden Datenkörper verknüpft sind, auszuführen. Diese beziehen sich vor allem auf die Verwendung der Begriffe "Modus" und "Kombinationsmodus". Tucker benutzt den Begriff eines Modus unter verschiedenen Bedeutungen. Zunächst umschreibt der Term "eine Menge von Indizes, durch welche Daten klassifiziert werden können" (Tucker, 1966, S. 280), und zur Identifikation der drei Klassifikationsmöglichkeiten werden die Symbole  $i$ ,  $j$  und  $k$  gewählt. Diesem Gebrauch entspricht z.B. die Aussage "Modus  $i$  kennzeichnet die untersuchten Personen". Darüberhinaus werden die Symbole  $i$ ,  $j$  und  $k$  als Laufindizes verwendet, indem etwa die  $i$ -te Person des Modus  $i$  mit  $P_i$  gekennzeichnet wird. Ferner kommen die Elemente der drei Indextmengen (etwa Personen, Skalen, Zeitpunkte), von Lohmöller als "Konzepte" bezeichnet, in zwei "Typen" vor: einmal im "observational type" als empirische Konzepte mit der Indizierung  $i$ ,  $j$ ,  $k$  und zweitens im "derivational mode" als idealisierte Konzepte mit der Indizierung  $m$ ,  $p$ ,  $q$  (Lohmöller, 1979, S. 138). Durch die Darstellung des Zusammenhangs zwischen empirischen und idealisierten Konzep-

ten in Form von Ladungskoeffizienten, z.B.  $a_{im}$ , wird die dritte Verwendung des Terminus "Modus" deutlich. Er dient hier der Kennzeichnung von Zellen einer Matrix.

Dreimodale Datenkörper werden somit als ein dreifachindizierter Quader mit  $i$  Zeilen,  $j$  Spalten und  $k$  Lagen dargestellt. Zur algebraischen Behandlung werden diese Quader in Matrizen oder Zweiwegtafeln angeordnet unter Verwendung von Kombinationsmodi. Der Kombinationsmodus ist das kartesische Produkt von zwei Modi und wird gekennzeichnet durch kombinierte Indizes wie beispielsweise  $ij$ . Es existieren insgesamt drei Möglichkeiten, die Daten in Matrizenform darzustellen, wobei jeweils eine Supermatrix entsteht. Die Elemente einer Supermatrix werden gekennzeichnet durch einen einfachen Index und einen kombinierten Index. Die Supermatrix  ${}_iX_{jk}$  besteht z.B. aus  $j$  nebeneinandergeordneten Teilmatrizen  ${}_iX_k$ , die Supermatrix  ${}_kX_{ij}$  aus  $i$  nebeneinandergeordneten Teilmatrizen  ${}_kX_j$  usw. (Bartussek, 1973, S. 175). Zur Identifikation einer spezifischen Anordnung werden zwei Subskripte verwendet; das Subskript vor der Matrix bezeichnet dabei die Zeilen der Matrix, das Subskript nach der Matrix die Spalten. Weitere Symbole, die in der folgenden Darstellung benutzt werden, und ihre Bedeutung sind in Tabelle 1 aufgeführt. Ferner wird die Tuckersche Indizierung von Matrizen nur dann eingesetzt, wenn dies aus Gründen der Eindeutigkeit erforderlich ist. In Tabelle 2 ist die im folgenden verwendete Darstellung der Modelle aufgeführt.  $A$  und  ${}_iA_f$  bezeichnen die gleiche Matrix. Die Matrizen  $A$  und  $\bar{A}$  bezeichnen jeweils die Faktormatrizen in Modus  $i$ , analog  $B$ ,  $\bar{B}$  die Faktormatrizen in Modus  $j$  und  $C$ ,  $\bar{C}$  die Faktormatrizen in Modus  $k$ . Zur Unterscheidung der Faktormatrizen aus den verschiedenen Modellen wurde das Zusatzkennzeichen "-" verwendet, so daß die Faktormatrizen des Tuckerschen und des CANDECOMP-Modells mit  $\bar{A}$ ,  $\bar{B}$  und  $\bar{C}$ , die des SUMMAX-Modells mit  $A$ ,  $B$  und  $C$  bezeichnet werden.

Tabelle 1: Übersicht über die verwendete Notation

$ij$	kennzeichnet den Kombinationsmodus ( $ij$ ), d.h. das kartesische Produkt von zwei Modi $i$ und $j$ ; "i: äußere Schleife, j: innere Schleife" (vgl. Tucker, 1966)
$[A, B, C]$	Tripel-Produkt von drei Matrizen $A$ , $B$ und $C$
$f I_{ff}$	dreimodale Identitätsmatrix
$D = \text{diag}(Z)$	$\begin{cases} d_{ij} = z_{ij} & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$
$X = (x_{ijk})$	dreimodale Datenmatrix mit $i$ Zeilen, $j$ Spalten und $k$ Lagen
$i X_{jk}$	Anordnung der dreimodalen Matrix in zweimodaler Form mit $i$ Zeilen, $j \times k$ Spalten; $jk$ : Kombinationsmodus
$A, B, C, \bar{A}, \bar{B}, \bar{C}$	Faktormatrizen
$F, V, E$	Eigenvektormatrizen
$\lambda_i, \mu_i, \psi_i, \delta_i$	Eigenwerte
$\Psi, \Delta$	Diagonalmatrizen, deren Diagonalelemente aus Eigenwerten bestehen - nach absteigender Größe geordnet
$m^G_{pq}, \bar{m}^G_{pq}$	Kernmatrizen
$P_A, P_B, P_C$	Permutationsmatrizen
$i A_f$	Darstellung der Matrix $A$ in der Terminologie von Tucker; Subskript $i$ kennzeichnet die Zeilen, Subskript $f$ die Spalten der Matrix $A$
$A', A^T$	Transponierte einer Matrix $A$
$\dim \text{col}(A)$	Dimension des Vektorraums, der durch die Spalten von $A$ aufgespannt wird
$A \times B = H$	Kronecker-Produkt aus zwei Matrizen $A$ , $B$ der Ordnung $(i, m)$ und $(j, p)$ . Es entsteht eine Übermatrix der Ordnung $(ij, mp)$ , deren Zellen gleich $h_{ij, mp} = a_{im} b_{jp}$ sind.

Tabelle 2: Übersicht über die Darstellung der Modelle

## 2a) CANDECOMP/PARAFAC:

$$x_{ijk} = \sum_r \bar{a}_{ir} \bar{b}_{jr} \bar{c}_{kr}$$

$$i^X_{jk} = i^{\bar{A}}_r r^I_{rr} (j^{\bar{B}}_r' \times k^{\bar{C}}_r') = [\bar{A}, \bar{B}, \bar{C}]$$

## 2b) Tuckersches Modell:

$$x_{ijk} = \sum_m \sum_p \sum_q \bar{a}_{im} \bar{b}_{jp} \bar{c}_{kq} \bar{g}_{mpq}$$

$$i^X_{jk} = i^{\bar{A}}_m m^{\bar{G}}_{pq} (j^{\bar{B}}_p' \times k^{\bar{C}}_q') = \bar{A} \bar{G} (\bar{B}' \times \bar{C}')$$

mit den Hauptachsentransformationen:

$$i^X_{jk} i^X_{jk}' = \bar{A} \Psi_A \bar{A}'$$

$$j^X_{ik} j^X_{ik}' = \bar{B} \Psi_B \bar{B}'$$

$$k^X_{ij} k^X_{ij}' = \bar{C} \Psi_C \bar{C}'$$

## 2c) SUMMAX:

$$x_{ijk} = \sum_f a_{if} b_{jf} c_{kf}$$

$$i^X_{jk} = i^A_f f^I_{ff} (j^B_f' \times k^C_f') = [A, B, C]$$

## Fortsetzung der Tabelle 2

nach Rotation:

$$x_{ijk} = \sum_m \sum_p \sum_q a_{im} b_{jp} c_{kq} g_{mpq}$$

$$i^X_{jk} = i^A_m m G_{pq} (j^B_p{}' \times k^C_q{}') = A G (B' \times C')$$

mit den Hauptachsentransformationen:

$$i^A_m = i^A_f T = F \Delta_A^{1/2}$$

$$j^B_p = j^B_f W = V \Delta_B^{1/2}$$

$$k^C_q = k^C_f H = E \Delta_C^{1/2}$$

Anmerkung: die Matrizen A und  $\bar{A}$  bezeichnen jeweils die Faktormatrizen in Modus i, analog B,  $\bar{B}$  die Faktormatrizen in Modus j und C,  $\bar{C}$  die Faktormatrizen in Modus k. Zur Unterscheidung der Faktormatrizen aus den verschiedenen Modellen wurde das Zusatzkennzeichen "-" verwendet.

### 3.1. Das Tuckersche Modell

#### 3.1.1. Die Grundgleichungen des Modells:

Die Grundgleichungen der dreimodalen Faktorenanalyse nach Tucker bestehen aus der Approximation der gemessenen oder beobachteten Werte  $x_{ijk}$  durch eine Kernmatrix, die mit den Faktorladungen der drei Modi gewichtet wird (Tucker, 1964a, 1964b, 1966):

$$(1) \quad x_{ijk} = \hat{x}_{ijk} + e_{ijk}$$

$$(2) \quad \hat{x}_{ijk} = \sum_m \sum_p \sum_q \bar{a}_{im} \bar{b}_{jp} \bar{c}_{kq} \bar{g}_{mpq} \cdot$$

In Matrixschreibweise entspricht Gleichung (2) der Darstellung

$$(3) \quad {}_i \hat{X}_{jk} = {}_i \bar{A}_m {}_m \bar{G}_{pq} ({}_j \bar{B}'_p \times {}_k \bar{C}'_q) \\ = \bar{A} \bar{G} (\bar{B}' \times \bar{C}')$$

Dabei enthalten die Matrizen  $\bar{A}$ ,  $\bar{B}$  und  $\bar{C}$  die Faktorladungen oder Gewichtszahlen aus Gleichung (2) und  $\bar{G}$  ist die Kernmatrix, die die drei Faktormatrizen miteinander verknüpft.  $\hat{X}$  kennzeichnet den Anteil an den beobachteten Werten, der durch das Modell approximiert wird. Durch Gleichung (1) wird die Möglichkeit eröffnet, solche Faktoren in der Analyse zu vernachlässigen, die nur einen geringen Beitrag zur Aufklärung der gesamten Varianz der Daten leisten. Wird hingegen der volle Rang für jede Faktormatrix betrachtet, so hat das Modell die Form

$$(4) \quad {}_i X_{jk} = \bar{A} \bar{G} (\bar{B}' \times \bar{C}'),$$

d.h. der Datenkörper  $X$  wird vollständig durch die modellspezifischen Parametermatrizen erklärt.

Analog der klassischen zweimodalen Hauptkomponentenanalyse werden zur Bestimmung der Faktormatrizen zunächst die Skalarproduktmatrizen für jeden Modus aus der Datenmatrix ge-

bildet und anschließend auf Hauptachsen transformiert:

$$(5a) \quad \sum_i X_{jk} \sum_i X_{jk}' = \bar{A} \Psi_A \bar{A}'$$

$$(5b) \quad \sum_j X_{ik} \sum_j X_{ik}' = \bar{B} \Psi_B \bar{B}'$$

$$(5c) \quad \sum_k X_{ij} \sum_k X_{ij}' = \bar{C} \Psi_C \bar{C}' \quad .$$

Die Faktormatrizen  $\bar{A}$ ,  $\bar{B}$  und  $\bar{C}$  enthalten jeweils die ortho-normalen Eigenvektormatrizen aus den drei möglichen Skalarproduktmatrizen, und  $\Psi$  ist die zugehörige Diagonalmatrix mit den Eigenwerten. Wegen der (spaltenweisen) Orthonormalität der Faktormatrizen gilt:

$$(6) \quad \bar{A}'\bar{A} = {}_m I_m; \quad \bar{B}'\bar{B} = {}_p I_p; \quad \bar{C}'\bar{C} = {}_q I_q \quad ,$$

wobei  $I$  die Identitätsmatrix bedeutet. Nach Extraktion der Faktoren in jedem Modus entsprechend Gleichungen (5), die hier die Eigenvektoren sind, wird die Kernmatrix bestimmt, indem die Datenmatrix mit den Faktormatrizen prä- bzw. postmultipliziert wird:

$$(7) \quad \bar{G} = \bar{A}' \sum_i X_{jk} (\bar{B} \times \bar{C}) \quad .$$

Wird der volle Rang für jede Faktormatrix betrachtet, so ist die Transponierte der Faktormatrix gleich ihrer Inversen, und Gleichung (7) entspricht der Multiplikation der Datenmatrix mit den inversen Faktormatrizen. Dann gilt für die Kernmatrix folgende Aussage wie von Tucker (1966) gezeigt wurde:

$$(8) \quad \sum_i \sum_j \sum_k x_{ijk}^2 = \sum_m \sum_p \sum_q \bar{g}_{mpq}^2 \quad .$$

D.h. die Kernmatrix stellt eine kondensierte Form der Datenmatrix dar, in welcher die empirischen Konzepte durch idealisierte Konzepte mit der Indizierung  $m$ ,  $p$  und  $q$  ersetzt wurden. Aus Gleichung (6) und (8) wird zusätzlich deutlich, daß die Ränge der drei Skalarproduktmatrizen -  $m$ ,  $p$ ,  $q$  - ,

die aus der Datenmatrix gebildet werden können, nicht notwendig gleich sind und daher die Anzahl der Faktoren zwischen den Modi variiert. Eine weitere Eigenschaft des Tuckerschen Modells beinhaltet die Möglichkeit anschließender Rotationen der extrahierten Faktormatrizen. Dazu seien die Rotationstransformationen in folgender Form gegeben:

$$(9a) \quad {}_i\bar{A}_{m \ m}^T {}_m1 = {}_i\bar{A}_{m1}$$

$$(9b) \quad {}_j\bar{B}_{p \ p}^T {}_p1 = {}_j\bar{B}_{p1}$$

$$(9c) \quad {}_k\bar{C}_{q \ q}^T {}_q1 = {}_k\bar{C}_{q1} \ .$$

Sind die Transformationsmatrizen quadratisch und regulär, so kann das vollständige Modell dargestellt werden wie in Gleichungen (10) und (11):

$$(10) \quad {}_m1\bar{G}_{p1q1} = ({}_mT_{m1})^{-1} {}_m\bar{G}_{pq} ({}_pT_{p1}')^{-1} \times ({}_qT_{q1}')^{-1}$$

$$(11) \quad {}_iX_{jk} = {}_i\bar{A}_{m1} {}_m1\bar{G}_{p1q1} ({}_j\bar{B}_{p1}' \times {}_k\bar{C}_{q1}') \ .$$

Dabei zieht Tucker (1966, 1972) orthogonale wie auch oblique Rotationstransformationen nach Kriterien der Einfachstruktur modelltheoretisch in Betracht, aber er weist gleichzeitig auf bisher ungelöste Probleme der Faktorrotation hin, insbesondere für den Fall, daß die rotierten Matrizen keinen vollen Rang aufweisen: "The existence of the flexibility afforded by the freedom of transformation ... is the source of many problems yet to be solved. The practical problems of determination of these transformations are yet to be solved" (Tucker, 1966, S. 291). Kroonenberg (1983b, 1984) diskutiert das Problem der Rotationstransformation der Faktormatrizen unter Bezugnahme auf die Kernmatrix. Die Kernmatrix ist zwar das zentrale Konzept des Tuckerschen Modells, aber die abstrakten Kernrelationen erweisen sich als sehr komplex hinsichtlich ihrer inhaltlichen Interpretierbarkeit. Zur Vereinfachung der Struktur der Kernmatrix, etwa in Form einer Diagonalmatrix, ist es nach Kroonenberg eher angebracht,

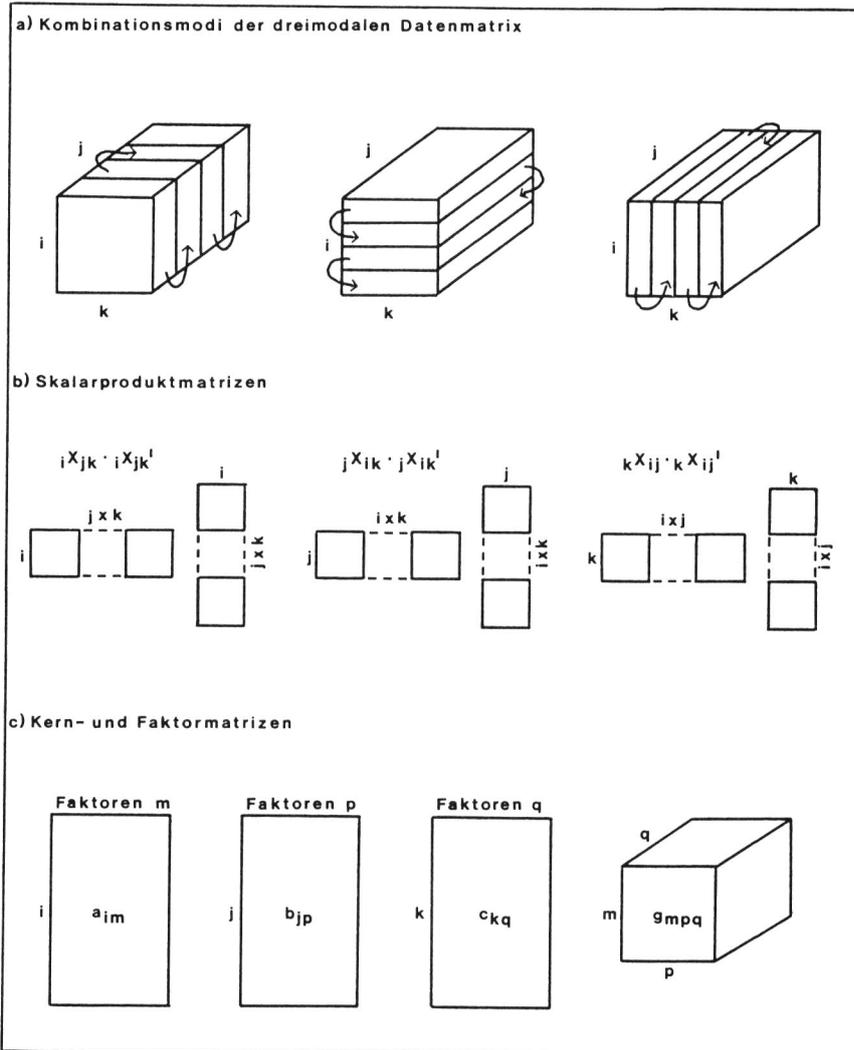
durch entsprechende Rotationen die Kernmatrix in eine Einfachstruktur zu überführen (vgl. Kroonenberg, 1983b, S. 107 - 124, zu korrespondierenden Techniken der Rotation der Kernmatrix). Zudem wurde in einer Reihe von Anwendungen bereits eine Einfachstruktur in den unrotierten Faktormatrizen deutlich, und die Struktur der Kernmatrix wurde weiter kompliziert, wenn Rotationen der Faktormatrizen vorgenommen wurden (Kroonenberg, 1984). Im Gegensatz zur faktorenanalytischen Tradition ist daher unter inhaltlich-interpretativen Gesichtspunkten eine Rotation der Faktormatrizen im Tucker-schen Modell nicht notwendig die Methode der Wahl.

### 3.1.2. Der Tuckersche Algorithmus:

Zur Berechnung der Kern- und Faktormatrizen wurden von Tucker (1966) drei verschiedene Algorithmen vorgeschlagen, wobei Methode II und III spezifisch der Analyse von Datenmatrizen dienen, die einen Personenmodus mit 300 oder mehr Probanden enthalten, und mit mehreren Kombinationsmodi arbeiten. Methode I ist dagegen in direkter Entsprechung zu den Grundgleichungen des Modells aufgebaut und umfaßt folgende Schritte, die schematisch in Abbildung 4 dargestellt sind:

- a) zunächst werden aus der Datenmatrix die Skalarproduktmatrizen für die drei Modi,  $i^{X_{jk}} i^{X_{jk}'}$ ,  $j^{X_{ik}} j^{X_{ik}'}$  und  $k^{X_{ij}} k^{X_{ij}'}$ , ermittelt
- b) für jede Skalarproduktmatrix werden die Eigenwerte und Eigenvektoren nach Gleichungen (5a) bis (5c) bestimmt
- c) zur Festlegung der angemessenen Dimensionalität der Faktormatrizen wird für jeden Modus getrennt die Höhe der Eigenwerte gegen die Anzahl der Eigenwerte aufgetragen und mittels des Scree-Tests (vgl. hierzu Revenstorf, 1980, S. 76f) beurteilt. In jedem Fall sollten Eigenvektoren, die sehr kleinen Eigenwerten (d.h. Eigenwerte nahe null) zugeordnet sind, in der weiteren Analyse nicht beibehalten werden (Tucker, 1966, S. 297)
- d) die Eigenvektoren, deren Eigenwerte nach dem Scree-Test substantiell und nach abfallender Größe geordnet sind, bilden die Spalten der Faktormatrizen  $\bar{A}$ ,  $\bar{B}$  bzw.  $\bar{C}$

Abb. 4: Die dreimodale Hauptkomponentenanalyse nach Tucker (1966)



e) die Kernmatrix wird nach Gleichung (7) gebildet, indem die Datenmatrix mit den Faktormatrizen prä- bzw. postmultipliziert wird.

Dieser Algorithmus, wie auch Methode II und III (vgl. auch Tucker, 1967), führt im allgemeinen nicht zu einer Approximation der Datenmatrix nach dem Least-Squares-Kriterium. Im Gegensatz zu Gleichung (8) für das vollständige Modell kann im Fall von dimensional reduzierten Faktormatrizen nicht bestimmt werden, wie hoch der Prozentsatz der Varianzaufklärung einer Konfiguration oder spezifischer Faktorkombinationen ist. Ferner ist auch nicht geklärt, welchen Einfluß die Elimination von Eigenvektoren auf die Parameter der Kernmatrix hat. Insbesondere Methode II und III, auf die an dieser Stelle nicht weiter eingegangen wird, führen zu Schätzwerten für die Eigenwerte und -vektoren, die weit von denen der Skalarproduktmatrizen abweichen können, wenn nur einer der Modi eine reduzierte von Faktoren umfaßt (Tucker, 1966; Kroonenberg, 1983b). Kroonenberg und De Leeuw (1980) haben einen Algorithmus zum Tuckerschen Modell entwickelt, der diese Nachteile weitgehend vermeidet.

### 3.1.3. Der Algorithmus von Kroonenberg und De Leeuw:

Zur Entwicklung einer Methode der Parameterschätzung, die auch dann zu einer Least-Squares-Lösung führt, wenn eine Konfiguration aus möglichst wenig Faktoren gewünscht ist, gehen Kroonenberg und De Leeuw (1980) von folgender Funktion aus:

$$(12) \quad g(\bar{A}, \bar{B}, \bar{C}, \bar{G}) = \text{Spur}({}_i X_{jk} - {}_i \hat{X}_{jk})({}_i X_{jk} - {}_i \hat{X}_{jk})'$$

$$\text{mit } {}_i \hat{X}_{jk} = \bar{A} \bar{G} (\bar{B}' \times \bar{C}'),$$

deren Minimum zu bestimmen ist. Unter Beachtung, daß die Minimierung der Funktion  $g$  nur von den Matrizen  $\bar{A}$ ,  $\bar{B}$  und  $\bar{C}$  abhängig ist, kann das Problem (12) nach einigen Umformungen auch als Maximierungsaufgabe formuliert werden (Kroonenberg & De Leeuw, 1980, S. 71f):

$$(13) \quad p(\bar{A}, \bar{B}, \bar{C}) = \text{Spur}(\bar{A}' \cdot X_{jk} (\bar{B} \bar{B}' \times \bar{C} \bar{C}') \cdot X_{jk}' \bar{A}) := \max .$$

Die Parameter aus  $\bar{A}$ ,  $\bar{B}$  und  $\bar{C}$  werden unter der Verwendung der Iterationsmethode von Bauer-Rutishauser nach folgendem Schema bestimmt, welches in Abbildung 5 graphisch veranschaulicht ist:

zunächst werden zwei Startkonfigurationen für  $\bar{A}$  und  $\bar{C}$  festgelegt. Aus diesen Startkonfigurationen wird  $\bar{B}$  geschätzt. Im nächsten Schritt wird  $\bar{A}$  und  $\bar{B}$  festgelegt und neue Schätzwerte für  $\bar{C}$  errechnet. Aus  $\bar{B}$  und der neu gebildeten Matrix  $\bar{C}$  wird  $\bar{A}$  bestimmt usw. Die Iterationsfolge wird schließlich abgebrochen, wenn zwei aufeinanderfolgende Iterationszyklen keine wesentliche Veränderung der Gesamtkonfiguration mehr ergeben, wobei ein Zyklus jeweils die Schätzung von  $\bar{A}$ ,  $\bar{B}$  und  $\bar{C}$  umfaßt.

Dieser Algorithmus wird allgemein als "Alternating-Least-Squares"-Ansatz bezeichnet und wird in seinem Grundprinzip auch zur Bestimmung der CANDECOMP-Konfigurationen von Carroll und Chang (1970) verwendet. Das wesentliche Element des Algorithmus, die Bauer-Rutishauser-Methode, geht dabei von einer symmetrischen und positiv definiten Datenmatrix A aus, deren p größte Eigenwerte und Eigenvektoren durch iteratives Durchlaufen der Schritte (ii) bis (v) bestimmt werden:

- (i) eine willkürliche, spaltenweise orthonormale Startkonfiguration  $X_0$  mit p Spalten wird gewählt. Dabei ist p die Anzahl der gewünschten Eigenvektoren, die zuvor festzusetzen ist.
- (ii) nach l Iterationen:  

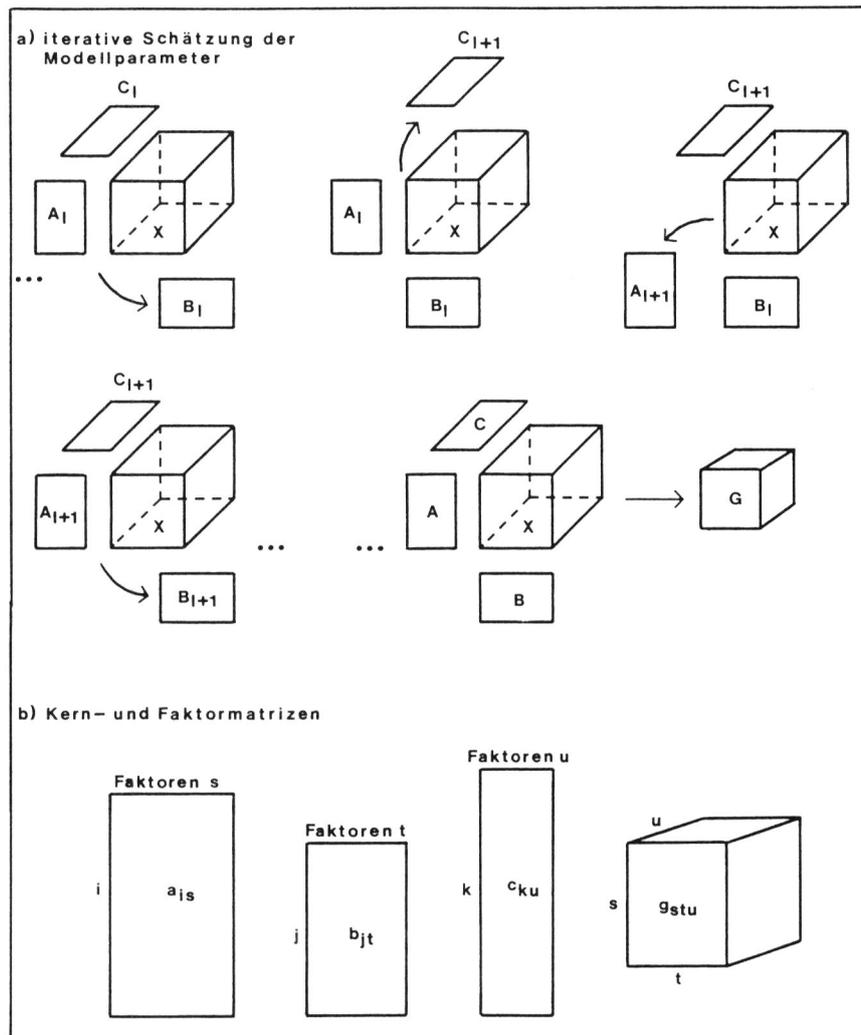
$$Y_1 = A X_1$$
- (iii)  $B_1 = Y_1' Y_1$
- (iv)  $T_1$  sei die Eigenvektormatrix von  $B_1$  und  $L_1$  die zugehörige Diagonalmatrix der Eigenwerte  $l_d$  mit  

$$l_1 \geq l_2 \geq \dots l_d \geq \dots l_p.$$
- (v)  $X_{1+1} = Y_1 T_1 L_1^{-1/2}$

Die Spalten der Matrix  $X_1$  und die Diagonalelemente von  $L_1$  konvergieren dann gegen die Eigenvektoren und Eigenwerte von

Abb.5: Schematische Darstellung der ALS-Methode von Kroonenberg &amp;

De Leeuw (1980)



A, wenn (Kroonenberg & De Leeuw, 1980, S. 75):

- a) die Anzahl der Iterationen  $l$  gegen unendlich geht
- b) die Spalten von  $X$  nicht orthogonal zu einem oder mehreren Eigenvektoren sind
- c) die Eigenwerte verschieden sind.

Zur Übertragung dieses Prinzips auf den dreimodalen Fall formen Kroonenberg und De Leeuw Gleichung (v) nach Postmultiplikation mit  $T_1$  um in

$$(14) \quad X_{1+1} = Y_1 T_1 L_1^{-1/2} T_1' = A X_1 (X_1' A^2 X_1)^{-1/2},$$

woraus sich folgender Algorithmus zur Bestimmung der Eigenvektoren der dreimodalen Datenmatrix  ${}_i X_{jk}$  ergibt:

- (i) die Anzahl der gewünschten Eigenvektoren  $m$ ,  $p$  und  $q$  für die Matrizen  $\bar{A}$ ,  $\bar{B}$  und  $\bar{C}$  sind festzusetzen, und Startkonfigurationen für  $\bar{B}$  und  $\bar{C}$  sind auszuwählen
- (ii) nach  $l$  Iterationen:
 
$$P_1 = {}_i X_{jk} (\bar{B}_1 \bar{B}_1' \times \bar{C}_1 \bar{C}_1') {}_i X_{jk}'$$

$$\bar{A}_{1+1} = P_1 \bar{A}_1 (\bar{A}_1' P_1^2 \bar{A}_1)^{-1/2} \quad (\text{"A-Teilschritt"})$$
- (iii)  $Q_1 = {}_j X_{ki} (\bar{C}_1 \bar{C}_1' \times \bar{A}_{1+1} \bar{A}_{1+1}') {}_j X_{ki}'$ 

$$\bar{B}_{1+1} = Q_1 \bar{B}_1 (\bar{B}_1' Q_1^2 \bar{B}_1)^{-1/2} \quad (\text{"B-Teilschritt"})$$
- (iv)  $R_1 = {}_k X_{ij} (\bar{A}_{1+1} \bar{A}_{1+1}' \times \bar{B}_{1+1} \bar{B}_{1+1}') {}_k X_{ij}'$ 

$$\bar{C}_{1+1} = R_1 \bar{C}_1 (\bar{C}_1' R_1^2 \bar{C}_1)^{-1/2} \quad (\text{"C-Teilschritt"})$$

Zur Initialisierung in Schritt (i) werden die Eigenvektoren aus Methode I nach Tucker (vgl. Abschnitt 3.1.2) eingesetzt. Kroonenberg und De Leeuw haben gezeigt, daß für jeden Iterationszyklus gilt:

$$p(\bar{A}_{1+1}, \bar{B}_{1+1}, \bar{C}_{1+1}) \geq p(\bar{A}_1, \bar{B}_1, \bar{C}_1).$$

Bei Gleichheit wird die Iterationsfolge abgebrochen, und aus den zuletzt errechneten Faktormatrizen wird die Kernmatrix

entsprechend Gleichung (7) bestimmt. Allerdings kann nicht gezeigt werden, daß mit dieser Methode das globale Maximum der Funktion (13) auch aufgefunden wird. Der Algorithmus ist mit dergleichen Problemen behaftet wie etwa Kruskals Konzept zur nonmetrischen multidimensionalen Skalierung und kann durchaus in einem lokalen Maximum konvergieren, welches nicht mit dem globalen Maximum identisch ist. Für eine spezifische Lösung ist daher nicht gesichert, daß sie dem eingangs geforderten Optimum auch entspricht.

#### 3.1.4. Multidimensionale Skalierung individueller Differenzen nach Tucker als Sonderfall der dreimodalen Faktorenanalyse:

Ahrens (1974) interpretiert das Tuckersche Modell (1972) zur multidimensionalen Skalierung (MDS) individueller Differenzen als Revision des von Tucker und Messick (1963) entwickelten "Points-of-View"-Ansatzes unter Verwendung der dreimodalen Faktorenanalyse. Die zu analysierenden Daten bestehen aus den Ähnlichkeitsurteilen mehrerer Versuchspersonen über alle Paare aus einer gegebenen Menge von Objekten; der Datenkörper variiert zwar in drei Richtungen, Versuchspersonen  $\times$  Objekte  $\times$  Objekte, umfaßt aber nur zwei Modi und wird daher u.a. von Carroll und Arabie (1980) als "two-mode three-way data" bezeichnet. Während im "Points-of-View"-Modell idealisierte Personen postuliert werden, für die jeweils ein separater Objektraum konstruiert wird, wird im Tuckerschen Modell zur MDS die Annahme verschiedener Reizräume für die idealisierten Personen aufgegeben. Vielmehr wird im Tuckerschen Modell die Existenz eines allen Versuchspersonen gemeinsamen, subjektiven Reizraumes zugrunde gelegt, und individuelle Differenzen werden in Form unterschiedlicher Gewichte abgebildet, mit denen für verschiedene Versuchspersonen die Dimensionen des gemeinsamen Reizraums in "individuelle Reizräume" transformiert werden. Diese Grundidee ist allerdings nicht spezifisch für das Tuckersche Modell, sondern gilt in gleicher Weise für alle Modelle, die im vorliegenden Abschnitt diskutiert werden. Die Verknüpfung mit der dreimodalen Faktorenanalyse wird

unmittelbar deutlich an der Grundgleichung des Tuckerschen MDS-Ansatzes:

$$(15) \quad x_{jj',i} = (\bar{B} \times \bar{B}) \bar{G} \bar{A}' \quad .$$

In Summenschreibweise:

$$(16) \quad x_{jj',i} = \sum_m \sum_p \sum_{p'} \bar{b}_{jp} \bar{b}_{j'p'} \bar{g}_{pp'm} \bar{a}_{mi}$$

Nach Transposition der Gleichung und Identifikation der Faktormatrizen für den zweiten und dritten Modus ergibt sich (15) direkt als Sonderfall der allgemeinen Grundgleichung (4). Dabei entsprechen die Spalten der Datenmatrix den Versuchspersonen  $i$ , und die Zeilen werden als Kombinationsmodus aus den Skalarprodukten für alle Reizpaare  $j, j'$  gebildet. Zur Herleitung der Skalarprodukte werden die Distanzschätzwerte  $d(j, j')$  für jede Versuchsperson  $i$  getrennt nach der Formel von Torgerson (vgl. auch Borg, 1981, S. 390):

$$(17) \quad x_{jj',i} = \frac{-1}{2} (d(j, j')^2 - \frac{1}{J} \sum_j d(j, j')^2 - \frac{1}{J} \sum_{j'} d(j, j')^2 + \frac{1}{J^2} \sum_j \sum_{j'} d(j, j')^2)$$

transformiert, wobei  $J$  die Anzahl der Objekte bezeichnet. Aus der (transformierten) Datenmatrix  $X$  können nun analog dem Vorgehen in der dreimodalen Faktorenanalyse die Eigenvektoren und Eigenwerte für jeden Modus bestimmt werden, wobei hier zwei der drei Eigenvektormatrizen übereinstimmen. Die Dimensionen des gemeinsamen Reizraumes sind die Spalten der Eigenvektormatrix  $\bar{B}$ , die individuellen Gewichte für die Versuchspersonen sind in der Matrix  $\bar{A}$  als Zeilen angeordnet. Die Kernmatrix  $\bar{G}$  ist symmetrisch, d.h.  $\bar{g}_{pp'm} = \bar{g}_{p'pm}$ , und weist bei einer angemessenen Repräsentation in der Diagonalen keine negativen Werte auf, d.h.  $\bar{g}_{ppm} \geq 0$ . Jede Teilmatrix  $\bar{G}_m$ , die in  $\bar{G}$  in Form einer Spalte  $m$  repräsentiert ist, wird als Beitrag der  $m$ -ten Dimension des Personenraumes für die subjektiven Räume der repräsentierten Personen interpretiert (Tucker, 1972; Ahrens, 1974). Im Vorgriff auf die

Interpretation des Tuckerschen Modells sei an dieser Stelle erwähnt, daß die Dimensionen des Personenraums als "idealisierte Personen" angesehen werden. Dann gibt die Matrix  $\bar{G}_m$  Auskunft über den Gebrauch der gemeinsamen Reizdimensionen durch die idealisierte Person  $m$  unter zwei Aspekten:

- a) die Diagonalelemente zeigen nach Radizierung  $\bar{g}_{ppm}^{1/2}$  differentielle Stauchungen bzw. Streckungen der Reizdimensionen  $p$  an
- b) die Werte außerhalb der Diagonalen liefern nach der Normierung  $\bar{g}_{pp'm} / (\bar{g}_{ppm} \bar{g}_{p'p'm})^{1/2}$  Informationen zu den Winkeln zwischen den Reizachsen  $p$  und  $p'$  für die idealisierte Person  $m$  (Kroonenberg, 1983b).

In gleicher Weise werden die Zellen der erweiterten Kernmatrix  $\bar{H}$

$$(18) \quad \bar{H} = \bar{G} \bar{A}$$

unter Bezugnahme auf die (realen) Versuchspersonen  $i$  interpretiert. Jede Teilmatrix  $\bar{H}_i$ , die in  $\bar{H}$  als Spalte  $i$  angeordnet ist mit

$$(19) \quad \bar{H}_i = \sum_m \bar{G}_m \bar{a}_{mi} ,$$

enthält Angaben zu der relativen Streckung/Stauchung der Reizdimensionen und zu den Winkeln zwischen den Reizdimensionen bezüglich des Urteilssystems der Versuchsperson  $i$ . Damit entsprechen die Spalten der mit den personenspezifischen Gewichten multiplizierten Kernmatrix den Transformationen, durch welche der gemeinsame Reizraum in den Reizraum des  $i$ -ten Subjektes abgebildet wird (Kroonenberg, 1983b).

3.1.5. Die viermodale Erweiterung des Tuckerschen Modells: Lastovicka (1981) hat das Tuckersche Modell der dreimodalen Faktorenanalyse erweitert zur Analyse viermodaler Datenkörper der Form  $x_{ijkl}$ . Die Grundgleichung der viermodalen Faktorenanalyse besteht aus der Approximation der Werte  $x_{ijkl}$  durch eine Kernmatrix, die nun mit vier Faktormatrizen ge-

wichtet wird:

$$(20) \quad \hat{x}_{ijkl} = \sum_m \sum_p \sum_q \sum_r \bar{a}_{im} \bar{b}_{jp} \bar{c}_{kq} \bar{d}_{lr} \bar{g}_{mpqr} \cdot$$

In Matrixschreibweise entspricht Gleichung (20) der Darstellung

$$(21) \quad \hat{X}_{ijkl} = \bar{A} \bar{G} (\bar{B}' \times \bar{C}' \times \bar{D}'),$$

wobei der Kombinationsmodus  $jkl$  das kartesische Produkt von drei Modi ist. Zur Berechnung der Faktormatrizen werden die vier möglichen Skalarproduktmatrizen

$$\begin{array}{l} i^{X_{jkl}} \quad i^{X_{jkl}'}, \quad j^{X_{ikl}} \quad j^{X_{ikl}'}, \\ k^{X_{ijl}} \quad k^{X_{ijl}'}, \quad l^{X_{ijk}} \quad l^{X_{ijk}'} \end{array}$$

gebildet, deren Eigenvektoren die Spalten von  $\bar{A}$ ,  $\bar{B}$ ,  $\bar{C}$  und  $\bar{D}$  sind. Da auch im viermodalen Fall die Faktormatrizen (spaltenweise) orthonormal sind, läßt sich die Kernmatrix in folgender Weise bestimmen:

$$(22) \quad \bar{G} = \bar{A}' \quad i^{X_{jkl}} \quad (\bar{B} \times \bar{C} \times \bar{D}) \cdot$$

Die Generalisierung dieses Modells auf den  $n$ -modalen Fall wurde 1986 von Kapteyn, Neudecker und Wansbeek formuliert, wobei die Faktormatrizen aus weiteren Modi analog zu Gleichung (21) durch das Kroneckerprodukt verknüpft werden.

### 3.1.6. Interpretation der Kern- und Faktormatrizen des dreimodalen Modells:

Die von Tucker (1966) diskutierte Interpretation des dreimodalen Modells unterscheidet nicht nur unter formalen Gesichtspunkten zwischen der Kern- und den Faktormatrizen, sondern auch hinsichtlich des inhaltlichen Zugangs. Die in jedem Modus extrahierten Faktoren werden als "idealisierte Elemente" aufgefaßt (vgl. Abschnitt 1), die für den betrachteten Gegenstandsbereich fundamentalere Konzepte darstellen als die erhobenen Einzelbeobachtungen (Tucker, 1966). Wurden beispielsweise Meßwerte einer Reihe von Versuchspersonen auf

mehreren Variablen in unterschiedlichen Untersuchungssituationen erhoben, so resultiert aus der Analyse jeweils eine Faktormatrix für die Personen, die Variablen und die Situationen. Für die Faktoren der drei Modi wird angenommen, daß sie als grundlegende Konstrukte idealisierte Personen, latente Variablen bzw. prototypische Situationen definieren (Kroonenberg, 1983b). Die Faktormatrizen sind implizite Definitionen der idealisierten Konzepte (Lohmöller, 1979), und die begriffliche Interpretation der Faktoren wird vorgenommen aufgrund der gemeinsamen inhaltlichen Bedeutung der Elemente eines jeden Modus, die hohe Ladungen auf den Faktoren aufweisen. Hinsichtlich der inhaltlichen Deutung der Faktoren als latente Konstrukte stimmt diese Auffassung mit der traditionellen faktorenanalytischen Sichtweise überein; im Unterschied zu der klassischen Faktorenanalyse wird nicht nur eine Menge latenter Einflußgrößen identifiziert, sondern jedem Modus wird eine andersgeartete latente Struktur unterlegt (Snyder, Law & Hattie, 1984). Der Bezug zwischen den idealisierten Konzepten wird in der Kernmatrix hergestellt, die mindestens drei verschiedene Interpretationen zuläßt. Die erste Form der inhaltlichen Deutung der Kernmatrix wurde von Tucker (1964a, 1966) und Levin (1963, 1965) vorgeschlagen und hat sich weitgehend durchgesetzt (vgl. Lohmöller, 1979; Kroonenberg, 1983b):

die Werte der Kernmatrix, auch als Kernrelationen bezeichnet, als Faktorwerte der idealisierten Einflußgrößen. Nach dieser Auffassung zeigen die Kernrelationen die Ausprägungen der idealisierten Personen auf den latenten Variablen in unterschiedlichen prototypischen Situationen an. Ein Beispiel wäre eine Aussage wie "die idealtypische Person  $m$  zeichnet sich aus durch hohe Werte auf dem Intelligenzfaktor  $p$  in streßfreien Situationen  $q$ ". Dies impliziert allerdings die vorhergehende Zuordnung von Begriffen zu den Faktoren der drei Modi (Kroonenberg, 1983b). Werden jedoch die Faktoren lediglich als ökonomische Zusammenfassungen der Einzelbeobachtungen betrachtet, deren Ziel die Erfassung der systematischen Variation in den Daten ist, so kann die Interpretation der Zellen der Kernmatrix als Faktorwerte

nicht aufrechterhalten werden. Auch Bartussek (1973) problematisiert diesen inhaltlichen Zugang zu der Kernmatrix unter der Perspektive, daß die Kernmatrix mit den Varianzbeiträgen der Faktoren konfundiert ist. D.h. in der Kernmatrix sind mehrere Informationen miteinander verknüpft, und ihre Zellen sind in ihrer numerischen Größe untereinander nicht vergleichbar. Durch die Normierung der Faktoren wäre nach Bartussek eine denkbare Interpretation des Modells reziprok zu der gegenwärtigen Vorgehensweise, indem die Faktormatrizen nicht Faktorladungen, sondern Faktorwerte enthalten und vice versa die Kernmatrix nicht Faktorwerte, sondern Faktorladungen wiedergibt. Unter Bezugnahme auf Gleichung (8) wird alternativ vorgeschlagen, die an der Gesamtvarianz der Daten normierten Kernrelationen, d.h.

$$\bar{g}_{mpq}^2 / \sum_i \sum_j \sum_k x_{ijk}^2 ,$$

als Beitrag der Faktorkombination m, p, q an der Variabilität der Rohdaten zu interpretieren. In dieser Interpretation repräsentiert die Kernmatrix eine faktorspezifische Aufteilung der Varianzkomponenten in direkter Analogie zu der Singulärwert-Zerlegung. Ferner werden hierdurch zwar numerisch bedeutsame Faktorkombinationen hervorgehoben, über die Art ihres Zusammenwirkens wird aber keine Aussage getroffen.

Eine dritte Möglichkeit besteht in der inhaltlichen Deutung der Kernrelationen als dreimodale Interaktionen. Neben der relativen Größe der Kernrelationen wird in dieser Interpretation vor allem das Vorzeichen der Zellen der Kernmatrix beachtet, wodurch die "Richtung" des Beziehungsgeflechtes zwischen den Faktoren in den Vordergrund gestellt wird. So führt eine Kernrelation, die den Bezug zwischen einer idealtypischen Person m, einem Intelligenzfaktor p und einem Streßfaktor q herstellt und mit einem positiven Vorzeichen versehen ist, zu der Schlußfolgerung, daß die Person m auf dem Intelligenzfaktor p hohe Werte aufweist, wenn sie sich in einer streßbehafteten Situation q befindet. Dabei ist in diesem Beispiel impliziert, daß die Faktorladungen der Ele-

mente, die die Grundlage zur Benennung der Faktoren bilden, ebenfalls in jedem Modus positive Vorzeichen haben. Ist die gleiche Konstellation gegeben, nur mit negativem Vorzeichen für eine andere idealtypische Person  $m'$ , so bedeutet dies, daß  $m'$  auf dem Intelligenzfaktor  $p$  geringe Werte aufweist in einer mit Streß behafteten Situation  $q$ . Allerdings können auch Vorzeichenkombinationen der Kernrelationen und zugeordneten Faktorladungen auftreten wie  $(+,+,-,-)$ ,  $(-,+,-,-)$ ,  $(+,-,+,-)$  usw. Die Interpretation solcher Kombinationen ist wesentlich schwieriger, da abwechselnd die positiven bzw. negativen Ladungen der zu den Vorzeichen korrespondierenden Faktoren zu beachten sind. Der wesentliche Unterschied zu der Auffassung der Kernrelationen als Faktorwerte besteht darin, daß hier Interaktionen zwischen Gruppen von Faktorladungen herangezogen werden, während im Konzept der Faktorwerte die den Faktoren zugeordneten Begriffe die Grundlage der Interpretation bilden (Kroonenberg, 1983b). Schließlich ergibt sich eine vierte Form der inhaltlichen Deutung der Kernmatrix im Rahmen des dreimodalen MDS-Modells, die zuvor bereits dargestellt wurde: die Kernrelation  $\bar{g}_{pp'm}$  als Kosinus des Winkels zwischen den Dimensionen  $p$  und  $p'$ .

Nach Kroonenberg (1984) ist die Wahl der spezifischen Interpretation der Kernmatrix abhängig vom untersuchten Gegenstandsbereich, von der Interpretierbarkeit der Faktoren und von der zugrunde liegenden Fragestellung. Durch das Tucker'sche Modell selbst werden keine zusätzlichen Kriterien impliziert, die eine der vier Interpretationsmöglichkeiten zwingend erscheinen lassen. In einer Reihe von Anwendungen wurden Schwierigkeiten in der Interpretation des Modells deutlich, bedingt durch die abstrakten Kernrelationen (Snyder, Law & Hattie, 1984). Um den Zugang zu der Kernmatrix zu vereinfachen, wurde versucht, durch Rotationstransformationen die Kernmatrix in eine einfacher strukturierte Diagonalform zu überführen. Dieser Ansatz ist allerdings an sehr restriktive Bedingungen gebunden, und entsprechende Algorithmen zur Auffindung angemessener Transformationen sind noch in der Entwicklung befindlich.

### 3.2. Das CANDECOMP/PARAFAC-Modell

#### 3.2.1. Die Grundgleichungen des Modells:

Die folgende Darstellung des CANDECOMP-Modells (für CANonical DECOMPosition of N-way Tables) ist eng orientiert an Carroll und Changs Publikation (1970). Im Gegensatz zum Tuckerschen Modell, in dem die Verknüpfung der Faktoren aus den verschiedenen Modi in der Kernmatrix vorgenommen wird, wird in CANDECOMP auf eine komplexe Verflechtung der Faktormatrizen verzichtet. Algebraisch entspricht die Kernmatrix in CANDECOMP dem dreimodalen Analogon einer Identitätsmatrix (Tucker, 1972; Kruskal, 1976; Harshman & Lundy, 1984), so daß die Grundgleichung des Modells folgende Form hat:

$$(23) \quad x_{ijk} = \hat{x}_{ijk} + e_{ijk} \quad \text{mit}$$

$$(24) \quad \hat{x}_{ijk} = \sum_r \bar{a}_{ir} \bar{b}_{jr} \bar{c}_{kr} \quad .$$

In Matrizennotation:

$$(25) \quad {}_i\hat{X}_{jk} = {}_i\bar{A}_r {}_rI_{rr} ({}_j\bar{B}_r' \times {}_k\bar{C}_r') = [\bar{A}, \bar{B}, \bar{C}] \quad ,$$

wobei die kompakte Schreibweise  $[\ ]$  Kruskals (1977) Terminologie zur Darstellung des Tripelproduktes von Matrizen A, B, C entspricht.  $\hat{X}$  ist der durch das Modell approximierte Anteil an den beobachteten Werten X und  $e_{ijk}$  kennzeichnet einen Fehlerterm, der nicht näher spezifiziert ist (Carroll & Pruzansky, 1984). Ein weiterer Unterschied zu dem Tuckerschen Modell wird ebenso an der CANDECOMP-Grundgleichung deutlich und bezieht sich auf die Anzahl der extrahierten Faktoren in den drei Modi; während das Tuckersche Modell eine unterschiedliche Dimensionalität in jedem Modus zuläßt, ist CANDECOMP restriktiver, indem die Zahl der Faktoren r zwischen den Modi übereinstimmen muß. Wie auch im Tuckerschen Modell ist eine exakte Zerlegung der Datenmatrix X in zugrunde liegende CANDECOMP-Faktoren möglich (Kruskal, 1976, 1977)

$$(26) \quad x_{ijk} = \sum_r \bar{a}_{ir} \bar{b}_{jr} \bar{c}_{kr} \quad ,$$

die zu einer vollständigen Aufklärung der Varianz der Daten führt. Kruskal konnte hier zeigen, daß sich die Eigenschaften der Faktormatrizen  $\bar{A}$ ,  $\bar{B}$  und  $\bar{C}$  wesentlich von den Eigenschaften klassischer faktorenanalytischer Zerlegungen unterscheidet, die aus der Faktorisierung einer zweimodalen Matrix nach dem Theorem von Eckart und Young (1936) resultieren. Das wichtigste Charakteristikum des trilinearen Modells ist die Elimination des Rotationsproblems, von Kruskal als "intrinsic axis property" bezeichnet:

im Gegensatz zu den klassischen faktorenanalytischen Modellen und dem Tuckerschen Modell ist die Orientierung der Faktoren in CANDECOMP festgelegt, so daß keine zusätzlichen Faktorrotationen durchgeführt werden können. Bezeichnen  $i$ ,  $j$  und  $k$  die Anzahl der Elemente in jedem der drei Modi, so konnte Kruskal darüberhinaus nachweisen, daß die benötigte Zahl der Faktoren zur Reproduktion der Datenmatrix nicht gleich  $\min(i,j,k)$  ist, sondern  $\max(i,j,k)$  übersteigt. Als Konsequenz dieser Rangeigenschaft dreimodaler Matrizen können sogar linear abhängige Faktoren in jedem Modus auftreten; die Orthogonalität der Faktoren in den klassischen Verfahren kann hier nicht aufrechterhalten werden.

Carroll und Chang (1970) heben vor allem die eindeutige Orientierung der Referenzachsen als Desideratum ihres Modells hervor und ziehen aus dieser Eigenschaft die Schlußfolgerung, daß den Faktoren eine unmittelbare psychologische Bedeutung zukommt. Auch Harshman und Lundy (1984) sehen in der "intrinsic axis property" ein besonders bedeutungsvolles Spezifikum des Modells, das sie mit zwei Aspekten begründen. Einmal erfordert das trilineare Modell keine Annahmen zur Auffindung einer Konfiguration, welche über die Modellgleichung hinausgehen. Die traditionelle Faktorenanalyse dagegen wirft nach Harshman und Lundy das Problem auf, neben dem Modell zur Faktorextraktion zusätzliche Kriterien wie etwa eine Einfachstruktur zu definieren, die keinen wesentlichen Bezug zu den Modellparametern aufweisen. Zweitens beansprucht CANDECOMP nicht die Allgemeingültigkeit traditionel-

ler Verfahren, die immer zur Herleitung von Faktoren aus beobachteten Variablen führen (Revenstorf, 1980). Vielmehr kann nach Harshman und Lundy in empirischen Analysen abgeschätzt werden, ob einem Datenkörper fixierte Achsen zugrunde liegen oder nicht, da bei Verletzung dieser Annahme eine degenerierte Konfiguration erzeugt wird, die keinen konsistenten interpretativen Zugang ermöglicht und sich durch sehr hohe Korrelationen zwischen den Faktoren innerhalb der drei Modi auszeichnet.

Während Carroll und Chang das CANDECOMP-Modell eher zur theoretischen Untermauerung ihres INDSCAL-Verfahrens konzipiert haben und daher eigentlich an Fragestellungen der multidimensionalen Skalierung interessiert waren (Carroll & Arabie, 1980), begründet Harshman (1976; Harshman & Berenbaum, 1981) die Entwicklung seines PARAFAC-Modells (für PARAllel FACtors factor analysis) mit der angestrebten Überwindung des Rotationsproblems in der klassischen Faktorenanalyse. Harshman ging dabei von Cattells "principle of proportional profiles" zur Auswahl einer spezifischen Rotationstransformation aus, welches er auf dreimodale Matrizen übertrug. Obwohl beide Modelle unter einem unterschiedlichen Rationale formuliert wurden, sind sie algebraisch völlig äquivalent (Kruskal, 1984; Carroll & Pruzansky, 1984; Harshman & Berenbaum, 1981) und verwenden identische Algorithmen zur Faktorextraktion (Harshman & Lundy, 1984). Allerdings sind in diesem Rahmen zwei Aspekte erwähnenswert, die zu einer Differenzierung der Verfahren führen. Erstens wurde PARAFAC spezifisch zur Analyse dreimodaler Datenkörper konzipiert, während CANDECOMP allgemeiner ist und eine Erweiterung für den n-modalen Fall enthält. Zweitens umfaßt Harshmans Vorgehen zusätzlich eine Reihe von Prozeduren zur angemessenen Normierung oder Standardisierung des Datenkörpers vor der faktorenanalytischen Dekomposition der Daten. Diese Prozeduren sind zwar nicht in PARAFAC selbst verankert und können direkt in der "preprocessing"-Phase aller Modelle eingesetzt werden, sie werden aber allgemein als das wesentliche Charakteristikum von Harshmans Zugang betrachtet

(Kruskal, 1984; Snyder, Law & Hattie, 1984).

CANDECOMP/PARAFAC zeichnet sich zwar durch die "intrinsic axis property" aus, jedoch bleibt neben möglichen Permutationen der Faktoren eine Unterdeterminiertheit des Modells hinsichtlich der Skalierung der Achsen bestehen (Kruskal, 1984). D.h. es ist immer möglich, die Achsen in ein oder zwei Modi zu strecken oder zu stauchen, wenn im verbleibenden Modus die inverse Operation durchgeführt wird. Enthält Modus  $i$  etwa die Versuchspersonen, so werden allgemein die Faktoren von Modus  $j$  und  $k$  auf die Länge eins normiert mit der reziproken Skalierung der Faktoren in Modus  $i$ . Dieses Vorgehen ist nicht zwingend, impliziert aber nach Harshman und Lundy (1984) hinsichtlich der Interpretation der skalierten Konfiguration wesentliche Vorteile, die in Abschnitt 3.2.5. aufgegriffen werden.

CANDECOMP/PARAFAC enthält in seiner allgemeinen Form nach Gleichungen (24) bis (26) eine direkte Faktorisierung des zugrunde liegenden Datenkörpers. Im Gegensatz zum Tucker-schen Modell, in dem eine vorhergehende Bestimmung von Skalarproduktmatrizen erforderlich ist, erfolgt hier eine Zerlegung der Datenmatrix ohne weitere Zwischenschritte. Daneben existiert das symmetrische trilineare CANDECOMP/PARAFAC-Modell als Sonderfall zur indirekten Faktorisierung von Matrizen. Dieser Ansatz nimmt insofern eine indirekte Zerlegung vor, als eine Menge von Kovarianzmatrizen derselben Variablen  $i, i'$  betrachtet wird, die jeweils unter verschiedenen (experimentellen) Bedingungen  $k$  erhoben wurden. Die Kovarianzen  $x_{ii'k}$  werden über die Elemente des Modus  $j$  gebildet, so daß ein "three-way two-mode"-Datenkörper entsteht. Anschließend wird die latente Faktorstruktur in den Modi  $i$  und  $k$  bestimmt, die nach folgender Gleichung definiert ist:

$$(27) \quad x_{ii'k} = \sum_r \bar{c}_{kr} \bar{a}_{ir} \bar{a}_{i'r} \quad .$$

Die Werte  $\bar{a}_{ir}$  sind dabei Faktorladungen der Variablen  $i$ , und

die Koeffizienten  $\bar{c}_{kr}$  werden als Indikatoren für den Einfluß des Faktors  $r$  auf die Bedingung  $k$  interpretiert. Das Modell zur Analyse von Kovarianzmatrizen ist eng verknüpft mit der INDSICAL-Methode der multidimensionalen Skalierung individueller Differenzen, welche in Abschnitt 3.2.3. diskutiert wird.

### 3.2.2. Der ALS-Algorithmus zum CANDECOMP/PARAFAC-Modell:

Der "Alternating-Least-Squares"-Ansatz, der bereits im Rahmen des Algorithmus von Kroonenberg und De Leeuw skizziert wurde, wird in CANDECOMP/PARAFAC zur Bestimmung des Minimums der Funktion

$$(28) \quad g(\bar{A}, \bar{B}, \bar{C}) = \text{Spur}({}_i X_{jk} - {}_i \hat{X}_{jk})({}_i X_{jk} - {}_i \hat{X}_{jk})'$$

$$\text{mit } {}_i \hat{X}_{jk} = [\bar{A}, \bar{B}, \bar{C}]$$

eingesetzt. Zunächst muß die Anzahl der gewünschten Faktoren  $r$  bestimmt werden. Unter Verwendung der Methode von Bauer-Rutishauser werden folgende Schritte iterativ durchlaufen, die in Abbildung 6 graphisch veranschaulicht sind:

(i) für  $\bar{B}$  und  $\bar{C}$  sind zwei Startkonfigurationen  $\bar{B}_1$  und  $\bar{C}_1$  auszuwählen

(ii) nach 1 Iterationen:

$$P_1 = {}_i X_{jk} (\bar{B}_1 \times \bar{C}_1)$$

$$Q_1 = \bar{B}_1' \bar{B}_1 \times \bar{C}_1' \bar{C}_1$$

$$\bar{A}_{1+1} = P_1 Q_1^{-1} \quad (\text{"A-Teilschritt"})$$

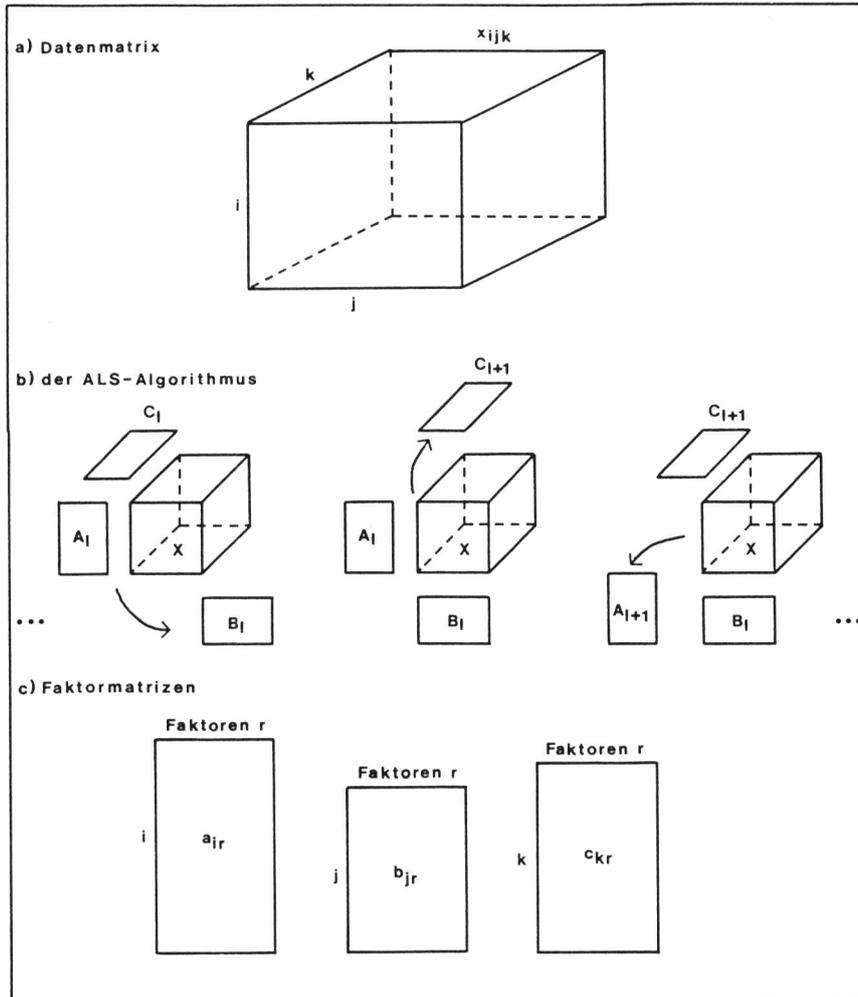
(iii)  $P_1 = {}_j X_{ik} (\bar{A}_{1+1} \times \bar{C}_1)$

$$Q_1 = \bar{A}_{1+1}' \bar{A}_{1+1} \times \bar{C}_1' \bar{C}_1$$

$$\bar{B}_{1+1} = P_1 Q_1^{-1} \quad (\text{"B-Teilschritt"})$$

(iv)  $P_1 = {}_k X_{ij} (\bar{A}_{1+1} \times \bar{B}_{1+1})$

Abb. 6: Das CANDECOP-Verfahren von Carroll &amp; Chang (1970)



$$Q_1 = \bar{A}_{1+1}' \bar{A}_{1+1} \times \bar{B}_{1+1}' \bar{B}_{1+1})$$

$$\bar{C}_{1+1} = P_1 Q_1^{-1} \quad (\text{"C-Teilschritt"})$$

Dabei wird im Gegensatz zu dem Vorgehen von Kroonenberg und De Leeuw hier nicht gefordert, daß die Matrizen A, B und C aus orthonormalen Spalten aufgebaut sein müssen. Auch der CANDECOMP/PARAFAC-Algorithmus ist mit dem Problem behaftet, daß möglicherweise nur ein lokales Optimum der Funktion (28) aufgefunden wird, das mit dem globalen Optimum nicht übereinstimmt. Carroll und Pruzansky (1984, S. 379) bemerken dazu: " In practice, the method seems almost always to converge to the global optimum solution. There is a mild local minimum problem... This problem seems to be very slight in comparison with that in non-metric two-way MDS...". Borg (1981, S. 509) beurteilt hingegen die Optimierungseigenschaften als undurchsichtig und stellt fest, daß zusätzlich keine rationale Startkonfiguration bisher definiert wurde. Neben den unterschiedlichen Auffassungen zu der Qualität des ALS-Algorithmus erscheint in diesem Zusammenhang erwähnenswert, daß die Faktoren in allen Modi simultan extrahiert werden. Im Unterschied zu den klassischen faktorenanalytischen Verfahren, die eine sukzessive Extraktion vornehmen, muß daher zum Vergleich von Konfigurationen unterschiedlicher Dimensionalität jeweils eine separate Analyse durchgeführt werden. Generell kann nicht davon ausgegangen werden, daß etwa die Faktoren einer vierdimensionalen Lösung sich unverändert auch in der fünfdimensionalen Lösung wiederfinden (Harshman & Berenbaum, 1981; Harshman, Laderfeged & Goldstein, 1977). Dies gilt in gleicher Weise für den Algorithmus von Kroonenberg und De Leeuw, der zusätzliche Kombinationsmöglichkeiten in der Wahl der Parameter m, p und q für die Dimensionalität der Faktormatrizen eröffnet. Aufgrund dieser Charakteristika der ALS-Methode wurden neue Kriterien zur Beurteilung der Angemessenheit einer Konfiguration erforderlich. Die wesentlichen Aspekte dieser Kriterien werden in Abschnitt 6 diskutiert.

### 3.2.3. Multidimensionale Skalierung individueller Differenzen als Sonderfall von CANDECOMP/PARAFAC:

Der wohl am meisten angewendete Vertreter aus der Klasse der Modelle individueller Differenzen ist die ebenfalls von Carroll und Chang (1970) begründete INDSCAL-Methode (für Individual Differences multidimensional SCALing). Die Skalarproduktform von INDSCAL ist identisch mit Gleichung (27) zur Analyse von Kovarianzmatrizen. Diese Verknüpfung der MDS mit dem faktorenanalytischen Ansatz wird, wie auch schon im Tuckerschen Modell, ermöglicht durch Torgersons Prozedur zur Herleitung von Skalarprodukten  $x_{jj',i}$  aus den Distanzschätzwerten  $d(j,j')$  für jede Person  $i$  entsprechend Gleichung (17). INDSCAL umfaßt drei Arten von Räumen:

- a) den Personenraum ("subject space"),
- b) den Reizraum ("group stimulus space") und
- c)  $i$  individuelle Räume ("private perceptual spaces").

In der Skalarproduktform des Modells

$$x_{jj',i} = \sum_r \bar{b}_{jr} \bar{b}_{j',r} \bar{a}_{ir}$$

repräsentiert  $\bar{A} = (\bar{a}_{ir})$  den Personenraum und  $\bar{B} = (\bar{b}_{jr})$  den Reizraum der beurteilten Objekte. Die Koeffizienten  $\bar{a}_{ir}$  werden als personenspezifische Gewichte oder "salience" interpretiert, die den  $r$ -dimensionalen Reizraum, der für alle Personen gemeinsam ist, in die individuellen Räume  $\bar{z}_{jr,i}$  transformiert:

$$(29) \quad \bar{z}_{jr,i} = \bar{a}_{ir}^{1/2} \bar{b}_{jr} \quad (\text{individueller Raum für Person } i)$$

Die Orientierung der Referenzachsen in den interindividuell unterscheidbaren subjektiven Reizräumen ist durch  $\bar{B}$  eindeutig festgelegt, während Unterschiede zwischen den Urteilssystemen der Personen durch differentielle Streckungen/Stauchungen der Achsen erfaßt werden. Dabei werden die personenspezifischen Gewichte als Indikatoren der Bedeutsamkeit der Dimensionen  $r$  für die Person  $i$  interpretiert und weisen bei einer angemessenen Repräsentation keine negativen Werte auf (Carroll & Chang, 1970). Zwei Personen verfügen nach dem

INDSCAL-Ansatz über völlig verschiedene individuelle Räume, wenn die Koeffizienten einer Person auf den Dimensionen nahe null sind, für die hohe Gewichte der anderen Person vorliegen. Im Extremfall existieren für zwei oder mehrere Gruppen von Personen keine gemeinsamen Referenzachsen, so daß der übergeordnete Reizraum, repräsentiert durch  $\bar{B}$ , aus der Summe der Dimensionen der individuellen Räume besteht. Für den allgemeinen Fall wird jedoch angenommen, daß die Personen  $i$  zumindest teilweise Gemeinsamkeiten in ihren Wahrnehmungsräumen aufweisen. Die Distanzschätzwerte für eine Person  $i$ ,  $d_{jj'i}$ , ergeben sich nach INDSCAL aus

$$(30) \quad d_{jj'i} = \left( \sum_r \bar{a}_{ir} (\bar{b}_{jr} - \bar{b}_{j'r})^2 \right)^{1/2} ,$$

die als "weighted Euclidean distances" bezeichnet werden (Young, 1987). Gegenüber etwa der nonmetrischen MDS nach Kruskal wird in INDSCAL nur die euklidische Metrik zugrunde gelegt, und ferner wird angenommen, daß die erhobenen Reiz-(un)ähnlichkeiten  $\delta_{jj'i}$  durch eine lineare Funktion  $f$  mit den euklidischen Distanzschätzwerten verknüpft sind:

$$(31) \quad f_i(\delta_{jj'i}) \cong d_{jj'i} ,$$

wobei generell für jede Person  $i$  eine separate Funktion  $f$  postuliert wird.

Wie auch in CANDECOMP/PARAFAC werden die Dimensionen des gemeinsamen Reizraumes auf die Länge eins normiert, so daß nach Anwendung der inversen Operation auf den Personenraum die Länge eines Vektors  $\bar{a}_i$ , der die Koordinaten einer Person  $i$  repräsentiert, als Maß für die Varianzaufklärung der Daten von Person  $i$  durch die INDSCAL-Konfiguration interpretiert wird (Carroll & Pruzansky, 1984).

Während sich INDSCAL formal vom Tuckerschen Modell dadurch unterscheidet, daß die Kernmatrix auf eine Diagonalform in allen  $r$  Lagen eingeschränkt wird verknüpft mit der Unmöglichkeit von Faktorrotationen, läßt das dritte von Carroll und Chang konzipierte Modell - IDIOSCAL (Individual Differences In Orientation SCALing, vgl. Carroll & Wish, 1974) -

neben differentiellen Gewichtungen der Reizdimensionen auch orthogonale Rotationen zu. Die Grundgleichung des Modells

$$(32) \quad X_i = \bar{B} \bar{A}_i \bar{B} \quad ,$$

in der  $X_i$  die Skalarproduktmatrix für Person  $i$  und  $\bar{B}$  den Reizraum der beurteilten Objekte bezeichnet, wird weiter zerlegt in:

$$(33) \quad \bar{A}_i = T_i C_i T_i' \quad .$$

$T_i$  ist eine orthogonale Matrix, die idiosynkratische Rotationen des Reizraums anzeigt;  $C_i$  ist diagonal und enthält die individuellen Dimensionsgewichte. Gleichung (32) entspricht formal dem Tuckerschen MDS-Modell bei Betrachtung der erweiterten Kernmatrix. Allerdings ist  $\bar{A}_i$  im Tuckerschen Modell (vgl. Abschnitt 3.1.4) das Produkt aus einer (spaltenweise) orthogonalen Eigenvektormatrix und einer Kernmatrix, die allgemein keine Orthogonalitätseigenschaften aufweist.

#### 3.2.4. Die n-modale Erweiterung von CANDECOMP/PARAFAC:

Während PARAFAC spezifisch zur Analyse dreimodaler Datenkörper entwickelt wurde (Harshman & Lundy, 1984), ist CANDECOMP das Akronym für ein generelles multilineares Modell, welches eine Analyse von Daten aus beliebig vielen Modi gestattet. Werden die Modi mit  $1, 2, \dots, n$  und ihre Elemente mit  $i_1, i_2, \dots, i_n$  bezeichnet, so ergibt sich die n-modale Grundgleichung (Carroll & Chang, 1970)

$$(34) \quad \hat{x}_{i_1 i_2 \dots i_n} = \sum_r \bar{a}_{i_1 r} \bar{a}_{i_2 r} \dots \bar{a}_{i_n r} \quad .$$

$\bar{A}_n$  ist die Faktormatrix des Modus  $n$ , und aus der Setzung  $\bar{a}_{i_r} = \bar{a}_{i_1 r}$ ,  $\bar{b}_{j_r} = \bar{a}_{i_2 r}$ ,  $\bar{c}_{k_r} = \bar{a}_{i_3 r}$  ergibt sich das dreimodale Modell als Sonderfall des allgemeinen Ansatzes. Zur Berechnung der Faktormatrizen wird der ALS-Algorithmus verwendet, wobei aus dem Produkt von  $n-1$  Parametermatrizen die  $n$ -te Matrix iterativ geschätzt wird. Der Algorithmus wurde zur

Analyse von maximal siebenmodalen Datenkörpern implementiert. Bisherige Anwendungen des allgemeinen Modells umfassen Least-Squares-Schätzungen von Kontingenztafeln nach dem "latent class model" von Lazarsfeld (Carroll & Pruzansy, 1984).

### 3.2.5. Interpretation des dreimodalen CANDECOMP/PARAFAC-Modells:

Die Interpretation des CANDECOMP/PARAFAC-Modells unterscheidet sich in mehreren Aspekten von der Interpretation des Tuckerschen Modells und gestattet durch das Fehlen komplexer intermodaler Faktorverknüpfungen einen direkten Zugang. Im Gegensatz zum Tuckerschen Modell, das mit den formal unterscheidbaren Kern- und Faktormatrizen verschiedene Interpretationen verbindet, besteht CANDECOMP/PARAFAC aus drei konzeptuell gleichartigen Konfigurationen. Für jede Konfiguration wird angenommen, daß ihre zugrunde liegenden Dimensionen latente Konstrukte darstellen. Die Koeffizienten der Faktormatrizen werden in jedem Modus als Faktorladungen aufgefaßt, deren direktes Zusammenwirken die Einzelbeobachtungen bestimmt.

Während im Tuckerschen Modell jedem Modus eine andersgertete latente Struktur unterlegt wird, stimmt CANDECOMP/PARAFAC mit der traditionellen faktorenanalytischen Sichtweise insofern überein, als Faktoren identifiziert werden, die zwischen den Modi einander zugeordnet sind und in denen sich jeweils eine latente Einflußgröße  $r$  manifestiert. D.h. zur Beschreibung des Datenkörpers wird jeweils ein Tripel von Faktoren herangezogen, welches als eine Einflußgröße betrachtet wird und eine Menge zusammengehöriger Konstrukte definiert. Der Anteil einer Komponente  $r$  an der Varianz der Daten ergibt sich aus der Multiplikation der zugeordneten Faktorladungen. Ferner werden aufgrund dieser Grundform von CANDECOMP/PARAFAC multiple Interaktionen der Faktoren aus verschiedenen Modi ausgeschlossen, so daß inhaltliche Bezüge nur innerhalb eines Tripels hergeleitet werden können. Es resultiere beispielsweise aus einer Untersuchung konnotativer Begriffsbedeutungen die Struktur des semantischen Raumes

mit den Referenzachsen "evaluation", "activity" und "potency" im Skalenmodus, und es seien auf den zugeordneten Faktoren im Personenmodus jeweils den gleichen Probanden hohe Ladungen zugeordnet. Dann zeichnen sich nach CANDECOMP/PARAFAC für die auf den entsprechenden Personenfaktoren repräsentierten Probanden nur solche Begriffe durch evaluative Komponenten aus, die hohe Faktorladungen in dem Faktortripel aufweisen, welches die Bewertungsdimension umfaßt; alle übrigen Begriffe enthalten keine evaluativen Aspekte. Hierin liegt ein wesentlicher Unterschied zum Tuckerschen Modell (Harshman & Lundy, 1984), das durch das Konzept der Kernmatrix Zusammenhänge mit anderen Begriffsdimensionen zulassen würde. Innerhalb eines Modus werden den CANDECOMP/PARAFAC-Faktoren spezifische Interpretationen aufgrund der gemeinsamen inhaltlichen Bedeutung der Elemente mit hohen Faktorladungen zugeordnet. Zur inhaltlichen Deutung des Zusammenhangs der Faktoren zwischen den Modi, d.h. der Faktortripel, kann in Abhängigkeit vom untersuchten Gegenstandsbereich einer der Modi als bedeutungsdefinierend in den Vordergrund gestellt werden - etwa der Skalenmodus, der Modus der Variablen oder der Testitems (Harshman & De Sarbo, 1984).

Dieser allgemeine Zugang zur Interpretation des CANDECOMP/PARAFAC-Modells erfährt verschiedene Erweiterungen bzw. Abwandlungen in Abhängigkeit von der Normierung des Datenkörpers und der Faktormatrizen. Wie bereits angemerkt wurde, werden allgemein die Faktoren der Modi  $j$  und  $k$  auf die Länge eins normiert mit der inversen Skalierung in Modus  $i$ . Diese Vorgehensweise impliziert zwei Möglichkeiten der Interpretation der normierten Konfiguration. Einmal können die Varianzbeiträge der Faktoren in Modus  $i$  direkt als Anteil an der Varianzaufklärung der Daten aufgefaßt werden. Zum zweiten stellen die Faktorladungen  $\bar{a}_{ir}$  spezifische Gewichte oder "saliances" dar, d.h. sie sind Indikatoren der Bedeutsamkeit der  $r$ -ten Faktorkombination für das  $i$ -te Element. Nach dieser Auffassung werden in Modus  $i$  einzelne Faktorladungen herangezogen, die auf Variationen in der interindividuellen Bedeutsamkeit hindeuten. Sind in Modus  $i$  z.B. die Faktorla-

dungen unterschiedlicher Versuchsbedingungen angeordnet, so zeigen die Koeffizienten  $\bar{a}_{ir}$  das Ausmaß des Einflusses der Kombination  $\bar{b}_r$  und  $\bar{c}_r$  auf die Bedingung  $i$  an (Carroll & Pruzansky, 1984; Kroonenberg, 1983b).

Harshman und Lundy (1984) sehen in dieser Normierungskonvention die direkte Erweiterung der Singulärwert-Zerlegung und fassen die standardisierten Koeffizienten der Modi  $j$  und  $k$  als Faktorwerte auf. Verschiedene weitere Interpretationsmöglichkeiten werden von Harshman und Lundy für spezifische Skalierungen der Datenmatrix und der Faktormatrix diskutiert. Beispielhaft sei folgende Konstellation herausgegriffen:

die Datenmatrix sei über Modus  $j$  und  $k$  mittelwertszentriert und über Modus  $i$  standardisiert. Werden zusätzlich die Faktoren der Modi  $j$  und  $k$  standardisiert, so können die Koeffizienten  $\bar{b}_{jr}$  und  $\bar{c}_{kr}$  als durchschnittliche Faktorwerte interpretiert werden. Wie von Kroonenberg (1983b) hervorgehoben wurde, existieren insgesamt 14 verschiedene Möglichkeiten der Normierung ( $i$ -,  $j$ -,  $ij$ -,  $jk$ -Normierung usw.) und ebenso viele Möglichkeiten der Standardisierung des Datenkörpers, verknüpft mit sieben alternativen Strategien der Normierung der Faktormatrizen. Dementsprechend gestattet die Vorgehensweise von Harshman und Lundy sicherlich differenzierte inhaltliche Aufschlüsse, beinhaltet aber kein allgemeingültiges Konzept in der Interpretation dreimodaler Konfigurationen.

## 3.3. Das SUMMAX-Modell

## 3.3.1. Die Grundgleichungen des Modells:

Während in den bisher dargestellten Ansätzen eine Übertragung der "Eckart-Young-Decomposition" zweimodaler Matrizen auf den dreimodalen Fall vorgenommen wurde, enthält das von Orlik (1980) begründete SUMMAX-Verfahren eine Generalisierung des Prinzips der Centroid-Analyse von Thurstone (1947). Sowohl in CANDECOMP/PARAFAC als auch in SUMMAX besteht die Grundgleichung aus der Approximation der beobachteten Werte durch drei Faktormatrizen gleicher Spaltenordnung  $f$ :

$$(35) \quad \hat{x}_{ijk} = \sum_f a_{if} \sum_{f'} b_{jf'} \sum_{k'} c_{kf'} = [A, B, C] \quad ,$$

in Summenschreibweise:

$$(36) \quad \hat{x}_{ijk} = \sum_f a_{if} b_{jf} c_{kf} \quad .$$

Im Unterschied zu CANDECOMP/PARAFAC wird in SUMMAX sukzessive ein Tripel von Faktoren  $f'$ ,  $f' = 1, \dots, f$ , bestimmt, so daß gilt:

$$(37) \quad \sum_i \sum_j \sum_k z_{if'} z_{jf'} z_{kf'} x_{ijkf'} = s_{f'} := \max \quad .$$

*wat is de verbetering?*

$x_{ijkf'}$  sind die Elemente der Datenmatrix bei Extraktion des ersten Tripels, d.h.  $f' = 1$ , und die Elemente der jeweiligen Residualmatrix für  $f' > 1$ . Die Vektoren  $z_i, z_j, z_k$ , in Analogie zu dem Term Eigenvektor als "Eigenpattern" bezeichnet, enthalten Werte von +1 oder -1 und stellen die Grundlage zur Definition der Faktormatrizen dar. Entsprechend dem Rationale der Centroid-Analyse sind die drei Eigenpattern als Indikatoren der Elemente des Datenkörpers interpretierbar, die zu reflektieren sind, damit "die GesamtSUMme aller Werte der dreimodalen Matrix MAXimal wird" (Orlik, 1980, S. 192). Wird nach diesem Maximierungsprinzip etwa die Gesamtsumme der Werte  $x_{ijk}$  durch Vorzeichenumpolung des  $i$ -ten Elementes erhöht, so enthält die  $i$ -te Zelle des Eigenpatterns  $z_{if'}$  den Wert -1; analog werden die Eigenpattern für  $j$  und  $k$  be-

stimmt. Sind die Eigenpattern für zwei Modi nach Gleichung (37) bestimmt worden, so definiert die Multiplikation der Datenmatrix mit ihrem Kroneckerprodukt jeweils das Eigenpattern für den dritten Modus. Dazu seien folgende Definitionen betrachtet:

$$(38a) \quad \sum_j \sum_k z_{jf'} z_{kf'} x_{ijkf'} = u_{if'}$$

$$(38b) \quad \sum_i \sum_k z_{if'} z_{kf'} x_{ijkf'} = v_{jf'}$$

$$(38c) \quad \sum_i \sum_j z_{if'} z_{jf'} x_{ijkf'} = w_{kf'} \quad .$$

Die drei aus Gleichung (38) resultierenden Vektoren  $u$ ,  $v$  und  $w$  enthalten die Randsommen der Datenmatrix nach Reflektion der Elemente in je zwei Modi. Wie von Orlik (1980) gezeigt wurde, gelten folgende Beziehungen, die den wechselseitigen Zusammenhang zwischen  $u$ ,  $v$  und  $w$  verdeutlichen:

$$(39a) \quad \sum_i z_{if'} u_{if'} = s_{f'}$$

$$(39b) \quad \sum_j z_{jf'} v_{jf'} = s_{f'}$$

$$(39c) \quad \sum_k z_{kf'} w_{kf'} = s_{f'} \quad .$$

Aus (38) und (39) ergibt sich die Definition der Faktorladungen für ein Tripel  $f'$ :

$$(40a) \quad a_{if'} = u_{if'} s_{f'}^{-2/3}$$

$$(40b) \quad b_{jf'} = v_{jf'} s_{f'}^{-2/3}$$

$$(40c) \quad c_{kf'} = w_{kf'} s_{f'}^{-2/3} \quad .$$

is dit  
proces  
optimaal  
?

Durch Subtraktion des Tripelproduktes  $a_{if'}, b_{jf'}, c_{kf'}$  von der Datenmatrix wird die Residualmatrix bestimmt, aus der anschließend ein neues Faktortripel extrahiert wird. Die sukzessive extrahierten Faktoren bilden die Spalten der SUMMAX-Faktormatrizen A, B und C.

Neben der Übereinstimmung der Anzahl der Faktoren  $f$  in den drei Modi teilt die Grundform des SUMMAX-Modells eine Reihe von Eigenschaften mit CANDECOMP/PARAFAC, die an dieser Stelle nur erwähnt seien und erst in Abschnitt 4 systematisch dargestellt werden:

- das Verfahren gestattet eine direkte Faktorisierung des Datenkörpers ohne vorhergehende Bestimmung von Skalarproduktmatrizen
- die Faktoren innerhalb eines Modus sind im allgemeinen nicht orthogonal
- die Faktoren zeichnen sich durch die "intrinsic axis property" aus.

Zur Kompensation der im Vergleich mit der Hauptachsentransformation suboptimalen Eigenschaften der Faktormatrizen und zur Vermeidung von Interpretationsproblemen durch Parameterrestriktionen wie in CANDECOMP/PARAFAC wird die Grundform des SUMMAX-Modells erweitert um eine Kernmatrix, die die Möglichkeit von Faktorrotationen eröffnet. Standardmäßig werden die Faktormatrizen für jeden Modus getrennt auf Hauptachsen transformiert. Das Modell, im folgenden als erweitertes SUMMAX-Modell bezeichnet, kann durch Gleichungen (41) und (42) dargestellt werden:

$$(41) \quad iX_{jk} = iA_m mG_{pq} (jB_p' \times kC_q') = A G (B' \times C')$$

$$(42) \quad mG_{pq} = fT_m' fI_{ff} (fW_p \times fH_q) = T' I (W \times H)$$

mit

$$(43a) \quad \widetilde{iA}_m = iA_f T = F \Delta_A^{1/2}$$

$$(43b) \quad \widetilde{jB}_p = jB_f W = V \Delta_B^{1/2}$$

$$(43c) \quad \widetilde{kC}_q = kC_f H = E \Delta_C^{1/2} .$$

$T$ ,  $W$  und  $H$  sind die orthogonalen Rotationen, die die Faktoren der drei Modi auf Hauptachsen transformieren.  $\Delta$  kenn-

*we kann das als T orthogonal zu A mit dem hoch A orthogonal?*

zeichnet drei Diagonalmatrizen, die die zugehörigen Eigenwerte enthalten, und  $F$ ,  $V$ ,  $E$  sind die auf Einheitslänge normierten Hauptkomponenten. Existieren in den Faktormatrizen der Grundgleichung (35) linear abhängige Faktoren, so resultieren aus diesem Vorgehen weniger Hauptachsen  $m$ ,  $p$ ,  $q$  als unrotierte Faktoren  $f$ . Dabei impliziert das Modell nicht nur die Möglichkeit der Rotation auf Hauptachsen, sondern orthogonale wie auch oblique Rotationstransformationen etwa nach Kriterien der Einfachstruktur oder Procrustes-Rotationen können in Abhängigkeit von der zugrunde liegenden Fragestellung angewendet werden. Bei obliquen Rotationen sind in der Definition der Kernmatrix die Transponierten der Transformationsmatrizen durch die inversen Matrizen zu ersetzen. Werden weitere Rotationen im Anschluß an die Hauptachsentransformation durchgeführt, so ist die Darstellung des vollständigen Modells formal identisch mit Gleichungen (9) bis (11) des Tuckerschen Ansatzes, wobei die Regularität der Transformationsmatrizen vorausgesetzt wird.

Im Unterschied zum Tuckerschen Modell gehen die Eigenwerte aus der Hauptachsentransformation in die rotierten Faktormatrizen  ${}_i A_m$ ,  ${}_j B_p$  und  ${}_k C_q$  ein. Daher ist die SUMMAX-Kernmatrix nicht mit den Varianzbeiträgen der Faktoren konfundiert; vielmehr verknüpft die Kernmatrix die in den verschiedenen Modi getrennt durchgeführten Rotationen (Orlik, 1980).

Weitere Unterschiede zwischen dem Konzept der Kernmatrix in SUMMAX und im Tuckerschen Modell ergeben sich bei Betrachtung der Kernrelationen als Trilinearformen, die als Erweiterung einer allgemeinen Bilinearform dargestellt werden können. Wird der Vektorraum  $\mathbb{R}^n$  betrachtet, so ist mit der Abbildung

$$B : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \quad ,$$

die in jedem Argument linear ist, eine Bilinearform gegeben. Dieser bilinearen Funktion kann eine Matrix  $A$  zugeordnet werden mit (vgl. Lamprecht, 1978, 1983; Betten, 1977)

$$B(x,y) = A_{ij} x_i y_j \quad .$$

Für den Sonderfall  $B(x,y) = x_i y_j$  erhält man das Skalarprodukt, das häufig als Maß für die Interrelation der Vektoren

$x$  und  $y$  interpretiert wird. Die der bilinearen Funktion entsprechende Trilinearform ist definiert als (Betten, 1977)

$$F(x,y,z) = A_{ijk} x_i y_j z_k .$$

Dieser Ausdruck entspricht der Darstellung der Tuckerschen Kernmatrix nach Gleichung (7), wenn  $x$ ,  $y$  und  $z$  jeweils als eine Spalte der Faktormatrizen  $\bar{A}$ ,  $\bar{B}$  und  $\bar{C}$  aufgefaßt werden. Damit sind die Kernrelationen im Tuckerschen Modell als allgemeine Trilinearformen interpretierbar. Dagegen erhält man die SUMMAX-Kernmatrix

$$g_{mpq} = t_m w_p h_q$$

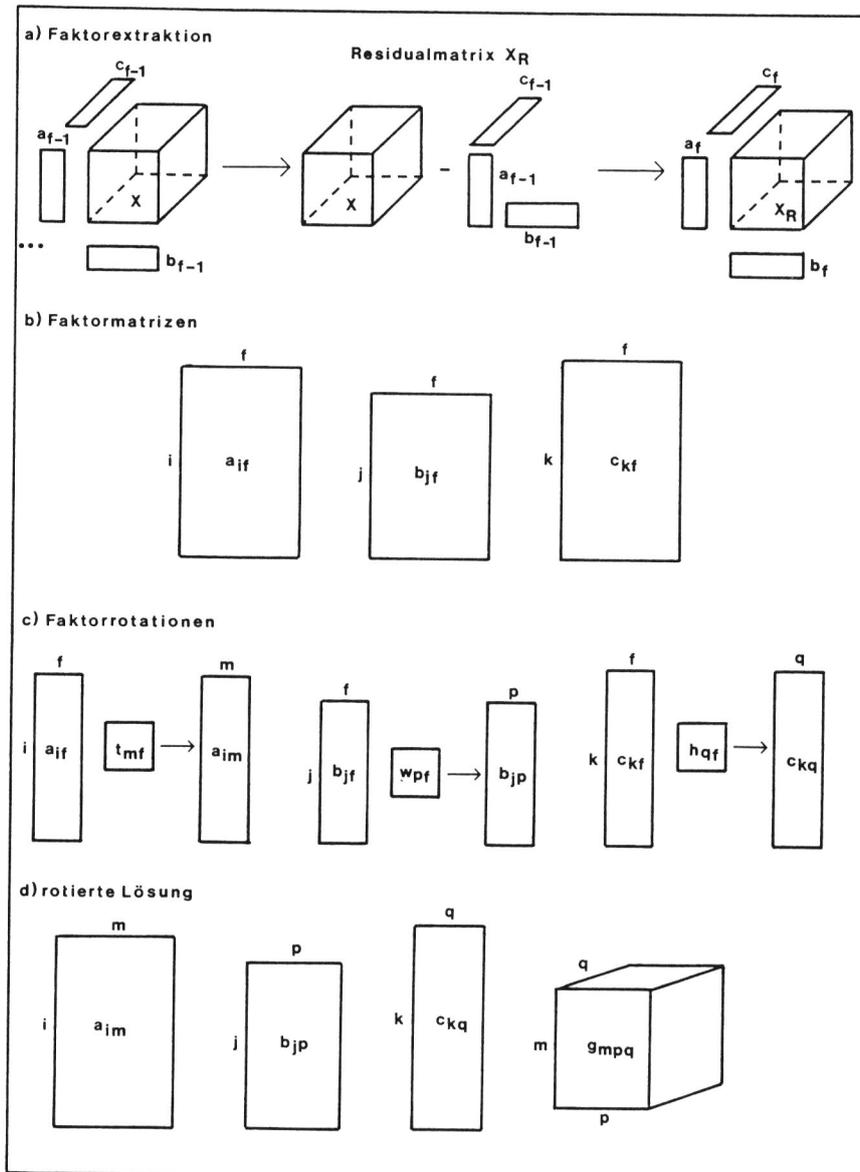
als den Sonderfall der allgemeinen Trilinearform, der als direkte Erweiterung des Skalarproduktes interpretierbar ist (vgl. auch Kowalsky, 1970). Daher impliziert die Definition der SUMMAX-Kernmatrix eine im Vergleich zum Tuckerschen Modell spezifischere Interpretation als Indikator für die Interrelationen der Vektoren  $t$ ,  $w$  und  $h$ .

### 3.3.2. Der SUMMAX-Algorithmus:

Wie auch der ALS-Algorithmus des CANDECOMP/PARAFAC-Modells schätzt der SUMMAX-Algorithmus iterativ aus zwei Eigenpattern die Vorzeichen für das Eigenpattern des dritten Modus. In Gleichung (39) wurde bereits der zyklische Zusammenhang zwischen den Eigenpattern der verschiedenen Modi (Orlik, 1980, S. 194) deutlich, der im SUMMAX-Verfahren zur Extraktion der Faktoren eingesetzt wird:

die Datenmatrix wird im Wechsel mit je zwei Eigenpattern multipliziert bis eine Stabilisierung aufeinander folgender Approximationen eintritt. Im Unterschied zu der ALS-Methode ist die vorhergehende Festsetzung der Anzahl der Faktoren in SUMMAX nicht notwendig. Vielmehr wird jeweils sukzessive ein Faktortripel extrahiert, und anschließend wird die Residualmatrix errechnet, aus der ein neues Tripel bestimmt wird, wie in Abbildung 7 graphisch veranschaulicht ist. Die Anzahl substantieller Faktoren wird mittels des Scree-Tests festgelegt; alternativ kann vor Durchführung der Analyse der Prozentsatz der gewünschten Varianzaufklärung bestimmt werden. Wie in den klassischen faktorenanalytischen Verfahren bleiben die bereits extrahierten Faktoren bei der Gewinnung

Abb.7: Schematische Darstellung der SUMMAX-Methode von Orlik (1980)



weiterer Faktortripel konstant, so daß zum Vergleich von Konfigurationen unterschiedlicher Dimensionalität keine separaten Analysen durchgeführt werden müssen. Im einzelnen durchläuft der SUMMAX-Algorithmus iterativ folgende Schritte zur Extraktion eines Tripels  $f'$  (Kohler, 1980):

- (i) für  $z_{if'1}$  und  $z_{jf'1}$  sind zwei Startkonfigurationen auszuwählen  
(ii) nach  $l$  Iterationen:

$$s_{kl} = \sum_i \sum_j z_{if'1} z_{jf'1} x_{ijkf'}$$

$$z_{kf',l+1} = s_{kl} / |s_{kl}|$$

(iii)  $s_{j1} = \sum_i \sum_k z_{if'1} z_{kf',l+1} x_{ijkf'}$

$$z_{jf',l+1} = s_{j1} / |s_{j1}|$$

(iv)  $s_{i1} = \sum_j \sum_k z_{jf',l+1} z_{kf',l+1} x_{ijkf'}$

$$z_{if',l+1} = s_{i1} / |s_{i1}| .$$

Die Iterationsfolge wird abgebrochen, wenn die Eigenpattern aus zwei Zyklen  $l$  und  $l+1$  übereinstimmen, und die Faktorladungen werden anschließend nach Gleichungen (39) und (40) errechnet. Für die Auswahl einer Startkonfiguration  $z_{if'1}^{(i)}$  wird aus allen  $i$  Teilmatrizen  $jX_k$  diejenige Matrix  $jX_k^{(i)}$  bestimmt, deren Werte vom Betrage die größte Summe ergeben. Das Vorzeichen der Summe aller Werte aus dem Ausdruck

$$\left| jX_k^{(i1)} + jX_k^{(i)} \right| - \left| jX_k^{(i1)} - jX_k^{(i)} \right|$$

ist dann das  $i$ -te Vorzeichen des Eigenpatterns  $z_{if'1}$ . Analog wird das Eigenpattern  $z_{jf'1}$  aus dem Ausdruck

$$\left| z_{if'1} iX_k^{(j1)} + z_{if'1} iX_k^{(j)} \right| - \left| z_{if'1} iX_k^{(j1)} - z_{if'1} iX_k^{(j)} \right|$$

festgesetzt, wobei zur Bestimmung von  $j1$  das Maximum von  $\sum_k \sum_i |z_{if'1} x_{ijk}|$  aufgesucht wird.

- convergenzie unbebed  
- lokale minima.

Auch der SUMMAX-Algorithmus ist mit dem Problem behaftet, daß möglicherweise nur ein lokales Maximum der Funktion (37) aufgefunden wird; nach Orlik (1980) entsprechen die Konvergenzeigenschaften des Verfahrens denen der klassischen Centroid-Analyse; diese sind aber für die dreimodale Generalisierung bisher nicht bewiesen.

### 3.3.3. Multidimensionale Skalierung individueller Differenzen als Sonderfall von SUMMAX:

Analog dem CANDECOMP-Modell läßt sich ein Ansatz zur multidimensionalen Skalierung individueller Differenzen aus der SUMMAX-Grundgleichung herleiten. Werden zwei der drei Modi miteinander identifiziert, so ergibt sich zunächst die Skalarproduktform des Modells

$$(44) \quad x_{jj',i} = \sum_f b_{jf} b_{j',f} a_{if} \quad .$$

Wie auch im Tuckerschen Ansatz und in CANDECOMP wird die Verknüpfung der MDS mit der dreimodalen Faktorenanalyse hergestellt durch Torgersons Transformation der Distanzschätzwerte  $d(j,j')$  in Skalarprodukte  $x_{jj',i}$  für jede Person  $i$ . Die Skalarproduktmatrizen lassen sich nach Gleichung (44) zerlegen in einen Reizraum  $B$  der beurteilten Objekte  $j$  und einen Personenraum  $A$ . Die Koeffizienten  $a_{if}$  sind die personenspezifischen Gewichte, durch welche der gemeinsame Reizraum  $B$  in die individuellen Reizräume

$$(45) \quad y_{jf,i} = a_{if}^{1/2} b_{jf}$$

transformiert wird. Unter der Annahme, daß nur euklidische Distanzschätzwerte  $d_{jj',i}$  betrachtet werden, die jeweils aus den erhobenen Reiz(un)ähnlichkeiten  $\delta_{jj',i}$  einer Person  $i$  durch eine lineare Funktion  $f_i$  hervorgehen, ergibt sich das Distanzmodell

$$(46) \quad d_{jj',i} = \left( \sum_f (y_{jf,i} - y_{j',f,i})^2 \right)^{1/2} \quad .$$

Nach Einsetzen von (45) in (46) resultiert die "weighted

Euclidean Distance"-Sonderform von SUMMAX:

$$(47) \quad d_{jj',i} = \left( \sum_f a_{if} (b_{jf} - b_{j',f})^2 \right)^{1/2} .$$

Die zunächst festgelegte Orientierung der Referenzachsen wird aufgehoben im erweiterten Modell, dessen Skalarproduktform nach Hauptachsentransformation von  ${}_i A_f$  und  ${}_j B_f$  gegeben ist durch

$$(48) \quad {}_i X_{jj'} = {}_i A_m {}_m G_{pp} ({}_j B_p' \times {}_j B_p')$$

$$\text{mit } {}_m G_{pp} = T' I (W \times W).$$

Nach Vertauschung des ersten und dritten Modus, d.h.  ${}_j X_{ji} = {}_j B_p {}_p G_{pm} ({}_j B_p' \times {}_i A_m')$ , kann die Skalarproduktmatrix für eine Person  $i$  dargestellt werden als

$$(49) \quad X_i = {}_j B_p \left( \sum_m a_{im} G_m \right) {}_j B_p' \\ = {}_j B_p H_i {}_j B_p' .$$

In die Matrix  $H_i$  gehen die individuellen Gewichtungen und auch individuelle Rotationen des gemeinsamen Reizraums ein. Da die Kernmatrix  $G_m$  das Skalarprodukt der normierten Transformationsvektoren  $t_m$  und  $w_p$  enthält, zeigen die Werte außerhalb der Diagonalen von  $H_i$  die Winkel zwischen den Reizachsen  $p$  und  $p'$  für Person  $i$  an, und die Diagonalelemente von  $H_i$  sind als Indikatoren differentieller Streckungen/Stauchungen der Reizdimensionen  $p$  interpretierbar.

#### 3.3.4. Die n-modale Erweiterung des SUMMAX-Modells:

Die Grundgleichung des SUMMAX-Modells gestattet die Erweiterung zu einem generellen multilinearen Ansatz für die direkte Faktorisierung n-modaler Datenmatrizen. Dazu seien wie in CANDECOMP/PARAFAC die Modi mit  $1, 2, \dots, n$  und ihre Elemente mit  $i_1, i_2, \dots, i_n$  bezeichnet. Das allgemeine Modell besteht aus der Approximation der beobachteten Werte durch  $n$  Faktormatrizen  $(a_{ijf})$ ,  $j = 1, \dots, n$

$$(50) \quad \hat{x}_{i1 i2 \dots in} = \sum_f a_{i1f} a_{i2f} \dots a_{inf}$$

Dabei werden sukzessive n-Tupel von Eigenpatternen  $f'$  bestimmt, die nach Prä- und Postmultiplikation mit der Datenmatrix die Gesamtsumme aller Werte aus  $X$  bzw. aus der Residualmatrix  $X_f$ , maximieren. Werden die Randsummen der Datenmatrix analog Gleichung (38) mit  $u_{ijf'}$ , bezeichnet

$$(51) \quad \sum_{i1} \sum_{ij'} \sum_{in} z_{i1f'} z_{ij'f'} x_{i1 i2 \dots inf'} = u_{ijf'} \quad ,$$

$j \neq j'$ , so gilt für jeden Vektor  $u_{ijf'}$ ,  $j = 1, \dots, n$ :

$$(52) \quad \sum_{ij} z_{ijf'} u_{ijf'} = s_{f'} \quad .$$

Die Faktorladungen sind im allgemeinen Modell definiert als (P. Orlik, persönliche Mitteilung, 1989):

$$(53) \quad a_{ijf'} = u_{ijf'} s_{f'}^{-(n-1)/n} \quad .$$

Zur Berechnung der Faktormatrizen wird ein Zyklus (ii) bis (iv) des SUMMAX-Algorithmus erweitert auf  $n$  Schritte. Dabei wird im Wechsel aus der Prä- und Postmultiplikation der Datenmatrix mit  $n-1$  Eigenpatternen das  $n$ -te Eigenpattern geschätzt. Die Erweiterung des dreimodalen SUMMAX-Algorithmus zur Analyse von Datenkörpern beliebig vieler Modi ist bisher nicht implementiert.

### 3.3.5. Die Interpretation des dreimodalen SUMMAX-Modells:

Wie bereits an den Grundgleichungen deutlich wurde, umfaßt das SUMMAX-Modell zwei unterscheidbare Zugänge zur Analyse dreimodaler Daten. Wird von der Möglichkeit der Faktorrotationen abgesehen, so werden die Faktormatrizen der drei Modi direkt miteinander verknüpft, und das Modell weist formale Bezüge zu CANDECOMP/PARAFAC auf. Die Grundform des Modells enthält die Annahme von drei konzeptuell gleichartigen Konfigurationen, durch deren Zusammenwirken die Einzelbeobachtungen beschrieben werden können. Im wesentlichen wird eine Menge latenter Einflußgrößen  $f$  postuliert, die den Konfigu-

rationen der drei Modi zugrunde liegt und jeweils ein Tripel zusammengehöriger Konstrukte definiert. Multiple Interaktionen der Konstrukte aus verschiedenen Modi werden in der Grundform des SUMMAX-Modells ausgeschlossen. Daher existieren Zusammenhänge zwischen den Konstrukten, operationalisiert durch hohe Faktorladungen auf jeweils einer Referenzachse, nur innerhalb eines Faktortripels. Ferner besteht analog CANDECOMP/PARAFAC die Möglichkeit, die Achsen in ein oder zwei Modi zu strecken bzw. zu stauchen, wenn im verbleibenden Modus die inverse Operation durchgeführt wird. Dieser Transformation entspricht die Einführung von drei Diagonalmatrizen in Gleichung (35), aus deren Multiplikation die Identitätsmatrix resultiert. Werden die Faktoren der Modi  $j$  und  $k$  auf die Länge eins normiert mit einer reziproken Skalierung in Modus  $i$ , so sind die Varianzbeiträge der Faktoren in Modus  $i$  als Anteil an der Varianzaufklärung der Daten interpretierbar. Darüber hinaus stellen die Koeffizienten  $a_{if}$  in einer normierten Lösung Indikatoren der individuellen Bedeutsamkeit der  $f$ -ten Faktorkombination für das  $i$ -te Element dar.

Während die formale Identität der SUMMAX-Grundform und des CANDECOMP-Modells eine direkte Übertragung des interpretativen Zugangs gestattet, gelten entsprechende inhaltliche Zusammenhänge für das erweiterte SUMMAX-Modell und Tuckers Ansatz nur bedingt. Das erweiterte Modell enthält zunächst in Übereinstimmung mit der Tuckerschen Konzeption eine zusätzliche Kernmatrix, die die Möglichkeit von Faktorrotationen eröffnet. Durch die standardmäßig durchgeführten Hauptachsentransformationen in SUMMAX werden die in allen Modi gleichartig wirksamen Einflußgrößen der Grundform aufgegliedert in drei verschiedene latente Strukturen, die den Modi unterlegt werden. Innerhalb eines Modus können die in den Faktormatrizen enthaltenen Konfigurationen in Übereinstimmung mit der klassischen faktorenanalytischen Sichtweise als "idealisierte Elemente" aufgefaßt werden. Aus einer latenten Referenzgröße entstehen  $m$  idealtypische Personen,  $p$  latente Variablen und  $q$  prototypische Situationen. Ferner

läßt die intermodale Verknüpfung der Faktoren in der Kernmatrix multiple Zusammenhänge zwischen den idealtypischen Elementen der verschiedenen Modi zu. So kann beispielsweise eine latente Variable enge Bezüge zu mehreren prototypischen Situationen aufweisen.

Obwohl das erweiterte SUMMAX-Modell und das Tuckersche Modell hinsichtlich der Implikationen für die Interpretation der Faktormatrizen und der Möglichkeit komplexer Verflechtungen der Faktoren aus den drei Modi übereinstimmen, sind sie aufgrund der unterschiedlichen Definition der Kernmatrix nicht äquivalent. Zwei Kriterien, die zu einer Differenzierung der Kernrelationen zwischen beiden Modellen führen, wurden bereits genannt:

erstens gehen die Varianzbeiträge der SUMMAX-Faktoren in die Faktormatrizen ein, während sie im Tuckerschen Modell mit der Kernmatrix verknüpft werden. Daher stellt die Tuckersche Kernmatrix ein Konglomerat aus verschiedenen Informationen dar. Ferner ergibt sich durch die Tuckersche Konzeption der Faktormatrizen als Eigenvektormatrizen die Notwendigkeit, bereits bei der Faktorextraktion eine Kernmatrix einzuführen. Die Definition der Kernmatrix im SUMMAX-Modell stellt sich in anderer Weise dar, da sie formal die Rotationstransformationen aller Modi verbindet. Die Kernrelationen des Tuckerschen Ansatzes können allgemein als Trilinearformen aufgefaßt werden, wobei die trilineare Funktion  $F(x,y,z)$  nicht in spezifischer Weise eingeschränkt ist und vier verschiedene inhaltliche Zugänge zuläßt. Dagegen ist die der trilinearen Funktion  $F(x,y,z)$  zugeordnete Matrix in SUMMAX die dreimodale Identitätsmatrix, so daß die Kernrelationen das trilineare Analogon des Skalarproduktes darstellen und als Maß für die Interrelation jeweils eines Tripels von Faktoren aufgefaßt werden können. Damit die Kernrelationen in ihrer numerischen Größe vergleichbar sind, führt Orlik (1980) folgende Normierungsvorschrift ein:

es seien die Modi  $i$  und  $j$  zu einem Kombinationsmodus zusammengefaßt. Dann ergeben sich die normierten Kernrelationen

$$\tilde{g}_{mpq} \text{ aus der Kernmatrix } G \text{ nach} \\ (54) \quad \tilde{g}_{mpq} = g_{mpq} / (\sum_f t_{mf}^2 w_{pf}^2)^{1/2}$$

Die normierten Kernrelationen nehmen Werte zwischen -1 und +1 an und geben nach Orlik (1980) die Interrelation zwischen einem Faktor des Modus  $k$  und einem Kombinationsfaktor aus Modus  $i$  und  $j$  wieder. Setzt sich der Kombinationsmodus aus dem kartesischen Produkt von  $j$ ,  $k$  oder  $i$ ,  $k$  zusammen, so ist der Nenner von (54) durch die zugeordneten Transformationsvektoren zu ersetzen. Dabei ist in Abhängigkeit vom untersuchten Gegenstandsbereich festzusetzen, aufgrund von welchem Kombinationsmodus die Normierung durchgeführt werden soll. Eine allgemeine Vorgehensweise besteht darin, dem Modus, der die erhobenen Merkmale, Skalen oder Tests umfaßt, die Kombinationsfaktoren der beiden übrigen Modi gegenüberzustellen (Orlik, 1980, S. 203). Die normierte Kernmatrix enthält in diesem Fall etwa die Zusammenhänge zwischen den latenten Merkmalen und den prototypischen Situationen für die verschiedenen idealtypischen Personen. Alternative Möglichkeiten der Normierung der Kernrelationen werden von Orlik (1980, 1981) diskutiert, wenn die Beschreibung der Konstrukte eines Modus aufgrund der Kombination der beiden weiteren Modi im Mittelpunkt der Analyse steht. Insgesamt impliziert die Definition der Kernmatrix im SUMMAX-Modell unabhängig von der Wahl einer spezifischen Normierung die Interpretation als multilineares Skalarprodukt.

Die Modelle der dreimodalen Faktorenanalyse, die in dem vorliegenden Abschnitt dargestellt wurden, weisen zahlreiche Bezüge in ihren formalen Grundgleichungen und Interpretationsmöglichkeiten auf. Dagegen sind die Grundgleichungen der Modelle mit unterschiedlichen algebraischen Eigenschaften verbunden. Wie im folgenden Abschnitt gezeigt wird, lassen sich die Modelle zwei unterscheidbaren Klassen, der Klasse der trilinearen und der quadrilinearen Ansätze, zuordnen. Mit dieser Klassifikation ist eine Abgrenzung der Modelle hinsichtlich ihrer unterscheidbaren Eigenschaften verknüpft.

4.

Formale Eigenschaften und Spezifika  
der Modelle der dreimodalen Faktor-  
analyse

Die Modelle der dreimodalen Faktorenanalyse, die im vorangehenden Abschnitt dargestellt wurden, können entsprechend der Taxonomie multidimensionaler Methoden der Datenanalyse von Carroll und Arabie (1980) nach der Anzahl der postulierten Konfigurationen oder "Räume" und ihrer Zusammenhänge, nach dem Typus des zugrunde liegenden geometrischen Modells (Distanzmodelle, Skalarproduktmodelle) und nach Restriktionen der Modellparameter ("internal models", Restriktion ausgewählter Parameter, "external models"; in den "external models" werden alle Parameter a priori festgelegt. Es wird dann überprüft, ob die Faktorladungen in allen Modi eine lineare Funktion der festgesetzten Parameter sind. Restriktionen ausgewählter Parameter betreffen nur eine Teilmenge aller Parameter, und in den "internal models" wird keiner der Modellparameter a priori festgelegt) klassifiziert werden. Den Modellen ist gemeinsam, daß sie keine Restriktionen der Parameter enthalten; sie sind als "internal spatial models" zu bezeichnen, die von Carroll und Arabie den Skalarproduktmodellen zugeordnet werden. Auf der Datenebene werden nach der Taxonomie von Carroll und Arabie vollständige, metrische, monadische bzw. dyadische Meßwerte oder Beobachtungen vorausgesetzt. Das Klassifikationskriterium der Anzahl der Konfigurationen und ihrer Zusammenhänge bezieht sich auf die im Modell angenommene Multilinearität und führt zu einer Unterteilung der Modelle in zwei Klassen:

a) trilineare Modelle:

es werden drei verschiedene Konfigurationen angenommen, deren zugeordnete Faktormatrizen direkt miteinander verknüpft werden. In diese Klasse gehören das CANDECOMP-Modell (Carroll & Chang, 1970), PARAFAC (Harshman, 1976) und SUMMAX (Orlik, 1980)

b) quadrilineare Modelle:

die quadrilinearen Modelle enthalten neben den Faktormatrizen eine Kernmatrix, die den Zusammenhang zwischen den drei Konfigurationen beschreibt. Dieser Klasse gehören der von Tucker (1964a) entwickelte Ansatz und das erweiterte SUMMAX-Modell (Orlik, 1980) an.

Das Summax-Modell kann beiden Klassen zugeordnet werden, da

eine trilineare Zerlegung zur Faktorextraktion vorgenommen wird, die durch die anschließende Einführung einer Kernmatrix und der damit verbundenen Möglichkeit der Faktorenrotation in eine quadrilineare Grundgleichung übergeht.

Mit der Klassifikation der dreimodalen Analyseverfahren in trilineare und quadrilineare Modelle sind eine Reihe von Eigenschaften verbunden, die beide Klassen voneinander unterscheiden und innerhalb einer Klasse allen Modellen gemeinsam sind. Grundlegend sind verschiedene Transformationseigenschaften der Faktoren, die Rotationstransformationen nur innerhalb der quadrilinearen Modellklasse umfassen, und verschiedene Eigenschaften der Ränge der Konfigurationen, die orthogonale Faktoren für quadrilineare Zerlegungen zulassen, jedoch schiefwinklige und als Sonderfall linear abhängige Faktoren für trilineare Zerlegungen implizieren.

#### 4.1. Eigenschaften der trilinearen Modelle:

Carroll und Chang (1970) begründen das CANDECOMP-Modell mit der  $n$ -modalen Generalisierung des klassischen Theorems von Eckart und Young (1936), dessen Gegenstand die Zerlegung einer bimodalen Matrix in ihre Eigenwerte und Eigenvektoren ist. Die "Eckart-Young-Zerlegung" führt dabei zu einer Least-Squares-Approximation der bimodalen Matrix  $X$  aus  $r$  Faktoren mit  $r \leq \min(I, J)$ , wobei  $I$  die Anzahl der Zeilen und  $J$  die Anzahl der Spalten von  $X$  ist. D.h. für die Analyse bimodaler Matrizen ist immer gewährleistet, daß die Anzahl der extrahierten Faktoren, die zur Reproduktion der Datenmatrix benötigt wird, höchstens gleich  $\min(I, J)$  ist.

Eine analoge Aussage gilt weder für die Generalisierung der "Eckart-Young-Zerlegung" auf den dreimodalen Fall im CANDECOMP-Modell noch für PARAFAC und SUMMAX. Eine Least-Squares-Approximation einer dreimodalen Datenmatrix  $X$  benötigt eine Anzahl an Faktortripeln  $r$ , die die Anzahl der Zeilen  $I$ , der Spalten  $J$  oder der Lagen  $K$  von  $X$  übersteigt:  $r > \min(I, J, K)$ . Kruskal (1976) weist auf fehlerfreie numerische Beispiele

hin, in denen die Datenmatrix aus  $r > \max(I, J, K)$  Faktortripeln aufgebaut ist. So wurden etwa zur Zerlegung einer Matrix der Ordnung  $8 \times 8 \times 8$  mindestens 11 und höchstens 64 Faktortripel benötigt. Aus den trilinearen Analyseverfahren resultieren "more factors than subjects, tests and treatments" (Kruskal, 1976).

Ergebnisse zu den Rangeigenschaften dreimodaler Matrizen derart, daß wesentlich mehr Faktoren extrahierbar sind als es unter der konventionellen zweimodalen Perspektive möglich erscheint, wurden zunächst von Harshman (1976; vgl. Harshman & Lundy, 1984) beschrieben. Auf dem Hintergrund dieser Spezifika publizierte Kruskal (1976, 1977) die umfassendsten Abhandlungen (Harshman & Lundy, 1984) zu den Eigenschaften dreimodaler Matrizen. Die wichtigsten Theoreme Kruskals zu den Rängen und zugelassenen Transformationen der Faktormatrizen, die für die Diskussion der trilinearen Modelle relevant sind, werden im folgenden dargestellt. Dabei unterscheidet sich das von Kruskal diskutierte Rangkonzept bei dreimodalen Matrizen von der Definition des Ranges einer zweimodalen Matrix; während im zweimodalen Fall der Rang einer Matrix gleich der Anzahl linear unabhängiger Zeilen bzw. Spalten ist, führt Kruskal den Rang einer dreimodalen Matrix auf die Anzahl der Faktortripel zurück, die zur Reproduktion der Matrix notwendig sind. Ferner ist der Rang einer dreimodalen Matrix im Gegensatz zum zweimodalen Fall keine eindeutig bestimmte Größe.

Es sei eine dreimodale  $I \times J \times K$ -Matrix  $X$  mit der Grundgleichung der trilinearen Modelle

$$x_{ijk} = \sum_r a_{ir} b_{jr} c_{kr}$$

gegeben. Der Rang von  $X$  ist definiert als die minimale Anzahl  $r$  von Faktortripeln, die benötigt werden, damit die Gleichung erfüllt ist (Kruskal, 1977). Wie im vorangegangenen Abschnitt wird  $X$  als Supermatrix aufgefaßt, die aus  $i$  nebeneinander geordneten Teilmatrizen  ${}_j X_k$  besteht. Diese Teilmatrizen seien mit  $X_i$ ,  $i = 1, \dots, I$ , bezeichnet. Die

Dimensionalität des Raumes, der durch die Gesamtheit der  $X_i$  erzeugt wird, wird mit  $\dim_1(X)$  bezeichnet. Ist  $\dim_1(X)$  gleich der Anzahl  $I$  der Teilmatrizen, so wird  $X$  1-nondegenerate genannt (Kruskal, 1977).

Im Gegensatz zu dem bimodalen Fall kann der Rang einer dreimodalen Matrix bisher nicht exakt bestimmt werden, vielmehr werden Grenzwerte zur Abschätzung der Ranggröße spezifiziert. Es gelten folgende Aussagen:

$$\max_i \text{Rang}(X_i) \leq \text{Rang}(X) \leq \sum_i \text{Rang}(X_i) \quad ,$$

$$\text{Rang}(X) \leq JK \quad .$$

Der untere Grenzwert für den Rang von  $X$  wird weiter präzisiert in folgendem Korollar:

es sei  $X$  1-nondegenerate und es sei  $T = \left\{ u / u \neq 0 \right\}$ , wobei  $u$  ein Vektor ist. Dann gilt:

$$\text{Rang}(X) \geq \min_{u \text{ in } T} \text{Rang}(uX) + \dim_1(X) - 1 \quad .$$

Wird beispielsweise eine dreimodale Matrix  $X$  der Ordnung  $I \times J \times K$  mit  $I < J < K$  betrachtet und besitzen alle Matrizen  $uX$  vollen Rang, so ist  $X$  in mindestens  $J + I - 1$  Faktortripel und höchstens in  $JK$  Faktortripel zerlegbar. Würde die gleiche Matrix  $X$  als zweimodale Anordnung mit  $I$  Zeilen und  $JK$  Spalten aufgefaßt, so wären höchstens  $I$  Faktoren extrahierbar. Unter Beachtung des Zusammenhangs zwischen  $uX$  und  $[A, B, C]$ :

$$uX = [uA, B, C] = B \text{diag}(uA) C'$$

kann ferner zur Abschätzung des Ausdrucks  $\min \text{Rang}(uX)$  das Rang-Lemma von Kruskal (1976) eingesetzt werden:

es sei  $\text{Rang}(B) = J_0$ ,  $\text{Rang}(C) = K_0$ . Es seien alle  $J_0$  Spalten von  $B$  linear unabhängig, alle  $K_0$  Spalten von  $C$  linear unabhängig und  $I_1 = \min(J_0, K_0, J_0 + K_0 - r)$ . Für eine beliebige Diagonalmatrix  $D$  gilt:

$$\text{Rang}(BDC') \begin{cases} = \text{Rang}(D) & , \quad \text{wenn } \text{Rang}(D) \leq I_1 \\ \geq I_1 & , \quad \text{sonst} \quad . \end{cases}$$

Da der Rang einer dreimodalen Matrix aufgrund der Anzahl der extrahierten Faktortripel definiert ist, enthalten die beschriebenen Rangeigenschaften unmittelbar Aussagen zu den Charakteristika trilinearer Verfahren: trilineare Zerlegungen lassen Faktormatrizen A, B und C zu, die linear abhängige Faktoren enthalten. D.h. ein Faktor  $a_r$  in Modus i läßt sich als Linearkombination weiterer Faktoren aus A darstellen und enthält damit Informationen, die vollständig aus den übrigen Faktoren abgeleitet werden können. Analog gilt die gleiche Eigenschaft für die Faktormatrizen B und C. Als Sonderfall folgt aus Kruskals Befunden, daß ein Faktor in allen drei Modi gleichzeitig linear abhängig sein kann (Harshman & Lundy, 1984). Hierin unterscheiden sich die trilinearen Ansätze wesentlich von den klassischen faktorenanalytischen Verfahren, etwa der Hauptkomponentenanalyse, die die Orthogonalität der extrahierten Faktoren gewährleisten. Erfüllen die Faktoren aus den trilinearen Zerlegungen bereits das Kriterium der linearen Unabhängigkeit nicht, so muß im allgemeinen Fall die noch stärkere Orthogonalitätsbedingung aufgegeben werden; es entstehen Konfigurationen mit obliquen Referenzachsen. Jedoch geht mit der im Vergleich zu der klassischen Hauptkomponentenanalyse suboptimalen Faktorenstruktur (Orlik, 1980) der Vorteil der trilinearen Ansätze einher, daß eine gegebene Datenmatrix durch die Faktormatrizen vollständig zerlegt werden kann; es verbleibt bei genügend hoher Anzahl an Faktortripeln keine "Restvarianz" der Datenmatrix wie im Tuckerschen Modell, welches die gesamte Varianz des Datenkörpers durch die Kernmatrix erfaßt.

ist  
war?

~~~~~

Neben den Rangeigenschaften ist das wichtigste Spezifikum, das die trilinearen Modelle sowohl von den zweimodalen Ansätzen als auch von den quadrilinearen Modellen unterscheidet, die "intrinsic axis property": einzig in den trilinearen Verfahren ist die Orientierung der Faktoren festgelegt,

so daß keine zusätzlichen Faktorrotationen durchgeführt werden können. Carroll und Chang (1970) ziehen aus dieser Eigenschaft die Schlußfolgerung, daß den Faktoren eine unmittelbare psychologische Bedeutung zukommt. Diese Schlußfolgerung ist von mehreren Autoren in Frage gestellt worden (vgl. Ahrens, 1974; Borg, 1981), denn es werden hier aus den rein formalen Eigenschaften eines Verfahrens Aussagen zu der empirischen Validität und Stabilität der Ergebnisse abgeleitet. Ferner wird eingewendet (vgl. Harshman & Lundy, 1984), daß auch klassische Ansätze bedeutsame Grunddimensionen liefern, wenn eine analytische oder hypothesenorientierte Rotation der Faktorenextraktion angeschlossen wird. Harshman und Lundy (1984) diskutieren die begriffliche Interpretation einer CANDECOMP/PARAFAC-Konfiguration zunächst als Arbeitshypothese zu dem untersuchten Gegenstandsbereich; der Nachweis der Stabilität der Ergebnisse ist dabei gesondert zu erbringen, und erst die Replikation der faktorenanalytischen Befunde in weiteren Untersuchungen rechtfertigt die Annahme der Bedeutsamkeit der aufgefundenen Strukturen.

Harshman und Lundy (1984) sehen die Elimination des klassischen "Rotationsproblems" in der simultanen Analyse von I verschiedenen Datenmatrizen der Ordnung  $J \times K$  begründet, während in den zweimodalen Verfahren nur eine einzige Matrix betrachtet wird. Im Vergleich zu den quadrilinearen Modellen enthalten die trilinearen Zerlegungen zusätzliche Restriktionen, indem die Kernmatrix auf das dreimodale Analogon der Identitätsmatrix eingeschränkt wird, und aus diesen Restriktionen folgen die spezifischen Charakteristika der trilinearen Modelle (Kruskal, 1976).

Zur Darstellung der unterschiedlichen Transformationsmöglichkeiten in den klassischen faktorenanalytischen Modellen und in den trilinearen Modellen sei eine zweimodale Matrix  $X$  der Ordnung  $I \times J$  gegeben. Es sei  $X = AB' = \bar{A}\bar{B}'$ , wobei  $A$ ,  $\bar{A}$ ,  $B$  und  $\bar{B}$  jeweils  $r$  Spalten enthalten,  $r \leq \text{Rang}(X)$ . Dann existiert eine reguläre  $r \times r$ -Matrix  $T$ , so daß (Kruskal,

1976)

$$\bar{A} = AT \quad \text{und} \quad \bar{B} = B(T')^{-1} .$$

Bei der Zerlegung einer zweimodalen Matrix besteht insofern eine Unbestimmtheit, als eine beliebige reguläre Transformation der Lösung  $AB'$  die Ausgangsdaten ebenso gut reproduziert. Durch  $T$  wird eine orthogonale oder oblique Rotation der  $r$  Referenzachsen spezifiziert. Dabei wird für praktische Anwendungen zu der allgemeinen Bedingung der Regularität der Rotationsmatrix  $T$  zusätzlich gefordert, daß durch  $T$  eine Einfachstruktur nach den Kriterien von Thurstone (vgl. Revenstorf, 1980) erzeugt wird. Die Einfachstruktur als Rotationskriterium soll eine sinnvolle Interpretation der Faktoren gewährleisten (Pawlik, 1968).

Dagegen besteht in den trilinearen Modellen eine restriktivere Form der Unbestimmtheit der Faktormatrizen  $A, B, C$ : es sei  $X$  eine dreimodale Matrix der Ordnung  $I \times J \times K$  mit  $X = [A, B, C] = [\bar{A}, \bar{B}, \bar{C}]$ . Es enthalte  $\tilde{B}$  eine beliebige Menge von Spalten aus  $B$ , und  $\tilde{C}$  enthalte die korrespondierenden Spalten aus  $C$ . Dann ist  $H_{BC}$  definiert als (analog  $H_{AB}, H_{AC}$ ):

$$H_{BC}(\delta) = \min_{\text{card}(\tilde{B})=\delta} (\text{Rang}(\tilde{B}) + \text{Rang}(\tilde{C}) - \delta).$$

Theorem (Kruskal, 1977):

Es enthalte jede der sechs Matrizen  $A, B, C, \bar{A}, \bar{B}, \bar{C}$   $r$  Spalten. Es sei  $\text{Rang}(B) = J_0, \text{Rang}(C) = K_0$  mit

$$\begin{aligned} H_{AB}(\delta) &\geq \min(\delta, r - K_0 + 2) \\ H_{AC}(\delta) &\geq \min(\delta, r - J_0 + 2) \\ H_{BC}(\delta) &\geq \min(\delta, 1). \end{aligned}$$

Dann existiert eine Permutationsmatrix  $P$  und Diagonalmatrizen  $D, M, N$  so daß

$$\bar{A} = APD, \quad \bar{B} = BPM, \quad \bar{C} = CPN \quad \text{und} \quad DMN = I.$$

Das Theorem läßt beispielsweise eine Konfiguration zu, deren Referenzachsen hinsichtlich ihrer Orientierung eindeutig bestimmt sind und die einen Faktor enthält, der in allen drei Modi aus einer Linearkombination von vier weiteren Faktoren besteht (Harshman & Lundy, 1984).

Der wesentliche Unterschied zu dem zweimodalen Fall besteht darin, daß die trilinearen Modelle nicht eine beliebige reguläre Transformation zulassen, sondern nur eine Permutation der Faktortripel, und zusätzlich können die Referenzachsen eines Modus gestreckt oder gestaucht werden, wenn in den beiden übrigen Modi die inverse Skalierung vorgenommen wird. Harshman und Lundy (1984) bezeichnen diese Form der Unbestimmtheit als trivial, da sie nicht zu Unterschieden in der Interpretation einer gegebenen Lösung führt.

Im vorangegangenen Abschnitt wurden bereits einige weitere Spezifika der trilinearen Modelle diskutiert, die zu Implikationen in der Interpretation der Faktormatrizen führen. Abschließend seien die Eigenschaften, die die Modelle der trilinearen Klasse gemeinsam haben und von den quadrilinearen Verfahren unterscheiden, zusammengefaßt:

- a) die Zahl der extrahierten Faktoren stimmt zwischen den Modi überein
- b) es wird eine Menge latenter Einflußgrößen  $r$  identifiziert, die allen Modi zugrunde liegen und zusammengehörige Konstrukte definieren. Dabei stellen die Faktorladungen eines ausgewählten Modus Indikatoren der Bedeutsamkeit der latenten Einflußgrößen dar
- c) multiple Interaktionen der Faktoren aus verschiedenen Modi werden ausgeschlossen; Zusammenhänge zwischen den Modi können nur innerhalb eines Faktortripels hergeleitet werden
- d) trilineare Verfahren liefern schiefwinklige und linear abhängige Faktoren; die Zahl der extrahierbaren Faktoren ist größer als die Zahl der Entitäten eines Modus
- e) die Orientierung der Faktoren ist in allen Modi festgelegt, so daß keine zusätzlichen Faktorrotationen durchge-

führt werden können.

#### 4.2. Eigenschaften der quadrilinearen Modelle:

Aus dem Einsatz der Hauptachsentransformation, im Tuckerschen Ansatz zur Faktorextraktion und im erweiterten SUMMAX-Modell zur Faktorrotation, ergeben sich Eigenschaften der quadrilinearen Klasse, die enge Bezüge zu den Charakteristika der klassischen zweimodalen Verfahren aufweisen. Wie im vorangegangenen Abschnitt bereits dargestellt wurde, wird in den Grundgleichungen des Tuckerschen Modells von den Skalarproduktmatrizen des Datenkörpers,  $i^{X_{jk}} i^{X_{jk}'}$ ,  $j^{X_{ik}} j^{X_{ik}'}$  und  $k^{X_{ij}} k^{X_{ij}'}$ , ausgegangen, deren Eigenvektoren die Faktoren der drei Modi sind. Auch im erweiterten SUMMAX-Modell wird standardmäßig die Hauptachsentransformation eingesetzt, indem die Faktormatrizen der SUMMAX-Grundform in jedem Modus auf Hauptkomponenten rotiert werden. Unabhängig von der Wahl der zugrunde liegenden Datenmatrizen teilen beide Modelle Eigenschaften der "Eckart-Young-Zerlegung" bimodaler Datensätze: sie führen zu wechselseitig orthogonalen Faktoren, und die Anzahl der Faktoren eines Modus übersteigt die Zahl der Entitäten des Modus nicht. Aus der Bestimmung der Skalarprodukte eines dreimodalen Datenkörpers der Ordnung  $I \times J \times K$  resultieren drei symmetrische Matrizen mit  $\text{Rang}(i^{X_{jk}} i^{X_{jk}'}) \leq I$ ,  $\text{Rang}(j^{X_{ik}} j^{X_{ik}'}) \leq J$  und  $\text{Rang}(k^{X_{ij}} k^{X_{ij}'}) \leq K$ . Da die Anzahl der Eigenvektoren, deren Eigenwerte ungleich null sind, in jedem Modus mit dem Rang der Skalarproduktmatrix identisch ist, gilt für die Dimensionalität der drei Konfigurationen im Tuckerschen Modell:

$$\dim(\bar{A}) = m \leq I, \quad \dim(\bar{B}) = p \leq J \quad \text{und} \quad \dim(\bar{C}) = q \leq K.$$

Analog enthalten die Faktormatrizen des erweiterten SUMMAX-Modells  $m$ ,  $p$  und  $q$  Faktoren mit

$$\dim(A) = m \leq \min(I, f), \quad \dim(B) = p \leq \min(J, f) \quad \text{und} \\ \dim(C) = q \leq \min(K, f),$$

wobei  $f$  die Anzahl der unrotierten Faktortripel bezeichnet. Umfassen die unrotierten Faktormatrizen nach SUMMAX linear

abhängige Faktoren, so führt die Hauptachsentransformation zu einer Reduktion der Anzahl der Referenzachsen. Die schiefwinkligen bzw. linear abhängigen Faktoren der SUMMAX-Grundform werden durch orthogonale Bezugsachsen ersetzt, die eine ökonomischere Beschreibung der Konfigurationen in den drei Modi gestatten. Ebenso werden im Tuckerschen Modell die "idealisierten Elemente" eines Modus durch wechselseitig orthogonale Faktoren beschrieben. Im Unterschied zu SUMMAX sind die Varianzbeiträge der Faktoren nicht mit den Eigenwerten der Skalarproduktmatrizen identisch, sondern werden auf den Betrag eins normiert, während die Eigenwerte mit der Kernmatrix verknüpft werden.

Obwohl beide quadrilinearen Verfahren in Übereinstimmung mit der "Eckart-Young-Zerlegung" gewährleisten, daß die Faktoren ein orthogonales Bezugssystem in jedem Modus aufspannen und ihre Anzahl nicht größer ist als die Zahl der Zeilen, Spalten oder Lagen des zugrunde liegenden Datenkörpers, existiert keine den klassischen Verfahren entsprechende Aussage zu der Gleichheit der Dimensionalität der drei Konfigurationen. Die quadrilinearen Modelle führen im allgemeinen nicht zu einer identischen Anzahl an Faktoren in allen Modi. Vielmehr wird jedem Modus eine andere dimensionale Struktur unterlegt, so daß eine latente Einflußgröße  $f$  der SUMMAX-Grundform in  $m$  idealtypische Personen,  $p$  latente Variablen und  $q$  prototypische Situationen aufgegliedert wird. Dabei ist die Variation der Zahl der SUMMAX-Faktoren,  $m$ ,  $p$  und  $q$ , verknüpft mit der Struktur der extrahierten Faktormatrizen  ${}_iA_f$ ,  ${}_jB_f$  und  ${}_kC_f$ . In der Grundform des Modells wird zwar für jeden Modus eine übereinstimmende Zahl an Referenzachsen vorausgesetzt, woraus jedoch keine Entsprechungen in der Anzahl der unabhängigen Faktoren folgen. D.h. ist ein Faktor in einem Modus darstellbar als Linearkombination weiterer Faktoren, so kann der gleiche Faktor in den übrigen Modi linear unabhängig sein. Ferner gestattet die Bestimmung des Ranges einer dreimodalen Matrix keine Rückschlüsse auf die Ränge der zugeordneten Skalarprodukte. Während der Übergang von der Datenmatrix  $X$  zu den Skalarpro-

duktmatrizen  $XX'$  und  $X'X$  in den zweimodalen Verfahren keine Rangänderungen induziert, weisen dreimodale Anordnungen keine derartigen Invarianzeigenschaften auf.

Quadrilineare Modelle umfassen die Möglichkeit von Faktorrotationen und sind aufgrund dieser Eigenschaft als direkte Erweiterung der "Eckart-Young-Zerlegung" interpretiert worden (Kroonenberg, 1983b; Orlik, 1980). Zur Darstellung der Analogie zwischen den klassischen Verfahren und den quadrilinearen Zerlegungen sei eine zweimodale Matrix  $X$  mit  $X = AB'$  und den rotierten Faktormatrizen  $AT_1 = A_1$ ,  $BT_2 = B_2$  gegeben. Die inversen Transformationsmatrizen sind in der Form  $G = (T_1)^{-1} (T_2')^{-1}$  darstellbar mit der Modellgleichung  $X = A_1 G B_2'$ . In den quadrilinearen Modellen wird die zweimodale Modellgleichung generalisiert durch drei rotierte Matrizen der Form

$${}_i A_m {}_m T_{m1} = {}_i A_{m1}, \quad {}_j B_p {}_p T_{p1} = {}_j B_{p1} \quad \text{und} \quad {}_k C_q {}_q T_{q1} = {}_k C_{q1}.$$

Aus der Verknüpfung der Rotationsmatrizen mit der Kernmatrix

$${}_m G_{p1q1} = ({}_m T_{m1})^{-1} {}_m G_{pq} (({}_p T_{p1}')^{-1} \times ({}_q T_{q1}')^{-1})$$

ergibt sich eine der bimodalen Modellgleichung entsprechende Repräsentation mit

$$X = {}_i A_{m1} {}_m G_{p1q1} ({}_j B_{p1}' \times {}_k C_{q1}').$$

Analog den zweimodalen Verfahren sind beliebige reguläre Transformationen  ${}_m T_{m1}$ ,  ${}_p T_{p1}$  und  ${}_q T_{q1}$  zugelassen, so daß orthogonale wie auch oblique Rotationen der Faktormatrizen durchgeführt werden können.

Insgesamt folgen aus der zusätzlichen Definition der Kernmatrix in den quadrilinearen Modellen eine Reihe von Eigenschaften, die sich von den Charakteristika trilinearzerlegungen grundlegend unterscheiden. Insbesondere wird die "intrinsic axis property" aufgehoben und linear abhängige

Faktoren werden eliminiert, während die Anzahl der Faktoren zwischen den Modi variiert und Faktorkombinationen der Form  $a_m b_p c_q$  eingeführt werden. Dagegen sind die quadrilinearen Modelle den trilinearen Verfahren äquivalent, wenn die Kernmatrix auf das dreimodale Analogon einer Identitätsmatrix mit

$$g_{mpq} \begin{cases} = 1 & \text{für } m = p = q \\ = 0 & \text{sonst} \end{cases}$$

*λ is owl  
mym  
owl goal*

eingeschränkt werden kann bzw. wenn die Kernmatrix in allen  $q$  Lagen simultan diagonalisierbar ist, d.h.  $g_{mpq} = 0$  für  $m \neq p$  (Carroll & Chang, 1970; Harshman & Lundy, 1984). Daher wurde zur Analyse des Zusammenhangs beider Modellklassen versucht, unter Anwendung des ALS-Algorithmus die Kernmatrix des Tuckerschen Modells in eine Diagonalform zu transformieren.

Allerdings gestattet diese Vorgehensweise keine allgemeinen Aussagen zu den Bezügen der verschiedenen dreimodalen Modelle, denn es werden sehr restriktive Bedingungen bezüglich der Eigenschaften der Datenmatrix vorausgesetzt. Harshman und Lundy (1984) interpretieren eine Konfiguration nach dem Tuckerschen Modell bereits dann als eine näherungsweise CANDECOMP/PARAFAC-Repräsentation, wenn die Kernrelationen außerhalb der Diagonalen im Vergleich zu den Diagonalzellen der Kernmatrix relativ gering sind. Dabei konnte in einer Reihe empirischer Datenanalysen keine nur approximative Diagonalisierung der Kernmatrix erreicht werden (Kroonenberg, 1983b). Die Schwierigkeiten und Restriktionen, die mit den Verfahren zur Diagonalisierung der Kernmatrix verknüpft sind, werden in der Diskussion erörtert.

Ogleich die Anwendungen verschiedener ALS-Verfahren auf das Tuckersche Modell nicht sehr erfolgreich waren, hat sich diese Vorgehensweise zur Herleitung des Zusammenhangs zwischen den dargestellten Modellen und zur besseren Interpretierbarkeit der abstrakten Kernrelationen durchgesetzt (vgl.

Kroonenberg, 1983b; Harshman & Lundy, 1984). So bemerkt Tucker (1972, S. 26) zu seinem Modell: "while ... the models used by Carroll and Chang (1970) appear to be special cases of the present model, the precise relations deserve further study". Hinsichtlich der Problematik, die mit der Anwendung der ALS-Algorithmen auf quadrilineare Modelle einhergeht, sind die präzisen Relationen bisher nicht aufgezeigt worden. Dagegen existieren allgemeine Zusammenhänge, die keine spezifischen Annahmen zu der Struktur der Kernmatrix implizieren, auf der Grundlage des SUMMAX-Modells. Diese generellen Bezüge, die die Ableitung der quadrilinearen Ansätze aus den trilinearen Modellen ohne notwendige Restriktionen der Kernmatrizen umfassen, sind im folgenden Abschnitt dargestellt.

## 5.

Zusammenhänge und theoretische Bezüge zwischen den Modellen der dreimodalen Faktorenanalyse

Gegenstand des vorliegenden Abschnitts ist die Analyse der Zusammenhänge zwischen den Modellen der dreimodalen Faktorenanalyse. Obwohl sich die trilinearen Modelle von den quadrilinearen in ihren Eigenschaften grundlegend unterscheiden und daher enge Verbindungen nicht offensichtlich sind, können allgemeine Bezüge der beiden Klassen auf der Grundlage des SUMMAX-Modells abgeleitet werden. Es gibt bislang zwar kein übergeordnetes Modell, welches durch Restriktion spezifischer Parameter die betrachteten Verfahren als Sonderfälle umfaßt. Jedoch ergeben sich unter bestimmten Transformationen sowohl CANDECOMP/PARAFAC als auch der Tuckersche Ansatz aus den beiden Grundformen des SUMMAX-Modells. Wie im folgenden gezeigt wird, ist SUMMAX mit den übrigen Verfahren der trilinearen Klasse durch Permutationen und Reskalierungen der Faktoren verknüpft. Darüber hinaus gestattet die Betrachtung orthogonaler Rotationstransformationen der Faktormatrizen bzw. der Kernmatrix die Überführung des erweiterten SUMMAX-Modells in die Tuckersche dreimodale Hauptkomponentenanalyse.

### 5.1. Der Bezug des SUMMAX-Modells zu CANDECOMP/PARAFAC:

Im folgenden wird von der CANDECOMP/PARAFAC-Modellgleichung

$$iX_{jk} = i\bar{A}_r r I_{rr} (j\bar{B}_r' \times k\bar{C}_r') = [\bar{A}, \bar{B}, \bar{C}]$$

und dem trilinearen SUMMAX-Modell mit

$$iX_{jk} = iA_f f I_{ff} (jB_f' \times kC_f') = [A, B, C]$$

ausgegangen. Es wird eine dreimodale Matrix  $X$  zugrunde gelegt, die nicht gleich der Nullmatrix ist, und es wird vorausgesetzt, daß  $\dim_1(X) > 1$  und  $\dim_1(X_f) > 1$ .

Die Repräsentation  $X = [\bar{A}, \bar{B}, \bar{C}]$  besitze einen minimalen Rang  $r$ , so daß  $r = \text{Rang}(X)$ .

Theorem 1:

Es sei  $X = [\bar{A}, \bar{B}, \bar{C}]$  das Tripelprodukt aus  $\bar{A}$ ,  $\bar{B}$  und  $\bar{C}$ , das einen minimalen Rang besitzt, so daß  $r = \text{Rang}(X)$ . Dann gilt für den Zusammenhang zwischen  $[\bar{A}, \bar{B}, \bar{C}]$  und  $[A, B, C]$ :  $r = f$ , und es existieren drei Permutationsmatrizen  $P_A$ ,  $P_B$ ,  $P_C$  und nicht-singuläre Diagonalmatrizen  $D$ ,  $M$ ,  $N$ , so daß

$$(57) \quad \bar{A} = A P_A D, \quad \bar{B} = B P_B M, \quad \bar{C} = C P_C N \quad \text{mit}$$

$$(58) \quad [P_A D, P_B M, P_C N] = I,$$

wobei  $I$  die Identitätsmatrix bezeichnet.

D.h. die Faktormatrizen eines jeden Modus haben die gleichen Spalten bis auf eine Permutation und Multiplikation mit Skalaren ungleich null. Die Anzahl der Faktoren stimmt überein, und die Wirkungen der Permutations- und Diagonalmatrizen heben sich wechselseitig auf.

5.1.1. Beweis des Zusammenhangs zwischen SUMMAX und CANDECOMP/PARAFAC:

Es sei  $\text{Rang}(A) = I_0$ ,  $\text{Rang}(\bar{A}) = \bar{I}_0$ . Es bezeichne  $w(y)$  die Anzahl der Komponenten eines beliebigen Vektors  $y$ , die ungleich null sind. Es sei  $\text{col}(A)$  der lineare Teilraum des  $I$ -dimensionalen Vektorraums  $\mathbb{R}^I$ , der durch die Spalten von  $A$  aufgespannt wird. Das orthogonale Komplement von  $\text{col}(A)$  sei mit  $\text{Kern}(A)$  bezeichnet.

Grundlegend für den Beweis des Zusammenhangs zwischen SUMMAX und CANDECOMP/PARAFAC ist das Permutationslemma (Kruskal, 1977, S.133), das folgende Aussage enthält:

es seien zwei Matrizen  $A$  und  $\bar{A}$  der Ordnung  $I \times r$ ,  $I \times \bar{r}$  mit  $r \geq \bar{r}$  gegeben.  $A$  enthalte keine Spalten, die vollständig aus Nullen bestehen. Für einen Vektor  $x$  mit

$$w(x\bar{A}) \leq \bar{r} - \text{Rang}(\bar{A}) + 1$$

gelte:

$$w(xA) \leq w(x\bar{A}).$$

Dann folgt:

$r = \bar{r}$ , und es existiert eine Permutationsmatrix  $P_A$  und eine nicht-singuläre Diagonalmatrix, so daß

$$(59) \quad \bar{A} = A P_A D.$$

Ist zusätzlich jedes Paar von Spalten aus  $A$  linear unabhängig, so sind  $P_A$  und  $D$  eindeutig bestimmt.

Es wird gezeigt, daß die Voraussetzungen des Permutationslemmas erfüllt sind, so daß  $A$ ,  $\bar{A}$  bzw.  $B$ ,  $\bar{B}$  und  $C$ ,  $\bar{C}$  durch Permutationen und Multiplikationen mit Skalaren auseinander hervorgehen.

Nach Voraussetzung besitzt die Repräsentation  $[\bar{A}, \bar{B}, \bar{C}]$  einen minimalen Rang  $r$ . Da  $f$  nicht kleiner als  $r$  sein kann und  $r$  die Anzahl der Spalten in  $\bar{A}$  ist, ist  $f \geq r$ .

Es enthalte  $A$  eine Spalte  $a_{if'}$ , die vollständig aus Nullen besteht. Aus

$$(61) \quad a_{if'} = u_{if'} s_{f'}^{-2/3} = 0$$

$$= \left( \sum_j \sum_k z_{jf'} z_{kf'} x_{ijkf'} \right) \left( \sum_i \sum_j \sum_k z_{if'} z_{jf'} z_{kf'} x_{ijkf'} \right)^{-2/3}$$

und der Definition der Vektoren  $z_{if'}$ ,  $z_{jf'}$ ,  $z_{kf'}$  folgt: die Elemente  $x_{ijkf'}$  sind bis auf ihr Vorzeichen identisch oder  $X$  ist die Nullmatrix. Die Nullmatrix ist nach Voraussetzung ausgeschlossen, und aus  $\dim_1(X) > 1$  folgt, daß es mindestens zwei linear unabhängige  $X_i$  und  $X_i$  geben muß. Also enthält  $A$  keine Spalten, die vollständig aus Nullen bestehen.

Um zu zeigen, daß auch die dritte Bedingung des Permutationslemmas erfüllt ist, wird von der Zerlegung des Vektorraums  $\mathbb{R}^I$  in  $\text{col}(A)$  und  $\text{Kern}(A)$  ausgegangen. Für diese Zerlegung gelten folgende Aussagen (Lamprecht, 1980, Kapitel I; Koecher, 1985, Kapitel 1, Paragraph 8):

der Vektorraum  $\mathbb{R}^I$  mit der Dimension  $I$  ist die direkte Summe aus  $\text{col}(A)$  und  $\text{Kern}(A)$  mit  $I = \dim \text{Kern}(A) + \dim \text{col}(A)$ .

Mit der analogen Zerlegung in  $\text{col}(\bar{A})$  und  $\text{Kern}(\bar{A})$  ist zu zeigen, daß alle Spalten von  $A$  und ihre Linearkombinationen in  $\text{col}(\bar{A})$  enthalten sind.

Für den (trivialen) Fall, daß  $\text{Kern}(\bar{A}) = \text{Kern}(A) = 0$ , folgt sofort:

wenn  $x \in \text{Kern}(\bar{A})$ , dann  $w(x\bar{A}) = w(xA) = 0$ . In diesem Fall stimmen  $\text{col}(\bar{A})$  und  $\text{col}(A)$  überein mit  $\dim \text{col}(\bar{A}) = \dim \text{col}(A) = I$ , und es sind alle Spalten aus  $A$  und ihre Linearkombinationen in  $\text{col}(\bar{A})$  enthalten.

Für den (nicht trivialen) Fall sei  $\dim \text{col}(\bar{A}) < I$  und  $\dim \text{col}(A) < I$ . Daraus folgt für die orthogonalen Komplemente, daß  $\dim \text{Kern}(\bar{A}) = I - \bar{I}_0 > 0$  und  $\dim \text{Kern}(A) = I - I_0 > 0$ . D.h. es gibt mindestens einen Vektor  $x$  aus  $\text{Kern}(\bar{A})$  mit  $x \neq 0$  und mindestens einen Vektor  $y$  aus  $\text{Kern}(A)$  mit  $y \neq 0$ . Ohne Beschränkung der Allgemeinheit kann vorausgesetzt werden, daß nicht alle Spalten aus  $A$  und  $\bar{A}$  wechselseitig orthogonal zueinander sind. Denn sonst kann eine beliebige Spalte  $a_i$  aus  $A$  gewählt werden, so daß

$$[a_i A, B, C] = B \text{diag}(a_i A) C' = a_i X \neq 0 \quad \text{und}$$

$$a_i X = [a_i \bar{A}, \bar{B}, \bar{C}] = \bar{B} \text{diag}(a_i \bar{A}) \bar{C}' = 0 \quad ,$$

was einen Widerspruch erzeugt. Daher muß mindestens eine Spalte  $a_i$  aus  $A$  existieren, die zu mindestens einer Spalte  $\bar{a}_j$  aus  $\bar{A}$  nicht orthogonal ist. Dann aber folgt, daß  $a_i \bar{A} \neq 0$ . D.h.  $a_i$  ist nicht in  $\text{Kern}(\bar{A})$  enthalten, also muß  $a_i$  ein Element aus  $\text{col}(\bar{A})$  sein. Die Schnittmenge von  $\text{col}(A)$  und  $\text{col}(\bar{A})$  ist nicht leer und  $\dim(\text{col}(A) \cap \text{col}(\bar{A})) \geq 1$ , da  $a_i$  nicht der Nullvektor ist.

In diesem Fall gilt für den Orthogonalraum der Schnittmenge:

$$(62) \quad \dim (\operatorname{col}(A) \cap \operatorname{col}(\bar{A}))^\perp < I \quad \text{und}$$

$$(63) \quad (\operatorname{col}(A) \cap \operatorname{col}(\bar{A}))^\perp = \operatorname{Kern}(A) + \operatorname{Kern}(\bar{A}) \quad .$$

Ist die Schnittmenge von  $\operatorname{Kern}(A)$  und  $\operatorname{Kern}(\bar{A})$  der Nullraum, dann folgt (Lamprecht, 1978, Satz 5.16, S. 108):

$$(64) \quad \begin{aligned} \dim (\operatorname{col}(A) \cap \operatorname{col}(\bar{A}))^\perp &= \dim \operatorname{Kern}(A) + \dim \operatorname{Kern}(\bar{A}) \\ &= 2I - (I_0 + \bar{I}_0) \\ \text{und} \quad I_0 + \bar{I}_0 &> I \quad . \end{aligned}$$

D.h. der gesamte Vektorraum  $\mathbb{R}^I$  kann nicht die direkte Summe aus  $\operatorname{Kern}(A)$  und  $\operatorname{Kern}(\bar{A})$  sein, und  $\operatorname{Kern}(A)$  muß mindestens einen Vektor  $x_0 \neq 0$  enthalten, der auch in  $\operatorname{Kern}(\bar{A})$  enthalten ist. Damit sind mindestens zwei Vektoren,  $x_0$  und der Nullvektor  $0$  gefunden, die Elemente von beiden orthogonalen Komplementen sind.

Aus der Basis von  $\operatorname{Kern}(\bar{A})$  können  $I - \bar{I}_0$  Vektoren  $x_i$  gewählt werden, die nach Hinzufügen eines  $(I - \bar{I}_0 + 1)$ -ten Vektors aus  $\operatorname{Kern}(\bar{A})$  den Nullvektor erzeugen mit

$$(65) \quad 0 = \sum_i d_i x_i, \quad d_i \neq 0, \quad i = 1, \dots, I - \bar{I}_0 + 1.$$

Diese Linearkombination ist auch in  $\operatorname{Kern}(A)$  enthalten, wobei (Lamprecht, 1978, S.117, II.6.3b - II.6.3d) nach Einsetzen von (65):

$$(66) \quad \begin{aligned} 0 A &= \left( \sum_i d_i x_i \right) A = \sum_i d_i x_i A = \sum_i d_i (x_i A) \\ &= \sum_i d_i z_i = 0. \end{aligned}$$

Entweder existieren auch Linearkombinationen der Ausdrücke  $x_i A$  mit  $x_i A \neq 0$ , die den Nullvektor erzeugen, oder für alle  $x_i A$  gilt, daß  $x_i A = 0$ . Im zweiten Fall folgt wegen  $x_i A = 0$ , daß alle Vektoren  $x_i$  in  $\operatorname{Kern}(A)$  enthalten sind. Dann sind auch die  $I - \bar{I}_0$  Basisvektoren in  $\operatorname{Kern}(A)$  enthalten sowie alle Linearkombinationen der Basisvektoren, und es folgt:

$$(67) \quad \text{Kern}(\bar{A}) \subset \text{Kern}(A) \quad ,$$

$$(68) \quad \text{col}(A) \subset \text{col}(\bar{A}) \quad \text{und} \quad I_0 \cong \bar{I}_0 \quad .$$

D.h. in diesem Fall sind alle Spalten von  $A$  und ihre Linearkombinationen in  $\text{col}(\bar{A})$  enthalten.

Wenn aber  $x_i A \neq 0$  gilt, dann enthält die Lösungsmenge von  $\sum_i d_i z_i = 0$  sicherlich die bereits gefundene Lösung  $d = (d_1, d_2, \dots, d_i, \dots)$ , und die Basisvektoren  $x_i$  aus  $\text{Kern}(\bar{A})$  sind in  $\text{col}(A)$  enthalten. Hieraus folgt für den Vektor  $x_0$  aus der Schnittmenge von  $\text{Kern}(A)$  und  $\text{Kern}(\bar{A})$ :

entweder  $x_0$  ist selbst Basisvektor. Dann muß einer der Vektoren  $x_i$  aus  $\text{Kern}(\bar{A})$ , die in  $\text{col}(A)$  enthalten sind, identisch mit  $x_0$  sein. Wegen der Abgeschlossenheit der linearen Teilräume  $\text{Kern}(A)$  und  $\text{col}(A)$  ist dann  $x_0$  in  $\text{col}(A)$  enthalten, wodurch ein Widerspruch erzeugt wird.

Ist  $x_0$  kein Basisvektor, dann muß  $x_0$  eine Linearkombination der Basisvektoren  $x_i$  aus  $\text{Kern}(\bar{A})$  sein, die in  $\text{col}(A)$  enthalten sind. In diesem Fall gibt es Skalare  $m_i$ , so daß  $\sum_i m_i x_i = x_0$ . Wegen der Abgeschlossenheit von  $\text{Kern}(A)$  und  $\text{col}(A)$  muß  $x_0$  auch in diesem Fall Element von  $\text{col}(A)$  sein. Es ergibt sich erneut ein Widerspruch. Also folgt für die Basisvektoren  $x_i$  aus  $\text{Kern}(\bar{A})$ , daß sie auch in  $\text{Kern}(A)$  enthalten sind. Dann sind aber auch alle Linearkombinationen der  $x_i$  in  $\text{Kern}(A)$ , und es gilt wie in (67) und (68):

$$\text{Kern}(\bar{A}) \subset \text{Kern}(A) \quad ,$$

$$\text{col}(A) \subset \text{col}(\bar{A}) \quad \text{und} \quad I_0 \cong \bar{I}_0 .$$

D.h. für einen beliebigen Vektor  $x$  mit  $x\bar{A} = 0$  und  $w(x\bar{A}) = 0$  folgt:

$$(69) \quad xA = 0 \quad \text{und} \quad w(xA) = 0 .$$

Wegen  $\dim \text{col}(\bar{A}) = \bar{I}_0$  gibt es  $\bar{I}_0$  Spalten aus  $\bar{A}$ , die eine Basis von  $\text{col}(\bar{A})$  bilden. Unter Beachtung, daß  $\text{col}(A) \subset \text{col}(\bar{A})$ , ist jede Spalte  $a_j$  aus  $A$ ,  $j = 1, \dots, f$ , darstellbar

als Linearkombination

$$(70) \quad a_j = \sum_i m_i \bar{a}_i, \quad ,$$

wobei  $m_i \neq 0$  Skalare sind. Ist eine Spalte  $a_j$  aus  $A$  eine Linearkombination aus  $\bar{a}_i$  mit  $x \bar{a}_i = 0$ , dann folgt sofort:

$$(71) \quad x a_j = 0.$$

Existiert ein beliebiges  $a_j$  mit  $x a_j \neq 0$ , dann folgt wegen  $x a_j = x(\sum_i m_i \bar{a}_i) \neq 0$ :  $x$  ist nicht aus  $\text{Kern}(\bar{A})$ , also ist  $x$  aus  $\text{col}(\bar{A})$ . Dann muß es aber zu einem  $a_j$  mindestens ein  $\bar{a}_i$  geben mit  $x \bar{a}_i \neq 0$ . D.h.

$$(72) \quad \text{Rang } \text{diag}(x\bar{A}) \geq \text{Rang } \text{diag}(xA).$$

Wegen  $\text{Rang } \text{diag}(x\bar{A}) = w(x\bar{A})$  (Kruskal, 1977) folgt:

$$w(x\bar{A}) \geq w(xA).$$

Damit sind alle Voraussetzungen des Permutationslemmas erfüllt, und es gilt die Aussage des Lemmas in Gleichung (59):

$$f = r, \quad \bar{A} = A P_A D,$$

wobei  $P_A$  eine Permutationsmatrix und  $D$  eine nicht-singuläre Diagonalmatrix bezeichnet. In analoger Weise können die beiden übrigen Aussagen in Theorem 1 gezeigt werden:

$$(57') \quad \bar{B} = B P_B M \quad \text{und} \quad \bar{C} = C P_C N.$$

Aus  $X = [A, B, C] = [\bar{A}, \bar{B}, \bar{C}]$  und Einsetzen von (59) und (57') folgt:

$$\begin{aligned} X &= A P_A D I ((M P_B' B') x (N P_C' C')) \\ &= A P_A D I (M x N)(P_B' x P_C')(B' x C') \\ &= A I (B' x C') \end{aligned}$$

Also folgt der in Theorem 1 behauptete Zusammenhang:

$$(58) \quad P_A D I (M \times N) (P_B' \times P_C') = [P_A D, P_B M, P_C N] = I \quad .$$

### 5.1.2. Implikationen des Bezuges des SUMMAX-Modells zu CANDECOMP/PARAFAC:

CANDECOMP/PARAFAC teilt nicht nur eine Reihe von formalen Eigenschaften mit SUMMAX, sondern ist dem trilinearen SUMMAX-Modell direkt äquivalent; werden die nach der Grundform des SUMMAX-Modells extrahierten Faktoren entsprechend dem CANDECOMP/PARAFAC-Rationale skaliert bzw. angeordnet, so resultieren identische Konfigurationen in jedem Modus. Dieser enge Zusammenhang impliziert ferner, daß auch das INDSCAL-Modell aus der SUMMAX-Grundgleichung zur multidimensionalen Skalierung individueller Differenzen ableitbar ist.

*folgt  
daraus  
wohl?*

Werden zusätzlich personenspezifische Rotationstransformationen des Objektraums aus INDSCAL einbezogen, so umfaßt der SUMMAX-Ansatz zusätzlich das IDIOSCAL-Modell. Damit können die Modelle, die trilineare Zerlegungen der zugrunde liegenden "two-mode three-way" bzw. "three-mode three-way" Datenkörper vornehmen, als eine konzeptuell und formal einheitliche Klasse angesehen werden. Aus den Konfigurationen nach dem SUMMAX-Modell ergibt sich CANDECOMP/PARAFAC, indem drei Permutationsmatrizen  $P_A$ ,  $P_B$ ,  $P_C$  der Ordnung  $r \times r$  und drei Diagonalmatrizen  $D$ ,  $M$ ,  $N$  gleicher Ordnung eingeführt werden. Zur Ableitung des INDSCAL- und IDIOSCAL-Modells genügen jeweils zwei Diagonal- und Permutationsmatrizen. CANDECOMP/PARAFAC läßt sich in folgender Weise dem SUMMAX-Modell zuordnen:

$$i\bar{A}_r = iA_r P_A D$$

$$j\bar{B}_r = jB_r P_B M$$

$$k\bar{C}_r = kC_r P_C N \quad .$$

Damit besitzen beide Konfigurationen die gleiche Anzahl an Faktortripeln, und die Faktoren eines jeden Modus stimmen überein bis auf eine Permutation ihrer Abfolge und einer

Multiplikation mit Skalaren ungleich null. Die Permutations- und Diagonalmatrizen weisen folgende Relation zueinander auf:

$$[P_A D, P_B M, P_C N] = I \quad .$$

D.h. die Wirkungen der Permutationen und Skalierungen der Varianz der Bezugsachsen heben sich wechselseitig auf. Die Bestimmung der Diagonalmatrizen kann analog der allgemein üblichen Vorgehensweise in CANDECOMP/PARAFAC vorgenommen werden, indem die Varianzaufklärungen der Faktoren in zwei Modi auf den Betrag eins normiert werden mit der inversen Skalierung im übrigen Modus.

*incorrect -  
verlaanede  
vrijantie  
is alleen  
bepaald bij  
orthogonale  
oplossingen*

Unter der Perspektive des erweiterten SUMMAX-Modells, das durch die Einführung einer Kernmatrix zusätzliche Rotations- transformationen der Faktormatrizen gestattet, erscheint CANDECOMP/PARAFAC als ein Sonderfall von SUMMAX, welcher durch zwei definierte Klassen von Transformationen aus der allgemeinen trilinearen Grundgleichung hervorgeht. Werden diese Transformationen auf orthogonale Rotationen der Faktormatrizen erweitert, so ist das Tuckersche Modell aus der quadrilinearen Zerlegung nach SUMMAX ableitbar, wie im folgenden gezeigt wird.

#### 5.2. Der Bezug von SUMMAX zu dem Tuckerschen Modell:

Zur Darstellung des Zusammenhangs zwischen SUMMAX und dem Tuckerschen Modell wird von Tuckers Konzeption der dreimodalen Faktorenanalyse in der in Abschnitt 3 beschriebenen Form ausgegangen (Numerierung der Gleichungen wie in Abschnitt 3):

$$(4) \quad {}_i X_{jk} = \bar{A} \bar{G} (\bar{B}' \times \bar{C}') \quad \text{mit}$$

$$(5a) \quad {}_i X_{jk} {}_i X_{jk}' = \bar{A} \Psi_A \bar{A}'$$

$$(5b) \quad {}_j X_{ik} {}_j X_{ik}' = \bar{B} \Psi_B \bar{B}'$$

$$(5c) \quad {}_k X_{ij} {}_k X_{ij}' = \bar{C} \Psi_C \bar{C}' \quad .$$

Ferner wird die quadrilineare SUMMAX-Modellgleichung zugrunde gelegt

$$(41) \quad {}_i X_{jk} = A G (B' \times C') ,$$

deren Faktormatrizen A, B und C aus der Hauptachsentransformation der trilinearen Zerlegung hervorgehen:

$$(43a) \quad A = {}_i A_f \quad T = F \Delta_A^{1/2}$$

$$(43b) \quad B = {}_j B_f \quad W = V \Delta_B^{1/2}$$

$$(43c) \quad C = {}_k C_f \quad H = E \Delta_C^{1/2} \quad .$$

Dabei bezeichnen  $\Delta_A$ ,  $\Delta_B$  und  $\Delta_C$  drei Diagonalmatrizen, deren Diagonalen die Eigenwerte der Faktormatrizen aus den drei Modi enthalten, und die Spalten von F, V und E sind die zugeordneten Eigenvektoren.

Zwischen SUMMAX und dem Tuckerschen Modell können zwei Bezüge aufgezeigt werden. Im allgemeinen Fall ist eine Zerlegung der mit den Varianzen der Faktoren verknüpften SUMMAX-Kernmatrix in Eigenvektoren notwendig, um das Tuckersche Modell abzuleiten. Dabei wird angenommen, daß die dreimodale Matrix X nicht gleich der Nullmatrix ist. Die zweite Verknüpfung beider Modelle gründet auf der zusätzlichen Voraussetzung, daß die SUMMAX-Kernmatrix eine Struktur analog den Thurstoneschen Kriterien der Einfachstruktur in allen q Lagen aufweist. In diesem Fall ist keine interponierte Analyse der SUMMAX-Kernmatrix notwendig, und das Tuckersche Modell ergibt sich direkt aus den Konfigurationen nach SUMMAX. Theorem 2 spezifiziert den allgemeinen Zusammenhang zwischen beiden Modellen, und Theorem 3 enthält den Sonderfall des Bezuges der Modelle bei Einfachstruktur der SUMMAX-

Kernmatrix.

Theorem 2:

Es sei

$$(73a) \quad \Delta_A^{1/2} \quad {}_m G_{pq} \quad (\Delta_B \times \Delta_C) \quad {}_m G_{pq}' \quad \Delta_A^{1/2} = U_A \quad \Psi_1 \quad U_A'$$

$$(73b) \quad \Delta_B^{1/2} \quad {}_p G_{mq} \quad (\Delta_A \times \Delta_C) \quad {}_p G_{mq}' \quad \Delta_B^{1/2} = U_B \quad \Psi_2 \quad U_B'$$

$$(73c) \quad \Delta_C^{1/2} \quad {}_q G_{mp} \quad (\Delta_A \times \Delta_B) \quad {}_q G_{mp}' \quad \Delta_C^{1/2} = U_C \quad \Psi_3 \quad U_C' \quad ,$$

wobei  $\Psi_1, \Psi_2, \Psi_3$  drei Diagonalmatrizen und  $U_A, U_B, U_C$  drei orthonormale Matrizen bezeichnen. Dann folgt:

$$(74) \quad \bar{A} = F U_A \quad , \quad \bar{B} = V U_B \quad , \quad \bar{C} = E U_C \quad \text{und}$$

$$(75) \quad {}_m \bar{G}_{pq} = (U_A' \quad \Delta_A^{1/2}) \quad {}_m G_{pq} \quad ((\Delta_B^{1/2} U_B) \times (\Delta_C^{1/2} U_C)) \quad .$$

D.h. werden die Kernrelationen  $g_{mpq}$  mit den Eigenwerten der Faktoren aus A, B, C gewichtet, so sind die spaltenweise angeordneten Eigenvektoren der gewichteten Kernmatrix die Transformationsmatrizen, die die SUMMAX-Faktoren in die Tuckerschen Faktoren überführen. Die Tuckersche Kernmatrix ergibt sich dann aus der Multiplikation der gewichteten Matrix  ${}_m G_{pq}$  mit den Transformationsmatrizen.

Theorem 3:

Wenn für beliebige  $g_{ijk} \neq 0$  zusätzlich gilt:

$$(76) \quad g_{i'j'k'} = 0 \quad , \quad \begin{array}{l} i' = 1, \dots, i-1, i+1, \dots, m \quad , \\ j' = 1, \dots, j-1, j+1, \dots, p \quad , \\ k' = 1, \dots, k-1, k+1, \dots, q \quad , \end{array}$$

dann folgt:

$$(77) \quad \bar{A} = F \quad , \quad \bar{B} = V \quad , \quad \bar{C} = E \quad \text{und}$$

$$(78) \quad {}_m \bar{G}_{pq} = \Delta_A^{1/2} \quad {}_m G_{pq} \quad (\Delta_B \times \Delta_C)^{1/2} \quad .$$

In diesem Fall haben die Faktormatrizen eines jeden Modus die gleichen Spalten bis auf eine Multiplikation mit Skalaren ungleich null, welche die Faktoren aus A, B und C auf die Länge eins normiert. Darüber hinaus resultiert die Tuckersche Kernmatrix aus der Gewichtung der Kernrelationen  $g_{mpq}$  mit den Eigenwerten der Faktoren aus A, B und C.

5.2.1. Beweis des Zusammenhangs zwischen SUMMAX und dem Tuckerschen Modell:

Der Beweis des Bezuges beider Modelle ist in zwei Teile untergliedert. Im ersten Teil werden Aussagen über den Zusammenhang der Eigenwerte abgeleitet. Die entsprechenden Aussagen werden eingesetzt, um A, B, C in  $\bar{A}$ ,  $\bar{B}$ ,  $\bar{C}$  zu transformieren und den Zusammenhang der Matrizen  $m_{pq}^G$  und  $m_{pq}^{\bar{G}}$  aufzuzeigen. Abschließend wird Theorem 3 als Sonderfall der vorhergehenden Aussagen aufgezeigt. Im folgenden wird die Indizierung der Kernmatrix  $m_{pq}^G$  nur dann eingesetzt, wenn dies aus Gründen der Eindeutigkeit erforderlich ist. Ansonsten bezeichnet die alternative Darstellung G immer die Anordnung  $m_{pq}^G$ .

Nach Gleichung (41) gilt für die Skalarproduktmatrix von X:

$$(79) \quad iX_{jk} iX_{jk}' = A G (B'B \times C'C) G' A' \quad .$$

Einsetzen der rechten Seite der Gleichungen (43a) bis (43c) in Gleichung (79) ergibt:

$$\begin{aligned} iX_{jk} iX_{jk}' &= F \Delta_A^{1/2} G (\Delta_B^{1/2} V' V \Delta_B^{1/2} \times \Delta_C^{1/2} E' E \Delta_C^{1/2}) G' \Delta_A^{1/2} F' \\ &= F \Delta_A^{1/2} G (\Delta_B \times \Delta_C) G' \Delta_A^{1/2} F' \quad , \end{aligned}$$

da V und E orthonormal sind. Das charakteristische Polynom der Skalarproduktmatrix ist:

$$(80) \quad \varphi(\sigma) = \left| \begin{array}{cc} iX_{jk} & iX_{jk}' \\ & -\sigma I \end{array} \right| = 0$$

$$= (\sigma - \psi_{A1})(\sigma - \psi_{A2}) \dots (\sigma - \psi_{Am}) \quad .$$

Wegen  $F'F = I$  folgt dann (Lamprecht, 1978, Korollar 8.10, S.172 und Lemma 11.3, S.234) für das charakteristische Polynom in (80):

$$(\sigma - \psi_{A1})(\sigma - \psi_{A2}) \dots (\sigma - \psi_{Am}) =$$

$$\left| F \Delta_A^{1/2} G (\Delta_B \times \Delta_C) G' \Delta_A^{1/2} F' - \sigma I \right| =$$

$$\left| \Delta_A^{1/2} G (\Delta_B \times \Delta_C) G' \Delta_A^{1/2} - \sigma I \right| \quad .$$

Dann aber stimmen die Eigenwerte der Skalarproduktmatrix und der gewichteten Kernmatrix  $G$  überein. Werden die Eigenwerte nach abfallender Größe in den Diagonalen von  $\Psi_A$  und  $\Psi_1$  angeordnet, so folgt:

$$(81a) \quad \Psi_A = \Psi_1$$

und analog gilt:

$$(81b) \quad \Psi_B = \Psi_2 \quad ,$$

$$(81c) \quad \Psi_C = \Psi_3 \quad .$$

D.h. die Eigenwerte eines jeden Modus im Tuckerschen Modell sind identisch mit den Eigenwerten der Kernmatrix in SUMMAX, die mit den Varianzen der Faktoren gewichtet wurde.

Ist  $F$  nicht zeilenweise orthonormal, so können  $F$  und  $F'$  zyklisch vertauscht werden (Gröbner, 1966, Satz 4, S.169), so daß der Ausdruck

$$\left| \Delta_A^{1/2} G (\Delta_B \times \Delta_C) G' \Delta_A^{1/2} F'F - \sigma I \right|$$

resultiert, und hieraus folgt das gleiche Ergebnis für alle Eigenwerte ungleich null.

Es seien die Eigenvektoren  $u_{Ai}$  mit

$$(\Delta_A^{1/2} G (\Delta_B \times \Delta_C) G' \Delta_A^{1/2} - \psi_{Ai} I) u_{Ai} = 0$$

als Spalten in  $U_A$  angeordnet. Dann folgt für die Skalarproduktmatrix  $i^{X_{jk}} i^{X_{jk}'}$ :

$$\begin{aligned}
 (83) \quad F \Delta_A^{1/2} G (\Delta_B \times \Delta_C) G' \Delta_A^{1/2} F' &= F U_A \Psi_1 U_A' F' \\
 &= \bar{A} \Psi_A \bar{A}' \\
 &= \bar{A} \Psi_1 \bar{A}' \\
 &= (F U_A) \Psi_1 (F U_A)' .
 \end{aligned}$$

Hieraus folgt (Rao & Mitra, 1971, Complement 1, S.17) die in Gleichung (74) aus Theorem 2 behauptete Aussage:

$$(74a) \quad \bar{A} = F U_A$$

und analog gilt nach Vertauschung der Zeilen, Spalten und Lagen in  $m^{G_{pq}}$ :

$$(74b) \quad \bar{B} = V U_B ,$$

$$(74c) \quad \bar{C} = E U_C .$$

D.h. die als Spalten angeordneten Eigenvektoren der gewichteten SUMMAX-Kernmatrix spezifizieren die Transformation, welche die normierten Hauptkomponenten aus A, B und C auf die Tuckerschen Eigenvektoren in  $\bar{A}$ ,  $\bar{B}$  und  $\bar{C}$  abbildet.

Gleichung (41) des SUMMAX-Modells kann mit  $\bar{A} = F U_A$ ,  $\bar{B} = V U_B$  und  $\bar{C} = E U_C$  in folgender Form dargestellt werden:

$$\begin{aligned}
 i^{X_{jk}} &= F \Delta_A^{1/2} G (\Delta_B^{1/2} \times \Delta_C^{1/2}) (V' \times E') \\
 &= F U_A U_A' \Delta_A^{1/2} G (\Delta_B \times \Delta_C)^{1/2} ((V U_B U_B')' \times (E U_C U_C')') \\
 &= F U_A U_A' \Delta_A^{1/2} G (\Delta_B \times \Delta_C)^{1/2} (U_B \times U_C) (U_B' V' \times U_C' E') \\
 &= \bar{A} (U_A' \Delta_A^{1/2} G (\Delta_B \times \Delta_C)^{1/2} (U_B \times U_C)) (\bar{B}' \times \bar{C}')
 \end{aligned}$$

Aus  $\bar{G} = \bar{A}' \ i^X_{jk} \ (\bar{B} \times \bar{C})$  folgt die Aussage (75) in Theorem 2:

$$\begin{aligned} \bar{G} &= \bar{A}' \bar{A} \ (U_A' \ \Delta_A^{1/2} G (\Delta_B \times \Delta_C)^{1/2} (U_B \times U_C)) (\bar{B}' \times \bar{C}') (\bar{B} \times \bar{C}) \\ &= (U_A' \ \Delta_A^{1/2}) G (\Delta_B^{1/2} U_B \times \Delta_C^{1/2} U_C) \end{aligned}$$

D.h. die Tuckersche Kernmatrix ergibt sich aus der Verknüpfung der SUMMAX-Kernmatrix mit den Eigenwerten der SUMMAX-Faktoren und den Transformationsmatrizen  $U_A$ ,  $U_B$  und  $U_C$ .

Sind für die Kernmatrix  $G$  speziell die Voraussetzungen von Theorem 3 erfüllt, dann gilt:

$$(84) \quad \sum_p \sum_q g_{mpq} g_{m'pq} = 0 \quad \text{für } m \neq m'.$$

D.h.  $GG'$  ist selbst eine Diagonalmatrix. Da nach Voraussetzung jeder einzelne Term  $g_{mpq}g_{m'pq}$  gleich null ist, folgt ferner:

$$(85) \quad \sum_p \sum_q g_{mpq} g_{m'pq} \mu_{pq} = 0$$

für beliebige Skalare  $\mu_{pq}$ . Dann ist aber auch  $G(\Delta_B \times \Delta_C)G'$  und  $\Delta_A G(\Delta_B \times \Delta_C)G'$  diagonal. Es folgt (Lamprecht, 1978, Paragraph 11, S.230ff):

ein Eigenwert von  $\Delta_A^{1/2} G (\Delta_B \times \Delta_C) G \Delta_A^{1/2}$  stimmt mit einem Diagonalelement des Ausdrucks überein. Daher ist

$$(86) \quad \psi_{Am} = \delta_{Am} \sum_p \sum_q \delta_{Bp} \delta_{Cq} g_{mpq}^2$$

Eigenwert von  $i^X_{jk} i^X_{jk}'$ . Bei entsprechender Anordnung der Eigenwerte in  $\psi_A$ :

$$(87) \quad \psi_A = \Delta_A^{1/2} G (\Delta_B \times \Delta_C) G' \Delta_A^{1/2}$$

Aus  $i^X_{jk} i^X_{jk}' = \bar{A} \psi_A \bar{A}' = F \psi_A F'$  folgt die in Gleichung (77), Theorem 3, behauptete Identität:

$$(77a) \quad \bar{A} = F.$$

Analog gilt:  $\bar{B} = V$  und  $\bar{C} = E$ .

Und hieraus folgt sofort:

$$(78) \quad \bar{G} = \Delta_A^{1/2} G (\Delta_B \times \Delta_C)^{1/2} .$$

5.2.2. Implikationen des Bezuges zwischen SUMMAX und dem Tuckerschen Modell:

In Analogie zu dem Bezug zwischen CANDECOMP/PARAFAC und SUMMAX geht das Tuckersche Modell aus der quadrilinearen Zerlegung nach SUMMAX durch zwei Klassen von Transformationen, Reskalierung der SUMMAX-Konfiguration und anschließende orthogonale Rotationen, hervor. Dabei folgt die Reskalierung der SUMMAX-Konfiguration dem Tuckerschen Rationale, indem normierte Faktoren in den drei Modi zugrunde gelegt werden und die Kernmatrix mit den Eigenwerten der zugeordneten Faktoren gewichtet wird. Die Extraktion der Eigenvektoren der gewichteten Kernmatrix liefert die orthogonalen Transformationen, welche die normierten Hauptkomponenten aus SUMMAX in die Tuckerschen Faktoren überführen. Während die Faktormatrizen in CANDECOMP/PARAFAC und SUMMAX bis auf eine Multiplikation mit Skalaren und eine Permutation übereinstimmen, resultieren aus dem quadrilinearen SUMMAX-Modell und dem Tuckerschen Ansatz Bezugsachsen, die in jedem Modus durch eine orthogonale Rotation aufeinander abgebildet werden können:

$$\bar{A} = A \Delta_A^{-1/2} U_A ,$$

$$\bar{B} = B \Delta_B^{-1/2} U_B \quad \text{und}$$

$$\bar{C} = C \Delta_C^{-1/2} U_C .$$

Darüber hinaus sind die Transformationsmatrizen  $U_A$ ,  $U_B$  und  $U_C$  eindeutig bestimmt. Die Spalten von  $U_A$  ergeben sich aus der Hauptachsentransformation der gewichteten SUMMAX-Kernmatrix, verknüpft mit ihrer Transponierten. Wie auch im Rahmen der Reskalierung der Faktoren wird hier im wesentlichen das

Grundprinzip der Tuckerschen Methode angewendet, wobei die Skalarproduktmatrix des Datenkörpers,  $iX_{jk} iX_{jk}'$ , durch die Skalarprodukte der Kernrelationen ersetzt wird. In gleicher Weise werden  $U_B$  und  $U_C$  aus der gewichteten Kernmatrix nach SUMMAX abgeleitet, indem die beiden übrigen "derivational types" in der Kernmatrix jeweils als Zeilen angeordnet werden und nach Vertauschung der Zeilen, Spalten und Lagen von  ${}_m G_{pq}$  erneut die Eigenvektoren bestimmt werden. D.h. die Transformationsmatrizen  $U_B$  und  $U_C$  ergeben sich aus der Hauptachsentransformation von  ${}_p G_{mq}$  und  ${}_q G_{mp}$ , die zuvor mit den Eigenwerten der Faktoren aus A, B und C gewichtet wurden. Ferner geht die Kernmatrix des SUMMAX-Modells in die Tuckersche Kernmatrix über, wenn die gewichtete Matrix  ${}_m G_{pq}$  mit den Transformationsmatrizen multipliziert wird:

$${}_m \bar{G}_{pq} = (U_A' \Delta_A^{1/2}) {}_m G_{pq} ((\Delta_B^{1/2} U_B) \times (\Delta_C^{1/2} U_C)) .$$

Diese Verknüpfung ist äquivalent zu einer orthogonalen Rotation der Kernmatrix, die entgegengesetzt zu der Transformation der Faktormatrizen gerichtet ist.

Während im allgemeinen Fall eine Konfiguration nach SUMMAX in die entsprechende Lösung des Tuckerschen Modells überführt werden kann, ist der Rückschluß in umgekehrter Richtung nicht möglich; die Tuckerschen Eigenwerte und Eigenvektoren enthalten keine Informationen, aus denen die zugeordneten Eigenwerte in  $\Delta_A$ ,  $\Delta_B$  und  $\Delta_C$  abgeleitet werden können, und damit sind die SUMMAX-Faktoren in A, B und C nicht rekonstruierbar.

Neben diesem generellen Bezug des SUMMAX-Modells zu dem Tuckerschen Verfahren existieren weitere Zusammenhänge, die auf direktem Wege abgeleitet werden können unter der Voraussetzung, daß die SUMMAX-Kernmatrix den in Theorem 3 formulierten Bedingungen genügt. In diesem speziellen Fall wird angenommen, daß  ${}_m G_{pq}$  simultan in allen q Lagen analog den Thurstoneschen Kriterien der Einfachstruktur aufgebaut ist. Dabei wird das Prinzip der Einfachstruktur innerhalb jeder Teilmatrix  ${}_m G_p$  als auch zwischen allen q Anordnungen  ${}_m G_p$

zugrunde gelegt. Die geforderte Restriktion entspricht einer Invertierung der Kriterien der Diagonalisierungsverfahren, die im vorangegangenen Abschnitt angesprochen wurden. Während in den ALS-Verfahren zur Diagonalisierung einer Kernmatrix identische Strukturen in jeder der  $q$  Lagen von  ${}_m G_{pq}$  angestrebt werden, führt die Anwendung des Thurstoneschen Prinzips zu Teilmatrizen  ${}_m G_p$ , die zwischen den  $q$  Lagen keine gemeinsamen hohen Werte aufweisen. Hierdurch werden multiple Interaktionen der Faktoren eliminiert, so daß jeweils eine einzige Kombination der Form  $g_{mpq} = t_m w_p h_q$  ungleich null existiert. Ist diese Voraussetzung erfüllt, so stimmen die Faktoren aus SUMMAX und dem Tuckerschen Modell bis auf eine Multiplikation mit Skalaren überein:

$$\begin{aligned}\bar{A} &= A \Delta_A^{-1/2} \quad , \\ \bar{B} &= B \Delta_B^{-1/2} \quad \text{und} \\ \bar{C} &= C \Delta_C^{-1/2} \quad .\end{aligned}$$

Die Skalare sind in diesem Fall mit den Varianzaufklärungen der SUMMAX-Faktoren identisch, und die Gewichtung der SUMMAX-Kernmatrix entsprechend dem Tuckerschen Vorgehen erzeugt die Tuckerschen Kernrelationen:

$${}_m \bar{G}_{pq} = \Delta_A^{1/2} {}_m G_{pq} (\Delta_B \times \Delta_C)^{1/2} \quad .$$

Die Reskalierung der SUMMAX-Konfiguration überführt die Kern- und Faktormatrizen bereits in das Tuckersche Modell, so daß eine zusätzliche Analyse der SUMMAX-Kernmatrix entfällt.

### 5.3. Die Rekonstruktion der dreimodalen Modelle aus SUMMAX:

Obwohl sich die trilinearen und quadrilinearen Modelle in ihren Eigenschaften grundlegend unterscheiden, existieren allgemeine Zusammenhänge zwischen beiden Modellklassen.

Dabei kann die Verbindung zwischen den verschiedenen dreimodalen Verfahren durch das SUMMAX-Modell hergestellt werden unter Betrachtung von drei Transformationen: Reskalierungen, Permutationen und orthogonale Rotationen. Innerhalb der trilinearen Klasse wurde gezeigt, daß das CANDECOMP/PARAFAC-Modell zu Konfigurationen führt, die aus der trilinearen Zerlegung nach SUMMAX rekonstruierbar sind. Werden die Grunddimensionen nach SUMMAX entsprechend dem CANDECOMP/PARAFAC-Rationale reskaliert, bzw. werden die Bezugsachsen permutiert, so resultieren identische Lösungen. Aus dieser direkten Äquivalenz der trilinearen Verfahren folgt darüber hinaus, daß auch das INDSCAL- und IDIOSCAL-Modell aus der SUMMAX-Grundgleichung zur multidimensionalen Skalierung individueller Differenzen ableitbar ist. Der Übergang zu den quadrilinearen Modellen erfolgt aufgrund der Hauptachsentransformation der trilinearen Konfigurationen und der zugeordneten Definition der Kernmatrix. In der Hauptachsenform der extrahierten Faktortripel sind sowohl die grundsätzlich anderen Eigenschaften der quadrilinearen Modelle als auch die Verbindungen zum Tuckerschen Modell begründet. Wird die quadrilineare SUMMAX-Konfiguration in jedem Modus orthogonal rotiert, so ergibt sich das Tuckersche Modell, wobei die zugeordneten Transformationsmatrizen eindeutig bestimmt sind und aus der SUMMAX-Kernmatrix direkt abgeleitet werden können. Allerdings muß der Übergang von den trilinearen Verfahren zum Tuckerschen Ansatz über das SUMMAX-Modell vollzogen werden. Durch die feste Skalierung der Faktoren auf Einheitslänge in jeweils zwei Modi gestattet CANDECOMP/PARAFAC keinen analogen Rückschluß; wie in Abschnitt 5.2.1. gezeigt wurde, ist nach Durchführung der Hauptachsentransformation eine Gewichtung der Kernmatrix in Form von

$$\Delta_A^{1/2} G (\Delta_B \times \Delta_C)^{1/2}$$

notwendig zur Ableitung des Tuckerschen Modells. Diese spezielle Gewichtung kann aber nicht erreicht werden, wenn bereits vor der Rotation auf Hauptkomponenten die Varianzaufklärungen in zwei Modi normiert wurden. Daher kann im allgemeinen Fall nicht von der skalierten CANDECOMP/PARAFAC-Konfiguration auf das Tuckersche Modell geschlossen werden.

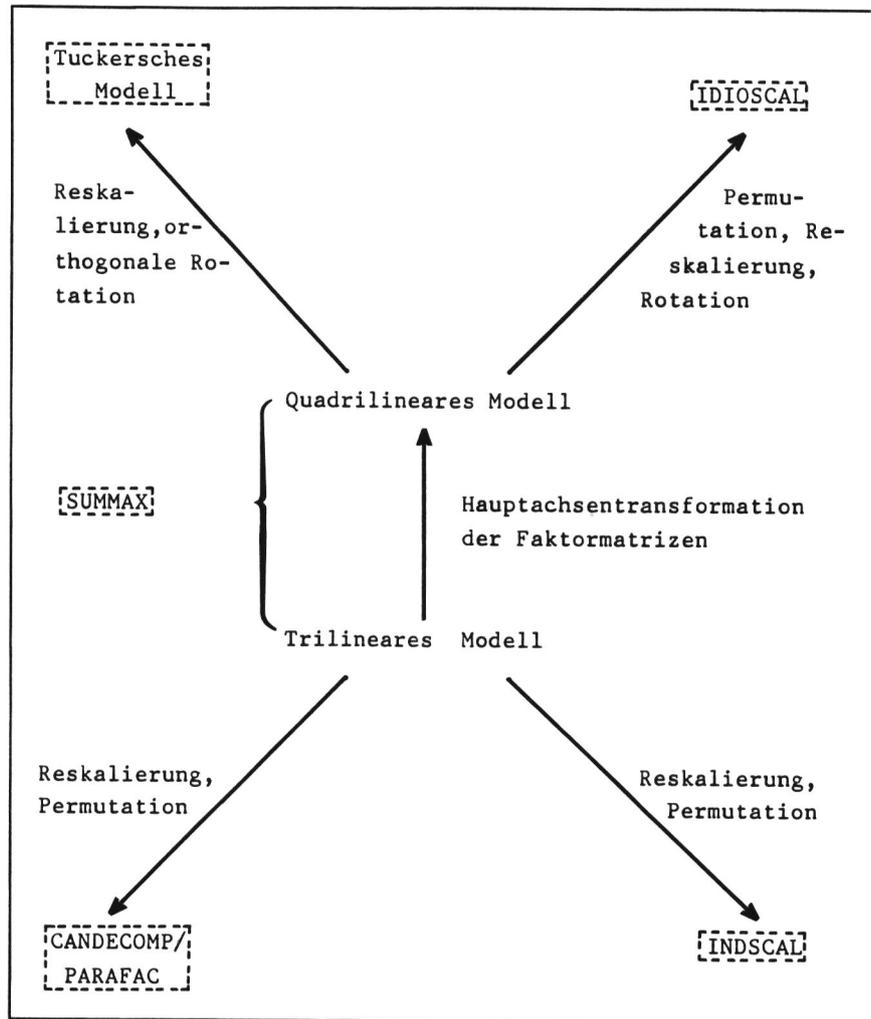


Abbildung 7.1: Schematische Darstellung des Bezuges von SUMMAX zu dem Tuckerschen Modell, CANDECOMP/PARAFAC, INDSCAL UND IDIOSCAL

Nach den Ergebnissen des vorliegenden Abschnitts sind die betrachteten Modelle der dreimodalen Faktorenanalyse durch einfache Transformationen aus SUMMAX analytisch ableitbar; aus der Kenntnis der SUMMAX-Konfiguration können die Konfigurationen nach CANDECOMP/PRARFAC und dem Tuckerschen Modell rekonstruiert werden. Zur Begründung der modelltheoretischen Zusammenhänge wurden keine Bedingungen vorausgesetzt, die zu einer wesentlichen Einschränkung der Aussagen führen. Auf der Datenebene wurde eine dreimodale Matrix zugrunde gelegt, die nicht mit der Nullmatrix identisch ist und die mehr als eine linear unabhängige Teilmatrix  $X_i$  enthält. Im Rahmen empirischer Datenanalysen ist die erste Voraussetzung trivial erfüllt, und eine Verletzung der zweiten Annahme impliziert einen Datenkörper, dessen Faktoren mittels klassischer zweimodaler Verfahren extrahierbar sind. Bezüglich der Modelle wurde vorausgesetzt, daß die CANDECOMP-Repräsentation  $X = [\bar{A}, \bar{B}, \bar{C}]$  einen minimalen Rang  $r$  besitzt mit  $r = \text{Rang}(X)$  und die Tuckerschen Faktormatrizen in jedem Modus den Rang der korrespondierenden Skalarproduktmatrizen aufweisen. Entsprechend der in Abschnitt 3 und 4 dargestellten Eigenschaften der Verfahren gewährleisten beide Modelle, daß diese Annahmen bei genügend hoher Anzahl an Faktoren erfüllt sind. Da ferner den dreimodalen Konfigurationen keine spezifischen Bedingungen unterlegt wurden, sind die gezeigten modelltheoretischen Zusammenhänge allgemein gültig. In Abschnitt 8 werden die diskutierten Bezüge exemplarisch an einem Datensatz aus dem Bereich der Farbwahrnehmung dargestellt.

6.

Anwendungen der Modelle der dreimodalen Faktorenanalyse: Vorbemerkungen

Der empirische Teil der vorliegenden Arbeit enthält eine Modelluntersuchung der betrachteten Modelle der dreimodalen Faktorenanalyse und Multidimensionalen Skalierung individueller Differenzen. Um einen Vergleich der Gemeinsamkeiten oder Unterschiede der Verfahren im Rahmen empirischer Datenanalysen zu ermöglichen, erschien es sinnvoll, Datensätze aus Methodenexperimenten auszuwählen, deren Struktur bereits bekannt ist und mithilfe anderer Verfahren analysiert wurde (vgl. Orlik, 1967; Borg, 1981). In den folgenden Abschnitten werden die Daten aus vier Untersuchungen analysiert, deren zugrunde liegender Versuchsaufbau jeweils eine Vorhersage der relevanten faktorenanalytischen Befunde gestattet. Dabei führten die zugeordneten experimentellen Bedingungen zu deutlichen Unterschieden in den Beziehungen der korrespondierenden Stimuli, Personen und Variablen. Die Hypothesen, die aus den Konfigurationen der verschiedenen Verfahren resultieren, werden hinsichtlich der Reproduktion der bekannten Struktureigenschaften der Daten diskutiert.

Ein wesentlicher Unterschied zwischen den modelltheoretischen Betrachtungen und Anwendungen der Verfahren im Rahmen empirischer Datenanalysen ist bedingt durch die Konvergenzeigenschaften der Algorithmen. Wie in Abschnitt 3 ausgeführt wurde, enthält CANDECOMP/PARAFAC, der ALS-Ansatz von Kroonenberg und De Leeuw und SUMMAX jeweils eine Parameterfunktion, deren globales Optimum iterativ zu bestimmen ist. Dabei ist jedoch eine Konvergenz in ein lokales Maximum bzw. Minimum nicht ausgeschlossen, und für eine spezifische Lösung ist daher nicht gesichert, daß sie den geforderten Kriterien eines Verfahrens auch entspricht. Im Rahmen der "Alternating-Least-Squares"-Ansätze ist ferner die a priori zu bestimmende Dimensionalität kritisch; Konfigurationen unterschiedlicher Dimensionalität können erhebliche Unterschiede in den Faktorladungen aufweisen und zu divergierenden Strukturhypothesen führen. Auch unabhängig von den Eigenschaften der Algorithmen implizieren die engen algebraischen Bezüge der verschiedenen dreimodalen Verfahren aufgrund der notwendigen Transformationen keine Identität der

entsprechenden Konfigurationen. Anhand der gewählten Datensätze wird exemplarisch gezeigt, welche divergierenden oder übereinstimmenden Strukturen die Verfahren erzeugen.

#### 6.1. Untersuchte Datensätze:

Entsprechend der Taxonomie von Carroll und Arabie (1980) lassen sich zwei der ausgewählten Datensätze als "three-way three mode data" charakterisieren; diese bilden die Grundlage des Vergleichs zwischen CANDECOMP/PARAFAC, SUMMAX und der Tuckerschen dreimodalen Faktorenanalyse. Zwei weitere Datensätze enthalten als "three-way two mode data" jeweils dyadische Meßwerte oder Beobachtungen für eine Reihe von Probanden; sie gestatten einen Vergleich des INDSCAL-Verfahrens mit dem Tuckerschen und dem SUMMAX-Ansatz zur Multidimensionalen Skalierung individueller Differenzen.

Abschnitt 7 enthält die Reanalyse eines Modellexperimentes von Orlik aus dem Gebiet der Farbwahrnehmung. In dieser Untersuchung wurde einer Gruppe von Versuchspersonen Papierquadrate vorgegeben, die nach Farbe und Größe variierten, während einer zweiten Gruppe von Probanden nur Größeninformationen dargeboten wurden. Die Stimuli waren jeweils auf ausgewählten Farb- und Größenskalen einzustufen ("three-way three mode data"). Nach dem Versuchsaufbau wird eine Differenzierung beider Personengruppen mit hohen bzw. niedrigen Kovariationen der Farbdimension mit den Farbskalen erwartet.

Abschnitt 8 enthält ebenfalls ein Beispiel aus dem Bereich der Farbwahrnehmung. Sowohl normal farbsichtigen Probanden als auch Versuchspersonen mit einer Deuteranomalie wurden Farbreize aus dem gesamten Bereich des Seh-Spektrums vorgegeben, die hinsichtlich ihrer Ähnlichkeit zu beurteilen waren ("three-way two mode data"). Für die normal farbsichtigen Personen wird ein homogener Farbkreis mit einer "Gelb-Blau"- und einer "Rot-Grün"-Dimension und für die farbenfehlsichtigen Probanden eine Deformation der "Rot-

Grün"-Dimension erwartet.

Die in Abschnitt 9 dargestellte Untersuchung von Rosenberg und Kim hat zum Gegenstand, in welcher Weise Varianten der Sortiermethode die Beachtung unterschiedlicher Organisationsaspekte genealogischer Begriffe induzieren ("three-way two mode data"). Es zeigten sich deutliche Unterschiede in der Verwendung des durch die Begriffe implizierten Geschlechtes als Sortiergrundlage je nachdem, ob Versuchspersonen Gelegenheit zu nur einer Sortierung oder zu mehreren Sortierungen erhielten.

Abschnitt 10 enthält eine Reanalyse der Studie von Osgood und Luria zu der Validität des Semantischen Differentials. Zur Untersuchung eines klinischen Falls von multipler Persönlichkeit wurden wichtige Beziehungen der Patientin in jedem Persönlichkeitszustand durch das Semantische Differential erhoben ("three-way three-mode data"). Der wichtigste Befund der Studie besteht in gegenläufigen Einstufungen sozial unerwünschter Konzepte wie "fraud" und "hatred" und des Selbstkonzeptes "me" auf den "Potency"- und Evaluations-skalen zwischen den verschiedenen Persönlichkeitszuständen.

#### 6.2. Algorithmen und deren Abbruchkriterien:

Der im CANDECOMP/PARAFAC-Modell verwendete "Alternating-Least-Squares"-Algorithmus (vgl. Lyttkens, 1966; Wold, 1966) zur Faktorenextraktion erfordert die Festlegung der Anzahl zu extrahierender Faktoren vor der Datenanalyse. Dabei werden die Faktoren für eine Konfiguration festgelegter Dimensionalität simultan bestimmt, und das Muster der Faktorladungen einer  $r$ -Faktorlösung kann sich durch Hinzunahme eines  $(r+1)$ -ten Faktors für die Gesamtkonfiguration deutlich verändern. Harshman (1984) stellte eine Liste von Kriterien zur Beurteilung der Angemessenheit einer Konfiguration vor, die eine Abschätzung der Konvergenz des Algorithmus, der Interpretierbarkeit und Stabilität verschiedener Lösungen umfaßt.

Zu den wichtigsten Kriterien gehört der Scree-Test des faktoriell aufgeklärten Varianzanteils der CANDECOMP/PARAFAC-Konfiguration über verschiedene Dimensionen. Aus den Lösungen unterschiedlicher Dimensionalität wird jeweils der Datenkörper rückgerechnet, und die rückgerechneten Daten werden mit den Ausgangsdaten korreliert. Die Höhe der quadrierten Korrelation wird gegen die Anzahl der extrahierten Faktortripel aufgetragen. Eine deutliche Veränderung des Kurvenverlaufs zeigt die angemessene Dimensionalität an, wobei nach Harshman (1984) die Lösung vor dem Kurvenanstieg auszuwählen ist.

Neben der Anwendung des Scree-Tests ist das Konvergenzverhalten des Algorithmus für jede Lösung zu überprüfen, um das Vorliegen einer instabilen Konfiguration auszuschließen. Dabei geht Harshman von mindestens 50 notwendigen Iterationen des Algorithmus aus. Die Faktorladungen in den drei Modi bzw. die Korrelation der Konfigurationen mit den Ausgangsdaten sollten sich zwischen den letzten  $n$  Iterationen und der  $(n-20)$ -ten oder  $(n-10)$ -ten Iteration nur geringfügig verändern. Harshman betrachtet Veränderungen, die weniger als 1% betragen, als Indikatoren der Stabilität einer Lösung. Wird dieses Kriterium wesentlich überschritten, so ist zu erwarten, daß sich die Faktorladungen bei Fortsetzung der Iterationsabfolge substantiell verändern und zu Unterschieden in der Interpretation führen.

Zusätzlich liefert die Höhe der Korrelationen der Faktoren innerhalb eines jeden Modus Hinweise auf die angemessene Dimensionalität. Korrelieren die Faktoren in einem Modus sehr hoch (Harshman diskutiert als Grenzwert eine Korrelation  $\geq 0.60$ ), so sollten sie nur geringe Werte in den übrigen Modi einer Konfiguration aufweisen. Ist dieses Kriterium nicht erfüllt, dann ist möglicherweise die Anzahl der Faktortripel zu hoch gewählt worden, und es sollte eine niedriger dimensionale Lösung in Betracht gezogen werden. Wenn auch die niedriger dimensional Konfigurationen keine Reduktion der Höhe der Korrelationen zwischen den Faktoren beinhalten, so ist nach Harshman fraglich, ob nicht eine degenerierte Lösung vorliegt und ein anderes Verfahren zur

Datenanalyse eingesetzt werden sollte.

Können nach den genannten Kriterien zwei oder mehrere Lösungen alternativ gewählt werden, so sollten die Konfigurationen hinsichtlich Übereinstimmungen und Unterschieden in der Interpretation überprüft werden. Für den Fall, daß etwa Faktoren aus der zwei- bzw. dreidimensionalen Lösung zu unterschiedlichen Interpretationen führen, aber eine höherdimensionale Konfiguration eine Kombination aus beiden Lösungen darstellt, geben die Konfigurationen aus zwei und drei Faktortripeln nur unvollständige Teilstrukturen des zugrunde liegenden Datenkörpers wieder. Wird hingegen die Information aufgrund eines Faktortripels in einer höherdimensionalen Lösung in zwei oder mehreren Dimensionen wiederholt, so ist die kompaktere Lösung, die zu der konsistenteren Interpretation führt, auszuwählen.

Ferner dienen die von Harshman besonders hervorgehobenen "many-fit-diagnostics" der Abschätzung der Zuverlässigkeit verschiedener Konfigurationen. Dabei wird der gesamte Datenkörper nach Split-Half-Methoden über den Modus der Personen unterteilt in partielle Datenkörper. Jeder Split-Half-Datensatz wird anschließend separat analysiert. Jede getrennt durchgeführte Analyse sollte nach Harshman zu einer identischen Interpretation führen. Zusätzlich können Korrelationen oder Faktor-Kongruenzkoeffizienten zwischen den Faktoren der partiellen Datensätze berechnet werden, die Hinweise auf stabile Strukturen der Gesamtkonfiguration liefern. Jeder Split-Half-Datenkörper sollte nach Harshman mindestens 10 bis 15 Personen umfassen.

Die Auswahl der angemessenen Konfiguration nach der Tucker-schen dreimodalen Hauptkomponentenanalyse ist an der in Abschnitt 3 beschriebenen Vorgehensweise orientiert. Im ersten Auswertungsschritt wird für jeden Modus die Skalarproduktmatrix aus dem Datenkörper bestimmt. Zur Beurteilung der Dimensionalität der Konfiguration werden die Eigenwerte jeder Skalarproduktmatrix gegen die Anzahl der extrahierten Faktoren aufgetragen. Der Kurvenverlauf der Eigenwerte wird in jedem Modus mittels des Scree-Tests beurteilt, um die

Anzahl substantieller Faktoren zu bestimmen. Diese Vorgehensweise entspricht den Kriterien, die im Rahmen der klassischen faktorenanalytischen Verfahren angewendet werden, wobei die Bestimmung bedeutsamer Faktoren in jedem Modus getrennt vorgenommen wird. Im Unterschied zu der traditionellen Hauptkomponentenanalyse sind die Varianzaufklärungen der Tuckerschen Faktoren allerdings nicht mit den Eigenwerten identisch, so daß die Dimensionalität der Faktormatrizen nach Informationen (den Eigenwerten) beurteilt wird, die in den Faktormatrizen nicht repräsentiert sind.

Im SUMMAX-Modell wird der Faktor-Ähnlichkeitskoeffizient  $R^2$  (Orlik, 1980, S.195)

$$R^2 = \frac{(\sum_i \sum_j \sum_k x_{ijk} \hat{x}_{ijk})^2}{(\sum_i \sum_j \sum_k x_{ijk}^2)(\sum_i \sum_j \sum_k \hat{x}_{ijk}^2)}$$

eingesetzt, um den Prozentsatz der Varianz zu bestimmen, der durch die extrahierten Faktortripel aufgeklärt wird. Der Kurvenverlauf der Faktor-Ähnlichkeitskoeffizienten in Abhängigkeit von der Anzahl der extrahierten Dimensionen kann zur Auswahl einer angemessenen Konfiguration mittels des Scree-Tests beurteilt werden. Alternativ empfiehlt Orlik (1980, S.200), vor der Analyse die gewünschte Höhe der Varianzaufklärung festzusetzen und die Faktorenextraktion bis zu dem gewählten Grenzwert fortzusetzen, da der Scree-Test bei einem schnellen Abfall des Kurvenverlaufs zu einer Vernachlässigung der Modi mit hoher Variablenzahl führen kann.

Nach der Hauptachsentransformation der Faktormatrizen in jedem Modus wird erneut der Scree-Test angewendet, indem die Höhe der Eigenwerte gegen die Anzahl der Hauptkomponenten aufgetragen wird und deutliche Veränderungen des Kurvenverlaufs Aufschlüsse über die Anzahl substantieller Faktoren geben. Liegen rotierte Faktormatrizen vor, so kann zur Bestimmung bedeutsamer Faktorladungen das von Orlik (1980, S.202) diskutierte Trennkriterium

$$|a_{im}| \geq \sqrt{\frac{\sum_m QS_m}{n_i n_m}}$$

angewendet werden. Dabei kennzeichnet  $QS_m$  die Quadratsummen der Faktorladungen auf dem  $m$ -ten Faktor,  $n_i$  ist die Anzahl der Elemente im jeweiligen Modus und  $n_m$  die Anzahl der Faktoren.

Im abschließenden Auswertungsschritt wird die SUMMAX-Kernmatrix zur besseren Interpretierbarkeit und zur Vergleichbarkeit der Kernrelationen in ihrer numerischen Größe nachträglich normiert. In diesem Zusammenhang sei auf die Interpretation des SUMMAX-Modells in Abschnitt 3 verwiesen, in dem das Grundprinzip der Normierung der Kernmatrix dargestellt wurde.

Dagegen wird die Kernmatrix im Tuckerschen Ansatz wie auch im Rahmen des ALS-Verfahrens von Kroonenberg und De Leeuw (1980; Kroonenberg, 1983b) nicht normiert. Wie in den folgenden Abschnitten deutlich wird, zeigt die Höhe der Kernrelationen eine erhebliche Schwankungsbreite, und die Auswahl bedeutsamer Kernrelationen bleibt willkürlich. Ferner ist im CANDECOMP/PARAFAC-Modell die Bestimmung hoher Ladungskoeffizienten durch die fehlende Diskussion entsprechender Trennkriterien zwischen substantiellen und niedrigen Faktorladungen erschwert, da die Varianzaufklärungen der Faktoren in den Modi unterschiedlich gesetzt werden. Nach Harshman und Lundy (1984) sind die normierten Ladungen interpretierbar wie  $z$ -standardisierte Variablen. Zur Bestimmung der Substantialität der Faktorladungen im nicht normierten Modus empfehlen Harshman und Lundy, externe Kriterien heranzuziehen und die Höhe der Faktorladungen mit der Ausprägung der entsprechenden Variablen auf den externen Kriterien zu vergleichen. Liegen jedoch keine externen Kriterien vor, so muß aufgrund der relativen Höhe der Ladungen einer gegebenen Lösung ein Grenzwert für die zu interpretierenden Faktorladungen festgesetzt werden.

?  
 warum da  
 nicht alles  
 normieren

In den folgenden Abschnitten werden jeweils die Konfigurationen nach dem trilinearen und dem quadrilinearen SUMMAX-Verfahren dargestellt, um einen Vergleich der Hypothesen, die die Lösungen aus den verschiedenen dreimodalen Modellen implizieren, innerhalb der beiden Modellklassen zu ermöglichen. Ebenso werden für das Tuckersche Modell zwei Lösungen diskutiert, die unrotierte Konfiguration aus den extrahierten Eigenvektoren und die Konfiguration nach Durchführung von Faktorrotationen, wenn Transformationen der Faktormatrizen indiziert sind. Dieses Vorgehen erschien angezeigt aufgrund der kontroversen Diskussion um die Angemessenheit von Faktorrotationen im Tuckerschen Modell. Wie bereits in Abschnitt 3 dargestellt wurde, empfiehlt Tucker trotz noch ungelöster Probleme bei Anwendungen von Rotationstransformationen orthogonale wie auch oblique Verfahren zur Gewinnung einer Einfachstruktur in den Faktormatrizen: "In some applications a simple structure might be meaningful and common techniques for rotation of axes might be appropriate... The only general principle is that the transformations should aid in the interpretation of the results" (Tucker, 1972, S.9ff). Dagegen ist nach Kroonenberg (1983b, 1984) unter dem Aspekt der Interpretierbarkeit der Kernmatrix eine Rotation der Faktormatrizen nicht notwendig angezeigt. Wenn die Kernmatrix nur eine geringe Anzahl von Werten nahe null aufweist, implizieren die Kernrelationen ein schwer interpretierbares, komplexes Geflecht von Bezügen zwischen den Faktoren der drei Modi. Eine alternative Vorgehensweise zu der Rotation der Faktormatrizen auf interpretierbare Strukturen besteht darin, den Zugang zu der Kernmatrix in den Vordergrund zu stellen, und die Kernrelationen in eine Einfachstruktur zu überführen. Empirisch wurde in einer Reihe von Anwendungen eine Kernmatrix vorgefunden, die in der unrotierten Konfiguration eine Einfachstruktur aufwies und nach Durchführung von Faktorrotationen in den drei Modi kompliziert wurde (Kroonenberg, 1984).

Die Berechnung der Konfigurationen nach SUMMAX und dem Tuckerschen Modell (Methode I, vgl. Abschnitt 3) wurden auf

einem IBM PC/XT kompatiblen Computer mit einem numerischen Koprozessor durchgeführt. Zur Faktorenextraktion nach dem SUMMAX-Verfahren wurde eine Pascal-Version des Fortran-Programms von Kohler (1980) verwendet. Die Hauptachsentransformationen für beide Modelle wurden jeweils mit dem Jacobi-Verfahren zur Berechnung sämtlicher Eigenwerte und Eigenvektoren einer symmetrischen Matrix aus der von Engeln-Müllges und Reutter (1985) veröffentlichten Programmsammlung der RWTH Aachen durchgeführt. Das verwendete Verfahren ist eine Pascal-Version der Algol-Prozedur "jacobi" von Rutishauser (Engeln-Müllges & Reutter, 1985, S.303). Das Programm zur Varimax-Rotation wurde dem Programmpaket Turbo Lader Statistics, herausgegeben von Lauer und Wallwitz (1986), entnommen. Zur Berechnung der CANDECOMP- und INDSCAL-Konfigurationen wurden die Programme "Candecomp" und "Indscal" aus dem MDS(X)-Programmpaket (Coxon, 1981), implementiert auf dem Siemens-Großrechner der Universität des Saarlandes, verwendet. Die freundlicherweise von Kroonenberg im Quelltext bereitgestellten Programme "TUCKALS2" und "TUCKALS3" (Kroonenberg, 1985) konnten mit den verfügbaren Fortran-Compilern aufgrund von Sprach- und Systeminkompatibilitäten nicht erfolgreich kompiliert werden.

7.

Reanalyse eines Modellexperimentes  
von Orlik zur Psychophysik des Po-  
laritätsprofils

Gegenstand des folgenden Abschnitts ist die Reanalyse eines Modellexperimentes von Orlik (1965, 1980) aus dem Gebiet der optischen Wahrnehmung. In dieser Untersuchung wurde die dem Semantischen Differential zugrunde liegende Annahme überprüft, daß sich die Struktur psychologischer Merkmalsräume auf Eigenschaftsprofilen abbilden läßt. Sechzehn Versuchspersonen hatten die Aufgabe, vier Papierquadrate auf zehn ausgewählten Skalen des Semantischen Differentials einzustufen. Acht Versuchspersonen wurden Quadrate vorgegeben, die nach Größe und Farbe variierten. Die Seitenlängen des Reizmaterials wurden abgestuft zwischen 1 cm (kleines Quadrat) und 3 cm (großes Quadrat); die dargebotenen Farben umfaßten die Eindrucksqualitäten "violett", "karminrot" und "zinnoberrot" (vgl. Tabelle 3). In dieser experimentellen Bedingung wurde die Beachtung der Farb- und Größeninformationen durch eine bivariate Reizvorlage induziert. Dagegen wurde eine zweite Gruppe von Versuchspersonen mit einer univariaten Reizvorlage konfrontiert, die nach der Reizdimension "Größe", nicht aber nach dem Faktor "Farbe" variierte. Zur Beurteilung der Wahrnehmungsreize wurden die in Tabelle 3b aufgeführten sieben-stufigen Ratingskalen vorgegeben. Dabei konnte gezeigt werden (Orlik, 1965), daß acht Skalen in folgender Beziehung zu den kontrollierten Reizdimensionen stehen:

die Eigenschaftspolaritäten "groß-klein", "wuchtig-zierlich", "kleinlich-großzügig" und "eng-weit" spiegeln die wahrgenommene Flächengröße wider. Farbunterschiede zwischen den Quadraten werden durch die Skalen "traurig-froh", "grell-zurückhaltend", "lebhaft-ruhig" und "erregend-beruhigend" erfaßt.

-----

Die Verfasserin dankt Herrn Prof.Dr. P. Orlik für die freundliche Überlassung der Daten des Modellexperimentes, die im Rahmen eines Seminars zur Faktorenanalyse an der Universität des Saarlandes erhoben wurden.

Tabelle 3: Skalen, Reizmaterial, Versuchspersonen und experimentelle Bedingungen des Modellelements von Orlik

3a) Reizmaterial:

| Farbton | Farbe         |   | Größe    |          |          |
|---------|---------------|---|----------|----------|----------|
|         | Dunkelstufe   |   | 1,0<br>k | 2,0<br>m | 3,0<br>g |
| 11      | 5 (violett)   | v | vk       |          | vg       |
| 8       | 4 (karminrot) | k |          | km       |          |
| 7       | 2 (zinnober)  | z |          |          | zg       |

aus P. Orlik (1965). Die Farbkennziffern entsprechen den Farbzeichen nach DIN 6164.

3b) Skalen:

|                         |     |                       |     |
|-------------------------|-----|-----------------------|-----|
| abstoßend - sympathisch |     | lebhaft - ruhig       | (F) |
| froh - traurig          | (F) | klein - groß          | (G) |
| zierlich - wuchtig      | (G) | erregend - beruhigend | (F) |
| grell - zurückhaltend   | (F) | eng - weit            | (G) |
| kleinlich - großzügig   | (G) | schwierig - einfach   |     |

aus P. Orlik (1980). F kennzeichnet die Farbskalen, G die Größenskalen.

3c) Experimentelle Bedingungen und Vpn:

| bivariate Reizvorlage,<br>Variation nach Größe u. Farbe | univariate Reizvorlage,<br>Variation nach Größe |
|---------------------------------------------------------|-------------------------------------------------|
| Bec, Eng, Koh, Kru,<br>Orl, Pli, Sch, Wac               | Bar, Bie, Bru, Dil,<br>Erm, Hey, Lie, Ork       |

Dieser Befund wurde auch in der Analyse der Ratingdaten der Personengruppe deutlich, die mit der bivariaten Reizkonfiguration konfrontiert wurde (Orlik, 1980, 1981). Erwartungsgemäß zeigte sich eine hohe Kovariation des Reizfaktors "Farbe" mit den Farbskalen und des Faktors "Größe" mit den Größenskalen. Die Homogenität der Urteilssysteme der Versuchspersonen hinsichtlich der vorgegebenen Merkmalsvariationen fand allerdings in der Valenzbeurteilung der Farbtöne keine Entsprechung. In der Einstufung der Wahrnehmungsreize auf der Skala "abstoßend-sympathisch" konnten individuell unterschiedliche Farbpräferenzen festgestellt werden.

Im folgenden werden die Strukturhypothesen der dreimodalen Modelle diskutiert, die sich aus der Analyse beider experimenteller Bedingungen ergaben. Entsprechend dem Versuchsaufbau des Experimentes wird eine hohe Kovariation zwischen den zusammengehörigen Reiz- und Skalendimensionen für die Versuchspersonen der bivariaten Bedingung erwartet. Zusätzlich wird eine Differenzierung der beiden Personengruppen erwartet, verbunden mit geringen Kovariationen zwischen der Reizdimension "Farbe" und den Farbskalen für die Versuchspersonen der univariaten Bedingung.

#### 7.1. Normierung des Datenkörpers:

Bezeichnet  $i$  den Modus der Versuchspersonen und  $j$  den Modus der Skalen, so ist nach dem Vorgehen von Lohmöller (1979) und Orlik (1980) eine  $ij$ -Zentrierung die angemessene Normierung bei Daten, die mit dem Semantischen Differential erhoben wurden. Dementsprechend wurde in diesem Datensatz eine Mittelwertszentrierung getrennt für jede Versuchsperson und Eigenschaftspolarität durchgeführt.

#### 7.2. Die CANDECOMP/PARAFAC-Lösung:

Zur Beurteilung der angemessenen Dimensionalität der CANDE-

COMP-Konfiguration konnten in diesem Datensatz die von Harshman (1984) empfohlenen Split-Half-Methoden nicht eingesetzt werden aufgrund der geringen Anzahl von Versuchspersonen.

#### 7.2.1. Verwendete Auswahlkriterien:

Scree-Test der quadrierten Korrelationen  $R^2$  zwischen den Ausgangsdaten und dem rückgerechneten Datenkörper:

In Abbildung 8 ist die Höhe von  $R^2$  dargestellt für die zwei- bis zehndimensionale Lösung (Programmbeschränkung: Anzahl der Faktortripel  $\leq 10$ ). Der Kurvenverlauf des faktoriell aufgeklärten Varianzanteils zeigt keine deutlichen Veränderungen, und daher liefert der Scree-Test keine Hinweise auf die angemessene Dimensionalität.

#### Konvergenz der Lösungen:

Die zweidimensionale Lösung konvergiert in elf Iterationen, während bei der dreidimensionalen Konfiguration lediglich eine Iteration durchgeführt wird. Die vier- und fünfdimensionale Lösung benötigt 92 bzw. 71 Iterationsschritte. Die sechs- bis zehndimensionalen Konfigurationen erreichen auch nach 100 Iterationen das programminterne Abbruchkriterium nicht. Daher kann angenommen werden, daß die drei- und sechs- bis zehndimensionalen Lösungen keine stabilen Konfigurationen enthalten.

#### Überprüfung der Faktorinterkorrelationen innerhalb eines Modus:

Tabelle 4 zeigt die Korrelationen der Faktoren innerhalb der Modi für die zwei-, drei- und vierdimensionale Lösung. Nur die zweidimensionale Lösung weist geringe Korrelationen für den Personen- und den Skalenmodus auf, während für die übrigen Lösungen nach den Kriterien von Harshman (1984; vgl. Abschnitt 6) nicht ausgeschlossen werden kann, daß degenerierte Konfigurationen vorliegen.

Nach den Konvergenzeigenschaften und den Faktorinterkorrelationen wurde die Lösung aus zwei Faktortripeln ausgewählt,

# Abb. 8 : CANDECOMP

$R^2$  zwischen Roh – u. geschätzten Daten

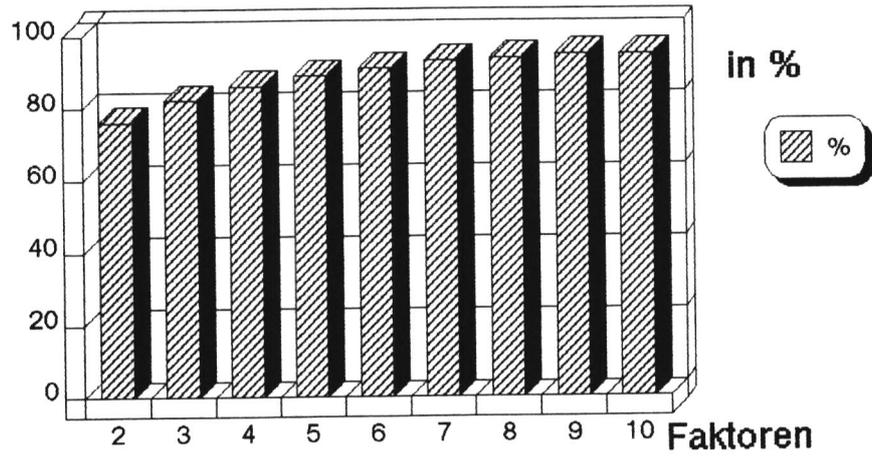


Tabelle 4: Korrelationen der CANDECOMP/PARAFAC-Faktoren

## 4a) zweidimensionale Lösung:

| Personenmodus: |      |      | Skalenmodus: |      |      | Modus der Reize: |      |      |
|----------------|------|------|--------------|------|------|------------------|------|------|
|                | 1    | 2    |              | 1    | 2    |                  | 1    | 2    |
| 2              | 0.09 | 1.00 | 2            | 0.02 | 1.00 | 2                | 0.62 | 1.00 |

## 4b) dreidimensionale Lösung:

| Personenmodus: |      |      | Skalenmodus: |      |       | Modus der Reize: |       |      |
|----------------|------|------|--------------|------|-------|------------------|-------|------|
|                | 1    | 2    |              | 1    | 2     |                  | 1     | 2    |
| 2              | 0.81 | 1.00 | 2            | 0.79 | 1.00  | 2                | -0.99 | 1.00 |
| 3              | 0.36 | 0.43 | 3            | 0.05 | -0.18 | 3                | -0.56 | 0.51 |

## 4c) vierdimensionale Lösung:

| Personenmodus: |       |       |       | Skalenmodus: |       |       |       |
|----------------|-------|-------|-------|--------------|-------|-------|-------|
|                | 1     | 2     | 3     |              | 1     | 2     | 3     |
| 2              | 0.83  | 1.00  |       | 2            | 0.80  | 1.00  |       |
| 3              | 0.57  | 0.71  | 1.00  | 3            | -0.07 | -0.38 | 1.00  |
| 4              | -0.57 | -0.65 | -0.71 | 4            | 0.25  | 0.40  | -0.86 |

## Modus der Reize:

|   | 1     | 2    | 3     |
|---|-------|------|-------|
| 2 | -0.99 | 1.00 |       |
| 3 | -0.63 | 0.56 | 1.00  |
| 4 | -0.31 | 0.41 | -0.51 |

die im folgenden dargestellt wird.

7.2.2. Interpretation der zweidimensionalen Konfiguration:  
Tabelle 5 zeigt die Faktoren der zweidimensionalen CANDECOMP-Lösung, die für den Skalenmodus und den Modus der Wahrnehmungsreize auf die Länge 1 normiert sind. Die Faktoren des Skalenmodus trennen angenähert die beiden Gruppen der vier Farbskalen und der vier Größenskalen. Die vorgegebenen Reizdimensionen "Farbe" und "Größe" finden ihre Entsprechung in den Faktoren des Reizmodus; die Korrelation des zweiten Reizfaktors mit den Reizcharakteristika der Tabelle 3a beträgt 0.99 für die Größe der Quadrate, die entsprechende Korrelation des ersten Faktors beträgt -0.71 für den Farbton und -0.93 für die Dunkelstufe der Reize:

|                | Farbton | Dunkelstufe | Größe |
|----------------|---------|-------------|-------|
| 1. Faktor (R1) | -0.71   | -0.93       | 0.64  |
| 2. Faktor (R2) | -0.30   | -0.43       | 0.99  |

Damit enthält die erste CANDECOMP-Dimension Informationen zu der Differenzierung der Farbtöne, während auf der zweiten Dimension Unterschiede zwischen den Größen der Quadrate abgebildet werden.

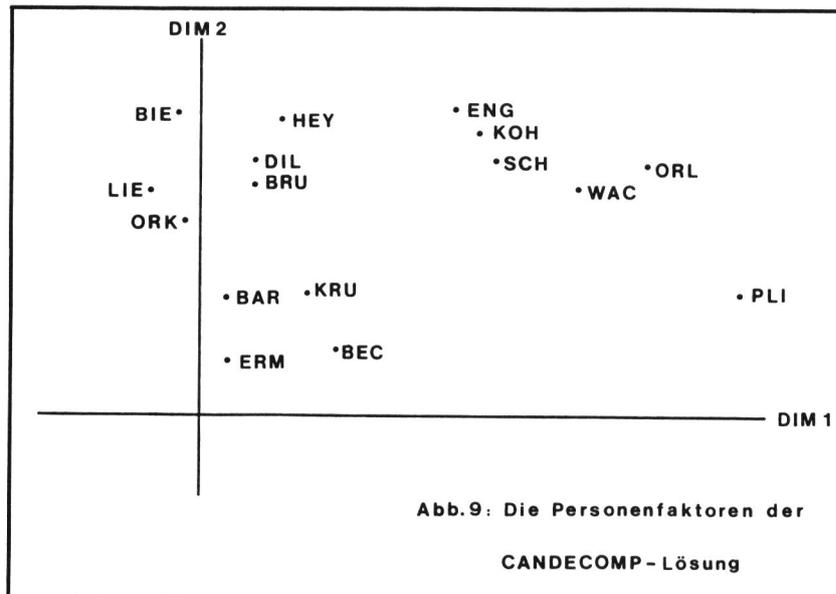
Abbildung 9 stellt die Konfiguration im Personenmodus dar. Der erste Faktor, der im Skalen- und Reizmodus der Abstufung der Farbtöne auf den Farbskalen entspricht, enthält nur angenähert eine Differenzierung der beiden Versuchsgruppen; zwei Versuchspersonen, denen Farb- und Größeninformationen vorliegen, werden hier nach der Höhe der Faktorladungen der univariaten Versuchsbedingung zugeordnet. Die Faktorladungen des zweiten Personenfaktors deuten daraufhin, daß auch die Größenbeurteilung der Wahrnehmungsreize von den Personen unterschiedlich vorgenommen wurde. Während die Mehrzahl der Personen nach der CANDECOMP-Lösung die Größenskalen zur Unterscheidung zwischen den Reizgrößen benutzen, differenzieren hier fünf Versuchspersonen in geringem Ausmaß.

Tabelle 5: zweidimensionale CANDECOMP-Lösung

| 5a) Vpn: |       |      | 5b) Skalen: |       |       |
|----------|-------|------|-------------|-------|-------|
|          | P1    | P2   |             | S1    | S2    |
| Bec      | 3.04  | 1.17 | absto       | -0.12 | 0.05  |
| Eng      | 4.71  | 5.82 | traur       | -0.42 | 0.16  |
| Koh      | 5.12  | 5.43 | zierl       | -0.01 | -0.59 |
| Kru      | 2.57  | 2.41 | grell       | 0.41  | 0.18  |
| Orl      | 8.76  | 4.67 | großz       | 0.15  | 0.24  |
| Pli      | 10.75 | 2.83 | ruhig       | -0.56 | 0.19  |
| Sch      | 5.53  | 4.78 | klein       | -0.09 | -0.63 |
| Wac      | 7.05  | 4.29 | erreg       | 0.42  | -0.08 |
| Bar      | 0.59  | 2.11 | weit        | 0.19  | 0.31  |
| Bie      | -0.31 | 5.62 | schwe       | 0.29  | 0.02  |
| Bru      | 1.03  | 4.24 |             |       |       |
| Dil      | 1.19  | 4.25 |             |       |       |
| Erm      | 0.53  | 1.09 |             |       |       |
| Hey      | 1.49  | 5.56 |             |       |       |
| Lie      | -1.07 | 4.36 |             |       |       |
| Ork      | -0.18 | 3.60 |             |       |       |

| 5c) Reize: |       |       |
|------------|-------|-------|
|            | R1    | R2    |
| vk         | -0.39 | -0.72 |
| km         | -0.30 | -0.22 |
| zg         | 0.86  | 0.43  |
| vg         | -0.17 | 0.51  |



### 7.2.3. Überprüfung der faktorenanalytischen Befunde auf der Datenebene:

Zur Überprüfung der faktorenanalytischen Befunde auf der Datenebene wurden die Reizbeurteilungen über die je Dimension zusammengehörigen Skalen ermittelt. Tabelle 6 zeigt die entsprechenden Mittelwerte getrennt nach den dimensionsspezifischen Personengruppen bzw. nach der Zuordnung der Personen zu den Versuchsbedingungen. Zusätzlich sind in Tabelle 6 die zugehörigen Varianzen der Ratingdaten für die Skalen- und Personengruppen aufgeführt. In Übereinstimmung mit dem Versuchsaufbau ist eine geringe Farbdifferenzierung für die Personen der univariaten experimentellen Bedingung auf der Datenebene nachweisbar (vgl. Tabelle 6a). Ferner zeigen sich in den Ratings einer Person aus der bivariaten Versuchsbedingung vergleichsweise geringere Differenzierungen zwischen den Farbtönen. Die Befunde auf der Datenebene, die der Klassifikation der Versuchspersonen nach der zweiten CANDECOMP-Dimension zugeordnet sind, sind in Tabelle 6b dargestellt. Nach den Urteilsmittelwerten und -varianzen ergeben sich zwar deutliche interindividuelle Unterschiede in den Größendifferenzierungen; jedoch enthält die Personengruppe, für die entsprechend dem zweiten Faktortripel eine nur geringfügige Unterscheidung zwischen den Reizgrößen erwartet wurde, eine Versuchsperson, deren Ratings die größte Kontrastierung zwischen kleinen und großen Quadraten aufweisen.

### 7.3. Die Lösung nach dem Tuckerschen Modell:

#### 7.3.1. Kriterien zur Bestimmung der Anzahl substantieller Faktoren:

Zur Bestimmung der Anzahl substantieller Faktoren wurde für jeden Modus getrennt die Höhe der Eigenwerte der Skalarproduktmatrizen gegen die Anzahl der Faktoren aufgetragen und mittels des Scree-Tests beurteilt. Abbildungen 10 und 11 zeigen die entsprechenden Kurvenverläufe, die im Personenmodus und im Modus der Skalen jeweils zur Auswahl einer drei-

Tabelle 6 : Ausgangsdaten : Urteilsmitelwerte pro Reiz auf den dimensionsentsprechenden Skalengruppen

## 6a) Ergebnisse zu den Farbskalen:

| $\bar{x}$ für<br>F-Skalen: | bivariate<br>Bedingung:<br>Bec, Eng, Koh<br>Orl, Pli, Sch<br>Wac, Kru | univariate<br>Bedingung:<br>Bar, Bie, Bru<br>Dil, Erm, Hey<br>Lie, Ork | einzelne<br>Personen: |       |       |
|----------------------------|-----------------------------------------------------------------------|------------------------------------------------------------------------|-----------------------|-------|-------|
|                            |                                                                       |                                                                        | Bec                   | Kru   | Lie   |
| vk                         | 0.85                                                                  | -0.24                                                                  | 1.19                  | 1.13  | -1.19 |
| km                         | 0.68                                                                  | 0.30                                                                   | -0.81                 | -0.37 | 0.69  |
| zg                         | -2.24                                                                 | 0.02                                                                   | -1.31                 | -1.12 | 0.56  |
| vg                         | 0.73                                                                  | -0.07                                                                  | 0.94                  | 0.37  | 0.31  |
| $s^2$                      | 2.83                                                                  | 0.67                                                                   | 2.02                  | 1.21  | 0.74  |

## 6b) Ergebnisse zu den Größenskalen:

| $\bar{x}$ für<br>G-Skalen: | Personen bzw. Personengruppen: |                           |                  |       |
|----------------------------|--------------------------------|---------------------------|------------------|-------|
|                            | Bec, Kru<br>Erm, Bar           | Eng, Koh, Orl<br>Sch, Wac | Bie, Dil,<br>Hey | Pli   |
| vk                         | -0.47                          | -1.91                     | -1.44            | -1.63 |
| km                         | -0.28                          | -0.66                     | -0.77            | -0.88 |
| zg                         | 0.41                           | 1.09                      | 1.15             | 2.88  |
| vg                         | 0.34                           | 1.49                      | 1.06             | -0.38 |
| $s^2$                      | 0.63                           | 2.76                      | 1.90             | 3.53  |

## 6c) Ergebnisse zu der Skala "abstoßend-sympathisch":

| $\bar{x}$ | Personen bzw. Personengruppen   |       |
|-----------|---------------------------------|-------|
|           | Bec, Eng, Koh,<br>Orl, Pli, Sch | Wac   |
| vk        | 0.46                            | -1.75 |
| km        | 0.63                            | -0.75 |
| zg        | -1.71                           | 4.25  |
| vg        | 0.63                            | -1.75 |

$\bar{x}$  ist der Mittelwert über die entsprechende Skalengruppe pro Reiz für die angegebene Person bzw. Personengruppe.  
 $s^2$  ist die Varianz über die entsprechenden Skalengruppen und Reize für die angegebene Person bzw. Personengruppe.

Abb.10: Tuckersche Loesung

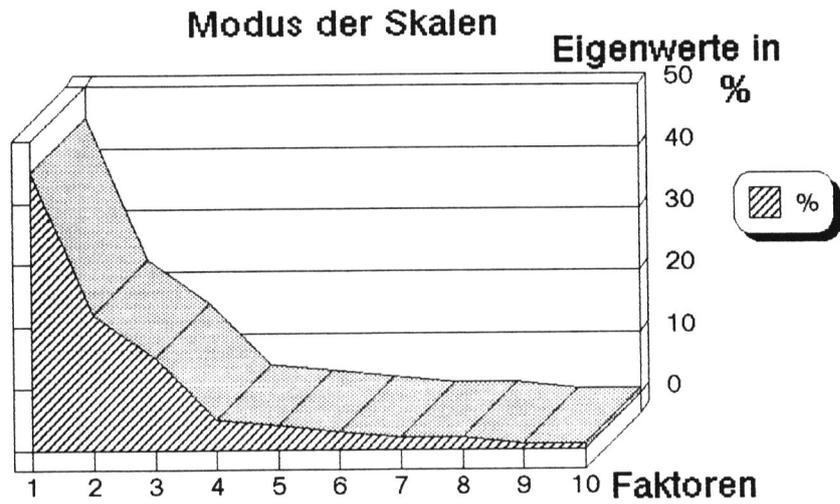
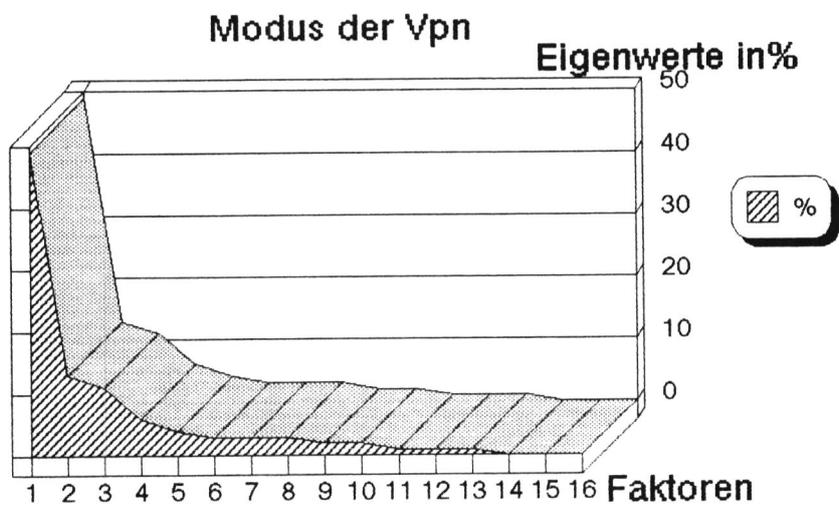


Abb.11: Tuckersche Loesung



dimensionalen Lösung führten. Im Reizmodus betragen die Eigenwerte in Prozent der Spur der Skalarproduktmatrix 64% und 23% für den ersten und zweiten Faktor, der dritte Faktor klärt noch 13% der Varianz auf, so daß alle Reizfaktoren in die Analyse einbezogen wurden.

Die Faktoren des Skalen- und Personenmodus wurden nach dem Varimax-Kriterium rotiert. Ferner wurde in Übereinstimmung mit dem Vorgehen von Orlik (1980) im Modus der Wahrnehmungsreize eine hypothesenorientierte Faktorenrotation gewählt. Dabei wurden die Reizfaktoren durch eine Procrustes-Analyse auf maximale Ähnlichkeit zu einer Referenzstruktur rotiert, die die Reizcharakteristika der Tabelle 3a enthielt. Durch die Auswahl der Reize ergibt sich ein schiefwinkliges Achsensystem, in dem vor allem die Reizfaktoren "Farbton" und "Dunkelstufe" in hohem Ausmaß korrelieren (Orlik, 1980). Zu der Möglichkeit obliquen Rotationstransformationen bemerkt Tucker (1972, S. 10): "First priority for these transformations should be given to establishing meaningful dimensions for the object space ... Note that the transformation is not restricted to being orthonormal, the concept of correlations among the factors not being relevant in the present context." Tabelle 7 zeigt die rotierte Tuckersche Lösung.

### 7.3.2. Interpretation der Lösung nach dem Tuckerschen Modell:

Für den Modus der Wahrnehmungsreize betragen die Korrelationen der Tuckerschen Faktoren mit den vorgegebenen Reizeigenschaften jeweils 1.0, so daß der erste Faktor den Farbton abbildet, der zweite Faktor der Dunkelstufe entspricht und der dritte Faktor als Größendimension interpretierbar ist. Die folgende Tabelle zeigt die Korrelationen der rotierten Faktoren:

|                  | Dunkelstufe (R2) | Größe (R3) |
|------------------|------------------|------------|
| Farbton (R1)     | 0.92             | -0.38      |
| Dunkelstufe (R2) |                  | -0.49      |

Tabelle 7 : die rotierte Lösung nach dem Tuckerschen Modell

## 7a) Vpn:

|     | P1    | P2    | P3    |
|-----|-------|-------|-------|
| Bec | 0.28  | 0.05  | -0.12 |
| Eng | 0.22  | -0.39 | -0.03 |
| Koh | 0.26  | -0.31 | -0.02 |
| Kru | 0.04  | 0.01  | 0.26  |
| Orl | 0.55  | -0.03 | 0.00  |
| Pli | 0.59  | 0.11  | 0.11  |
| Sch | 0.27  | -0.15 | 0.12  |
| Wac | 0.03  | 0.11  | 0.84  |
| Bar | 0.02  | -0.14 | -0.01 |
| Bie | -0.03 | -0.41 | -0.03 |
| Bru | -0.12 | -0.34 | 0.03  |
| Dil | -0.01 | -0.27 | 0.05  |
| Erm | 0.01  | -0.07 | 0.01  |
| Hey | -0.04 | -0.31 | 0.21  |
| Lie | -0.03 | -0.41 | -0.14 |
| Ork | -0.24 | -0.26 | 0.34  |

## 7b) Skalen:

|       | S1    | S2    | S3    |
|-------|-------|-------|-------|
| absto | 0.02  | -0.08 | -0.71 |
| traur | 0.12  | 0.36  | -0.41 |
| zierl | -0.55 | 0.04  | 0.25  |
| grell | 0.21  | -0.47 | -0.17 |
| großz | 0.27  | 0.05  | 0.41  |
| ruhig | 0.13  | 0.61  | -0.07 |
| klein | -0.63 | 0.00  | -0.07 |
| erreg | -0.03 | -0.48 | -0.12 |
| weit  | 0.39  | -0.03 | 0.21  |
| schwe | 0.08  | -0.19 | 0.09  |

## 7c) Reize:

|    | R1    | R2    | R3    |
|----|-------|-------|-------|
| vk | 0.49  | 0.41  | -0.75 |
| km | -0.35 | 0.00  | -0.15 |
| zg | -0.63 | -0.82 | 0.45  |
| vg | 0.49  | 0.41  | 0.45  |

## 7d) Kernmatrix:

|    | S1    |       |        | S2 |       |       | S3    |    |       |        |       |
|----|-------|-------|--------|----|-------|-------|-------|----|-------|--------|-------|
|    | R1    | R2    | R3     | R1 | R2    | R3    | R1    | R2 | R3    |        |       |
| P1 | 1.48  | -2.47 | 7.88   | P1 | -7.81 | 19.62 | -1.48 | P1 | 7.24  | -13.43 | 2.21  |
| P2 | -2.35 | 2.05  | -13.44 | P2 | 9.27  | -8.53 | -1.03 | P2 | -1.39 | 3.82   | 0.83  |
| P3 | 5.35  | -5.11 | 6.58   | P3 | -9.44 | 17.72 | -0.57 | P3 | -0.42 | 6.08   | -1.39 |

Im Modus der Skalen ergeben sich drei wechselseitig orthogonale Faktoren, die die Farbskalen von den Größenskalen trennen und einen zusätzlichen Evaluationsfaktor umfassen. Die Konfiguration des Personenmodus weist eine Einfachstruktur auf, deren erster Faktor höhere Ladungen für fünf Versuchspersonen der bivariaten Versuchsbedingung enthält. Auch der zweite Faktor zeigt keine eindeutige Gruppierung der Personen nach der Zuordnung zu den beiden experimentellen Bedingungen (vgl. Abbildung 12), und der dritte Faktor spiegelt Spezifika in den Urteilen einer einzelnen Person wider. Die Interpretation der zugehörigen Kernmatrix führt bei Betrachtung der Zellen vom Betrage  $\geq 6.0$  im wesentlichen zu drei Strukturhypothesen, die sich jeweils auf einen Skalenfaktor beziehen:

- a) die Kombination aus dem Reizfaktor "Größe" und dem Faktor der Größenskalen weist mit allen Personenfaktoren hohe Kovariationen auf und deutet daraufhin, daß die dimensionsentsprechenden Personengruppen in hohem Maße zwischen den Reizgrößen differenzieren
- b) auch für den Zusammenhang zwischen Farbton und den Farbskalen liegen hohe Kernrelationen mit allen Personenfaktoren vor
- c) die Kernrelationen, die den Evaluationsfaktor einbeziehen, legen die Hypothese nahe, daß für eine Person der bivariaten Versuchsbedingung die Präferenzordnung für die Farbabstufungen (d.h. der Dunkelstufen) gegenläufig ist zu den Ordnungen der anderen Personen.

Ein Vergleich mit den in Tabelle 6 aufgeführten Ergebnissen auf der Datenebene zeigt zunächst, daß die aufgrund der Kernmatrix erwarteten Größenbeurteilungen der Versuchspersonen mit den Ausgangsdaten übereinstimmen. Ebenso kann der faktorenanalytische Befund zu differentiellen Evaluationen der Wahrnehmungsreize nachgewiesen werden (vgl. Tabelle 6c). Die hinsichtlich des Aufbaus des Modellexperimentes wesentliche Strukturhypothese zur Beurteilung der Farbtöne muß jedoch zurückgewiesen werden, da die Tuckersche Konfiguration hier ausgeprägte Differenzierungen von Personen aus beiden Versuchsbedingungen erwarten läßt und damit die Va-

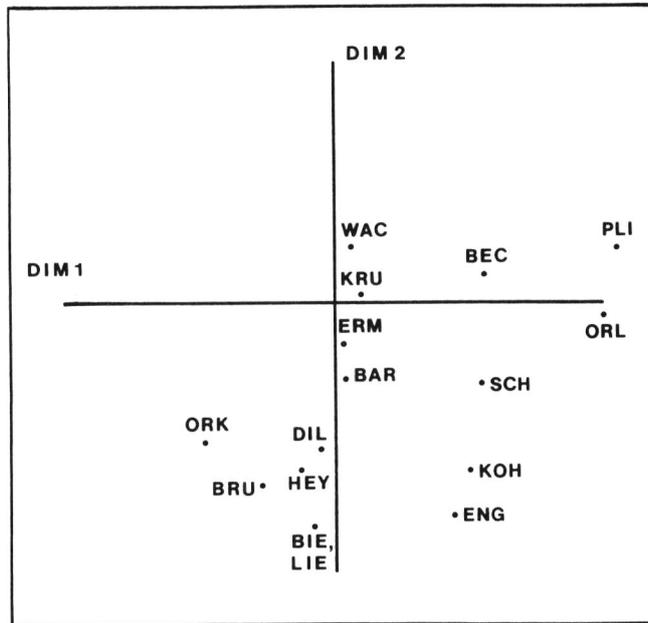


Abb. 12:  
Die Personen-  
faktoren der  
rotierten  
Tuckerschen  
Lösung

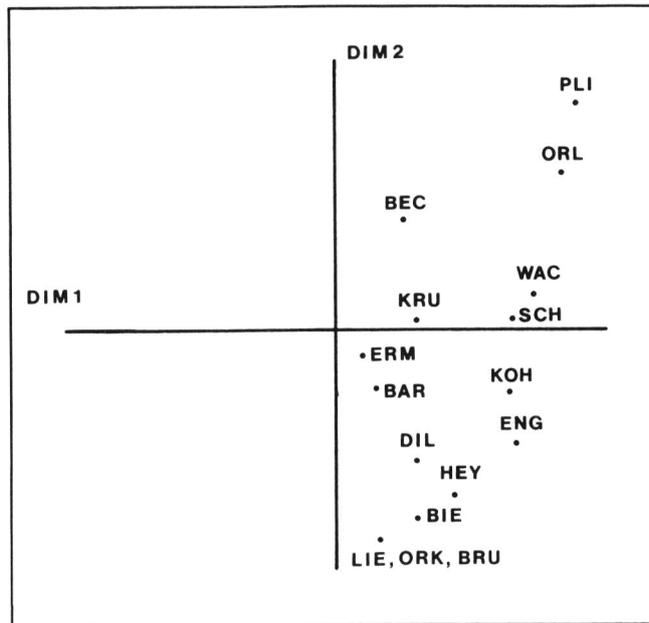


Abb. 13:  
Die Personen-  
faktoren der  
unrotierten  
Tuckerschen  
Lösung

riation der Farbinformationen nicht reproduziert.

Allerdings resultiert die fehlende Reproduktion der grundlegenden Unterschiede zwischen beiden Versuchsbedingungen nicht aus einer möglicherweise unangebrachten Faktorenrotation. Dieser Befund wird deutlich anhand der unrotierten Konfiguration nach dem Tuckerschen Modell, die in Tabelle 8 aufgeführt ist.

Im Modus der Skalen ergibt sich keine Trennung der Farb- und Größenskalen; der erste und der zweite Faktor umfassen jeweils Skalen aus beiden Bereichen. Lediglich der dritte Faktor ist durch die hohe Ladung der Skala "abstoßend-sympathisch" als Evaluationsfaktor interpretierbar. Für den Modus der Wahrnehmungsreize wurden die Korrelationen der Faktoren mit den Reizcharakteristika der Tabelle 3a ermittelt:

|               | Farbton | Dunkelstufe | Größe |
|---------------|---------|-------------|-------|
| 1.Faktor (R1) | -0.64   | -0.83       | 0.86  |
| 2.Faktor (R2) | 0.49    | 0.51        | 0.49  |
| 3.Faktor (R3) | -0.58   | -0.21       | 0.12  |

Nach der Höhe der Korrelationskoeffizienten enthält der erste Faktor eine Abstufung der Reize nach Größe und Dunkelstufe, während die übrigen Dimensionen keine deutlichen Zusammenhänge mit den vorgegebenen Reizeigenschaften aufweisen. Somit ermöglicht die unrotierte Lösung keine Zuordnung der Faktoren zu spezifischen Merkmalen der Skalen und der beurteilten Objekte analog der rotierten Konfiguration.

Der erste Faktor des Personenmodus zeigt angenähert eine Trennung der Versuchspersonen nach der Zuweisung zu den beiden Versuchsbedingungen (vgl. Abbildung 13). Der zweite Faktor enthält hohe Ladungskoeffizienten für vier Personen der univariaten experimentellen Anordnung, während der dritte Faktor erneut Spezifika in den Urteilen einer einzigen Person wiedergibt. Die Interpretation der Zellen der Kernmatrix vom Betrage  $\geq 9.0$  führt zu zwei wesentlichen Aussagen.

Tabelle 8 : die unrotierte Tuckersche Lösung

## 8a) Vpn:

|     | P1   | P2    | P3    |
|-----|------|-------|-------|
| Bec | 0.13 | 0.21  | -0.19 |
| Eng | 0.33 | -0.19 | -0.23 |
| Koh | 0.32 | -0.10 | -0.21 |
| Kru | 0.15 | 0.01  | 0.22  |
| Orl | 0.42 | 0.29  | -0.21 |
| Pli | 0.44 | 0.42  | -0.08 |
| Sch | 0.33 | 0.02  | -0.05 |
| Wac | 0.37 | 0.05  | 0.76  |
| Bar | 0.08 | -0.10 | -0.06 |
| Bie | 0.15 | -0.35 | -0.15 |
| Bru | 0.09 | -0.35 | -0.05 |
| Dil | 0.15 | -0.23 | -0.04 |
| Erm | 0.05 | -0.05 | -0.02 |
| Hey | 0.21 | -0.29 | 0.10  |
| Lie | 0.10 | -0.34 | -0.24 |
| Ork | 0.11 | -0.37 | 0.30  |

## 8b) Skalen:

|       | S1    | S2    | S3    |
|-------|-------|-------|-------|
| absto | -0.07 | 0.14  | -0.69 |
| traur | -0.24 | 0.43  | -0.27 |
| zierl | -0.37 | -0.42 | 0.23  |
| grell | 0.43  | -0.13 | -0.30 |
| großz | 0.24  | 0.12  | 0.41  |
| ruhig | -0.33 | 0.51  | 0.12  |
| klein | -0.46 | -0.42 | -0.09 |
| erreg | 0.28  | -0.32 | -0.26 |
| weit  | 0.34  | 0.20  | 0.20  |
| schwe | 0.20  | -0.10 | 0.03  |

## 8c) Reize:

|    | R1    | R2    | R3    |
|----|-------|-------|-------|
| vk | -0.59 | -0.38 | -0.51 |
| km | -0.25 | -0.06 | 0.83  |
| zg | 0.76  | -0.40 | -0.10 |
| vg | 0.08  | 0.83  | -0.22 |

## 8d) Kernmatrix:

| S1 |       |       | S2    |    |       | S3    |       |    |       |      |       |
|----|-------|-------|-------|----|-------|-------|-------|----|-------|------|-------|
| R1 | R2    | R3    | R1    | R2 | R3    | R1    | R2    | R3 |       |      |       |
| P1 | 20.45 | 0.76  | -1.99 | P1 | -0.14 | 9.47  | 2.22  | P1 | 0.10  | 0.89 | -0.52 |
| P2 | 0.97  | -4.49 | 1.88  | P2 | -9.15 | -0.92 | -1.99 | P2 | 1.56  | 0.63 | -1.38 |
| P3 | -0.19 | -0.78 | -1.31 | P3 | -0.83 | -0.17 | -0.79 | P3 | -9.44 | 3.58 | -0.61 |

Die beiden Kernrelationen, die einen hohen Zusammenhang zu dem ersten Personenfaktor aufweisen, legen etwa die Hypothese nahe, daß Personen aus der bivariaten Versuchsbedingung die Größen- und Farbskalen zur Unterscheidung der Reizgrößen benutzen. Zusätzlich impliziert die unrotierte Konfiguration eine analoge Aussage für vier Versuchspersonen der univariaten experimentellen Anordnung durch die hohe Kovariation des zweiten Personen- und Skalenfaktors mit dem ersten Reizfaktor. Damit findet der Versuchsplan des Modellexperimentes keine Entsprechung in der unrotierten Konfiguration; ferner entsprechen die Faktorladungen der Skalen nicht den Kriterien einer Einfachstruktur, und die Reizfaktoren bilden nur partiell die vorgegebenen Charakteristika der ausgewählten Farben und Größen ab.

#### 7.4. Die Lösung nach dem SUMMAX-Modell:

##### 7.4.1. Kriterien zur Auswahl einer Lösung:

Zur Bestimmung der Anzahl substantieller Faktoren wurde im ersten Auswertungsschritt der faktoriell aufgeklärte Varianzanteil nach sukzessiver Extraktion der SUMMAX-Faktortripel ermittelt. Die Kurve der Faktor-Ähnlichkeitskoeffizienten ist in Abbildung 14 in Abhängigkeit von der Anzahl der Faktortripel dargestellt. Eine deutliche Veränderung des Kurvenverlaufs zeigt sich nach dem zweiten Faktortripel, eine geringe Veränderung deutet sich nach Extraktion von sieben und neun Dimensionen an. Nach dem neunten Faktortripel wurde eine Varianzaufklärung von 80% erreicht, was als ausreichend angesehen wurde. Im folgenden werden zwei Konfigurationen diskutiert, die unrotierte Lösung aus zwei Faktortripeln und die rotierte Lösung nach Extraktion von neun Faktortripeln, um Vergleiche zu der CANDECOMP- und der Tuckerschen Lösung zu ermöglichen.

##### 7.4.2. Interpretation der unrotierten SUMMAX-Lösung aus zwei Faktortripeln:

Tabelle 9 zeigt die unrotierte SUMMAX-Konfiguration, die

## Abb. 14: SUMMAX

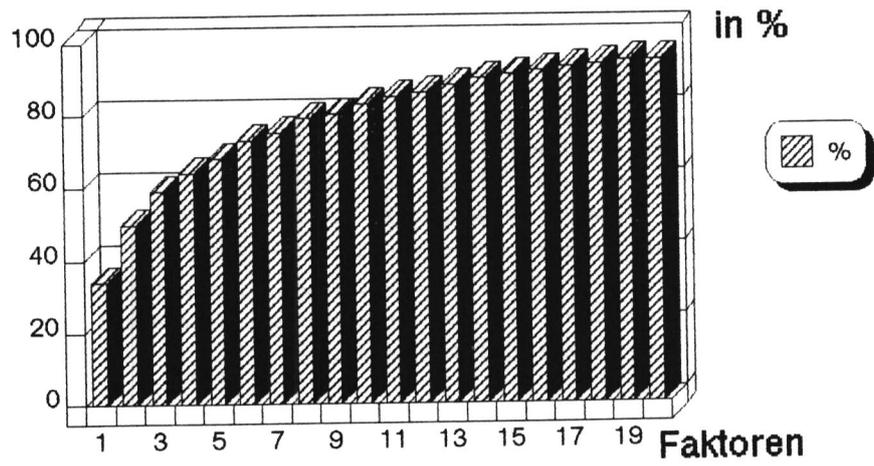
 $R^2$  zwischen Roh- u. geschätzten Daten

Tabelle 9 : die unrotierte SUMMAX-Lösung, reskaliert nach dem CANDECOMP-Rationale

## 9a) Vpn:

|     | P1    | P2    |
|-----|-------|-------|
| Bec | 1.31  | 2.91  |
| Eng | 8.07  | 2.73  |
| Koh | 8.73  | 3.44  |
| Kru | 2.18  | 0.18  |
| Orl | 8.51  | 7.96  |
| Pli | 12.00 | 6.57  |
| Sch | 8.07  | 6.78  |
| Wac | 6.77  | 5.78  |
| Bar | 2.18  | 0.16  |
| Bie | 3.92  | -1.06 |
| Bru | 0.87  | -2.32 |
| Dil | 4.04  | -0.47 |
| Erm | 1.09  | 0.81  |
| Hey | 5.13  | -0.28 |
| Lie | 2.62  | -1.84 |
| Ork | 1.53  | -0.54 |

## 9b) Skalen:

|       | S1    | S2    |
|-------|-------|-------|
| absto | 0.01  | 0.14  |
| traur | 0.08  | 0.44  |
| zierl | 0.53  | -0.43 |
| grell | -0.35 | -0.17 |
| großz | -0.29 | 0.05  |
| ruhig | 0.13  | 0.56  |
| klein | 0.58  | -0.36 |
| erreg | -0.14 | -0.31 |
| weit  | -0.35 | 0.11  |
| schwe | -0.13 | -0.18 |

## 9c) Reize:

|    | R1    | R2    |
|----|-------|-------|
| vk | 0.54  | 0.36  |
| km | 0.34  | 0.18  |
| zg | -0.75 | -0.86 |
| vg | -0.13 | 0.32  |

nach dem Rationale von CANDECOMP/PARAFAC reskaliert wurde. Dabei betragen die Längen der Skalen- und Reizfaktoren jeweils eins, während die Längen der Personenfaktoren die gesamte Varianz eines jeden Faktortripels wiedergeben. Nach der Höhe der Korrelationen der Reizfaktoren mit den vorgegebenen Reizcharakteristika aus Tabelle 3a ist der erste Faktor als Größenfaktor interpretierbar, und der zweite Faktor umfaßt den Farbton und die Dunkelstufe wie die folgende Tabelle verdeutlicht:

|                | Farbton | Dunkelstufe | Größe |
|----------------|---------|-------------|-------|
| 1. Faktor (R1) | 0.56    | 0.76        | -0.88 |
| 2. Faktor (R2) | 0.82    | 0.98        | -0.54 |

Wie auch in der CANDECOMP-Lösung trennen die Faktoren des Skalenmodus näherungsweise die Gruppe der vier Größenskalen von den vier Farbskalen. Die Faktoren des Personenmodus sind in Abbildung 15 dargestellt. Entsprechend der Höhe der Faktorladungen des ersten Personenfaktors ergeben sich zwei Gruppen von Versuchspersonen:

- a) sechs Versuchspersonen sind relativ niedrige "Gewichte" zugeordnet. In Zusammenhang mit dem ersten Skalen- und Reizfaktor resultiert für diese Personen die Hypothese, daß sie in geringerem Maße zwischen den Größen der Wahrnehmungsreize auf den Größenskalen differenzieren als
- b) die Versuchspersonen, denen höhere Gewichte zugeordnet sind.

Im Vergleich mit den in Tabelle 6b aufgeführten Ergebnissen auf der Datenebene führt das erste SUMMAX-Faktortripel zu einer Klassifikation der Versuchspersonen hinsichtlich der Größenbeurteilungen, die die Ausgangsdaten in angemessener Weise repräsentiert.

Der zweite Faktor, der im Skalen- und Reizmodus approximativ der Abstufung der Farbtöne auf den Farbskalen entspricht, impliziert eine Trennung der beiden Versuchsgruppen, wobei eine Person der univariaten Versuchsbedingung zugeordnet

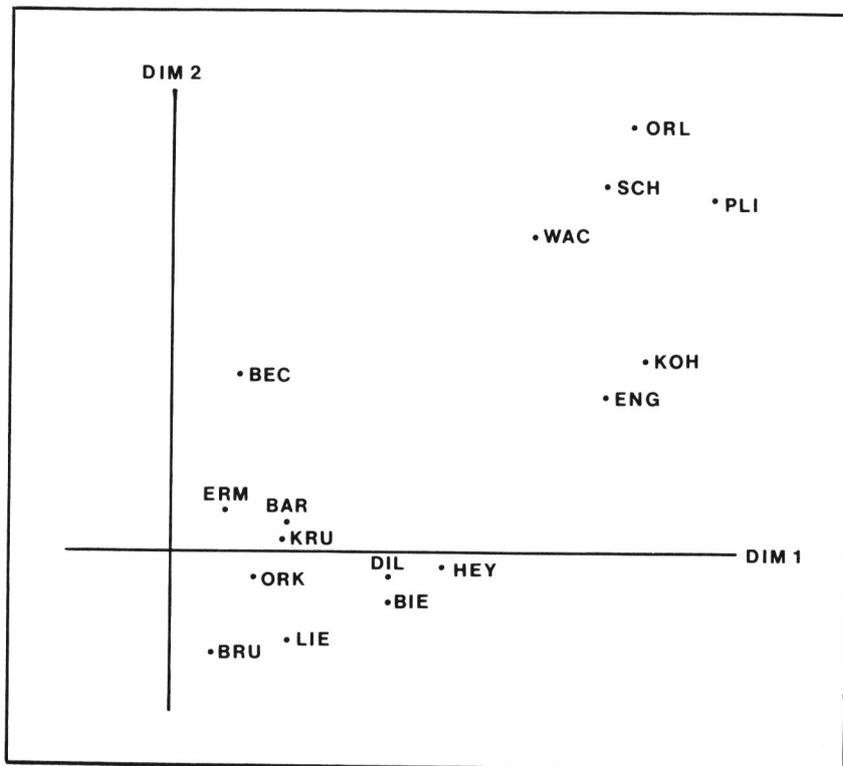


Abb. 15: Die Personenfaktoren der unrotierten SUMMAX-Lösung

wird, der Größen- und Farbinformationen vorlagen. Bereits in der Darstellung der CANDECOMP-Konfiguration konnte gezeigt werden, daß nach den Auswertungen auf der Datenebene die Versuchspersonen nur geringfügig zwischen den Farben der Quadrate auf den Farbskalen differenzieren, denen nur Größeninformationen vorlagen. Dagegen kommt die Abstufung der Farben in den Ratings der Versuchspersonen aus der bivariaten experimentellen Bedingung zum Ausdruck, wobei die Farbdifferenzierungen einer Person aus dieser Gruppe geringer waren, als aufgrund des Versuchsplans erwartet wurde.

#### 7.4.3. Interpretation der SUMMAX-Lösung aus neun Faktortriplets:

Im ersten Auswertungsschritt wurde die Konfiguration aus neun Dimensionen auf Hauptachsen transformiert. Der Kurvenverlauf der Eigenwerte der Faktormatrizen ist in den Abbildungen 16 und 17 für den Personen- und Skalenmodus dargestellt. Nach dem Scree-Test wurden drei Faktoren im Skalenmodus und im Personenmodus beibehalten; die Eigenwerte im Modus der Wahrnehmungsreize betragen 61%, 24% und 15% der Spur der Faktormatrix, und es wurden alle Reizfaktoren in die Analyse einbezogen. Tabelle 10 enthält die Faktormatrizen, die im Skalen- und Personenmodus nach der Rotation auf Hauptachsen anschließend nach dem Varimax-Kriterium weiterrotiert wurden. Analog dem Vorgehen von Orlik (1980) und den Auswertungen im Rahmen des Tuckerschen Modells wurde für den Modus der Reize eine schiefwinklige Procrustes-Rotation gewählt mit den in Tabelle 3a aufgeführten Reizcharakteristika als Zielvektoren. Dabei betrug die Korrelation zwischen den transformierten Reizfaktoren und den vorgegebenen Eigenschaften "Farbton", "Dunkelstufe" und "Größe" jeweils 1.0. Die folgende Tabelle zeigt die Korrelationen zwischen den rotierten Reizdimensionen:

Abb.16: SUMMAX

Modus der Skalen

Eigenwerte in

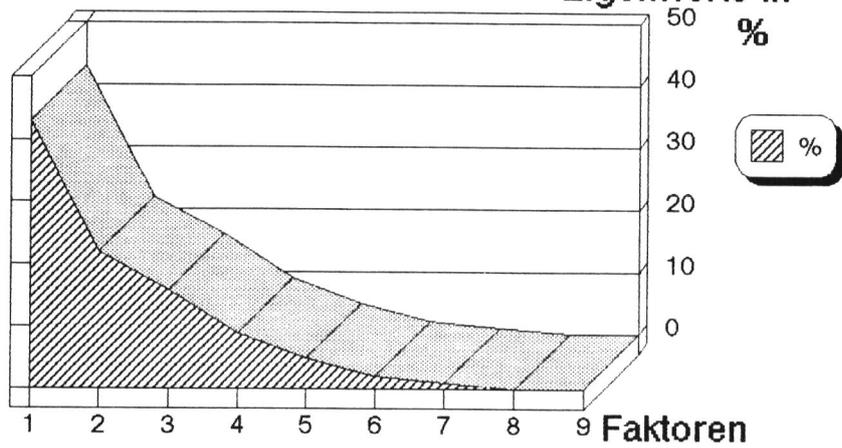


Abb.17: SUMMAX

Modus der Vpn

Eigenwerte in

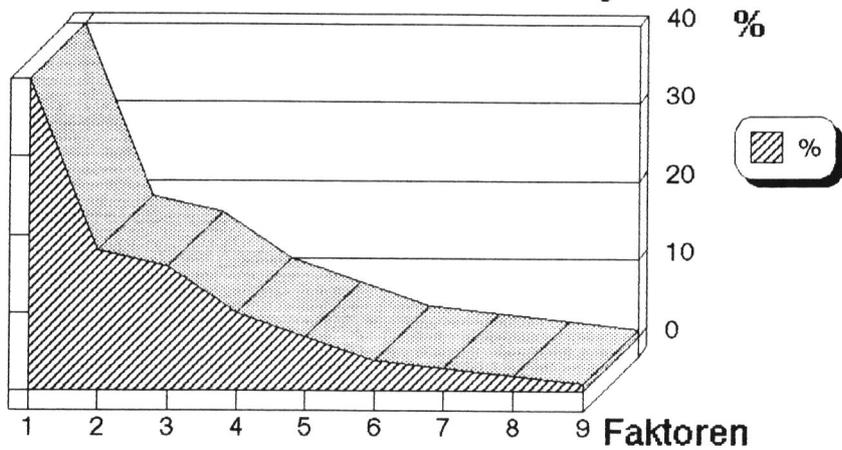


Tabelle 10 : die rotierte SUMMAX-Lösung

10a) Vpn:

|     | P1    | P2     | P3    |
|-----|-------|--------|-------|
| Bec | 0.48  | -0.21  | -0.44 |
| Eng | 0.66* | -0.53  | 0.71* |
| Koh | 1.18* | -0.37  | -0.09 |
| Kru | -0.24 | 0.11   | 0.22  |
| Orl | 1.40* | -0.06  | 0.17  |
| Pli | 1.75* | 0.82*  | -0.02 |
| Sch | 0.97* | -0.65* | 0.55  |
| Wac | 0.63* | 0.59   | 1.65* |
| Bar | 0.08  | -0.44  | -0.04 |
| Bie | 0.29  | -0.84* | 0.05  |
| Bru | -0.18 | -0.42  | 0.13  |
| Dil | 0.14  | -0.56  | -0.06 |
| Erm | 0.20  | -0.07  | 0.00  |
| Hey | 0.05  | -0.64* | 0.61  |
| Lie | 0.25  | -0.41  | -0.01 |
| Ork | -0.18 | -0.35  | 0.64* |

10b) Skalen:

|       | S1     | S2     | S3     |
|-------|--------|--------|--------|
| absto | -0.26  | -0.68  | -1.66* |
| traur | -0.27  | 0.69   | -1.28* |
| zierl | 2.22*  | 0.39   | 0.69   |
| grell | -1.26* | -1.46* | -0.05  |
| großz | -0.58  | 0.51   | 0.61   |
| ruhig | -0.17  | 1.75*  | -0.36  |
| klein | 2.27*  | -0.04  | -0.09  |
| erreg | -0.09  | -1.24* | -0.08  |
| weit  | -1.00* | 0.05   | 0.29   |
| schwe | 0.07   | 0.05   | -0.18  |

10c) Reize:

|    | R1     | R2     | R3     |
|----|--------|--------|--------|
| vk | 2.27   | 2.33   | -4.36* |
| km | -1.62  | 0.00   | -0.87  |
| zg | -2.92* | -4.67* | 2.61*  |
| vg | 2.27   | 2.33   | 2.61*  |

10d) Kernmatrix, normiert im Kombinationsmodus Vpn x Skalen:

| S1 |       |      | S2    |    |       | S3   |       |    |       |       |       |
|----|-------|------|-------|----|-------|------|-------|----|-------|-------|-------|
| R1 | R2    | R3   | R1    | R2 | R3    | R1   | R2    | R3 |       |       |       |
| P1 | -0.33 | 0.24 | -0.91 | P1 | -0.33 | 0.94 | 0.08  | P1 | 0.53  | -0.84 | 0.13  |
| P2 | 0.09  | 0.16 | 0.98  | P2 | 0.69  | 0.50 | -0.49 | P2 | -0.55 | 0.81  | 0.05  |
| P3 | -0.75 | 0.52 | -0.40 | P3 | -0.69 | 0.73 | 0.03  | P3 | 0.87  | 0.48  | -0.10 |

Substantielle Faktorladungen sind mit "\*" hervorgehoben.

|                  | Dunkelstufe (R2) | Größe (R3) |
|------------------|------------------|------------|
| Farbton (R1)     | 0.91             | -0.38      |
| Dunkelstufe (R2) |                  | -0.49      |

Die zugehörige Kernmatrix wurde im Kombinationsmodus (Versuchspersonen x Skalen) normiert (vgl. Orlik, 1980), und die in Tabelle 10 mit "\*" hervorgehobenen Faktorladungen sind substantiell nach dem von Orlik angegebenen Trennkriterium (1980, S. 202; vgl. Abschnitt 6). Im Modus der Skalen wird eine Einfachstruktur deutlich, die auch in dieser Lösung approximativ die Farb- und die Größenskalen voneinander trennt. Im Vergleich zu der Konfiguration aus zwei Faktortripeln tritt hier zusätzlich ein Evaluationsfaktor auf, der hauptsächlich die Skala "abstoßend-sympathisch" umfaßt. Die Faktoren des Personenmodus sind in Abbildung 18 dargestellt. Der erste Faktor enthält angenähert eine Trennung der beiden Versuchsgruppen. Auf dem zweiten Faktor sind jeweils drei Personen aus beiden experimentellen Bedingungen durch hohe Ladungen repräsentiert, und der dritte Faktor gibt im wesentlichen Urteilsspezifika einer Versuchsperson wieder. Die Interpretation der Kernmatrix führt zu drei Strukturhypothesen, die Aussagen zu der Größen- und Farbdifferenzierung wie auch zu der Evaluation der Wahrnehmungsreize implizieren. Der Faktor der Größenskalen weist hohe Kovariationen mit dem Reizfaktor "Größe" und zwei der Personenfaktoren auf. Diese Kernrelationen deuten daraufhin, daß die dimensionsentsprechenden Personengruppen ausgeprägte Unterscheidungen zwischen den Reizgrößen auf den Größenskalen vornehmen. Dergleiche faktorenanalytische Befund war bereits in der zweidimensionalen SUMMAX-Lösung enthalten und wird durch die in Tabelle 6b aufgeführten Ergebnisse auf der Datenebene bestätigt. Der Versuchsplan des Modellexperimentes kommt zum Ausdruck in den Kernrelationen, die den zweiten Reizfaktor mit dem Faktor der Farbskalen verknüpfen. Hier zeigt sich eine hohe Kovariation ausschließlich mit dem Personenfaktor, der die Personen der bivariaten Versuchsbedingung umfaßt.

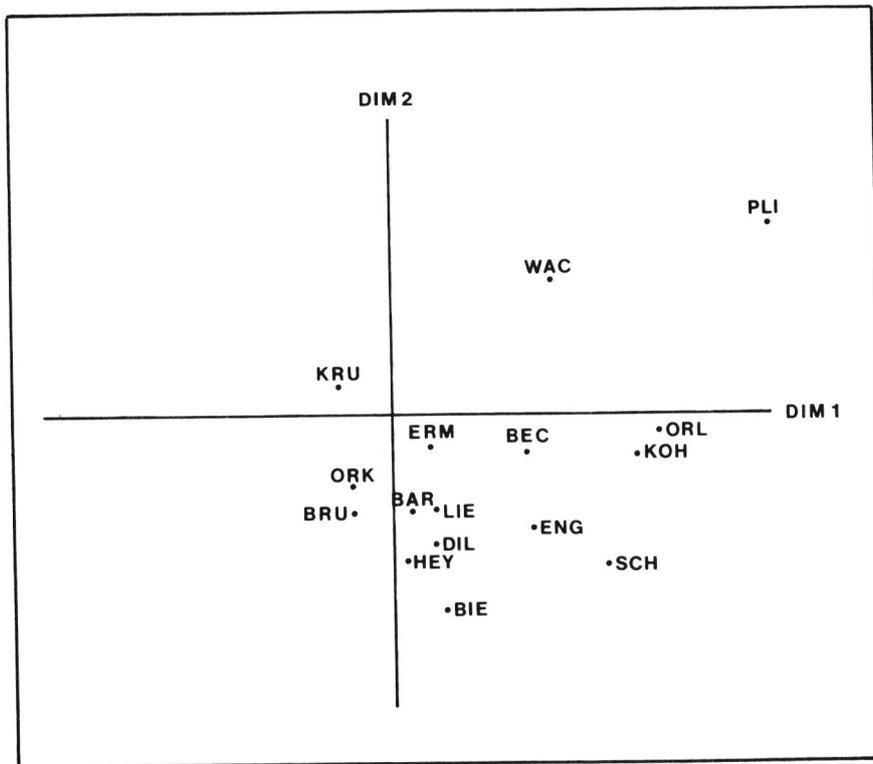


Abb.18: Die Personenfaktoren der rotierten SUMMAX-Lösung

Die rotierte Lösung führt damit zu Aussagen, die hinsichtlich der Größen- und Farbbeurteilungen mit der zweidimensionalen Konfiguration übereinstimmen. Unter Betrachtung des Evaluationsfaktors wird ein weiteres Merkmal der Datenstruktur deutlich, das aus der höheren Varianzaufklärung der neun-Faktor-Lösung resultiert. Die Kernrelationen, die das Gefüge der Personen- und Reizfaktoren aufgrund des Evaluationsfaktors wiedergeben, liefern zusätzlich Hypothesen über die individuellen Präferenzordnungen der Versuchspersonen:

da die Kernrelationen vom Betrage  $\geq 0.84$  für den ersten und dritten Personenfaktor entgegengesetzte Vorzeichen haben, wird hier nahegelegt, daß eine Versuchsperson der bivariaten experimentellen Bedingung die Farben der Quadrate gegenläufig zu den anderen Personen dieser Versuchsgruppe einstuft. Eine analoge Hypothese wurde auch in der Konfiguration des Tuckerschen Modells impliziert. In Tabelle 6c sind die den SUMMAX-Kernrelationen zugeordneten Ergebnisse zu der Evaluation der Wahrnehmungsreize aufgeführt, die den faktorenanalytischen Befund auf der Datenebene stützen.

#### 7.5. Die dreimodalen Konfigurationen im Vergleich:

Im Rahmen der vergleichenden Reanalyse des Modellexperimentes von Orlik reproduzierten die dreimodalen Modelle in sehr unterschiedlicher Weise charakteristische Strukturen des zugrunde liegenden Datensatzes. Dabei wies die unrotierte Tuckersche Konfiguration keine Bezüge zu der experimentellen Variation der dargebotenen Farbinformationen auf. Die Skalenfaktoren enthalten keine Trennung der Größen- und Farbskalen, und die Reizfaktoren korrespondieren nur teilweise zu den vorgegebenen Charakteristika des Reizmaterials. Nach anschließenden Varimax- bzw. hypothesenorientierten Rotationen der Tuckerschen Faktormatrizen wurde zwar der Zusammenhang der beiden Skalengruppen mit den dargebotenen Reizeigenschaften deutlich; die experimentelle Variation der dargebotenen Farbinformation fand dennoch keine Entsprechung

in der rotierten Konfiguration. Dabei ergab sich eine wesentlich höhere Übereinstimmung mit der Versuchsanordnung für die CANDECOMP-Lösung, die angenähert eine Trennung der Personengruppen nach unterschiedlicher Differenzierung der Reize auf den Farbskalen enthielt. Die Aussagen, die die CANDECOMP-Lösung hinsichtlich der Größenbeurteilungen nahelegte, konnten bis auf einen Einzelbefund durch die Ergebnisse auf der Datenebene gestützt werden. Ferner führte die zweidimensionale SUMMAX-Lösung zu einer angemessenen Klassifikation der Versuchspersonen bezüglich der Größen- und Farbbeurteilungen. Die Extraktion weiterer SUMMAX-Faktoren mit anschließenden Faktorrotationen erbrachte zunächst Aufschlüsse über die Struktur der Farb- und Größenbeurteilungen, die auch schon in der unrotierten Konfiguration aus zwei Faktortripeln enthalten sind. Im Gegensatz dazu trat ein zusätzlicher Evaluationsfaktor auf, der unterschiedliche Präferenzordnungen der Versuchspersonen aufzeigte.

8.

Dimensionen der Farbwahrnehmung: ein  
Vergleich der MDS-Modelle zur Struk-  
turanalyse individueller Differenzen

Die "three-way two mode multidimensional scaling"-Modelle stellen eine Erweiterung des Instrumentariums der klassischen Multidimensionalen Skalierung dar, indem sie eine Strukturanalyse individueller Differenzen zwischen Personen in der Wahrnehmung bzw. Beurteilung dargebotener Stimuli gestatten. Wie bereits in Abschnitt 2 und 3 ausgeführt wurde, ist in diesem Rahmen das INDSCAL-Modell von Carroll und Chang (1970) aus dem übergeordneten CANDECOMP-Ansatz entwickelt worden. Auch das Tuckersche MDS-Verfahren und SUMMAX erlauben eine strukturelle Darstellung der Ähnlichkeitsurteile mehrerer Personen. Unter Spezifikation einer Abbildung, welche die erhobenen Ähnlichkeitsurteile jeder Person mit euklidischen Distanzen verknüpft, und anschließender Transformation der Distanzen in Skalarprodukte ergibt sich eine Verbindung der MDS-Modelle zu den Verfahren der dreimodalen Faktorenanalyse, indem jeweils zwei Modi der dreimodalen Grundgleichungen miteinander identifiziert werden. Im folgenden werden die Konfigurationen der MDS-Modelle vergleichend diskutiert anhand der Daten einer von Helm (1964) durchgeführten Studie, die ebenfalls dem Gebiet der Farbwahrnehmung angehört.

Helm verwendete in seiner Untersuchung zur Struktur der Farbwahrnehmung zehn Farbreize aus dem gesamten Bereich des Seh-Spektrums, die sich hinsichtlich ihres Farbtones, aber nicht in Helligkeit und Sättigung unterschieden. Die ausgewählten Farben lassen sich nach ihrer Ähnlichkeit auf einem geschlossenen Kreis (der "Farbenkreis") anordnen, wobei jeder Farbton seinen beiden Nachbarn ähnlicher ist als jedem anderen Ton (Pohl, 1967; Helm, 1964). Jedes mögliche Tripel von Farbreizen wurde nacheinander der Versuchsperson dargeboten mit der Instruktion, die Stimuli hinsichtlich ihrer Ähnlichkeit zu beurteilen. Dabei hatte die Versuchsperson zunächst aus jedem Tripel die beiden Farben auszuwählen, die einander am wenigsten ähnlich waren und diese Farben auf die Punkte g und h einer der Abbildung 19 entsprechenden Tafel zu plazieren. Im nächsten Schritt wurde die dritte Farbe von der Versuchsperson innerhalb der Begrenzungen i und j so

Tabelle 11 : Farbreize und Versuchspersonen des Helmschen Experimentes

11a) Farbreize :

|                |                 |                 |
|----------------|-----------------|-----------------|
| Rot-Purpur (A) | Grün-Gelb 1 (G) | Blau (M)        |
| Rot-Orange (C) | Grün-Gelb 2 (I) | Purpur-Blau (O) |
| Gelb (E)       | Grün (K)        | Purpur 2 (Q)    |
|                |                 | Purpur 1 (S)    |

11b) Versuchspersonen :

Farbtüchtige Vpn :

|    |    |
|----|----|
| CB | RB |
| DT | RW |
| MS | HG |
| RS | SW |
| JH | MG |

Farbenfehlsichtige Vpn :

|    |
|----|
| CP |
| NF |
| JV |
| JN |

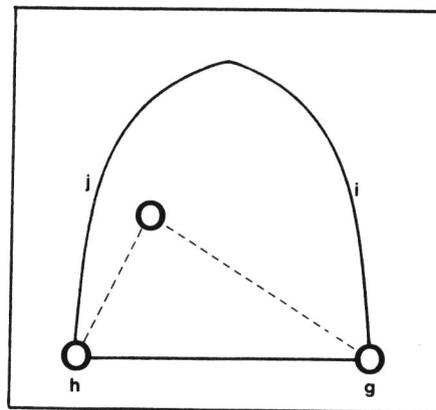


Abb.19:  
Schema der Helmschen Versuchsan-  
ordnung

lokalisiert, daß die Abstände zu den beiden zuerst gewählten Farben das Verhältnis der wahrgenommenen Ähnlichkeiten repräsentieren (vgl. Helm (1964, S.259) zu den Einzelheiten der Versuchsanordnung. Helm verwendete in seiner Untersuchung eine von Torgerson (1958, Kap. 11.5, S.293ff) vorgeschlagene Variante der "ratio-judgment"-Methoden zur Skalierung von Distanzen). An dem Experiment nahmen 14 Versuchspersonen teil, die zuvor mit den "pseudoisochromatischen Tafeln" auf ihre Farbtüchtigkeit untersucht wurden. Dabei waren zehn Probanden normal farbsichtig, während vier Vpn in unterschiedlichem Maße eine Deuteranomalie, d.h. eine herabgesetzte Farbempfindlichkeit mit Schwierigkeiten in der Differenzierung der Farben "Rot" und "Grün" zeigten (vgl. Tabelle 11). Die Versuchspersonen NF und CP wiesen eine mäßige Farbschwächung auf, die Personen JN und JV waren ausgeprägt farbenfehlsichtig. Ferner führten die Probanden NF und RB im Abstand von vier Wochen das Experiment zweimal durch. Die Distanz-Schätzwerte für alle Versuchspersonen und Reizpaare aus der Untersuchung wurden von Helm (1959, Appendix, Tabelle A.1) veröffentlicht.

Helm transformierte die Distanzen in Skalarprodukte für jeden Probanden und analysierte die resultierenden Matrizen für jede Versuchsperson separat mit der Hauptachsentransformation. Aus dieser Auswertung wie auch aus der Analyse durch das Points-of-Views-Modell (Helm & Tucker, 1962) ergaben sich deutliche Unterschiede zwischen den beiden Gruppen von Versuchspersonen. Erwartungsgemäß zeigte sich in den Konfigurationen der Farbreize für die normal farbsichtigen Probanden eine "Gelb-Blau"- und eine "Rot-Grün"-Dimension; das Muster der Farbtöne wies im Stimulusraum eine dem Farbkreis entsprechende Struktur auf. Dagegen zeigten die Konfigurationen der farbenfehlsichtigen Probanden eine dem Grad der Farbschädigung entsprechende Deformation der "Rot-Grün"-Dimension auf, verbunden mit schiefwinkligen Farbachsen: "... we have a more nearly one-dimensional structure. ... Thus, we observe a marked contraction of distances in the red-green direction of the stimulus-space" (Helm & Tucker, 1962, S. 443).

### 8.1. Normierung des Datenkörpers :

Entsprechend dem Vorgehen von Carroll und Chang (1970) und von Tucker (1972) wurde die Transformation der Distanz-Schätzwerte in Skalarprodukte für jede Versuchsperson getrennt nach der Formel von Torgerson (1958, S.257-258; vgl. Abschnitt 3) vorgenommen. Zusätzlich wurden die resultierenden Skalarproduktmatrizen nach den Empfehlungen von Carroll und Chang (1970) so normiert, daß die Quadratsummen der Werte für jede individuelle Matrix eins betragen (vgl. auch Dunn-Ranking, 1983; Kroonenberg, 1983b; Harshman & Lundy, 1984).

### 8.2. Die INDSCAL-Lösung :

Wie auch im CANDECOMP-Modell wird zur Extraktion der Dimensionen der "Alternating-Least-Squares"-Algorithmus verwendet, so daß die Beurteilung der Substantialität der Dimensionen anhand der von Harshman (1984) diskutierten Kriterien vorgenommen wurde. Aufgrund der geringen Anzahl von Vpn mußten Harshmans Split-Half-Methoden zur Abschätzung der Anpassungsgüte der INDSCAL-Lösungen außer acht gelassen werden.

#### 8.2.1. Verwendete Auswahlkriterien :

Scree-Test der quadrierten Korrelationen  $R^2$  zwischen den Ausgangsdaten und dem rückgerechneten Datenkörper :

Die Höhe von  $R^2$  beträgt für die zwei- bis fünfdimensionale Lösung 0,91, 0,93, 0,94 und 0,94 (Programmbeschränkung: Anzahl der Dimensionen  $\leq 5$ ). Der faktoriell aufgeklärte Varianzanteil beträgt bereits nach Extraktion von zwei Dimensionen 91% und erfährt durch Hinzunahme weiterer Dimensionen keine wesentliche Veränderung.

#### Konvergenz der Lösungen :

Die zwei- und dreidimensionalen Lösungen konvergieren in jeweils neun Iterationen, die vier- und fünfdimensionalen

Lösungen benötigen 10 bzw. 14 Iterationsschritte. In allen Lösungen ändert sich R zwischen der letzten und vorletzten Iteration um 0,001%. Somit ergibt sich hinsichtlich des Konvergenzverhaltens des Algorithmus kein Hinweis auf die Angemessenheit einer spezifischen Konfiguration.

Überprüfung der Faktorinterkorrelationen innerhalb eines Modus :

Tabelle 12 zeigt die Faktorinterkorrelationen für die zwei- bis vierdimensionale Lösung. Die drei- und vierdimensionalen Konfigurationen weisen relativ hohe Korrelationen zwischen der ersten und dritten Dimension in allen Modi auf, wobei sich die entsprechenden Werte in der hier nicht aufgeführten fünfdimensionalen Lösung nicht reduzieren. Da nach Harshman (1984) Dimensionen, die in einem Modus hoch korrelieren, in den übrigen Modi nur geringfügige Werte annehmen sollten, wurde die zweidimensionale Konfiguration ausgewählt, die auch von Carroll und Wish (1974) diskutiert wurde.

8.2.2. Interpretation der zweidimensionalen Konfiguration :

Tabelle 13 zeigt die INDSCAL-Dimensionen, die in den Abbildungen 20 und 21 graphisch dargestellt sind. Im Modus der Farbreize wird eine Struktur deutlich, die der Anordnung der Farbtöne auf dem Farbkreis entspricht. Dabei ergibt sich auf der ersten Dimension eine Differenzierung der Farben "Blau", "Purpur-Blau" versus "Gelb", "Gelb-Grün". Auf der zweiten Dimension zeigen die Farbreize "Rot", "Rot-Orange" versus "Grün", "Blau" hohe Ausprägungen; die Farbachsen sind angenähert interpretierbar als "Blau-Gelb"- und "Rot-Grün"-Dimensionen.

Die dimensionale Repräsentation der Versuchspersonen reflektiert das unterschiedliche Ausmaß der Farbenfehlsichtigkeit (vgl. Abbildung 21); während die farbtüchtigen Versuchspersonen auf beiden Dimensionen etwa gleich hohe Gewichte oder "salienes" aufweisen, wird die ausgeprägte Fehlsichtigkeit der Probanden JN und JV an den geringen Ladungen auf der zweiten Dimension deutlich. Auch die mäßig farbfehlsichtigen Versuchspersonen CP und NF zeigen entsprechend ihrer Gewich-

Tabelle 12 : Korrelationen der INDSCAL-Faktoren

12a) zweidimensionale Lösung :

| Personenmodus : |         |
|-----------------|---------|
| 1               | 2       |
| 1               | 1 -0.96 |
| 2               | 1       |

Farbmodus :

| Farbmodus : |         |
|-------------|---------|
| 1           | 2       |
| 1           | 1 -0.16 |
| 2           | 1       |

12b) dreidimensionale Lösung :

| Personenmodus : |              |   |
|-----------------|--------------|---|
| 1               | 2            | 3 |
| 1               | 1 -0.92 0.59 |   |
| 2               | 1 -0.78      |   |
| 3               | 1            |   |

Farbmodus :

| Farbmodus : |               |   |
|-------------|---------------|---|
| 1           | 2             | 3 |
| 1           | 1 -0.15 -0.51 |   |
| 2           | 1 0.13        |   |
| 3           | 1             |   |

12c) vierdimensionale Lösung :

| Personenmodus : |                    |   |   |
|-----------------|--------------------|---|---|
| 1               | 2                  | 3 | 4 |
| 1               | 1 -0.89 -0.89 0.76 |   |   |
| 2               | 1 0.67 -0.65       |   |   |
| 3               | 1 -0.23            |   |   |
| 4               | 1                  |   |   |

Farbmodus :

| Farbmodus : |                   |   |   |
|-------------|-------------------|---|---|
| 1           | 2                 | 3 | 4 |
| 1           | 1 0.01 0.69 -0.45 |   |   |
| 2           | 1 -0.59 0.14      |   |   |
| 3           | 1 -0.46           |   |   |
| 4           | 1                 |   |   |

Tabelle 13 : zweidimensionale INDSCAL-Lösung

13a) Vpn :

|     | P1   | P2   |
|-----|------|------|
| CB  | 0.79 | 0.51 |
| DT  | 0.63 | 0.74 |
| MS  | 0.69 | 0.67 |
| RS  | 0.73 | 0.62 |
| NF1 | 0.83 | 0.37 |
| JH  | 0.74 | 0.61 |
| RB1 | 0.54 | 0.77 |
| CP  | 0.86 | 0.32 |
| JN  | 0.96 | 0.08 |
| RW  | 0.72 | 0.61 |
| HG  | 0.74 | 0.60 |
| SW  | 0.66 | 0.69 |
| RB2 | 0.58 | 0.77 |
| NF2 | 0.85 | 0.37 |
| MG  | 0.79 | 0.46 |
| JV  | 0.93 | 0.09 |

13b) Farbreize :

|   | S1    | S2    |
|---|-------|-------|
| A | -0.26 | 0.40  |
| C | 0.22  | 0.48  |
| E | 0.44  | 0.21  |
| G | 0.38  | -0.17 |
| I | 0.31  | -0.34 |
| K | 0.14  | -0.40 |
| M | -0.14 | -0.37 |
| O | -0.35 | -0.19 |
| Q | -0.40 | 0.11  |
| S | -0.35 | 0.26  |

Abb.20:  
Konfiguration der Farbreize

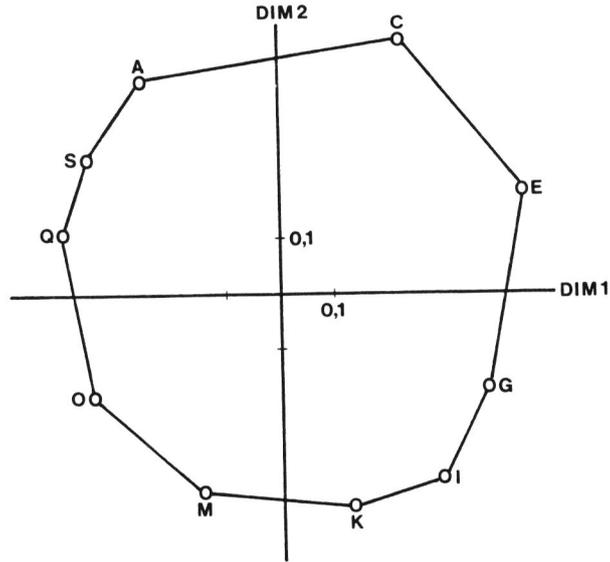


Abb.21:  
Konfiguration der Vpn

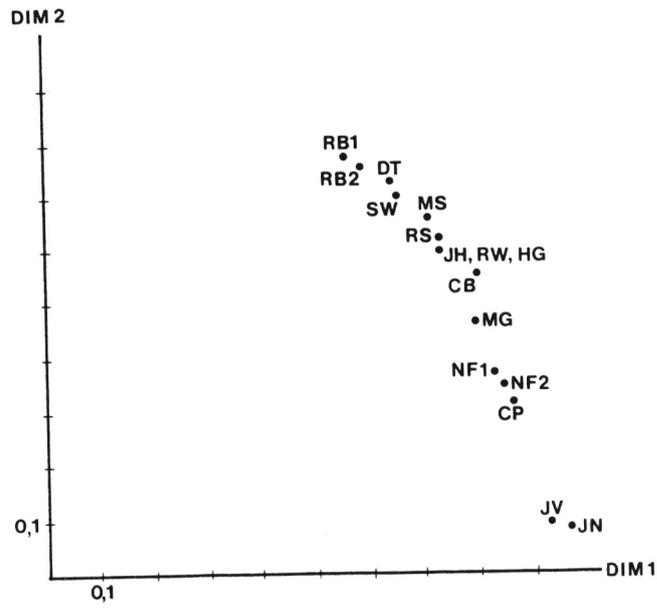


Abb. 22:  
Konfiguration der Vp JV

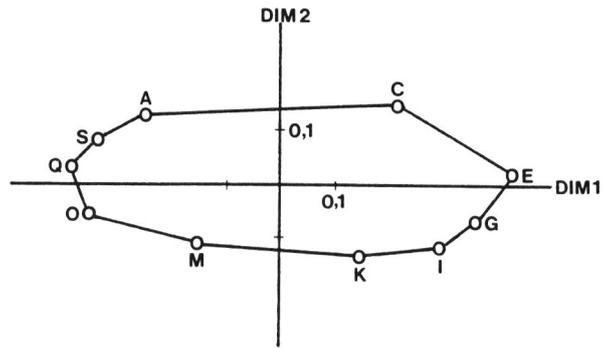
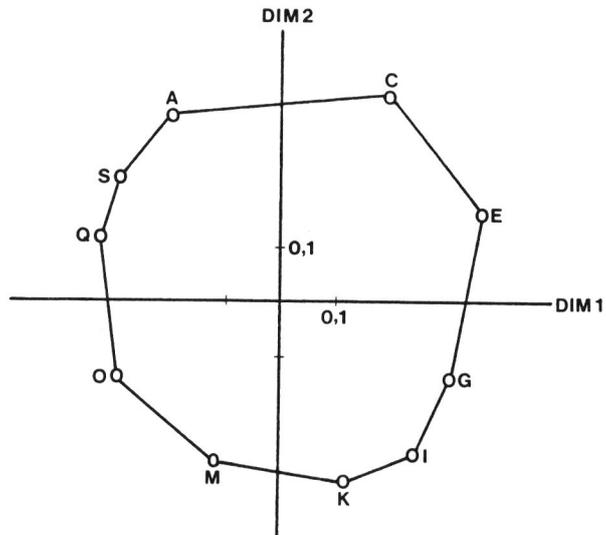


Abb. 23:  
Konfiguration der Vp MS



te eine Deformation der "Rot-Grün"-Dimension, jedoch in weniger extremer Ausprägung. Der Effekt der differentiellen Dimensionsgewichtung läßt sich an den individuellen Reizräumen ("private perceptual spaces", Carroll & Wish, 1974) aufzeigen, die sich aus der Gewichtung der Farbdimensionen mit den individuellen "saliences" ergeben (vgl. Abschnitt 3). Abbildung 23 stellt die entsprechende Konfiguration der normal farbsichtigen Versuchsperson MS dar. Beide Farbachsen dienen nach der INDSICAL-Analyse in gleicher Weise zur Differenzierung der Farbtöne. Für die farbenfehlsichtige Versuchsperson JV erscheinen die Rot- und Grüntöne wesentlich ähnlicher; die Konfiguration ist entlang der "Rot-Grün"-Dimension deutlich gestaucht (vgl. Abbildung 22).

Zusammenfassend werden in der zweidimensionalen INDSICAL-Lösung die aus dem Helmschen Untersuchungsdesign erwarteten Unterschiede zwischen den Probanden reproduziert. Im Personenmodus zeigt sich eine Abstufung der Höhe individueller "saliences", die dem Grad der Farbenfehlsichtigkeit entspricht. Die Ladungen der Farbreize induzieren eine Anordnung auf einem geschlossenen Kreis, wobei die zugrunde liegenden Achsen angenähert einer "Blau-Gelb"- und einer "Rot-Grün"-Dimension entsprechen.

### 8.3. Die Lösung nach dem Tuckerschen Modell :

Zur Bestimmung der angemessenen Dimensionalität der Konfiguration wurde für den Personenmodus die Höhe der Eigenwerte gegen die Anzahl der extrahierten Faktoren aufgetragen und mittels des Scree-Tests beurteilt (Tucker, 1966). Entsprechend Abbildung 24 wurde nach dem Kurvenverlauf der Eigenwerte eine zweidimensionale Lösung im Modus der Personen ausgewählt. Im Modus der Farbreize ergaben sich folgende Prozentwerte für die Eigenwerte: 62%, 30%, 3%, 2%, 1%, 0,7%, 0,5%, 0,4% und 0,3%. Eine deutliche Veränderung der Höhe der Eigenwerte zeigt sich nach der zweiten Dimension.

# Abb. 24: Tuckersche Loesung

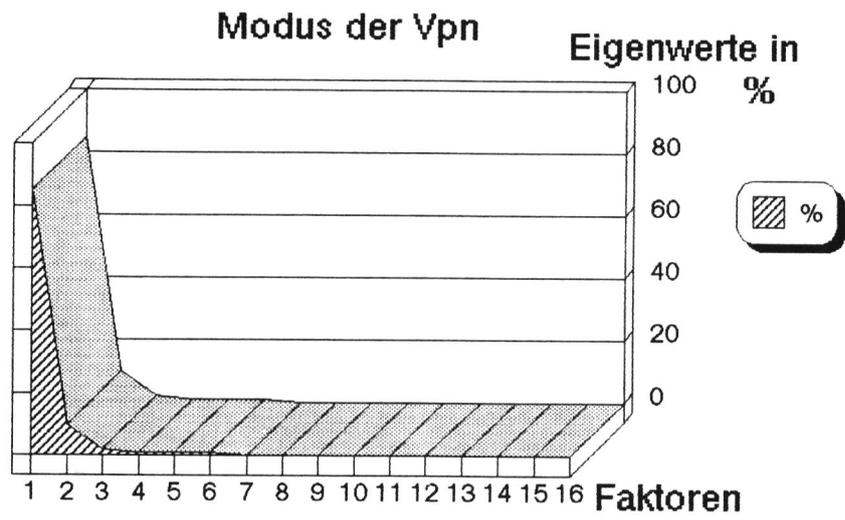


Tabelle 14 enthält die Lösung nach dem Tuckerschen Modell, wobei die angegebenen Werte im Personenmodus und im Modus der Farbreize die Ladungen auf den extrahierten Eigenvektoren sind. Die Eigenwerte wurden hier für einen späteren Auswertungsschritt zusätzlich aufgeführt. Die Kernmatrix aus dem Tuckerschen MDS-Modell zur Strukturanalyse individueller Differenzen weist symmetrische Zellen auf. Werden die Dimensionen des Personenmodus mit  $m$ , die Dimensionen des Farbmodus mit  $p$  gekennzeichnet, so gilt:  $g_{pp'm} = g_{p'pm}$ . Die Diagonalwerte der Kernmatrix werden vor der Interpretation radiziert und zeigen differentielle Stauchungen bzw. Streckungen der Reizdimensionen an (vgl. Abschnitt 3). Die Werte außerhalb der Diagonalen werden nach der Normierung  $g_{pp'm}/(g_{ppm} g_{p'p'm})^{1/2}$  als Winkel zwischen den Reizachsen  $b_p$  und  $b_{p'}$  aus der Perspektive der idealisierten Person  $m$  interpretiert. Tabelle 14c zeigt die normierte Kernmatrix. Drei Zellen sind nicht als reelle Werte darstellbar, da die Kernrelation, die die zweite Reizdimension mit dem zweiten Personenfaktor verknüpft, ein negatives Vorzeichen aufwies. Dabei sprechen negative bzw. nicht reelle Dimensionsgewichte prinzipiell dagegen, daß die vorliegende Konfiguration eine angemessene Repräsentation der Ausgangsdaten ist (Carroll & Chang, 1970; Kroonenberg, 1983b).

Tabelle 15 zeigt die Lösung nach dem Tuckerschen Modell, deren Farb- und Personenachsen nach dem Varimax-Kriterium rotiert wurden. Die zugeordnete Kernmatrix enthält positive Dimensionsgewichte und deutet zunächst daraufhin, daß die Struktur der unrotierten Dimensionen die mangelhafte Repräsentation verursachten. Allerdings führt die Interpretation der Kernmatrix aus der rotierten Konfiguration nicht zu einer Reproduktion der diagnostizierten Farbenfehlsichtigkeit der entsprechenden Personengruppe. Zunächst wird im Modus der Farbreize eine "Rot-Grün-Gelb"-Dimension und eine "Gelb-Orange-Blau"-Achse deutlich. Die Konfiguration im Personenmodus weist angenähert eine Einfachstruktur auf, deren erste Dimension hohe Ladungen für die normal farbsichtigen Versuchspersonen enthält, während der zweite Faktor

Tabelle 14 : die unrotierte Lösung nach dem Tuckerschen Modell

| 14a) Vpn :  |       |       | 14b) Farbreize : |       |       |
|-------------|-------|-------|------------------|-------|-------|
|             | P1    | P2    |                  | S1    | S2    |
| CB          | 0.25  | -0.04 | A                | 0.33  | 0.30  |
| DT          | 0.23  | -0.31 | C                | -0.10 | 0.54  |
| MS          | 0.25  | -0.14 | E                | -0.37 | 0.37  |
| RS          | 0.26  | -0.13 | G                | -0.40 | -0.02 |
| NF1         | 0.25  | 0.21  | I                | -0.37 | -0.23 |
| JH          | 0.26  | -0.08 | K                | -0.22 | -0.35 |
| RB1         | 0.22  | -0.37 | M                | 0.06  | -0.41 |
| CP          | 0.25  | 0.23  | O                | 0.29  | -0.31 |
| JN          | 0.24  | 0.44  | Q                | 0.39  | -0.03 |
| RW          | 0.25  | -0.08 | S                | 0.39  | 0.14  |
| HG          | 0.26  | -0.04 |                  |       |       |
| SW          | 0.25  | -0.15 |                  |       |       |
| RB2         | 0.24  | -0.33 |                  |       |       |
| NF2         | 0.26  | 0.22  |                  |       |       |
| MG          | 0.26  | 0.11  |                  |       |       |
| JV          | 0.23  | 0.46  |                  |       |       |
| Eigenwerte: | 14.03 | 1.06  |                  | 9.96  | 4.88  |

14c) die normierte Kernmatrix:

|    | S1   |       | S2 |       |      |
|----|------|-------|----|-------|------|
|    | S1   | S2    | S1 | S2    |      |
| P1 | 1.75 | -0.01 | P1 | -0.01 | 1.43 |
| P2 | 0.59 | /     | P2 | /     | /    |

Die mit "/" gekennzeichneten Zellen können aufgrund negativer Dimensionsgewichte nicht berechnet werden.

Tabelle 15 : die Lösung nach dem Tuckerschen Modell mit  
Varimax-rotierten Faktormatrizen

15a) Vpn :

|     | P1    | P2    |
|-----|-------|-------|
| CB  | 0.23  | 0.11  |
| DT  | 0.38  | -0.10 |
| MS  | 0.29  | 0.04  |
| RS  | 0.28  | 0.05  |
| NF1 | 0.07  | 0.33  |
| JH  | 0.25  | 0.10  |
| RB1 | 0.41  | -0.15 |
| CP  | 0.05  | 0.34  |
| JN  | -0.08 | 0.50  |
| RW  | 0.25  | 0.09  |
| HG  | 0.23  | 0.12  |
| SW  | 0.29  | 0.03  |
| RB2 | 0.40  | -0.12 |
| NF2 | 0.06  | 0.34  |
| MG  | 0.13  | 0.25  |
| JV  | -0.09 | 0.50  |

15b) Farbreize :

|   | S1    | S2    |
|---|-------|-------|
| A | 0.43  | 0.13  |
| C | 0.14  | 0.54  |
| E | -0.18 | 0.49  |
| G | -0.37 | 0.15  |
| I | -0.43 | -0.04 |
| K | -0.35 | -0.22 |
| M | -0.13 | -0.40 |
| O | 0.13  | -0.41 |
| Q | 0.35  | -0.20 |
| S | 0.41  | -0.04 |

15c) die normierte Kernmatrix :

|    | S1   |       |
|----|------|-------|
|    | S1   | S2    |
| P1 | 1.55 | 0.02  |
| P2 | 1.26 | -0.53 |

|    | S2    |      |
|----|-------|------|
|    | S1    | S2   |
| P1 | 0.02  | 1.37 |
| P2 | -0.53 | 1.11 |

die farbenfehlsichtigen Probanden umfaßt.

Die aufgrund der Kernmatrix implizierte Schiefwinkligkeit der Farbachsen für die fehlsichtige Personengruppe steht in Einklang zu den von Helm berichteten Ergebnissen. Der entsprechende Befund resultiert aus der Interpretation der Kernrelationen, die die Farbdimensionen miteinander verknüpfen. Im Gegensatz zu den farbtüchtigen Probanden, denen ein aus orthogonalen Dimensionen aufgespannter Farbkreis zugeordnet ist, liegt für die fehlsichtigen Probanden eine Korrelation von  $-0.53$  vor. Die Kernrelationen, die den dimensionsspezifischen Gewichten der Personen entsprechen, legen die Hypothese nahe, daß die farbenfehlsichtigen Personen eine geringeres Differenzierungsvermögen nicht nur bezüglich der "Rot-Grün"-Dimensionen besitzen, sondern auch Schwierigkeiten in der Unterscheidung zwischen "Gelb" und "Blau" haben. Der Farbkreis weist eine Längendeformation beider Achsen auf; daher müßte nach der vorliegenden Konfiguration eine Deuteranomalie zu einer Stauchung des gesamten Farbkreises führen, nachweisbar ist aber nur eine herabgesetzte Farbempfindlichkeit für rote und grüne Farbtöne.

#### 8.4. Die Lösung nach dem SUMMAX-Modell :

##### 8.4.1. Kriterien zur Auswahl einer Lösung :

Grundlage zur Bestimmung der Anzahl substantieller Faktoren war der Scree-Test der Kurve der Faktor-Ähnlichkeitskoeffizienten (Orlik, 1980), die in Abbildung 25 in Abhängigkeit von der Anzahl extrahierter Faktortripel dargestellt ist. Eine deutliche Veränderung des Kurvenverlaufs zeigt sich nach Extraktion des zweiten Faktortripels; der aufgeklärte Varianzanteil beträgt hier bereits 90% und wird durch Hinzunahme eines weiteren Faktortripels nicht wesentlich erhöht. Tabelle 16 enthält die zweidimensionale Lösung, die nach dem INDSCAL-Rationale reskaliert wurde.

## Abb.25: SUMMAX

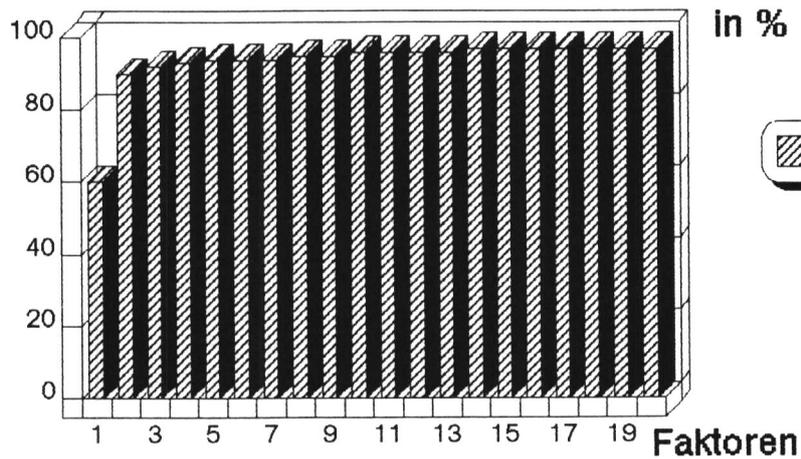
 $R^2$  zwischen Roh – u. geschätzten Daten


Tabelle 16: Reskalierung der unrotierten SUMMAX-Lösung aus  
2 Faktortripeln nach dem Rationale von INDSCAL

16a) Vpn :

|     | P1   | P2   |
|-----|------|------|
| CB  | 0.79 | 0.48 |
| DT  | 0.59 | 0.75 |
| MS  | 0.70 | 0.68 |
| RS  | 0.72 | 0.61 |
| NF1 | 0.83 | 0.34 |
| JH  | 0.74 | 0.62 |
| RB1 | 0.48 | 0.77 |
| CP  | 0.88 | 0.32 |
| JN  | 0.97 | 0.06 |
| RW  | 0.72 | 0.57 |
| HG  | 0.78 | 0.57 |
| SW  | 0.68 | 0.72 |
| RB2 | 0.55 | 0.78 |
| NF2 | 0.87 | 0.39 |
| MG  | 0.80 | 0.43 |
| JV  | 0.99 | 0.07 |

16b) Farbreize :

|   | S1    | S2    |
|---|-------|-------|
| A | -0.28 | 0.36  |
| C | 0.19  | 0.50  |
| E | 0.42  | 0.28  |
| G | 0.38  | -0.11 |
| I | 0.32  | -0.30 |
| K | 0.17  | -0.38 |
| M | -0.12 | -0.40 |
| O | -0.34 | -0.25 |
| Q | -0.40 | 0.07  |
| S | -0.36 | 0.22  |

#### 8.4.2. Interpretation der trilinearen SUMMAX-Lösung aus zwei Faktortripeln :

Ein Vergleich der SUMMAX-Lösung mit den entsprechenden INDSCAL-Dimensionen (vgl. Tabelle 13) zeigt, daß beide Lösungen in hohem Maße übereinstimmen. Geringe Abweichungen einzelner Ladungen resultieren aus den unterschiedlichen Algorithmen, die in beiden Modellen zur Extraktion der Dimensionen verwendet werden, bzw. aus der unterschiedlichen Rechengenauigkeit der korrespondierenden Programme. Bezüglich des Helmschen Datensatzes ist eine Reskalierung der SUMMAX-Faktoren ausreichend, um die Äquivalenz der Lösungen aufzuzeigen, und damit gibt die SUMMAX-Konfiguration in gleicher Weise Aufschluß über die Struktureigenschaften des Datenkörpers wie die INDSCAL-Lösung.

Um eine weitere Vergleichsmöglichkeit mit dem Tuckerschen Modell zu ermöglichen, wird im folgenden die SUMMAX-Konfiguration aufgrund von sieben Dimensionen mitgeteilt. Die Extraktion von fünf weiteren Faktortripeln, die zu einer Varianzaufklärung von 96% führt, wurde hier vorgenommen, um im letzten Auswertungsschritt den theoretischen Bezug zwischen beiden Modellen exemplarisch zu demonstrieren, welches eine hohe Ausschöpfung der Ausgangsdaten erfordert. Zunächst wird die SUMMAX-Lösung unter der Perspektive inhaltlicher Strukturhypothesen betrachtet. Der Kurvenverlauf der Eigenwerte der auf Hauptachsen transformierten Lösung ist in den Abbildungen 26 und 27 für jeden Modus dargestellt. Nach dem Scree-Test wurden zwei Faktoren jeweils im Personenmodus und im Modus der Farbreize beibehalten. Tabelle 17 enthält die Faktormatrizen, die nach dem Varimax-Kriterium weiterrotiert wurden. Die nach dem Trennkriterium von Orlik (1980) substantiellen Werte sind mit "\*" hervorgehoben. Die zugehörige Kernmatrix (vgl. Tabelle 17c) wurde nach der Vorgehensweise von Kroonenberg (1983b) normiert.

#### 8.4.3. Interpretation der rotierten SUMMAX-Lösung :

Die Konfiguration im Farbmodus weist auch hier eine Struktur auf, die der Anordnung der Farbtöne auf dem Farbkreis

Abb. 26: SUMMAX

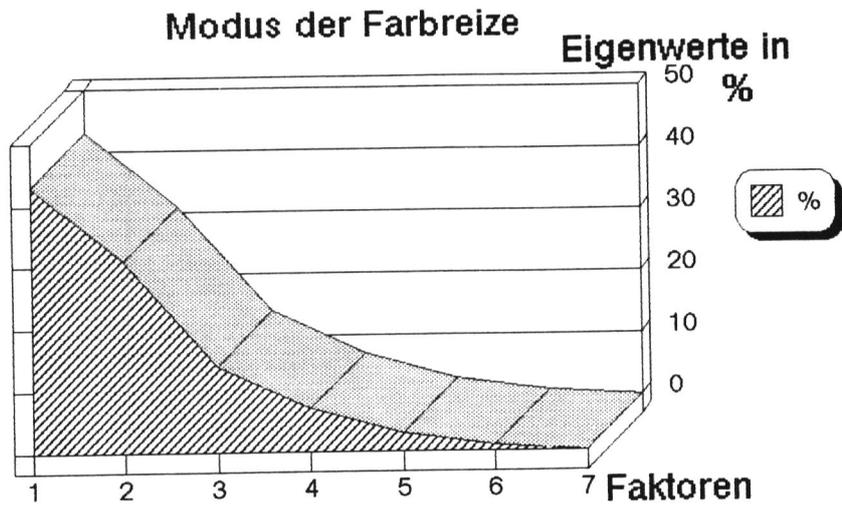


Abb. 27: SUMMAX

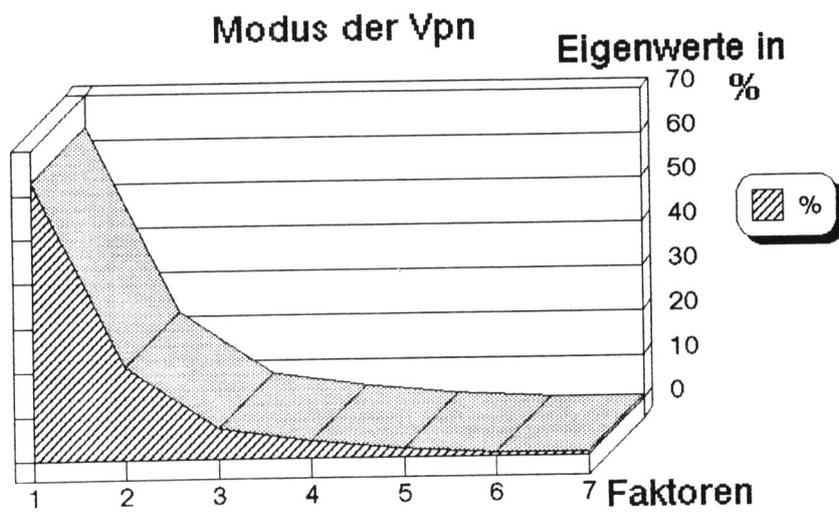


Tabelle 17 : auf Hauptachsen und VARIMAX-rotierte SUMMAX-Lösung nach Extraktion von 7 Dimensionen

17a) Vpn :

|     | P1    | P2    |
|-----|-------|-------|
| CB  | 0.23  | 0.29  |
| DT  | -0.15 | 0.58* |
| MS  | 0.27  | 0.42* |
| RS  | 0.17  | 0.44* |
| NF1 | 0.59* | 0.12  |
| JH  | 0.32  | 0.36* |
| RB1 | 0.04  | 0.51* |
| CP  | 0.48* | 0.16  |
| JN  | 0.55* | 0.01  |
| RW  | 0.34  | 0.36* |
| HG  | 0.40* | 0.21  |
| SW  | 0.25  | 0.43* |
| RB2 | 0.10  | 0.49* |
| NF2 | 0.58* | 0.16  |
| MG  | 0.57* | 0.15  |
| JV  | 0.46* | -0.06 |

17b) Farbreize :

|   | S1     | S2     |
|---|--------|--------|
| A | -0.73* | 0.15   |
| C | -0.44  | 0.77*  |
| E | 0.37   | 0.80*  |
| G | 0.74*  | 0.43   |
| I | 0.77*  | -0.01  |
| K | 0.52*  | -0.30  |
| M | 0.17   | -0.59* |
| O | -0.15  | -0.66* |
| Q | -0.52* | -0.45  |
| S | -0.72* | -0.14  |

17c) SUMMAX-Kernmatrix, normiert nach dem Rationale von Kroonenberg:

|    | S1   |       | S2    |      |
|----|------|-------|-------|------|
|    | S1   | S2    | S1    | S2   |
| P1 | 0.49 | 0.49  | 0.49  | 0.57 |
| P2 | 0.70 | -0.06 | -0.06 | 0.64 |

entspricht, mit den Grunddimensionen "Rot-Grün-Gelb" und "Gelb-Blau". Die Konfiguration im Personenmodus enthält angenähert eine Trennung der Probanden nach Farbenfehlsichtigkeit und normaler Farbempfindlichkeit mit Ausnahme der Personen HG und MG, die nach ophtalmologischen Untersuchungen keine Farbschwächung aufweisen.

Die Kernrelationen, die die beiden Farbdimensionen miteinander verknüpfen, entsprechen einem auch in der rotierten Tuckerschen Lösung angetroffenen Befund: für die farbtüchtigen Personen ergeben sich fast orthogonale Farbdimensionen (-0.06), während für die fehlsichtigen Probanden ein schiefwinkliges Achsensystem (0.49) vorliegt. Die Interpretation der dimensionsspezifischen Gewichte legt eine relative Stauung der "Rot-Grün"-Dimension für die farbenfehlsichtigen Versuchspersonen nahe.

Im wesentlichen impliziert die rotierte Konfiguration die gleichen Aussagen, die in der Lösung aus zwei Faktortripeln enthalten waren; eine Rotation der Faktormatrizen war nicht erforderlich, da der trilineare SUMMAX-Ansatz bereits klar interpretierbare Faktoren im Personenmodus wie auch im Modus der Farbreize erbrachte. Die zweite Auswertung wurde zusätzlich durchgeführt, um den Bezug des SUMMAX-Modells zu dem Tuckerschen Ansatz exemplarisch aufzuzeigen. Die entsprechenden Befunde werden im folgenden dargestellt.

Tabelle 18 zeigt die auf Hauptachsen transformierten SUMMAX-Dimensionen mit den zugehörigen Eigenwerten. Zur Überführung dieser Dimensionen in die Tuckersche Lösung werden die vollständigen Faktormatrizen zugrunde gelegt, die im ersten Auswertungsschritt spaltenweise auf die Länge eins normiert werden. Ferner werden die Kernrelationen, angeordnet in den beiden Kernmatrizen  $m_{pp}^G$  und  $p_{mp}^G$ , mit den Eigenwerten der korrespondierenden Dimensionen gewichtet. Auf die resultierenden gewichteten Matrizen wird erneut die Hauptachsen-transformation angewendet (vgl. Abschnitt 5), wobei sich direkt die Eigenwerte der Tuckerschen Lösung ergeben (vgl. Tabelle 19). Die Bestimmung dieser Eigenwerte ist nur approximativ möglich durch die programmintern beschränkte Rechen-

Tabelle 18 : auf Hauptachsen rotierte SUMMAX-Faktoren nach Extraktion von 7 Faktortripeln

| 18a) Vpn :  |      |       | 18b) Farbreize : |       |       |
|-------------|------|-------|------------------|-------|-------|
|             | P1   | P2    |                  | S1    | S2    |
| CB          | 0.36 | -0.08 | A                | -0.60 | 0.44  |
| DT          | 0.24 | -0.55 | C                | -0.08 | 0.88  |
| MS          | 0.46 | -0.17 | E                | 0.67  | 0.57  |
| RS          | 0.40 | -0.24 | G                | 0.85  | 0.08  |
| NF1         | 0.54 | 0.27  | I                | 0.69  | -0.33 |
| JH          | 0.47 | -0.09 | K                | 0.34  | -0.39 |
| RB1         | 0.35 | -0.38 | M                | -0.09 | -0.60 |
| CP          | 0.48 | 0.17  | O                | -0.41 | -0.53 |
| JN          | 0.42 | 0.34  | Q                | -0.66 | -0.19 |
| RW          | 0.49 | -0.07 | S                | -0.71 | 0.17  |
| HG          | 0.45 | 0.08  |                  |       |       |
| SW          | 0.46 | -0.19 |                  |       |       |
| RB2         | 0.38 | -0.33 |                  |       |       |
| NF2         | 0.55 | 0.22  |                  |       |       |
| MG          | 0.55 | 0.23  |                  |       |       |
| JV          | 0.40 | 0.22  |                  |       |       |
| Eigenwerte: | 3.16 | 1.07  |                  | 3.26  | 2.37  |

Tabelle 19 : die aus der SUMMAX-Lösung geschätzten Tucker-schen Ladungsmatrizen

| 19a) Vpn :  |       |       | 19b) Farbreize : |       |       |
|-------------|-------|-------|------------------|-------|-------|
|             | P1    | P2    |                  | S1    | S2    |
| CB          | 0.25  | -0.03 | A                | 0.32  | 0.30  |
| DT          | 0.20  | -0.32 | C                | -0.11 | 0.54  |
| MS          | 0.25  | -0.15 | E                | -0.37 | 0.38  |
| RS          | 0.23  | -0.16 | G                | -0.40 | -0.02 |
| NF1         | 0.26  | 0.14  | I                | -0.37 | -0.24 |
| JH          | 0.25  | -0.10 | K                | -0.21 | -0.35 |
| RB1         | 0.20  | -0.38 | M                | 0.07  | -0.41 |
| CP          | 0.25  | 0.19  | O                | 0.30  | -0.31 |
| JN          | 0.27  | 0.44  | Q                | 0.39  | -0.02 |
| RW          | 0.24  | -0.10 | S                | 0.39  | 0.15  |
| HG          | 0.28  | -0.04 |                  |       |       |
| SW          | 0.25  | -0.15 |                  |       |       |
| RB2         | 0.21  | -0.34 |                  |       |       |
| NF2         | 0.28  | 0.20  |                  |       |       |
| MG          | 0.26  | 0.08  |                  |       |       |
| JV          | 0.26  | 0.46  |                  |       |       |
| Eigenwerte: | 14.18 | 1.20  |                  | 10.01 | 4.90  |

genauigkeit und durch eine Varianzaufklärung der SUMMAX-Faktoren, die unter 100% beträgt. Entsprechend Tabelle 19 gehen etwa die Eigenwerte der SUMMAX-Lösung für die Farbdimensionen, 3.26 und 2.37, über in die Werte 10.01 und 4.90, die die Tuckerschen Eigenwerte recht gut approximieren (vgl. Tabelle 14). Zusätzlich sind die Eigenvektoren aus der Hauptachsentransformation, angeordnet als Spalten einer Transformationsmatrix, die Rotationsmatrizen, die die SUMMAX-Faktoren in die Tuckerschen Eigenvektoren überführen. Tabelle 19 verdeutlicht diesen Zusammenhang für den Personenmodus und den Modus der Farbreize.

##### 5. Schlußbemerkung :

Im Rahmen der vergleichenden Strukturanalyse des Helmschen Datensatzes führte die Anwendung des INDSCAL- und SUMMAX-Verfahrens zu Konfigurationen, die die variierende Farbempfindlichkeit der untersuchten Probanden deutlich widerspiegeln. Im Modus der Farbreize ergab sich eine Anordnung der ausgewählten Farbtöne entlang einem geschlossenen Kreis, der angenähert durch die Grunddimensionen "Rot-Grün" und "Gelb-Blau" aufgespannt ist. Auch die Konfiguration nach dem Tuckerschen Modell enthält eine Circumplex-Struktur für die Farbreize mit den beschriebenen Farbachsen; allerdings konnte die Tuckersche Kernmatrix zunächst nicht angemessen normiert werden. Durch Rotation nach Einfachstrukturkriterien wurde das Problem der angemessenen Normierung zwar aufgelöst; jedoch legte die rotierte Konfiguration Hypothesen zur differentiellen Farbwahrnehmung nahe, die keine Entsprechung in der diagnostizierten Farbschwächung der Versuchspersonen finden. Dabei enthält die Lösung nach dem Tuckerschen Modell Personen- und Farbdimensionen, die aus der quadrilinearen SUMMAX-Konfiguration nach Normierung der Dimensionen in beiden Modi durch orthogonale Rotationen hervorgehen, wie in einem abschließenden Auswertungsschritt gezeigt wurde.

9.

Varianten der Sortiertechnik: eine  
Methodenstudie von Rosenberg u. Kim

Gegenstand des folgenden Abschnitts ist die Reanalyse einer Methodenstudie von Rosenberg und Kim (1975) zu der Sortier-technik als Instrument der Erhebung von Ähnlichkeitsdaten. Rosenberg und Kim gingen in ihrer Untersuchung der Frage nach, in welcher Weise Varianten der Sortiermethode die Beachtung unterschiedlicher Organisationsaspekte genealogischer Begriffe induzieren. Das Versuchsmaterial bestand aus 15 Verwandtschaftsbegriffen, die in der englischen Sprache relativ häufig verwendet werden. Zur Erhebung der Ähnlichkeitsbeziehungen zwischen den Begriffen wurden zwei Sortier-techniken eingesetzt, von Rosenberg und Kim als "single-sort" und "multiple-sort" bezeichnet. Die Versuchsperson wurde mit der Aufgabe konfrontiert, die Begriffe nach subjektivem Empfinden in Klassen aufzuteilen, so daß Begriffe zur gleichen Klasse gehören, die hinsichtlich eines von der Versuchsperson gewählten Bedeutungsaspektes eine ähnliche Bedeutung haben, während einander unähnliche Begriffe unterschiedlichen Klassen zuzuordnen sind. Die von den Probanden gewählten Bedeutungsaspekte als Grundlage der Sortierungen wurden nicht erhoben. Unter der "single-sort"-Instruktion hatte die Versuchsperson nur Gelegenheit zu einer Sortierung, in der "multiple-sort"-Bedingung konnte sie mehrere Sortierungen durchführen, wobei jeweils ein neuer Bedeutungsaspekt auszuwählen war (Rosenberg & Kim, 1975, S.491). An der Untersuchung nahmen insgesamt 330 studentische Versuchspersonen teil, die entsprechend dem in Tabelle 20 aufgeführten Schema den Versuchsbedingungen zugewiesen waren. Für jede Instruktion und getrennt nach Geschlecht der Personen wurden Distanz-Schätzwerte der Begriffspaare ermittelt. Aus der "multiple-sort"-Bedingung wurden jeweils die erste und die zweite Sortierung in die weiteren Analysen einbezogen, so daß sechs verschiedene Gruppen entstanden: männliche Versuchspersonen - single-sort, weibliche Versuchspersonen - single-sort, männliche Versuchspersonen - multiple-sort 1, männliche Versuchspersonen - multiple-sort 2, weibliche Versuchspersonen - multiple-sort 1, weibliche Versuchspersonen - multiple-sort 2. Dabei wurde für jede Personengruppe zur Ermittlung der Distanz-Schätzwerte zwi-

Tabelle 20: Reizmaterial, Versuchsbedingungen und Versuchspersonen der Untersuchung von Rosenberg und Kim

-----  
 20a) Reizmaterial:  
 -----

|               |                 |                      |
|---------------|-----------------|----------------------|
| uncle (uncl)  | father (fath)   | son (son.)           |
| aunt (aunt)   | mother (moth)   | grandfather (grfa)   |
| cousin (cous) | sister (sist)   | grandmother (grmo)   |
| nephew (neph) | brother (brot)  | granddaughter (grda) |
| niece (niec)  | daughter (daug) | grandson (grso)      |

20b) Versuchsbedingungen und -personen:  
 -----

| Versuchsbedingungen | Versuchspersonen |               |
|---------------------|------------------|---------------|
|                     | männliche Vpn    | weibliche Vpn |
| single-sort:        | 85 Vpn (mals)    | 85 Vpn (fems) |
| multiple sort:      | 80 Vpn           | 80 Vpn        |
| 1. Sortierung       | (mal1)           | (fem1)        |
| 2. Sortierung       | (mal2)           | (fem2)        |

-----

Die in Klammern aufgeführten Abkürzungen werden in allen folgenden Tabellen verwendet.

-----

schen allen Begriffspaaren ausgezählt, wie häufig je zwei Begriffe unterschiedlichen Sortierklassen zugeordnet wurden. Die resultierenden Distanz-Schätzwerte erfüllen die Bedingungen einer Metrik (Miller, 1969) und sind in Übereinstimmung mit dem Vorgehen von Carroll und Arabie (1983) Grundlage aller folgenden Auswertungen. Die Distanz-Schätzwerte für jede Personengruppe und alle Begriffspaare aus der Untersuchung von Rosenberg und Kim finden sich u.a. in Arabie, Carroll und DeSarbo (1987, S. 62f, Tabelle 6).

Rosenberg und Kim berichten als Hauptergebnis ihrer Datenanalyse bedeutsame Unterschiede in der Klassifikation der Verwandtschaftsbegriffe in Abhängigkeit von der Sortieraufgabe. Nach Auszählung der Anzahl der Versuchspersonen, die unter den verschiedenen Instruktionen nach der Geschlechtszuordnung der Begriffe sortierten, wurde deutlich, daß die überwiegende Mehrzahl der Versuchspersonen aus der "single-sort"-Bedingung nicht nach der Geschlechtszugehörigkeit differenzierten. Hatten die Personen aber Gelegenheit zu mehrfachen Sortierungen, so wurde die Geschlechtsdimension zur Gruppierung der Begriffe verwendet, von den weiblichen Versuchspersonen der "multiple-sort"-Bedingung im ersten Versuchsdurchgang, von den männlichen Versuchspersonen vermehrt im zweiten Durchgang. Zwei weitere Grundlagen zur Gruppierung der Begriffe, die allerdings nicht zwischen den Instruktionen variierten, waren Differenzierungen nach der Stamm- versus Seitenlinie (mother, father, son ... versus uncle, aunt, cousin ...) und nach Aszendenten bzw. Deszendenten der zweiten Generation (grandfather, granddaughter usw. ). Nach Rosenberg und Kim sind mindestens zwei Organisationsaspekte in den Sortierungen nicht repräsentiert, die nach den Ergebnissen von Wallace und Atkins (1960) und dem anthropologischen Modell von Romney und D'Andrade (1964, zitiert nach Rosenberg & Kim) für das vorliegende Wortfeld relevant sind: die gesamte Generationenfolge und Aszendenten bzw. Deszendenten der ersten Generation versus Verwandte der Generation des Betrachters (father, mother, son ... versus brother, sister, cousin). In einer Reanalyse der Daten durch

ein hybrides Modell, welches eine dimensionale Darstellung mit einem clusteranalytischen Ansatz kombiniert, fanden Carroll und Pruzansky (1980) jedoch die gesamte Generationenabfolge deutlich repräsentiert (vgl. auch Carroll & Pruzansky, 1983, 1986; Carroll, 1976). Im folgenden werden wie im vorangegangenen Abschnitt die Modelle zur Multidimensionalen Skalierung verglichen anhand der Methodenstudie von Rosenberg und Kim. Dabei steht die dimensionale Repräsentation im Vordergrund, auf zusätzliche clusteranalytische Darstellungen wird verzichtet.

#### 9.1. Normierung des Datenkörpers:

Die bereits in der Auswertung des Helmschen Datensatzes beschriebenen Normierungen wurden auch auf die vorliegenden Daten angewendet. Für jede Personengruppe getrennt (vgl. Tabelle 20b) wurden die Distanz-Schätzwerte in Skalarprodukte transformiert, und die Quadratsumme jeder Skalarproduktmatrix wurde auf den Betrag eins normiert.

#### 9.2. Die INDSCAL-Lösung:

Zur Beurteilung der angemessenen Dimensionalität der INDSCAL-Konfiguration wurde die quadrierte Korrelation  $R^2$  der Ausgangsdaten mit dem geschätzten Datenkörper in Abhängigkeit von der Anzahl der Dimensionen ermittelt. Für die zwei- bis fünfdimensionale Lösung (Programmbeschränkung: Anzahl der Dimensionen  $\leq 5$ ) ergaben sich folgende Werte: 0,66, 0,85, 0,92 und 0,94. Deutliche Veränderungen zeigen sich nach Extraktion von zwei und drei Dimensionen, wobei eine vergleichsweise hohe Varianzaufklärung in der drei- und vierdimensionalen Lösung erreicht wird.

Die Anzahl der durchgeführten Iterationen variiert nur geringfügig zwischen den verschiedenen Konfigurationen; es werden jeweils zehn bzw. elf Iterationen benötigt. Die Höhe der Korrelationen zwischen den Dimensionen eines Modus ist

Tabelle 21: Interkorrelationen der INDSCAL-Dimensionen

-----  
 21a) zweidimensionale Lösung:  
 -----

Versuchsbedingungen:

|       | 1 | 2     |
|-------|---|-------|
| ----- |   |       |
| 1     | 1 | -0.96 |
| ----- |   |       |

Begriffe:

|       | 1 | 2     |
|-------|---|-------|
| ----- |   |       |
| 1     | 1 | -0.27 |
| ----- |   |       |

-----  
 21b) dreidimensionale Lösung:  
 -----

Versuchsbedingungen:

|       | 1 | 2     | 3     |
|-------|---|-------|-------|
| ----- |   |       |       |
| 1     | 1 | -0.95 | -0.98 |
| 2     |   | 1     | 0.96  |
| ----- |   |       |       |

Begriffe:

|       | 1 | 2     | 3    |
|-------|---|-------|------|
| ----- |   |       |      |
| 1     | 1 | -0.03 | 0.00 |
| 2     |   | 1     | 0.19 |
| ----- |   |       |      |

-----  
 21c) vierdimensionale Lösung:  
 -----

Versuchsbedingungen:

|       | 1 | 2     | 3     | 4     |
|-------|---|-------|-------|-------|
| ----- |   |       |       |       |
| 1     | 1 | -0.95 | -0.98 | -0.18 |
| 2     |   | 1     | 0.97  | 0.08  |
| 3     |   |       | 1     | 0.25  |
| ----- |   |       |       |       |

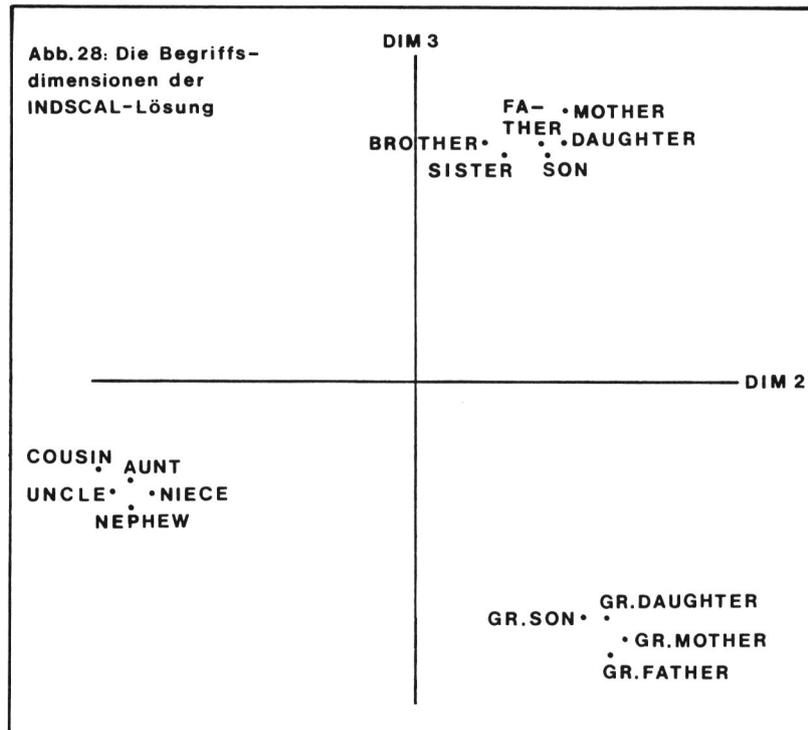
Begriffe:

|       | 1 | 2     | 3    | 4     |
|-------|---|-------|------|-------|
| ----- |   |       |      |       |
| 1     | 1 | -0.01 | 0.00 | -0.05 |
| 2     |   | 1     | 0.11 | 0.22  |
| 3     |   |       | 1    | 0.25  |
| ----- |   |       |      |       |

-----

Tabelle 22: Die dreidimensionale INDSCAL-Lösung

| 22a) Personengruppen: |      |      |      | 22b) Begriffe: |       |       |       |
|-----------------------|------|------|------|----------------|-------|-------|-------|
|                       | DIM1 | DIM2 | DIM3 |                | DIM1  | DIM2  | DIM3  |
| fems                  | 0.18 | 0.67 | 0.61 | aunt           | -0.27 | -0.35 | -0.12 |
| mals                  | 0.39 | 0.62 | 0.51 | brot           | 0.27  | 0.08  | 0.30  |
| fem1                  | 0.87 | 0.33 | 0.23 | cous           | 0.02  | -0.39 | -0.11 |
| fem2                  | 0.55 | 0.55 | 0.46 | daug           | -0.26 | 0.18  | 0.29  |
| mal1                  | 0.67 | 0.48 | 0.41 | fath           | 0.28  | 0.15  | 0.32  |
| mal2                  | 0.71 | 0.37 | 0.37 | grda           | -0.27 | 0.23  | -0.28 |
|                       |      |      |      | grfa           | 0.27  | 0.24  | -0.33 |
|                       |      |      |      | grmo           | -0.27 | 0.26  | -0.31 |
|                       |      |      |      | grso           | 0.27  | 0.21  | -0.29 |
|                       |      |      |      | moth           | -0.27 | 0.17  | 0.32  |
|                       |      |      |      | neph           | 0.25  | -0.35 | -0.14 |
|                       |      |      |      | niec           | -0.27 | -0.33 | -0.13 |
|                       |      |      |      | sist           | -0.26 | 0.11  | 0.30  |
|                       |      |      |      | son.           | 0.27  | 0.16  | 0.29  |
|                       |      |      |      | uncl           | 0.25  | -0.37 | -0.12 |



in Tabelle 21 für die Konfigurationen aus zwei, drei und vier Dimensionen aufgeführt. Die Korrelationen im Modus der Personengruppen zeigen sehr hohe Werte für jede der Lösungen, und die Interkorrelationen im Modus der Begriffe sind am geringsten für die dreidimensionale Konfiguration. Daher wurde unter zusätzlicher Beachtung der Varianzaufklärung die Lösung aus drei Faktoren ausgewählt.

#### 9.2.1. Interpretation der dreidimensionalen Lösung:

Tabelle 22 zeigt die Konfiguration aus drei Dimensionen, die für die zweite und dritte Dimension des Modus der Begriffe in Abbildung 28 graphisch dargestellt ist. Die erste Dimension ist als ein Geschlechtsfaktor interpretierbar, der die Verwandtschaftsbegriffe, die eine weibliche Geschlechtszugehörigkeit implizieren, von den männlichen Angehörigen trennt (vgl. Tabelle 22b). Der einzige Begriff, der nicht eindeutig einem Geschlecht zugeordnet werden kann ("cousin"), ist dabei nahe dem Ursprung lokalisiert. Der zweite Faktor ist interpretierbar als Stamm- versus Seitenlinie, und die dritte Dimension enthält eine Abgrenzung der Kernfamilie von den Aszendenten und Deszendenten der zweiten Generation. Ferner sind die kompakten Begriffsgruppierungen der Abbildung 28 identisch mit den aufgrund einer Complete-Linkage-Analyse gewonnenen Clustern der "single-sort"-Bedingungen (Rosenberg & Kim, 1975), und sie sind in der INDCLUS-Lösung von Carroll und Arabie (1983) enthalten.

Im Modus der Versuchsbedingungen bilden sich die von Rosenberg und Kim diskutierten Unterschiede zwischen den beiden Varianten der Sortiertechnik auf der ersten Dimension ab. Entsprechend der Höhe der "Gewichte" ist die Geschlechtsdimension für die weiblichen und männlichen Versuchspersonen der "single-sort"-Bedingung vergleichsweise gestaucht, wodurch eine geringere Differenzierung der Verwandtschaftsbegriffe nach der Geschlechtszugehörigkeit impliziert wird. Dagegen betonen nach der INDSCAL-Konfiguration weibliche Versuchspersonen, die Gelegenheit zu mehrfachen Sortierungen hatten, die Geschlechtsdimension im ersten Versuchsdurchgang und männliche Versuchspersonen bei der zweiten Sortierung.

Dieser Befund konnte von Rosenberg und Kim (1975, S.495) auch auf der Datenebene abgesichert werden. Auf der zweiten und dritten Dimension ergeben sich keine systematischen Unterschiede zwischen beiden Sortierbedingungen; die INDSCAL-Lösung legt hier die Hypothese nahe, daß jeweils eine Sortierung in jeder Versuchsbedingung nach Stamm- versus Seitenlinie und Kernfamilie versus Angehörige der zweiten Generation erfolgte. Insgesamt entspricht die Interpretation der INDSCAL-Konfiguration im wesentlichen den einführend dargestellten Ergebnissen; dabei ist allerdings eine Klassifikation des Reizmaterials nach der Abfolge Aszendenten der ersten Generation - Generation des Betrachters - Deszendenten der ersten Generation in der vorliegenden Lösung nicht enthalten.

### 9.3. Die Lösung nach dem Tuckerschen Modell:

Grundlage zur Bestimmung der Anzahl substantieller Dimensionen war die Höhe der Eigenwerte für den Modus der Begriffe und den Modus der Versuchsbedingungen in Abhängigkeit von der Anzahl der extrahierten Dimensionen. Es ergaben sich folgende Werte im Modus der Personengruppen, in Prozent der Spur der Faktormatrix: 89%, 9%, 0,9%, 0,4%, 0,2% und 0,1%. Der Kurvenverlauf der Eigenwerte für den Begriffsmodus ist in Abbildung 29 dargestellt. Im Modus der Personengruppen wurden zwei Dimensionen und im Modus der Begriffe drei Dimensionen als substantiell beurteilt. Tabelle 23 zeigt die Kern- und Ladungsmatrizen der unrotierten Lösung. Da sich eine weitere geringere Veränderung des Kurvenverlaufs der Eigenwerte nach der fünften Dimension für den Modus der Begriffe andeutet, sind in der Tabelle zusätzlich die vierte und fünfte Begriffsdimension aufgeführt.

#### 9.3.1. Interpretation der unrotierten Lösung:

Die Konfiguration des Begriffsmodus, die für die zweite und dritte Dimension in Abbildung 30 dargestellt ist, zeigt recht hohe Übereinstimmungen mit den entsprechenden INDSCAL-

# Abb. 29: Tuckersche Loesung

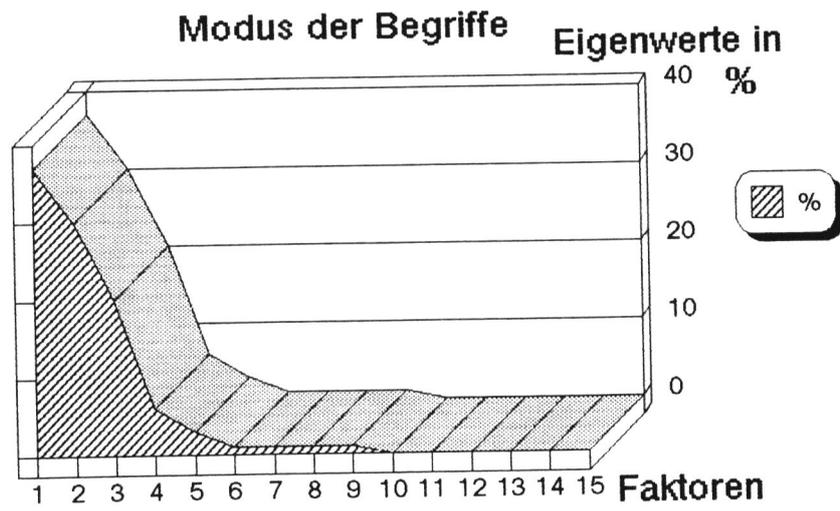


Tabelle 23: Die unrotierte Konfiguration nach dem Tuckerschen Modell

| 23a) Begriffe: |       |       |       |       |       | 23b) Personen-<br>gruppen: |            |
|----------------|-------|-------|-------|-------|-------|----------------------------|------------|
|                | P1    | P2    | P3    | P4    | P5    | M1                         | M2         |
| aunt           | 0.21  | -0.39 | -0.07 | -0.35 | -0.04 | fems                       | 0.38 -0.66 |
| brot           | -0.24 | 0.22  | -0.24 | 0.24  | -0.47 | mals                       | 0.42 -0.33 |
| cous           | -0.09 | -0.38 | -0.10 | 0.22  | -0.25 | fem1                       | 0.39 0.59  |
| daug           | 0.29  | 0.20  | -0.20 | 0.14  | 0.43  | fem2                       | 0.43 -0.06 |
| fath           | -0.23 | 0.29  | -0.21 | -0.37 | 0.03  | mal1                       | 0.43 0.14  |
| grda           | 0.29  | 0.06  | 0.38  | 0.24  | 0.06  | mal2                       | 0.41 0.29  |
| grfa           | -0.24 | 0.15  | 0.45  | -0.20 | -0.11 |                            |            |
| grmo           | 0.30  | 0.07  | 0.43  | -0.21 | -0.13 |                            |            |
| grso           | -0.24 | 0.14  | 0.40  | 0.23  | 0.07  |                            |            |
| moth           | 0.30  | 0.21  | -0.23 | -0.37 | 0.03  |                            |            |
| neph           | -0.31 | -0.30 | -0.04 | 0.19  | 0.24  |                            |            |
| niec           | 0.21  | -0.39 | -0.05 | 0.19  | 0.22  |                            |            |
| sist           | 0.29  | 0.14  | -0.25 | 0.26  | -0.45 |                            |            |
| son.           | -0.23 | 0.28  | -0.19 | 0.13  | 0.43  |                            |            |
| uncl           | -0.31 | -0.31 | -0.06 | -0.35 | -0.03 |                            |            |

| 23c) Kernmatrix, normiert nach dem Vorgehen von Kroonenberg: |      |      |      |      |      |
|--------------------------------------------------------------|------|------|------|------|------|
| M1                                                           | P1   | P2   | P3   | P4   | P5   |
| P1                                                           | 1.18 | 0.01 | 0.00 | 0.00 | 0.00 |
| P2                                                           | 0.01 | 1.14 | 0.00 | 0.00 | 0.00 |
| P3                                                           | 0.00 | 0.00 | 1.02 | 0.00 | 0.00 |
| P4                                                           | 0.00 | 0.00 | 0.00 | 0.75 | 0.00 |
| P5                                                           | 0.00 | 0.00 | 0.00 | 0.00 | 0.62 |

| M2 | P1   | P2 | P3 | P4 | P5 |
|----|------|----|----|----|----|
| P1 | 0.72 | /  | /  | /  | /  |
| P2 | /    | /  | /  | /  | /  |
| P3 | /    | /  | /  | /  | /  |
| P4 | /    | /  | /  | /  | /  |
| P5 | /    | /  | /  | /  | /  |

Die mit "/" gekennzeichneten Zellen der Kernmatrix konnten nicht berechnet werden.

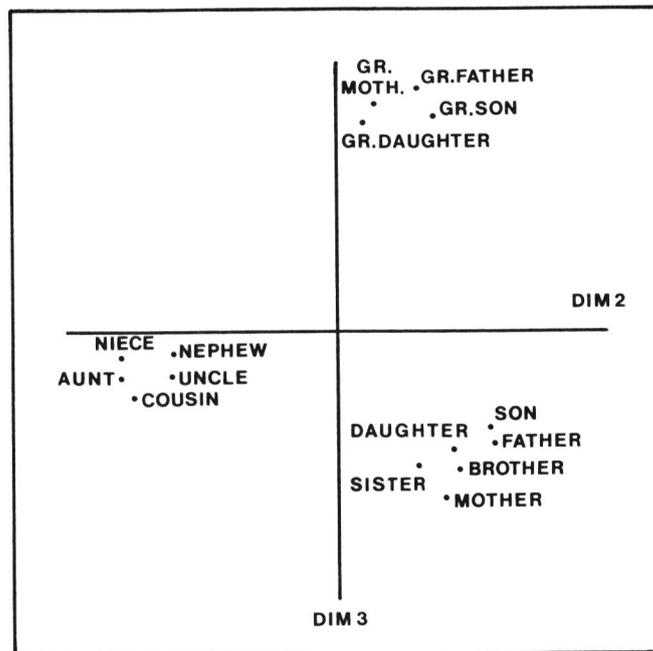


Abb.30: Die Begriffsdimensionen der Tuckerschen Lösung

Dimensionen. So ist auch in der Tuckerschen Lösung die erste Dimension als Geschlechtsfaktor interpretierbar (vgl. Tabelle 23a), die zweite Dimension enthält eine Differenzierung der Verwandtschaftsbegriffe in Stamm- und Seitenlinie, und die dritte Dimension impliziert eine Gruppierung nach der Zugehörigkeit zu der Kernfamilie oder zu den Angehörigen der zweiten Generation. Darüber hinaus ergeben sich aus der Lösung von fünf Dimensionen zusätzliche Differenzierungsaspekte, die nach dem Modell von Romney und D'Andrade relevante Strukturmerkmale enthalten: Aszendenten bzw. Deszendenten der ersten Generation versus Angehörige der Generation des Betrachters. Während die Dimensionen des Begriffsmodus klare Bezüge zu den bekannten genealogischen Relationen des vorliegenden Wortfeldes aufweisen, zeigt die Struktur des Modus der Personengruppen keine Entsprechungen zu den Ergebnissen von Rosenberg und Kim; dies wird insbesondere in Zusammenhang mit der normierten Tuckerschen Kernmatrix deutlich, die in Tabelle 23c aufgeführt ist. Die Kernrelationen, die den ersten Faktor der Personengruppen einbeziehen, legen hier die Hypothese nahe, daß unter allen Versuchsinstruktionen in gleicher Weise nach wechselseitig orthogonalen Begriffsdimensionen sortiert wurde, wobei die Bedeutsamkeit der Begriffsdimensionen sukzessive geringer wird. Für die Personengruppen des zweiten Faktors liegt offenbar keine angemessene Repräsentation der Verwandtschaftsbegriffe vor, da mehrere dimensionsspezifische Gewichte in der Diagonalen der Kernmatrix negative Vorzeichen aufweisen und daher nicht nach dem Vorgehen von Kroonenberg (1983b) und Tucker (1972) normiert werden können.

In einem weiteren Auswertungsschritt wurden die Faktoren der Personengruppen nach dem Varimax-Kriterium rotiert, und die Konfiguration des Begriffsmodus wurde beibehalten. Nach Rotation wiesen die Diagonalelemente der Kernmatrix nur positive Vorzeichen auf, so daß die Unangemessenheit der unrotierten Tuckerschen Lösung auf die Repräsentation im Modus der Personengruppen zurückführbar ist.

Tabelle 24: Die rotierte Konfiguration nach dem Tuckerschen Modell mit Varimax-rotierten Dimensionen im Modus der Personengruppen

| 24a) Begriffe: |       |       |       |       |       | 24b) Personen-<br>gruppen: |            |
|----------------|-------|-------|-------|-------|-------|----------------------------|------------|
|                | P1    | P2    | P3    | P4    | P5    | M1                         | M2         |
| aunt           | 0.21  | -0.39 | -0.07 | -0.35 | -0.04 | fems                       | -0.15 0.75 |
| brot           | -0.24 | 0.22  | -0.24 | 0.24  | -0.47 | mals                       | 0.10 0.52  |
| cous           | -0.09 | -0.38 | -0.10 | 0.22  | -0.25 | fem1                       | 0.68 -0.19 |
| daug           | 0.29  | 0.20  | -0.20 | 0.14  | 0.43  | fem2                       | 0.28 0.32  |
| fath           | -0.23 | 0.29  | -0.21 | -0.37 | 0.03  | mal1                       | 0.42 0.17  |
| grda           | 0.29  | 0.06  | 0.38  | 0.24  | 0.06  | mal2                       | 0.51 0.05  |
| grfa           | -0.24 | 0.15  | 0.45  | -0.20 | -0.11 |                            |            |
| grmo           | 0.30  | 0.07  | 0.43  | -0.21 | -0.13 |                            |            |
| grso           | -0.24 | 0.14  | 0.40  | 0.23  | 0.07  |                            |            |
| moth           | 0.30  | 0.21  | -0.23 | -0.37 | 0.03  |                            |            |
| neph           | -0.31 | -0.30 | -0.04 | 0.19  | 0.24  |                            |            |
| niec           | 0.21  | -0.39 | -0.05 | 0.19  | 0.22  |                            |            |
| sist           | 0.29  | 0.14  | -0.25 | 0.26  | -0.45 |                            |            |
| son.           | -0.23 | 0.28  | -0.19 | 0.13  | 0.43  |                            |            |
| uncl           | -0.31 | -0.31 | -0.06 | -0.35 | -0.03 |                            |            |

24c) Kernmatrix, normiert nach dem Vorgehen von Kroonenberg:

| M1 | P1    | P2    | P3    | P4    | P5   |
|----|-------|-------|-------|-------|------|
| P1 | 1.18  | -0.07 | -0.01 | -0.01 | 0.00 |
| P2 | -0.07 | 0.88  | 0.00  | -0.03 | 0.02 |
| P3 | -0.01 | 0.00  | 0.77  | -0.08 | 0.05 |
| P4 | -0.01 | -0.03 | -0.08 | 0.65  | 0.03 |
| P5 | 0.00  | 0.02  | 0.05  | 0.03  | 0.48 |

| M2 | P1   | P2   | P3    | P4    | P5    |
|----|------|------|-------|-------|-------|
| P1 | 0.72 | 0.13 | 0.03  | 0.02  | 0.02  |
| P2 | 0.13 | 1.04 | 0.00  | 0.03  | 0.00  |
| P3 | 0.03 | 0.00 | 0.95  | 0.05  | -0.02 |
| P4 | 0.02 | 0.03 | 0.05  | 0.68  | -0.06 |
| P5 | 0.02 | 0.00 | -0.02 | -0.06 | 0.56  |

### 9.3.2. Interpretation der rotierten Lösung:

Tabelle 24 zeigt die rotierte Konfiguration nach Normierung der Kernrelationen entsprechend dem von Kroonenberg (1983b) und Tucker (1972) diskutierten Vorgehen. Die Dimensionen des Modus der Personengruppen und Versuchsbedingungen führen angenähert zu einer Trennung der "multiple-sort"- von den "single-sort"-Bedingungen. Der wesentliche Unterschied zwischen den beiden Varianten der Sortiertechnik kommt zum Ausdruck in den Kernrelationen, die eine Gewichtung der Geschlechtsdimension der Begriffe enthalten. Für die Personengruppen, die nur Gelegenheit zu einer Sortierung hatten, und auch für den zweiten Versuchsdurchgang weiblicher Versuchspersonen aus der "multiple-sort"-Bedingung ist die Geschlechtsdimension gestaucht. Ferner impliziert die Kernmatrix auch hier für alle Personengruppen wechselseitig orthogonale Begriffsdimensionen, die mit Ausnahme des Geschlechtsfaktors hinsichtlich ihrer Bedeutsamkeit sukzessive geringer werden. Sowohl unter der "single-sort"- als auch unter der "multiple-sort"-Instruktion erfolgt eine Differenzierung nach Stamm- versus Seitenlinie, die in höherem Maße Beachtung findet als die Aspekte der Generationenfolge. Damit führt bei Betrachtung von drei Faktoren für die Verwandtschaftsbegriffe und einer Rotation im Modus der Personengruppen die Lösung nach dem Tuckerschen Modell zu Aussagen, die die wesentlichen Ergebnisse der Analyse von Rosenberg und Kim reproduzieren. Ferner weist die Kernmatrix angenähert eine Diagonalstruktur auf, wodurch die wechselseitige Orthogonalität der Begriffsdimensionen impliziert wurde. Unter Vernachlässigung der Kernrelationen außerhalb der Diagonalen, die nicht exakt gleich null sind, ist die vorliegende Lösung nach dem Tuckerschen Modell approximativ darstellbar in der INDSCAL-Form, indem die Konfiguration des Modus der Personengruppen mit den Diagonalwerten der Kernmatrix verknüpft wird (vgl. Harshman & Lundy, 1984; Kroonenberg, 1983b). Die entsprechende Multiplikation beider Matrizen, d.h.  $\sum_m a_{im} g_{mpp}$ , ergab folgende Schätzwerte für die Gewichte der INDSCAL-Lösung:

|      | DIM1 | DIM2 | DIM3 |
|------|------|------|------|
| fems | 0.18 | 0.69 | 0.59 |
| mals | 0.41 | 0.64 | 0.53 |
| fem1 | 0.84 | 0.32 | 0.23 |
| fem2 | 0.55 | 0.56 | 0.45 |
| mal1 | 0.67 | 0.51 | 0.40 |
| mal2 | 0.73 | 0.45 | 0.35 |

Ein Vergleich mit der in Tabelle 22 aufgeführten INDSCAL-Konfiguration zeigt die Einbettung der Lösung nach dem Tuckerschen Modell in die INDSCAL-Form; allerdings werden teilweise erhebliche Abweichungen der Ladungen beider Konfigurationen auf der dritten Begriffsdimension deutlich. In den ALS-Verfahren zur Transformation einer Konfiguration nach dem Tuckerschen Ansatz in die entsprechende INDSCAL- bzw. CANDECOMP-Lösung wird analog diesem Vorgehen zunächst eine näherungsweise Diagonalstruktur der Tuckerschen Kernmatrix aufgesucht, und deren Verknüpfung mit der Faktormatrix eines ausgewählten Modus liefert die zugeordnete trilineare Konfiguration. Weist die Tuckersche Kernmatrix angenähert eine Diagonalform auf wie im vorliegenden Beispiel, so werden beide Lösungen als äquivalente Repräsentationen aufgefaßt.

Darüber hinaus enthält die Konfiguration aus fünf Begriffsdimensionen zusätzliche Anhaltspunkte zur Struktur des untersuchten Wortfeldes. Neben Differenzierungen nach Geschlecht, Stamm- oder Seitenlinie und Abständen von zwei Generationen zum Betrachter ergibt sich hier in Übereinstimmung mit dem Modell von Romney und D'Andrade eine Abfolge der Angehörigen nach der ersten Generation.

#### 9.4. Die Lösung nach dem SUMMAX-Modell:

Zur Beurteilung der angemessenen Dimensionalität wurde der Faktor-Ähnlichkeitskoeffizient  $R^2$  zugrunde gelegt zum Ver-

gleich des aus den Faktortripeln geschätzten Datenkörpers mit den Ausgangsdaten (Orlik, 1980). Abbildung 31 stellt den Kurvenverlauf von  $R^2$  in Abhängigkeit von der Anzahl extrahierter Dimensionen dar. Eine deutliche Veränderung des Kurvenverlaufs zeigt sich nach Extraktion von drei Dimensionen, die bereits 84% der Varianz der Daten aufklären. Bei einer Konfiguration aus zehn Faktoren erhöht sich der faktoriell erklärte Varianzanteil auf 97%, was als ausreichend beurteilt wurde. Zum Vergleich mit INDSCAL und dem Tucker-schen Modell werden im folgenden beide Lösungen diskutiert, wobei nur die Konfiguration aus zehn Dimensionen auf Hauptachsen rotiert wurde. Die zugehörigen Eigenwerte für den Begriffsmodus sind in Abbildung 32 gegen die Anzahl der Dimensionen aufgetragen. Für den Modus der Personengruppen betragen die Eigenwerte in Prozent der Spur der Faktormatrix: 87%, 9%, 2%, 1% und 0,1%. Nach dem Scree-Test resultieren drei substantielle Begriffsdimensionen, und für die Personengruppen wurde eine Lösung aus zwei Dimensionen ausgewählt.

#### 9.4.1. Interpretation der unrotierten Konfiguration aus drei Dimensionen:

Tabelle 25 zeigt die unrotierte SUMMAX-Konfiguration, die zusätzlich für die zweite und dritte Begriffsdimension in Abbildung 33 dargestellt ist. Wie auch in der INDSCAL- und der Tucker-schen Lösung ist auf dem ersten Faktor des Begriffsmodus die Geschlechtszugehörigkeit der Familienmitglieder repräsentiert mit einer Trennung der weiblichen von den männlichen Angehörigen und dem Begriff "cousin" in der Nähe des Ursprungs. Die zweite Dimension differenziert zwischen der Stamm- und der Seitenlinie, und die dritte Dimension enthält die ebenfalls bereits bekannte Abgrenzung der Kernfamilie von den Aszendenten und Deszendenten der zweiten Generation (vgl. Abb. 33). Auch die Konfiguration im Modus der Personengruppen zeigt sehr hohe Übereinstimmungen mit den entsprechenden INDSCAL-Dimensionen und führt zu völlig identischen Aussagen. So bildet sich ebenso in der SUMMAX-Lösung der Unterschied zwischen den beiden Sortiertechniken

Abb. 31: SUMMAX

$R^2$  zwischen Roh – u. geschätzten Daten

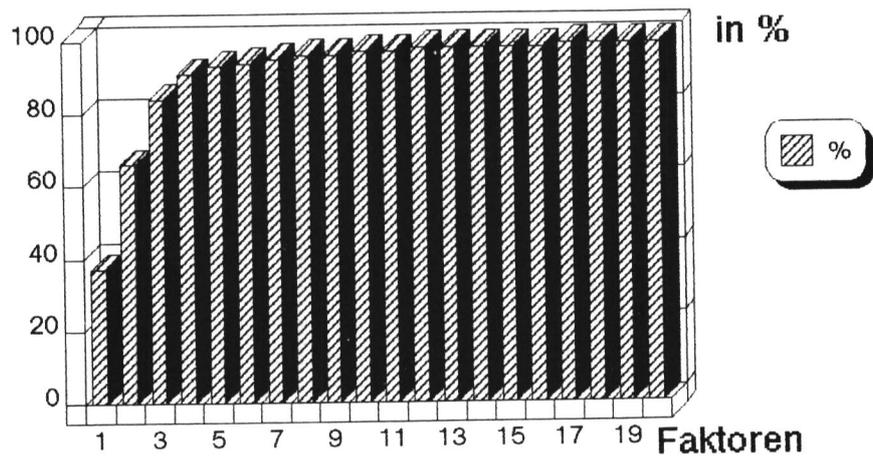


Abb. 32: SUMMAX

Modus der Begriffe

Eigenwerte in

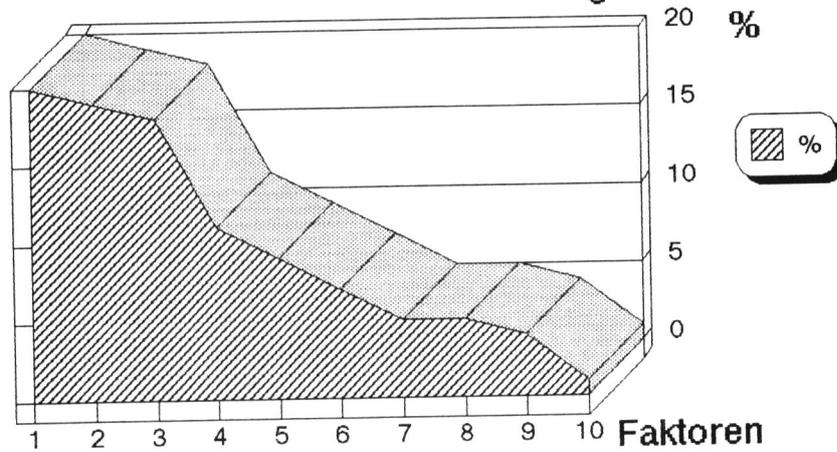
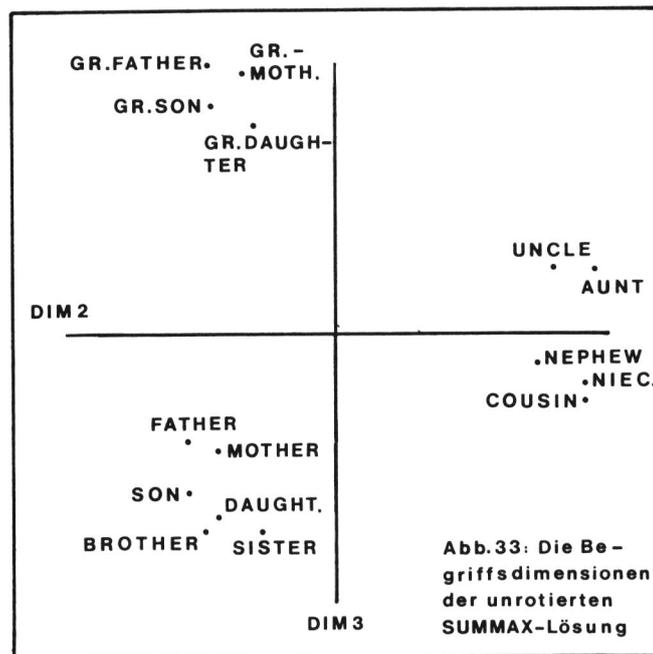


Tabelle 25: Die unrotierte SUMMAX-Lösung, normiert nach dem INDSICAL-Rationale

| 25a) Begriffe: |       |       |       | 25b) Personengruppen: |      |      |      |
|----------------|-------|-------|-------|-----------------------|------|------|------|
|                | DIM1  | DIM2  | DIM3  |                       | DIM1 | DIM2 | DIM3 |
| aunt           | 0.23  | 0.39  | 0.10  | fems                  | 0.18 | 0.71 | 0.52 |
| brot           | -0.26 | -0.19 | -0.29 | mals                  | 0.39 | 0.65 | 0.45 |
| cous           | -0.11 | 0.38  | -0.09 | fem1                  | 0.85 | 0.33 | 0.21 |
| daug           | 0.28  | -0.18 | -0.28 | fem2                  | 0.55 | 0.57 | 0.43 |
| fath           | -0.24 | -0.23 | -0.18 | mal1                  | 0.68 | 0.49 | 0.40 |
| grda           | 0.29  | -0.13 | 0.32  | mal2                  | 0.71 | 0.38 | 0.39 |
| grfa           | -0.25 | -0.20 | 0.43  |                       |      |      |      |
| grmo           | 0.30  | -0.15 | 0.42  |                       |      |      |      |
| grso           | -0.25 | -0.19 | 0.34  |                       |      |      |      |
| moth           | 0.30  | -0.18 | -0.19 |                       |      |      |      |
| neph           | -0.28 | 0.32  | -0.05 |                       |      |      |      |
| niec           | 0.23  | 0.38  | -0.06 |                       |      |      |      |
| sist           | 0.27  | -0.14 | -0.30 |                       |      |      |      |
| son.           | -0.24 | -0.23 | -0.26 |                       |      |      |      |
| uncl           | -0.28 | 0.34  | 0.10  |                       |      |      |      |



auf dem ersten Faktor ab, indem die Geschlechtsdimension für die "single-sort"-Bedingungen gestaucht ist. Ferner implizieren die Gewichte für die Personengruppen auf der zweiten und dritten Dimension keine systematischen Differenzen zwischen den Sortierbedingungen; unter jeder Instruktion erfolgt mindestens eine Sortierung nach Stamm- versus Seitenlinie und Kernfamilie versus Angehörige der zweiten Generation. Die SUMMAX-Konfiguration weist somit auf einer Auflösungsebene aus drei substantiellen Faktoren enge Bezüge zu den Befunden von Rosenberg und Kim auf. Dabei sind insbesondere Aspekte der Generationenfolge in der vorliegenden Lösung, wie auch schon im Rahmen der INDSCAL-Analyse, nur durch eine Dimension repräsentiert. Weitere Klassifikationsmerkmale des untersuchten Reizmaterials, die in Zusammenhang mit dem Modell von Romney und D'Andrade stehen, werden deutlich in der rotierten Konfiguration nach Extraktion von zehn Dimensionen, die in Tabelle 26 aufgeführt ist. Ebenso wie in der Lösung nach dem Tuckerschen Modell wurde aufgrund der klaren Interpretierbarkeit der vierte und fünfte Begriffsfaktor in die Tabelle aufgenommen. Die Dimensionen im Modus der Personengruppen wurden nach der Transformation auf Hauptachsen mittels des Varimax-Kriteriums weiterrotiert, während die Begriffsdimensionen mit der Hauptachsenform identisch sind. Die zugehörige Kernmatrix wurde nach dem Vorgehen von Kroonenberg (1983b) und Tucker (1972) normiert.

#### 9.4.2. Interpretation der rotierten Konfiguration:

Die in Tabelle 26 mit "\*" hervorgehobenen Werte sind substantiell nach dem von Orlik (1980) angegebenen Trennkriterium. Dabei wurde zur Ermittlung bedeutsamer Ladungen im Begriffsmodus eine dreidimensionale Lösung zugrunde gelegt, die für den ersten und dritten Faktor in Abbildung 34 dargestellt ist. Die Dimensionen der unrotierten Konfiguration werden auch hier wieder deutlich, allerdings in veränderter Reihenfolge und mit einer weniger kompakten Clusterstruktur. Die Differenzierung der Familienmitglieder nach ihrer Geschlechtszugehörigkeit erfolgt auf der zweiten Dimension,

Tabelle 26: Die rotierte SUMMAX-Lösung nach Extraktion von zehn Dimensionen

## 26a) Begriffe:

|      | P1     | P2     | P3     | P4    | P5    |
|------|--------|--------|--------|-------|-------|
| aunt | 0.02   | -0.07  | 0.43*  | -0.25 | 0.07  |
| brot | -0.26* | 0.17   | -0.16  | 0.30  | 0.27  |
| cous | 0.14   | 0.13   | 0.34*  | 0.22  | 0.15  |
| daug | -0.19  | -0.34* | -0.03  | 0.07  | -0.29 |
| fath | -0.41* | 0.16   | -0.22  | -0.20 | 0.03  |
| grda | 0.38*  | -0.29* | -0.11  | 0.09  | -0.07 |
| grfa | 0.30*  | 0.16   | -0.34* | -0.20 | 0.13  |
| grmo | 0.33*  | -0.30* | -0.12  | -0.22 | 0.10  |
| grso | 0.34*  | 0.17   | -0.33* | 0.11  | -0.04 |
| moth | -0.34* | -0.33* | 0.01   | -0.22 | 0.02  |
| neph | 0.07   | 0.39*  | 0.17   | 0.09  | -0.22 |
| niec | 0.17   | -0.08  | 0.37*  | 0.10  | -0.17 |
| sist | -0.20  | -0.30* | 0.05   | 0.29  | 0.24  |
| son. | -0.26* | 0.15   | -0.24  | 0.08  | -0.28 |
| uncl | -0.09  | 0.38*  | 0.22   | -0.25 | 0.06  |

26b) Personen-  
gruppen:

|      | M1    | M2     |
|------|-------|--------|
| fems | 0.44  | -1.36* |
| mals | 0.90  | -1.15* |
| fem1 | 1.06* | -0.46  |
| fem2 | 1.00* | -0.82  |
| mal1 | 1.15* | -0.82  |
| mal2 | 1.38* | -0.45  |

## 26c) normierte Kernmatrix:

| M1 | P1    | P2    | P3    | P4    | P5    |
|----|-------|-------|-------|-------|-------|
| P1 | 0.35  | -0.08 | -0.12 | 0.19  | -0.27 |
| P2 | -0.08 | 0.78  | 0.55  | 0.08  | 0.09  |
| P3 | -0.12 | 0.55  | 0.44  | 0.05  | 0.04  |
| P4 | 0.19  | 0.08  | 0.05  | 0.54  | -0.30 |
| P5 | -0.27 | 0.09  | 0.04  | -0.30 | 0.36  |

| M2 | P1    | P2    | P3    | P4    | P5    |
|----|-------|-------|-------|-------|-------|
| P1 | 0.60  | 0.54  | -0.14 | -0.42 | 0.14  |
| P2 | 0.54  | 0.01  | -0.32 | -0.37 | -0.45 |
| P3 | -0.14 | -0.32 | 0.67  | -0.04 | 0.05  |
| P4 | -0.42 | -0.37 | -0.04 | 0.42  | 0.16  |
| P5 | 0.14  | -0.45 | 0.05  | 0.16  | 0.52  |

Die mit "\*" gekennzeichneten Ladungen sind substantiell.

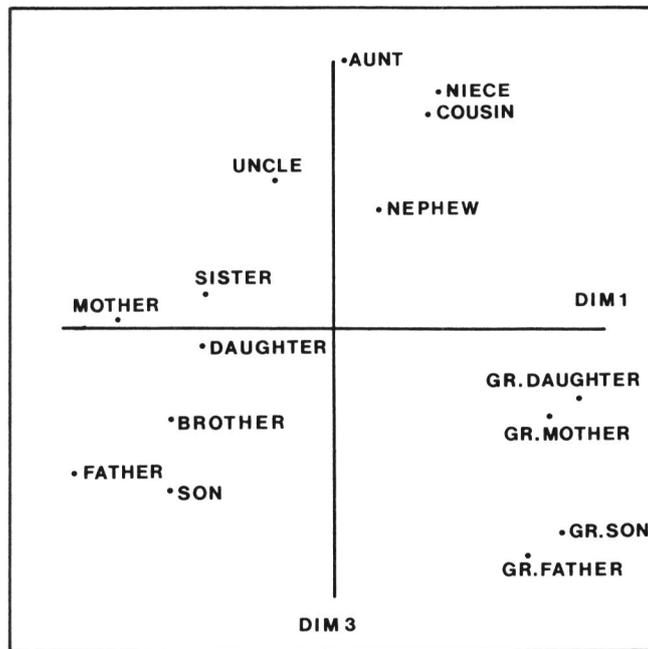


Abb. 34: Die Begriffsdimensionen der rotierten SUMMAX-Lösung

die beiden Pole der ersten Dimension werden gebildet von der Kernfamilie und den Angehörigen der zweiten Generation, und die dritte Dimension differenziert nach der Stamm- bzw. Seitenlinie. Die vierte und fünfte Dimension korrespondiert jeweils zu einem weiteren Aspekt der Generationenfolge: Angehörige der Generation des Betrachters werden Aszendenten der ersten Generation und Deszendenten der ersten Generation gegenübergestellt.

Für die Versuchsbedingungen wird eine Einfachstruktur deutlich, die die beiden Varianten der Sortiertechnik trennt. Nach den bedeutsamen Ladungen umfaßt die erste Dimension alle Personengruppen der "multiple-sort"-Bedingung, die zweite Dimension männliche und weibliche Versuchspersonen der "single-sort"-Bedingung.

Wesentliche Unterscheidungsmerkmale zwischen den Sortiervarianten zeigen sich aufgrund der Kernrelationen, die eine Gewichtung der Geschlechtsdimension enthalten. Während das dimensionsspezifische Gewicht auf höchste Priorität der Geschlechtsdifferenzierung in den multiplen Sortierungen hindeutet, wird die Geschlechtsdimension fast vollständig absorbiert für die Versuchspersonen, die nur Gelegenheit zu einer Sortierung hatten. Angemerkt sei zusätzlich, daß die hohen Zusammenhänge des zweiten Begriffsfaktors mit den übrigen Dimensionen für die "single-sort"-Bedingung aufgrund dieses geringen Dimensionsgewichts zustande kommen (die Kernrelation 0.54 beträgt z.B. vor der Normierung 0.004). Im Gegensatz zur Tuckerschen Konfiguration impliziert die SUM-MAX-Kernmatrix, daß aus der Perspektive der Personengruppen einige Begriffsdimensionen miteinander korreliert sind. So enthalten beispielsweise die Sortierungen nach Stamm- versus Seitenlinie auch Aspekte des Geschlechtes für die Versuchspersonen der "multiple-sort"-Bedingung, während für die Versuchspersonen aus der "single-sort"-Bedingung Gruppierungen nach Vor- und Nachfahren verschiedener Generationen einen Zusammenhang aufweisen. Dabei kann allerdings nach den vorliegenden Ausgangsdaten und den einführend dargestellten Analysen des Datenmaterials an dieser Stelle nicht entschieden werden, welches der beiden Modelle zu

einer angemesseneren Darstellung der Eigenschaften des Datenkörpers führt.

#### 9.5. Schlußbemerkung:

Im vorangegangenen Abschnitt erfolgte der Vergleich der Modelle zur Multidimensionalen Skalierung aufgrund von Ähnlichkeitsdaten, die mit dem klassischen Instrumentarium multivariater Methoden erhoben wurden. Dagegen wurde im vorliegenden Abschnitt ein Datentypus verwendet (zu der Unterscheidung vgl. Seber, 1984, S.347ff), dem eine andersgeartete Operationalisierung des Konzeptes "Ähnlichkeit" zugrunde liegt (Strube, 1984). Trotz des unterschiedlichen Datenmaterials kann eine recht hohe Übereinstimmung hinsichtlich der Schlußfolgerungen aus der Reanalyse der Studie von Rosenberg und Kim und aus der Helmschen Untersuchung festgestellt werden. INDSCAL und die unrotierte SUMMAX-Lösung führen zu einer identischen Hypothesenstruktur, die die drei wesentlichen Eigenschaften der Sortierdaten nach Rosenberg und Kim umfassen. Schwierigkeiten wurden erneut deutlich in der Normierung der Tuckerschen Kernmatrix nach Extraktion der zugehörigen Faktormatrizen; diese waren auf die dimensionale Repräsentation der Personengruppen zurückführbar und konnten ebenfalls durch Rotation nach dem Varimax-Kriterium aufgelöst werden. Im Unterschied zu der vergleichenden Reanalyse der Helmschen Untersuchung wurden im Rahmen der Sortierstudie vergleichbare Aussagen aus beiden quadrilinearen Modellen deutlich. Mit Ausnahme der Implikationen hinsichtlich korrelierter Begriffsdimensionen wurden in der jeweils rotierten Tuckerschen Lösung und SUMMAX-Konfiguration weitgehend übereinstimmende Strukturhypothesen nahegelegt. Neben dem Scree-Test zur Beurteilung der angemessenen Dimensionalität der Konfigurationen wurde in diesen beiden Lösungen das informelle Kriterium der Interpretierbarkeit der Begriffsdimensionen in Betracht gezogen, das in diesem Zusammenhang von Harshman (1984) explizit herangezogen wird. Gegenüber den unrotierten Lösungen erbrachte

dieses Vorgehen zusätzliche Hypothesen zu der Organisationsstruktur der untersuchten Verwandtschaftsbegriffe, die in Beziehung zu dem anthropologischen Modell von Romney und D'Andrade stehen.



10.

Osgood und Luria: ein Fall von multipler Persönlichkeit – ein empirischer Vergleich der dreimodalen Datenmodelle

Der folgende Abschnitt enthält abschließend eine Reanalyse der Untersuchung von Osgood und Luria (1954) zu einem klassischen Fall von multipler Persönlichkeit: Eve White, Jane und Eve Black (Thigpen & Cleckley, 1954). Osgood und Lurias Untersuchungsziel bestand in der Überprüfung der Validität des Semantischen Differentials als diagnostisches Instrument im Rahmen der klinischen Psychologie. In der Einzelfallstudie wurde einer Patientin, die nacheinander drei verschiedene Persönlichkeitszustände manifestierte, in jedem Zustand zweifach ein Semantisches Differential bestehend aus zehn Skalen vorgegeben. Dabei hatte die Patientin wichtige persönliche Beziehungen und Konzepte auf den sieben-stufigen, bipolaren Skalen einzustufen (vgl. Tabelle 27 zu den entsprechenden Skalen und Konzepten). Der von Osgood und Luria (1954) veröffentlichte Datensatz besteht insgesamt aus 3 Persönlichkeitszuständen x 2 Testzeitpunkten x 15 Konzepten x 10 Skalen, der in Anlehnung an Kroonenberg (1983b) durch Zusammenfassung der Persönlichkeiten und Testzeitpunkte als dreimodaler Datensatz behandelt wird. In Übereinstimmung mit dem Vorgehen von Kroonenberg, Osgood und Luria wurden die Skalen valuable - worthless, tasty - distasteful, deep - shallow und active - passive, die im Semantischen Differential in umgepolter Richtung vorgegeben wurden, rekodiert. Der Datenkörper wurde im Anschluß an Kroonenberg durch eine Mittelwertszentrierung der Ratingskalen über alle Persönlichkeiten normiert. Tabelle 28 zeigt einen Auszug aus den Ausgangsdaten. Einstufungen der Konzepte auf den Ratingskalen, die nach Osgood und Luria auf markante Unterschiede zwischen den Persönlichkeiten hinweisen und auf der Datenebene nachweisbar sind, wurden in der Tabelle hervorgehoben.

#### 10.1. Die CANDECOMP/PARAFAC-Lösung :

##### 10.1.1. Verwendete Auswahlkriterien :

Der Auswahl einer angemessenen Konfiguration wurde der Scree-Test der quadrierten Korrelationen zwischen den Ausgangsdaten und dem rückgerechneten Datenkörper, das Konver-

Tabelle 27: Konzepte, Skalen, Persönlichkeiten

-----  
27a) Konzepte :

|                  |                        |                     |
|------------------|------------------------|---------------------|
| love (love)      | mental sickness (sick) | self-control (cont) |
| child (chil)     | my mother (moth)       | hatred (hatr)       |
| my doctor (doct) | peace of mind (peac)   | my father (fath)    |
| me (me)          | fraud (frau)           | confusion (conf)    |
| my job (job)     | my spouse (spou)       | sex (sex)           |

## 27b) Skalen :

-----  
Evaluation :

valuable - worthless (valu)  
 clean - dirty (clea)  
 tasty - distasteful (tast)  
 relaxed - tense (rela)

## Activity:

fast - slow (fast)  
 active - passive (acti)  
 hot - cold (hot)

## Potency :

large - small (larg)  
 strong - weak (stro)  
 deep - shallow (deep)

## 27c) Persönlichkeiten :

-----  
 Eve White I (E.W. I)    Eve Black I (E.B. I)    Jane I (Jane I)  
 Eve White II (E.W. II)    Eve Black II (E.B. II)    Jane II (Jane II)  
 -----

Tabelle 28 : Auszug aus den Ausgangsdaten: Mittelwerte pro Konzept auf den Skalen der Studie von Osgood und Luria

| E.W.,             |      |      |      |      |      |      |      |                   |      |      |      |      |
|-------------------|------|------|------|------|------|------|------|-------------------|------|------|------|------|
| Jane              | Love | Chil | Doct | Job  | Sick | Moth | Peac | Frau <sup>c</sup> | Hatr | Fath | Conf | Sex  |
| hot <sup>a</sup>  | 1.1  | 1.1  | 1.4  | 0.8  | 0.3  | 0.0  | 1.0  | -2.8              | -2.9 | 1.9  | -0.6 | 0.0  |
| valu              | 3.0  | 3.0  | 3.0  | 3.0  | 1.8  | 3.0  | 3.0  | -3.0              | -2.8 | 3.0  | 0.3  | 1.8  |
| rela              | 1.1  | 1.3  | 3.0  | 1.9  | -3.0 | -1.9 | 2.9  | -2.6              | -2.8 | -1.3 | -2.9 | 1.3  |
| larg              | 2.8  | 0.9  | 3.0  | 2.4  | 2.8  | 2.1  | 3.0  | -2.0              | -1.9 | 3.0  | 1.5  | 1.5  |
| fast              | -0.5 | 1.8  | 2.0  | -1.9 | -2.5 | 1.9  | -2.5 | 2.4               | -1.3 | 1.5  | 2.4  | -0.3 |
| clea              | 3.0  | 3.0  | 3.0  | 3.0  | 2.5  | 3.0  | 3.0  | -2.8              | -2.5 | 3.0  | 0.6  | 1.6  |
| stro              | 2.5  | -1.0 | 3.0  | 2.1  | -0.8 | 1.4  | 3.0  | -2.8              | -2.8 | 3.0  | -1.4 | 1.0  |
| tast              | 2.8  | 3.0  | 3.0  | 1.8  | -1.3 | 2.8  | 3.0  | -3.0              | -3.0 | 3.0  | -2.6 | 1.0  |
| deep              | 2.8  | 2.5  | 2.8  | 2.3  | 2.9  | 2.3  | 2.9  | 1.3               | 2.0  | 2.6  | 2.9  | 1.5  |
| acti              | 2.1  | 3.0  | 3.0  | 0.8  | 3.0  | 3.0  | -0.6 | -0.1              | 2.1  | 2.6  | 2.4  | 0.4  |
| E.B.              |      |      |      |      |      |      |      |                   |      |      |      |      |
| Love              | Chil | Doct | Job  | Sick | Moth | Peac | Frau | Hatr              | Fath | Conf | Sex  |      |
| hot <sup>b</sup>  | -1.5 | -3.0 | 3.0  | -3.0 | -3.0 | -2.8 | 2.3  | 2.3               | 2.0  | 2.8  | -1.5 | -3.0 |
| valu              | -3.0 | -2.5 | 3.0  | -3.0 | -3.0 | 1.0  | 2.0  | 2.0               | 2.3  | 3.0  | -3.0 | -3.0 |
| rela              | -3.0 | -2.8 | 3.0  | -1.5 | -2.3 | -1.0 | 2.8  | 2.0               | 1.8  | -2.5 | -2.8 | -3.0 |
| larg              | -3.0 | 0.0  | 3.0  | -0.8 | -1.5 | 0.3  | 2.3  | 2.5               | 2.3  | -1.5 | -1.5 | -3.0 |
| fast              | 0.0  | 0.3  | 3.0  | -3.0 | 3.0  | 3.0  | 3.0  | 3.0               | 2.5  | 2.8  | 1.3  | 1.0  |
| clea              | 0.0  | -2.0 | 3.0  | -1.5 | -2.8 | 1.5  | 2.5  | 1.5               | 2.3  | 3.0  | -2.8 | -1.5 |
| stro              | -3.0 | 0.3  | 3.0  | -3.0 | -3.0 | 1.5  | 2.5  | 2.0               | 2.3  | -1.5 | -3.0 | -3.0 |
| tast              | -3.0 | -2.5 | 2.8  | -3.0 | -3.0 | -2.3 | 2.8  | 2.0               | 2.0  | 3.0  | -1.5 | -3.0 |
| deep              | -3.0 | -0.8 | 3.0  | -3.0 | -1.5 | -0.3 | 2.5  | 2.3               | 2.3  | -0.5 | 0.0  | -3.0 |
| acti              | -3.0 | -3.0 | 3.0  | -3.0 | -3.0 | 3.0  | 3.0  | 2.5               | 2.5  | 3.0  | -1.5 | -3.0 |
| Me                |      |      |      |      |      |      |      |                   |      |      |      |      |
| valu              | larg | clea | stro | tast |      |      |      |                   |      |      |      |      |
| E.W. <sup>b</sup> | -1.0 | -2.5 | 1.8  | -3.0 | -1.3 |      |      |                   |      |      |      |      |
| Jane              | 1.0  | 0.3  | 3.0  | 0.3  | 1.3  |      |      |                   |      |      |      |      |
| E.B.              | 3.0  | 1.5  | 3.0  | 3.0  | 3.0  |      |      |                   |      |      |      |      |

<sup>a</sup>Die angegebenen Werte sind die Mittelwerte über E.W. und Jane jeweils zu beiden Testzeitpunkten. <sup>b</sup>Die Werte sind die Mittelwerte über beide Testzeitpunkte in jedem Persönlichkeitszustand. <sup>c</sup>Markante Unterschiede zwischen den Persönlichkeitszuständen sind hervorgehoben.

genzverhalten des Algorithmus und die Höhe der Faktorinterrelationen innerhalb eines jeden Modus zugrunde gelegt. Die von Harshman (1984) diskutierten "many-fit-diagnostics" sind in diesem Datensatz nicht anwendbar.

Scree-Test der quadrierten Korrelationen  $R^2$  zwischen den Ausgangsdaten und dem rückgerechneten Datenkörper :  
für die zwei- bis zehndimensionale Lösung wurde die Höhe von  $R^2$  gegen die Dimensionalität aufgetragen (vgl. Abbildung 35). Eine deutliche Veränderung des Kurvenverlaufs ergibt sich nach dem dritten Faktor.

Konvergenz der Lösungen :

Für die zwei-, drei- und fünf- bis siebendimensionalen Lösungen wurden lediglich zwei Iterationen durchgeführt, während die Konfiguration aus vier Faktortripeln in 100 Iterationen konvergierte und sich R zwischen der 80ten und 100ten Iteration nur um 0.08% änderte. Nach den Kriterien von Harshman (1984) kann angenommen werden, daß nur die vierdimensionale Lösung eine zuverlässige und stabile Konfiguration enthält.

Überprüfung der Faktorinterkorrelationen innerhalb eines Modus :

Tabelle 29 zeigt die Korrelationen der Faktoren innerhalb der drei Modi aus der vierdimensionalen Lösung. Die Korrelationen zwischen dem ersten und zweiten Faktor weisen in allen Modi hohe Werte auf. Dabei zeigen allerdings die entsprechenden Interrelationen der Faktoren in der zwei- und dreidimensionalen Lösung, die hier nicht aufgeführt wurden, höhere Werte als in der vierdimensionalen Konfiguration.

Nach dem Scree-Test und den Konvergenzeigenschaften wurde die Konfiguration aus vier Faktortripeln ausgewählt, die im folgenden dargestellt wird.

# Abb. 35: CANDECOMP

$R^2$  zwischen Roh – u. geschätzten Daten

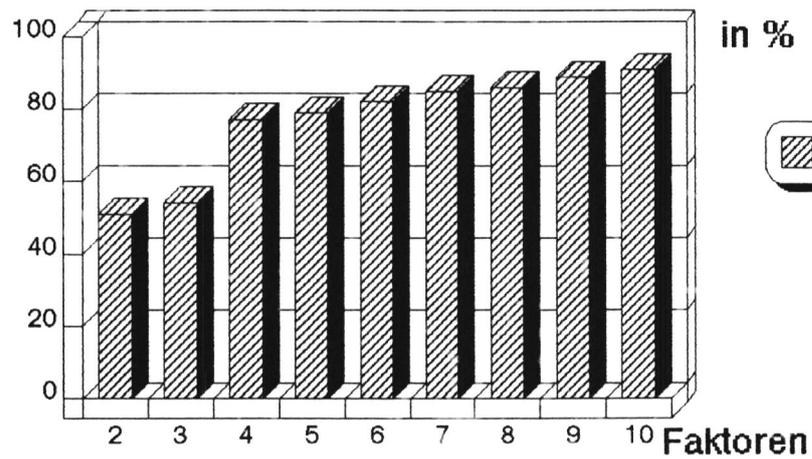


Tabelle 29: Korrelationen der CANDECOMP/PARAFAC-Faktoren

29a) Persönlichkeiten :

|   | 1 | 2    | 3    | 4     |
|---|---|------|------|-------|
| 1 | 1 | 0.97 | 0.58 | -0.98 |
| 2 |   | 1    | 0.58 | -0.98 |
| 3 |   |      | 1    | -0.69 |
| 4 |   |      |      | 1     |

29b) Skalen :

|   | 1 | 2    | 3     | 4     |
|---|---|------|-------|-------|
| 1 | 1 | 0.85 | -0.45 | -0.66 |
| 2 |   | 1    | -0.27 | -0.36 |
| 3 |   |      | 1     | -0.10 |
| 4 |   |      |       | 1     |

29c) Konzepte :

|   | 1 | 2     | 3     | 4    |
|---|---|-------|-------|------|
| 1 | 1 | -0.63 | 0.16  | 0.14 |
| 2 |   | 1     | -0.01 | 0.30 |
| 3 |   |       | 1     | 0.24 |
| 4 |   |       |       | 1    |

10.1.2. Interpretation der CANDECOMP/PARAFAC-Konfiguration: Tabelle 30 zeigt die CANDECOMP/PARAFAC-Faktoren der vierdimensionalen Lösung. Entsprechend der Interpretation der Konfigurationen nach CANDECOMP/PARAFAC (Carroll & Pruzansky, 1984; Harshman & Lundy, 1984) wird ein für alle Personen gemeinsamer Skalenraum und ein gemeinsamer Konzeptraum angenommen, wobei sich die Persönlichkeiten nur in den Gewichten oder "saliences" unterscheiden, die sie den gemeinsamen Konzept- und Skalenfaktoren zuweisen. Die Faktorladungen der Skalen weisen in diesem Datensatz keine Struktur analog den Grunddimensionen des semantischen Raums auf.

Die erste Dimension gibt hauptsächlich die Bedeutsamkeit der Konzepte "sickness", "fraud", "hatred" und "confusion" in Kombination mit dem negativen Pol der Bewertungsskalen sowie "weak" und "fast" für Eve White und Jane wieder, während Eve Black auf der ersten Dimension nicht repräsentiert ist. Die Interpretation der Ladungen auf der ersten Dimension legt somit die Hypothese nahe, daß Eve White und Jane sozial unerwünschte Konzepte ablehnen, während für Eve Black der erste Skalenfaktor kaum bedeutsam ist, um zwischen den Konzepten zu differenzieren.

Auch die zweite Dimension enthält wesentliche Gewichte für Eve White und Jane. Auf dem Konzeptfaktor K2 zeigen neben "love" und "job" erneut "sickness" und "hatred" relativ hohe Ausprägungen. Der Skalenfaktor S2 umfaßt hauptsächlich die Skalen "valuable", "fast", "clean" und "large". Durch das im Vergleich zur ersten Dimension umgepolte Vorzeichen von "sickness" und "hatred" ergeben sich inhaltliche Widersprüche:

während die Konfiguration der ersten Dimension nahelegt, Eve White und Jane stufen "sickness" und "hatred" als ablehnenswert und "fast" ein, folgt aus der zweiten Dimension die entgegengesetzte Hypothese.

Das dritte Faktortripel umfaßt gemeinsame Urteilsaspekte aller Persönlichkeiten. Entsprechend dem Konzeptfaktor K3 werden die Mitglieder der Familie, "child", "mother" und

Tabelle 30: Die CANDECOMP/PARAFAC-Lösung

## 30a) Persönlichkeiten :

|      |    | P1    | P2    | P3    | P4    |
|------|----|-------|-------|-------|-------|
| E.W. | I  | 28.15 | 22.80 | 15.37 | -3.23 |
| E.W. | II | 23.98 | 17.67 | 17.73 | -0.83 |
| E.B. | I  | 3.29  | 0.17  | 13.26 | 26.62 |
| E.B. | II | 2.70  | -5.50 | 15.92 | 29.94 |
| Jane | I  | 24.22 | 24.35 | 17.48 | -3.54 |
| Jane | II | 22.71 | 20.71 | 19.63 | -5.71 |

## 30b) Skalen :

|      | S1    | S2    | S3   | S4    |
|------|-------|-------|------|-------|
| hot  | 0.23  | 0.09  | 0.17 | -0.33 |
| valu | 0.35  | 0.33  | 0.36 | -0.38 |
| rela | 0.41  | 0.13  | 0.08 | -0.28 |
| larg | 0.29  | 0.34  | 0.28 | -0.27 |
| fast | -0.32 | -0.69 | 0.42 | -0.24 |
| clea | 0.31  | 0.32  | 0.39 | -0.28 |
| stro | 0.43  | 0.29  | 0.26 | -0.31 |
| tast | 0.43  | 0.24  | 0.36 | -0.36 |
| deep | 0.03  | 0.18  | 0.31 | -0.29 |
| acti | -0.09 | -0.04 | 0.36 | -0.39 |

## 30c) Konzepte :

|      | K1    | K2    | K3   | K4    |
|------|-------|-------|------|-------|
| love | -0.09 | 0.33  | 0.22 | 0.35  |
| chil | -0.10 | 0.10  | 0.34 | 0.35  |
| doct | 0.13  | -0.02 | 0.40 | -0.12 |
| me   | -0.22 | 0.20  | 0.17 | -0.24 |
| job  | -0.13 | 0.31  | 0.15 | 0.33  |
| sick | -0.53 | 0.60  | 0.26 | 0.31  |
| moth | -0.20 | 0.20  | 0.43 | 0.15  |
| peac | 0.14  | 0.18  | 0.16 | -0.20 |
| frau | -0.31 | 0.03  | 0.03 | -0.24 |
| spou | 0.05  | 0.03  | 0.10 | 0.28  |
| cont | -0.05 | 0.28  | 0.25 | 0.08  |
| hatr | -0.47 | 0.31  | 0.04 | -0.26 |
| fath | -0.09 | 0.16  | 0.43 | 0.06  |
| conf | -0.48 | 0.29  | 0.28 | 0.28  |
| sex  | -0.06 | 0.15  | 0.11 | 0.36  |

Die in der Tabelle angegebenen Werte sind die Faktorladungen der vierdimensionalen Lösung.

"father", wie auch der Therapeut übereinstimmend eingestuft in Richtung positive Evaluation und Activity. Nach den Faktorladungen auf der vierten Dimension, welche die Urteilspezifika von Eve Black enthält, steht das Konzept "child", wie auch "love", "job", "sickness" und "sex", in negativem Zusammenhang mit "valuable", "tasty", "active". Es ergibt sich ein Widerspruch zu der Einstufung von "child" entsprechend der dritten Dimension.

10.1.3. Die Überprüfung der faktorenanalytischen Befunde auf der Datenebene :

Die Hypothesen, die durch das erste Faktortripel nahegelegt wurden, können unter Rückgriff auf die in Tabelle 28 dargestellten Ausgangsdaten nicht bestätigt werden. Es kann weder eine nur geringfügige Differenzierung zwischen den Konzepten von Eve Black nachgewiesen werden noch eine konsistente Einstufung aller Konzepte von Eve White und Jane als "fast". Die Befunde aus der Interpretation des zweiten Faktortripels finden ihre Entsprechung in Tabelle 28. Aufgrund der Ausgangsdaten kann die Beurteilung der Konzepte "love", "job" und "sickness" als "slow" durch Eve White und Jane bestätigt werden, allerdings zeigen sich deutlich die negativen Bewertungen des Konzeptes "hatred".

Die nach der dritten Dimension gemeinsamen Urteilsaspekte aller Persönlichkeiten werden für die Konzepte "doctor", "mother" und "father" durch die Ausgangsdaten gestützt, während die Einstufungen von "child" deutliche Differenzen zwischen Eve Black und Eve White, Jane aufweisen.

Schließlich werden die für Eve Black spezifischen Hypothesen der vierten Dimension durch die Ausgangsdaten klar bestätigt.

Damit findet nur einer der faktorenanalytischen Befunde eine vollständige Entsprechung in den Ausgangsdaten, während die übrigen Hypothesen teilweise bereits in der Interpretation der Faktortripel deutliche Inkonsistenzen aufweisen. Nach den in Abschnitt 6 diskutierten Indikatoren für die Angemes-

senheit einer Konfiguration kann aufgrund der hohen Korrelationen der Faktoren nicht ausgeschlossen werden, daß eine degenerierte Lösung vorliegt.

#### 10.2. Die ALS-Lösung von Kroonenberg :

Die im folgenden diskutierte ALS-Lösung zu dem Tuckerschen Modell ist der Monographie von Kroonenberg (1983b) entnommen. Die Konfiguration besteht aus zwei Skalen-, drei Konzept- und zwei Persönlichkeitsfaktoren. Zugunsten dieser Lösung führt Kroonenberg an, daß die Hinzunahme eines weiteren Skalen- und Persönlichkeitsfaktors die gesamte Varianzaufklärung nur unwesentlich erhöht. Zusätzliche Entscheidungskriterien werden bezüglich der Analyse des Datensatzes von Osgood und Luria nicht mitgeteilt (vgl. auch Finkbeiner, 1984).

##### 10.2.1. Interpretation der Lösung :

Tabelle 31 zeigt die ALS-Lösung von Kroonenberg zu dem Tuckerschen Modell. Der erste Faktor im Modus der Persönlichkeiten repräsentiert Eve White und Jane, der zweite Faktor weist hohe Ladungen spezifisch für Eve Black auf. Die Skalenfaktoren weisen keine Struktur entsprechend den Grunddimensionen des semantischen Raumes auf und werden daher von Kroonenberg - wie auch die Konzeptfaktoren - nicht weiter interpretiert. Die Kernmatrix, die die Faktoren der verschiedenen Modi miteinander verknüpft, ist in Tabelle 31d dargestellt. Die gegensätzlichen Urteilsstrukturen von Eve Black im Kontrast zu Eve White und Jane drücken sich in den Kernrelationen aus, die den ersten Konzept- und Skalenfaktor verbinden: während Eve White und Jane "child", "love" und "job" positiv und als "stark" bewerten, "fraud" aber ablehnen, urteilt Eve Black in gegenläufiger Weise. Der nach Osgood und Luria markanteste Unterschied zwischen den Persönlichkeitszuständen wird in der ALS-Lösung reproduziert.

Tabelle 31: die ALS-Lösung nach Kroonenberg

## 31a) Persönlichkeiten :

|         | P1    | P2   |
|---------|-------|------|
| E.W. I  | 0.47  | 0.10 |
| E.W. II | 0.44  | 0.16 |
| E.B. I  | -0.11 | 0.62 |
| E.B. II | -0.13 | 0.76 |
| Jane I  | 0.51  | 0.06 |
| Jane II | 0.54  | 0.03 |

## 31b) Skalen :

|      | S1   | S2    |
|------|------|-------|
| valu | 0.42 | 0.01  |
| clea | 0.39 | 0.12  |
| tast | 0.38 | -0.23 |
| fast | 0.05 | 0.25  |
| acti | 0.31 | 0.45  |
| hot  | 0.25 | -0.18 |
| larg | 0.34 | 0.08  |
| stro | 0.33 | -0.26 |
| deep | 0.30 | 0.42  |
| rela | 0.21 | -0.62 |

## 31c) Konzepte :

|      | K1    | K2    | K3    |
|------|-------|-------|-------|
| doct | 0.23  | 0.53  | -0.01 |
| peac | 0.19  | 0.45  | 0.20  |
| fath | 0.26  | 0.31  | -0.21 |
| cont | 0.28  | 0.19  | 0.01  |
| moth | 0.24  | 0.17  | -0.32 |
| chil | 0.31  | -0.06 | -0.13 |
| love | 0.39  | -0.14 | 0.05  |
| job  | 0.30  | -0.16 | -0.08 |
| spou | 0.25  | -0.17 | 0.08  |
| sick | 0.23  | -0.23 | -0.38 |
| sex  | 0.26  | -0.25 | 0.05  |
| conf | 0.11  | -0.23 | -0.50 |
| me   | -0.08 | 0.30  | -0.24 |
| hatr | -0.27 | 0.11  | -0.43 |
| frau | -0.30 | 0.08  | -0.38 |

## 31d) Kernmatrix :

| E.W.,<br>Jane | S1 | S2 |
|---------------|----|----|
| K1            | 18 | 0  |
| K2            | 3  | -1 |
| K3            | 0  | -9 |

| E.B. | S1 | S2 |
|------|----|----|
| K1   | -4 | 2  |
| K2   | 13 | 1  |
| K3   | 0  | -4 |

Ferner werden Gemeinsamkeiten aller Persönlichkeiten durch die Kernrelationen zwischen dem ersten Skalenfaktor und dem zweiten Konzeptfaktor angezeigt; "doctor", "peace", "father" und "me" werden übereinstimmend eingestuft in Richtung positive Bewertung und Potency. Schließlich legen die Kernrelationen zwischen dem zweiten Skalenfaktor und dem dritten Konzeptfaktor die Hypothese nahe, daß ein zusätzlicher gemeinsamer Urteilsaspekt der Persönlichkeiten gegeben ist durch die Einstufung der Konzepte "mother", "sickness", "confusion", "hatred" und "fraud" als "active", "deep" und "tense".

#### 10.2.2. Überprüfung der faktorenanalytischen Befunde auf der Datenebene :

Die Strukturhypothese gegensätzlicher Bewertungen der Elemente des ersten Konzeptfaktors von Eve White, Jane und Eve Black entspricht den Einstufungen der Konzepte auf den korrespondierenden Skalengruppen (vgl. Tabelle 28). Auch die übereinstimmende Einstufung von "doctor", "peace of mind" und "father" kann bestätigt werden, jedoch zeigen sich Divergenzen bezüglich des Selbstkonzeptes "me". Während Jane und Eve Black "me" auf den Skalen des ersten Faktors in positiver Richtung einschätzen, besitzt für Eve White das Selbstkonzept eher negative Konnotationen; hierin sehen Osgood und Luria (1954, S.584) einen ausgeprägten Indikator "of disturbance in Eve White".

Die aus der Kombination des zweiten Skalenfaktors und des dritten Konzeptfaktors in der Kernmatrix resultierende Hypothese kann auf der Datenebene nicht vollständig nachgewiesen werden. Im Gegensatz zu der vorhergesagten Übereinstimmung zeigen sich Variationen zwischen den Persönlichkeiten bezüglich der Einstufungen von "sickness" und "confusion" auf den Skalen "active" und "deep" und der Einstufung von "hatred", "fraud" auf der Skala "relaxed".

### 10.3. Die Lösung nach dem Tuckerschen Modell :

Nach der Vorgehensweise von Tucker (1966, 1972) wurden im ersten Auswertungsschritt für jeden Modus die Skalarproduktmatrizen und die zugehörigen Eigenwerte und Eigenvektoren bestimmt. Abbildungen 36 und 37 stellen den Verlauf der Eigenwerte des Konzept- und Skalenmodus in Abhängigkeit von der Anzahl der Faktoren bzw. Eigenvektoren dar. Die Eigenwerte im Modus der Persönlichkeiten betragen in Prozent der Spur der Skalarproduktmatrix: 50%, 34%, 7%, 4%, 3% und 2%. Nach dem Scree-Test zur Beurteilung der angemessenen Dimensionalität (Tucker, 1966) ergibt sich eine zweidimensionale Lösung im Modus der Persönlichkeiten wie auch im Modus der Konzepte; eine Konfiguration aus drei Faktoren wurde für die Skalen angenommen. Tabelle 32 zeigt die in allen Modi Varimax-rotierten Faktormatrizen.

#### 10.3.1. Interpretation der Tuckerschen Lösung :

Im Modus der Persönlichkeiten wird eine Einfachstruktur deutlich, die im wesentlichen der ALS-Lösung von Kroonenberg entspricht. Schwierig ist erneut die Interpretation der Skalenfaktoren, die nicht die Struktur des semantischen Raumes aufweisen. Allerdings weisen der Skalenraum und der Konzeptraum nur geringe Bezüge zu den korrespondierenden Konfigurationen der ALS-Lösung auf.

Nach der in Tabelle 32d dargestellten Kernmatrix werden Unterschiede in den Urteilssystemen zwischen Eve Black und Eve White, Jane durch die Kernrelationen erfaßt, die den ersten Konzeptfaktor mit dem ersten und dritten Skalenfaktor verknüpfen. Demzufolge werden jeweils die Konzepte "love", "job", "sickness" und "sex" auf den Skalengruppen des ersten und dritten Skalenfaktors von Eve White und Jane gleichsinnig, aber von Eve Black gegensätzlich beurteilt.

Kernrelationen, die eine Kombination aus dem zweiten Konzeptfaktor mit dem ersten und dritten Skalenfaktor enthalten, deuten auf gemeinsame Urteilsaspekte aller Persönlichkeiten hin:

gegenläufig zu "doctor" und "peace" impliziert die Kern-

Abb. 36: Tuckersche Loesung

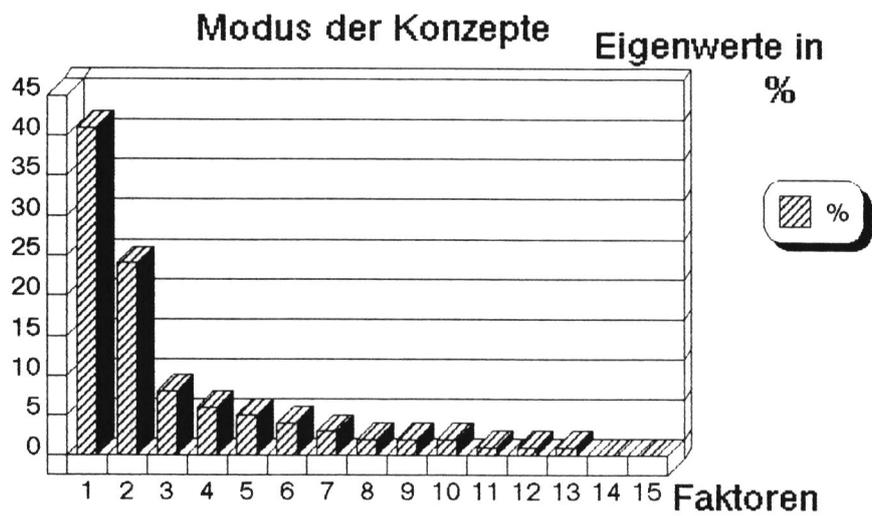


Abb. 37: Tuckersche Loesung

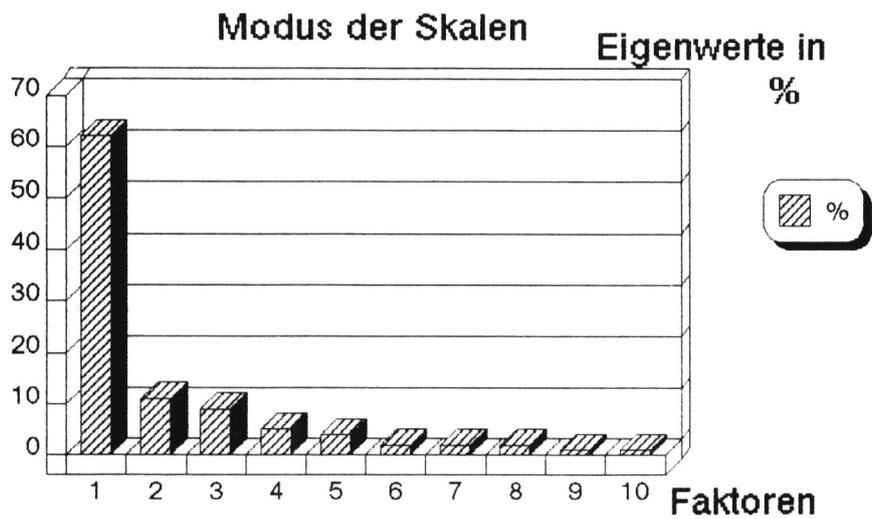


Tabelle 32 : die Tuckersche Lösung mit Varimax-rotierten Faktormatrizen

## 32a) Persönlichkeiten :

|         | P1    | P2    |
|---------|-------|-------|
| E.W. I  | 0.55  | 0.05  |
| E.W. II | 0.48  | 0.08  |
| E.B. I  | -0.01 | 0.67  |
| E.B. II | 0.01  | 0.73  |
| Jane I  | 0.49  | -0.05 |
| Jane II | 0.47  | -0.07 |

## 32b) Skalen :

|      | S1    | S2    | S3    |
|------|-------|-------|-------|
| hot  | 0.19  | -0.02 | 0.23  |
| valu | 0.35  | -0.11 | 0.26  |
| rela | -0.43 | 0.16  | 0.60  |
| larg | 0.16  | -0.11 | 0.29  |
| fast | 0.09  | 0.94  | 0.02  |
| clea | 0.28  | -0.19 | 0.23  |
| stro | -0.04 | 0.01  | 0.46  |
| tast | 0.12  | -0.01 | 0.41  |
| deep | 0.40  | 0.00  | 0.04  |
| acti | 0.61  | 0.18  | -0.05 |

## 32c) Konzepte :

|      | K1    | K2    |
|------|-------|-------|
| love | 0.38  | 0.16  |
| chil | 0.29  | 0.06  |
| doct | -0.18 | 0.41  |
| me   | -0.19 | -0.05 |
| job  | 0.38  | 0.03  |
| sick | 0.37  | -0.18 |
| moth | 0.07  | 0.15  |
| peac | -0.12 | 0.39  |
| frau | -0.21 | -0.43 |
| spou | 0.31  | 0.00  |
| cont | 0.11  | 0.24  |
| hatr | -0.18 | -0.40 |
| fath | -0.02 | 0.29  |
| conf | 0.27  | -0.33 |
| sex  | 0.38  | -0.05 |

## 32d) Kernmatrix:

| E.W.,<br>Jane | S1   | S2    | S3    |
|---------------|------|-------|-------|
| K1            | 9.27 | -6.31 | 10.27 |
| K2            | 9.13 | -5.21 | 32.15 |

| E.B. | S1     | S2    | S3     |
|------|--------|-------|--------|
| K1   | -24.55 | -3.09 | -29.81 |
| K2   | 3.32   | -0.1  | 3.46   |

matrix eine Einstufung von "fraud", "hatred" und "confusion" als ablehnenswerte und "schwache" Konzepte. Zusätzlich wird durch die Kernrelationen, die den Zusammenhang des ersten Konzept- und des zweiten Skalenfaktors aufgrund beider Dimensionen des Modus der Persönlichkeiten kennzeichnen, erwartet, daß die Einstufungen von "love", "job", "sickness" und "sex" auf der Skala "fast-slow" keine Variationen zwischen den Persönlichkeitszuständen aufweisen. Die gleiche Skala bildet schließlich nach der Tuckerschen Kernmatrix Spezifika in den Ratings von Eve White und Jane ab, indem "doctor" und "peace" entgegengesetzt zu "fraud", "hatred" und "confusion" als "fast" klassifiziert wird.

#### 10.3.2. Überprüfung der faktorenanalytischen Befunde auf der Datenebene :

Nach den in Tabelle 28 dargestellten Ergebnissen auf der Datenebene können gegenläufige Bewertungen der Konzepte "love", "job", "sickness" und "sex" zwischen Eve Black und Eve White, Jane für die Skalen nachgewiesen werden, die auf dem ersten Skalenfaktor hohe Ladungen aufweisen. Ferner zeigt die Bewertung dieser Konzepte mit Ausnahme von "sickness" auf dem dritten Skalenfaktor gegensätzliche Urteilsstrukturen auf, wie aus der Interpretation der Kernmatrix erwartet wurde. Nicht nachweisbar ist der Kontrast von "doctor", "peace" zu "fraud", "hatred" und "confusion" auf den Skalengruppen des ersten Skalenfaktors für Eve Black. Nach der oben diskutierten Hypothese sollte sich an dieser Stelle eine übereinstimmende Bewertung aller Persönlichkeiten ergeben, jedoch zeigen sich auf der Datenebene (vgl. auch Osgood & Luria, 1954, S. 584ff) gerade an der Beurteilung von "fraud" und "hatred" markante Unterschiede zwischen den Persönlichkeiten. Entsprechend der Tabelle 28 wird zusätzlich die Kontrastierung von "doctor", "peace" versus "fraud", "hatred" und "confusion" auf dem zweiten Skalenfaktor, der "fast - slow" umfaßt, nicht deutlich für Eve White und Jane.

Damit können die Befunde, die durch die Tuckersche Lösung impliziert werden, in einem wesentlichen Aspekt nicht durch entsprechende Ergebnisse auf der Datenebene gestützt werden. Zugleich folgen in diesem Datensatz aus der Interpretation der ALS-Lösung von Kroonenberg vergleichsweise konsistentere Aussagen insofern, als die grundlegenden Unterschiede zwischen den Persönlichkeitszuständen in der ALS-Lösung repräsentiert sind.

#### 10.4. Die Lösung nach dem SUMMAX-Modell :

##### 10.4.1. Kriterien zur Auswahl einer Lösung :

Der faktoriell aufgeklärte Varianzanteil nach sukzessiver Extraktion der SUMMAX-Faktortripel war Grundlage zur Bestimmung der Anzahl wesentlicher Faktoren. Abbildung 38 stellt die Kurve der Faktor-Ähnlichkeitskoeffizienten (Orlik, 1980, S. 195) in Abhängigkeit von der Anzahl der extrahierten Faktortripel dar. Es zeigt sich keine deutliche Veränderung des Kurvenverlaufs, geringe Veränderungen deuten sich nach dem zweiten, dritten und siebten Faktortripel an. Nach dem siebten Faktortripel wurde bereits eine Varianzaufklärung von 75% erreicht, was als ausreichend angesehen wurde.

Die Eigenwerte der auf Hauptachsen transformierten Faktormatrizen aus der Konfiguration, die sieben Faktortripel umfaßt, betragen in Prozent der Spur der unrotierten Faktormatrizen:

- a) im Skalenmodus:  
53%, 19%, 18%, 8%, 2%, 0,5% und 0,2%
- b) im Modus der Konzepte:  
32%, 25%, 12%, 11%, 8%, 7% und 5%
- c) im Modus der Persönlichkeiten:  
47%, 39%, 12%, 0,5%, und 0,2% .

Im Modus der Persönlichkeiten und der Konzepte wurden jeweils zwei Faktoren beibehalten; im Skalenmodus wurden drei Faktoren als substantiell beurteilt. Tabelle 33 enthält die quadrilineare SUMMAX-Lösung, deren Faktormatrizen über eine

## Abb. 38: SUMMAX

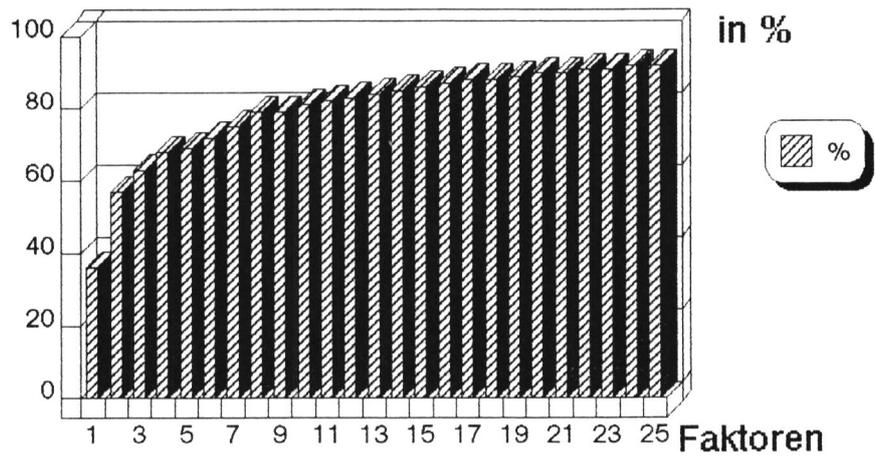
 $R^2$  zwischen Roh – u. geschätzten Daten

Tabelle 33 : auf Hauptachsen und VARIMAX-rotierte SUMMAX-Lösung  
nach Extraktion von 7 Faktortripeln

## 33a) Persönlichkeiten :

|         | P1    | P2    |
|---------|-------|-------|
| E.W. I  | 3.30* | 0.13  |
| E.W. II | 2.92* | -0.19 |
| E.B. I  | -0.35 | 3.69* |
| E.B. II | 0.18  | 3.97* |
| Jane I  | 2.79* | 0.02  |
| Jane II | 2.73* | -0.22 |

## 33b) Skalen :

|      | S1    | S2    | S3     |
|------|-------|-------|--------|
| hot  | 1.39* | 0.01  | -0.57  |
| valu | 2.14* | 0.52  | -0.47  |
| rela | 0.09  | -0.63 | -2.59* |
| larg | 0.84  | 0.67  | -1.39* |
| fast | -0.44 | 2.45* | -0.53  |
| clea | 2.18* | 0.08  | -0.18  |
| stro | 1.15  | 0.32  | -2.10* |
| tast | 2.13* | 0.14  | -1.09  |
| deep | 0.38  | 0.64  | -0.99  |
| acti | 0.88  | 1.55* | 0.24   |

## 33c) Konzepte :

|      | K1     | K2     |
|------|--------|--------|
| love | 1.80*  | -0.09  |
| chil | 0.68   | 0.24   |
| doct | -0.81  | 1.12*  |
| me   | -0.46  | 0.02   |
| job  | 1.45*  | -0.08  |
| sick | 1.26*  | -0.99* |
| moth | 0.02   | 0.47   |
| peac | -0.18  | 1.00*  |
| frau | -1.01* | -1.19* |
| spou | 0.63   | 0.09   |
| cont | 0.44   | 0.42   |
| hatr | -0.42  | -1.67* |
| fath | 0.40   | 0.44   |
| conf | 0.17   | -0.96* |
| sex  | 1.22*  | -0.24  |

33d) Kernmatrix, normiert im Kombinationsmodus Persönlichkeits-  
zustände x Skalen:

| E.W.,<br>Jane | K1    | K2    | E.B. | K1    | K2    |
|---------------|-------|-------|------|-------|-------|
| S1            | 0.57  | 0.82  | S1   | -0.95 | 0.31  |
| S2            | -0.91 | -0.43 | S2   | -0.97 | -0.25 |
| S3            | -0.22 | -0.97 | S3   | 0.99  | -0.16 |

Zwischentransformation auf Hauptachsen anschließend nach dem Varimax-Kriterium weiterrotiert wurden. Die zugehörige Kernmatrix wurde im Kombinationsmodus (Persönlichkeiten x Skalen) normiert (vgl. Orlik, 1980, Demonstrationsexperiment).

#### 10.4.2. Interpretation der Lösung :

Zur Bestimmung bedeutsamer Ladungen wurde das von Orlik (1980, S. 202; vgl. Abschnitt 6) diskutierte Trennkriterium angewendet, wobei substantielle Werte in Tabelle 33 mit "\*" hervorgehoben sind. Wie auch in der ALS-Konfiguration und der Lösung nach dem Tuckerschen Modell repräsentiert der erste Faktor im Modus der Persönlichkeiten Eve White und Jane, und der zweite Faktor weist hohe Ladungen für Eve Black auf. Der erste Skalenfaktor umfaßt im wesentlichen die Bewertungsskalen; der zweite und dritte Faktor enthält substantielle Faktorladungen für jeweils zwei Activity- und zwei Potency-Skalen. Im Modus der Konzepte ist auch hier eine Interpretation der Faktoren mit Schwierigkeiten verbunden.

Entsprechend den Kernrelationen in Tabelle 33d kovariiert für Eve Black der Evaluationsfaktor hoch negativ mit dem ersten Konzeptfaktor, während für Eve White und Jane ein positiver Zusammenhang besteht. Nach der SUMMAX-Analyse besteht damit der hauptsächlichste Gegensatz zwischen den Persönlichkeitszuständen in einer gegenläufigen Evaluation der entsprechenden Konzepte, die für Eve Black eine Ablehnung der Konzepte "love", "job", "sickness" und "sex" und gleichzeitig eine positive Bewertung des Konzeptes "fraud" hervorheben. Eine Gemeinsamkeit besteht zwischen den Persönlichkeitszuständen hinsichtlich der Einstufung der Konzepte des ersten Konzeptfaktors auf den Activity-Skalen; "fraud" ist "fast" und "active", während die Konzepte "love", "job" und "sex" tendenziell "slow" sind. Ferner impliziert die Kernmatrix eine starke Differenzierung zwischen den Konzepten auf den Potency-Skalen für Eve Black, indem sie "fraud" im Gegensatz zu der übrigen Gruppe des ersten Konzeptfaktors als "stark" einstuft.

Bezüglich des zweiten Konzeptfaktors zeigen sich hohe Kernrelationen zu den Skalenfaktoren für Eve White und Jane. Die Kernmatrix legt die Hypothese nahe, daß in beiden Persönlichkeitszuständen Kontraste akzentuiert werden zwischen den Konzepten "doctor", "peace", die als "gut" und "stark" eingestuft werden, versus "fraud", "hatred" und "confusion", die eher abgelehnt werden und "schwache" Konzepte darstellen.

#### 10.4.3. Überprüfung der faktorenanalytischen Befunde auf der Datenebene :

Nach den in Tabelle 28 dargestellten Ergebnissen auf der Datenebene kann die gegenläufige Kontrastierung von "love", "job", "sickness", "sex" und "fraud" auf der Skalengruppe des ersten Skalenfaktors zwischen Eve White, Jane und Eve Black deutlich nachgewiesen werden. Nach Osgood und Luria zeigen sich hier in besonderer Weise die Unterschiede zwischen den Persönlichkeiten; Eve Black's Urteilssystem "becomes completely disoriented from the norm" (Osgood & Luria, 1954, S.584), wie auch aus der faktorenanalytischen Konfiguration erwartet wurde. Ebenso findet die Hypothese, daß alle Persönlichkeiten in gleicher Richtung zwischen diesen Konzepten auf den Skalengruppen des zweiten Skalenfaktors differenzieren, ihre Entsprechung in den Ausgangsdaten. Ferner werden die in der SUMMAX-Lösung implizierten Aussagen zu den Einstufungen der Konzepte des zweiten Konzeptfaktors auf den Bewertungs- und Potency-Skalen für Eve White und Jane durch die Ausgangsdaten bestätigt. Um einen weiteren Vergleich mit der Konfiguration nach dem CANDECOMP/PARAFAC-Modell zu ermöglichen, wird im folgenden zusätzlich die unrotierte SUMMAX-Lösung aus drei Faktortripeln mitgeteilt.

#### 10.4.4. Interpretation der unrotierten SUMMAX-Lösung aus drei Faktortripeln :

Zur vergleichenden Analyse der Konfiguration aus dem trilinearen SUMMAX-Modell in Relation zu den Ergebnissen nach CANDECOMP/PARAFAC wurden die SUMMAX-Faktoren im Konzept- und Skalenmodus auf die Länge eins normiert mit der inversen Normierung im Modus der Persönlichkeitszustände. Tabelle 34

Tabelle 34 : Reskalierung der unrotierten SUMMAX-Lösung aus 3  
 Faktortripeln nach dem Rationale von CANDECOMP

-----  
 34a) Persönlichkeiten :

|         | P1     | P2    | P3   |
|---------|--------|-------|------|
| E.W. I  | -14.33 | 11.15 | 8.78 |
| E.W. II | -12.69 | 11.32 | 4.82 |
| E.B. I  | 16.29  | 16.91 | 5.88 |
| E.B. II | 21.98  | 22.55 | 6.33 |
| Jane I  | -15.84 | 6.48  | 9.92 |
| Jane II | -19.65 | 6.66  | 7.82 |

34b) Skalen :

|      | S1    | S2   | S3    |
|------|-------|------|-------|
| hot  | 0.25  | 0.28 | 0.10  |
| valu | 0.45  | 0.40 | 0.07  |
| rela | 0.25  | 0.23 | 0.37  |
| larg | 0.39  | 0.23 | 0.00  |
| fast | -0.06 | 0.13 | -0.62 |
| clea | 0.35  | 0.38 | 0.22  |
| stro | 0.34  | 0.41 | 0.28  |
| tast | 0.37  | 0.48 | 0.25  |
| deep | 0.28  | 0.17 | -0.20 |
| acti | 0.25  | 0.27 | -0.48 |

34c) Konzepte :

|      | K1    | K2    | K3    |
|------|-------|-------|-------|
| love | -0.42 | -0.17 | 0.34  |
| chil | -0.28 | -0.07 | -0.03 |
| doct | -0.11 | 0.54  | -0.14 |
| me   | 0.19  | 0.11  | 0.13  |
| job  | -0.29 | -0.27 | 0.34  |
| sick | -0.27 | -0.34 | -0.17 |
| moth | -0.16 | 0.12  | -0.26 |
| peac | -0.14 | 0.34  | 0.27  |
| frau | 0.42  | -0.13 | -0.32 |
| spou | -0.20 | -0.19 | 0.29  |
| cont | -0.27 | 0.09  | 0.06  |
| hatr | 0.33  | -0.16 | -0.28 |
| fath | -0.15 | 0.24  | -0.02 |
| conf | -0.05 | -0.29 | -0.45 |
| sex  | -0.26 | -0.32 | 0.31  |

-----

zeigt die nach dem CANDECOMP/PARAFAC-Rationale normierte Lösung.

Die Faktoren des Konzept- und Skalenmodus weisen keine Übereinstimmungen zwischen beiden Konfigurationen auf; im Modus der Persönlichkeiten läßt sich unter dem Aspekt der durch die Faktorladungen implizierten Interpretation der dritte SUMMAX-Faktor dem dritten CANDECOMP-Faktor zuordnen, während die beiden übrigen SUMMAX-Faktoren in der CANDECOMP-Lösung nicht repräsentiert sind. CANDECOMP/PARAFAC und SUMMAX legen dementsprechend unterschiedliche Hypothesen nahe, wie insbesondere an der ersten Dimension deutlich wird. Nach dem ersten SUMMAX-Faktortripel bestehen die wesentlichen Differenzen in den Ratings von Eve Black versus Jane, Eve White in einer gegenläufigen Einstufung der Konzepte "love", "hatred" und "fraud" auf den Bewertungsskalen und der Skala "large-small". Dieser Befund wurde bereits in der ALS-Lösung von Kroonenberg und in der rotierten SUMMAX-Konfiguration deutlich, während eine analoge Aussage in CANDECOMP/PARAFAC nicht enthalten war bzw. zu inhaltlichen Inkonsistenzen führte. Ein neuer Aspekt deutet sich in der Interpretation des zweiten Faktortripels an. Werden die Ladungen der Persönlichkeiten als dimensionsspezifische Gewichte aufgefaßt, so ergibt sich hier eine differentielle Abstufung der Persönlichkeiten nach dem Ausmaß, in dem hauptsächlich die Bewertungsskalen benutzt werden, um zwischen "doctor", "peace" und "sex" zu differenzieren. Für Eve Black bestehen starke Gegensätze zwischen diesen Konzepten, während Jane am wenigsten eine Polarisierung der Konzepte akzentuiert. Eine entsprechende Aufschlüsselung der Ausgangsdaten nach den Konzeptbewertungen für jeden Persönlichkeitszustand ergab folgende Mittelwerte pro Konzept auf der dimensionsentsprechenden Skalengruppe:

- a) Eve Black:  
     "doctor" : 2,9      "peace" : 2,4      "sickness" : -2,9
- b) Eve White:  
     "doctor" : 3,0      "peace" : 3,0      "sickness" : -0,1
- c) Jane:  
     "doctor" : 3,0      "peace" : 3,0      "sickness" : 1,3

Schließlich impliziert das dritte Faktortripel eine Aussage zu übereinstimmenden Urteilsstrukturen in allen Persönlichkeitszuständen, die auch in der rotierten Konfiguration aus SUMMAX enthalten ist, indem "love", "job", "sex" als "slow" und "passive" eingeordnet werden und auf diesen Skalen den Gegenpol zu "fraud" und "confusion" darstellen. Erneut zeigen sich keine Überschneidungen der faktorenanalytischen Befunde mit der CANDECOMP/PARAFAC-Konfiguration.

#### 10.5. Die faktorenanalytischen Konfigurationen im Vergleich:

Im Rahmen der vergleichenden Analyse des Datensatzes von Osgood und Luria zeigten sich deutliche Unterschiede zwischen den Konfigurationen nach CANDECOMP/PARAFAC, SUMMAX, dem ALS-Ansatz von Kroonenberg und dem Tuckerschen Modell hinsichtlich der Widerspruchsfreiheit der Interpretation und der Reproduktion wesentlicher Strukturen in den Ausgangsdaten. In der CANDECOMP/PARAFAC-Lösung zeichneten sich inhaltliche Widersprüche bei der Interpretation der Gesamtkonfiguration ab. Aussagen aufgrund eines Faktortripels wurden teilweise durch die Verknüpfung der Faktoren in einem weiteren Tripel in eine gegenteilige Hypothese gekehrt. Die Lösungen der übrigen Verfahren implizierten dagegen eine konsistente Hypothesenstruktur. Darüber hinaus mußten wesentliche Befunde der CANDECOMP/PARAFAC-Konfiguration nach ihrer Überprüfung anhand der Ausgangsdaten zurückgewiesen werden. In Zusammenhang mit den von Harshman (1984) angegebenen Kriterien für die Angemessenheit einer Lösung kann vermutet werden, daß eine degenerierte Konfiguration interpretiert wurde. Hinsichtlich der Reproduktion markanter Unterschiede zwischen den Persönlichkeitszuständen erwies sich auch die Lösung nach dem Tuckerschen Verfahren als problematisch, während die Anwendung des ALS-Algorithmus zu einer angemessenen Repräsentation der Struktur der Ausgangsdaten führte. Die aus der trilinearen, unrotierten SUMMAX-Lösung resultierenden Aussagen wie auch die Hypothesen aus

der rotierten Konfiguration entsprachen den wesentlichen Befunden auf der Ebene der Ausgangsdaten.

Wie in Abschnitt 6 bereits ausgeführt wurde, zeichnen sich die im Rahmen der empirischen Modelluntersuchungen ausgewählten Datensätze dadurch aus, daß sie aus Modellexperimenten stammen. Die Struktur der zugrunde liegenden Daten war in allen Beispielen bereits vor der Durchführung der Analysen bekannt, und es konnte gezeigt werden (vgl. Orlik, 1965, 1980; Helm, 1964; Rosenberg & Kim, 1975; Osgood & Luria, 1954), daß die Varianzquellen in den analysierten Daten im wesentlichen auf die eingeführten experimentellen Bedingungen zurückführbar sind. Analoge Voraussetzungen hinsichtlich der Variationsquellen dreimodaler Datenkörper sind sicherlich in einer Reihe von Anwendungen nicht gegeben, in denen die betrachteten Ansätze als hypothesen-generierende Verfahren eingesetzt werden. Entsprechende Studien, in denen die Modelle der dreimodalen Faktorenanalyse bzw. der Multidimensionalen Skalierung individueller Differenzen zur Auffindung von Strukturhypothesen in einem zunächst wenig bekannten Untersuchungsfeld verwendet wurden, führten zu Schlußfolgerungen, die nur teilweise in Einklang mit den Befunden der vorliegenden Modelluntersuchungen stehen. Der folgende Abschnitt enthält eine abschließende Diskussion der empirischen und theoretischen Bezüge der betrachteten Modelle auf dem Hintergrund bekannter Forschungsergebnisse.



11.

Diskussion

### 11.1. Implikationen der theoretischen Bezüge zwischen den Modellen im Vergleich zu den Diagonalisierungsansätzen:

Unter methodologischer Perspektive wird Tuckers Ansatz als übergeordnetes Modell betrachtet (Tucker, 1972; Kroonenberg, 1983b), welches durch eine Transformation der Kernmatrix in Diagonalform mit den trilinearen Verfahren verbunden ist; die Tuckersche dreimodale Faktorenanalyse ist den trilinearen Modellen äquivalent, wenn die Kernmatrix auf das dreimodale Analogon einer Identitätsmatrix mit

$$g_{mpq} \begin{cases} = 1 & \text{für } m = p = q \\ = 0 & \text{sonst} \end{cases}$$

eingeschränkt werden kann bzw. wenn die Kernmatrix in allen  $q$  Lagen simultan diagonalisierbar ist:  $g_{mpq} = 0$  für  $m \neq p$ . Entsprechende Verfahren zur Diagonalisierung der Tuckerschen Kernmatrix wurden ursprünglich im Kontext der Multidimensionalen Skalierung individueller Differenzen entwickelt (vgl. Harshman & Lundy, 1984).

Hierzu zählt der Algorithmus von MacCallum (1976), in welchem durch eine Gradientenmethode jene Transformationsmatrix bestimmt wird, die alle symmetrischen Lagen der Kernmatrix simultan in eine Diagonalform überführt. Für dieses Transformationsproblem existiert dann eine exakte Lösung, wenn  $m$  nicht größer als zwei ist, d.h. wenn eine Konfiguration nach dem Tuckerschen Verfahren nicht mehr als zwei Dimensionen im Personenraum enthält. Ist  $m \geq 3$ , so existiert im allgemeinen Fall keine exakte Diagonalform, und es wird mittels einer iterativen Prozedur eine Lösung bestimmt, die die geforderte Diagonalstruktur möglichst gut approximiert. Nach MacCallum (1976, S.397) ist die transformierte Tuckersche Lösung bereits dann zu der entsprechenden INDSCAL-Konfiguration nicht äquivalent, wenn eine einzige Kernrelation außerhalb der Diagonalen einen hohen Wert aufweist.

Den Verfahren, die auf den dreimodalen Fall mit einer nicht-symmetrischen Kernmatrix anwendbar sind, liegt das Prinzip des ALS-Algorithmus zugrunde (Kroonenberg, 1983b; Harshman & Lundy, 1984). Zur Darstellung der Grundidee der Diagonalisierungsansätze geht Kroonenberg von der Repräsentation des Tuckerschen Modells für jede Lage  $X_k$  des Datenkörpers aus:

$$(88) \quad X_k = A \left( \sum_q c_{kq} G_q \right) B' \quad .$$

Nur unter der Annahme, daß die Kernmatrix  $G$  in allen  $q$  Lagen  $G_q$  simultan diagonalisierbar ist, existieren nicht-singuläre Matrizen  $H$  und  $E$ , so daß alle  $q$  Teilmatrizen  $D_q$  mit

$$(89) \quad D_q = H G_q E'$$

gleichzeitig eine Diagonalform aufweisen. Hieraus ergibt sich folgende Modellgleichung für den Tuckerschen Ansatz (Kroonenberg, 1983b, S.57ff):

$$(90) \quad \begin{aligned} X_k &= A \left( \sum_q c_{kq} G_q \right) B' \\ &= A H^{-1} \left( \sum_q c_{kq} H G_q E' \right) (B E^{-1})' \\ &= U \left( \sum_q c_{kq} D_q \right) V' = U W_k V' \end{aligned}$$

mit  $U = A H^{-1}$  ,  $V = B E^{-1}$  ,  $W_k = \sum_q c_{kq} D_q$ .

Da in diesem Fall auch alle Teilmatrizen  $W_k$  diagonal sind, resultiert die äquivalente Repräsentation

$$(91) \quad x_{ijk} = \sum_m u_{im} v_{jm} w_{mmk} = \sum_m u_{im} v_{jm} w_{km} \quad ,$$

wobei die Zeilen der Matrix  $W = (w_{km})$  die Diagonalen der  $W_k$  enthalten. Diese Darstellung ist identisch mit der Grundgleichung von CANDECOMP/PARAFAC und gibt den bisher in der Literatur diskutierten Bezug der Modelle wieder (vgl. Harshman & Lundy, 1984). Wie bereits aufgeführt, ist die wesentliche Voraussetzung, daß die Kernmatrix exakt diagona-

lisierbar ist.

Eine solche exakte Lösung des Diagonalisierungsproblems existiert lediglich für folgende Sonderfälle (Kroonenberg, 1983b):

- a) der dritte Modus umfaßt nur einen Faktor, d.h.  $q = 1$
- b) in allen Modi liegen jeweils zwei Faktoren vor
- c) der erste Modus ist identisch mit dem zweiten Modus und  $q = 2$ .

Sind diese Bedingungen, die eine exakte Diagonalform der Kernmatrix garantieren, nicht erfüllt, so wird durch Anwendung eines "Alternating-Least-Squares"-Algorithmus eine approximative Lösung bestimmt. Dabei wird von der Funktion

$$(92) \quad f(H,E,D) = \sum_k \text{Spur} (G_k - H D_k E')' (G_k - H D_k E')$$

$$\text{mit } X_k = A \left( \sum_q c_{kq} G_q \right) B' = A G_k B'$$

ausgegangen, deren Minimum aufgesucht wird. Wie von Kroonenberg (1983b, S.108ff) gezeigt wurde, wird das Problem der simultanen Diagonalisierung der Teilmatrizen gelöst durch

$$(93) \quad H = U_A' V_A^{-1} \text{ mit } U_A = \sum_k D_k E' G_k', \quad V_A = \sum_k D_k E' E D_k,$$

$$(94) \quad E = U_B' V_B^{-1} \text{ mit } U_B = \sum_k D_k H' G_k, \quad V_B = \sum_k D_k H' H D_k,$$

$$(95) \quad D_k = (d_{1k}, \dots, d_{mk}) \text{ mit } d_{ik} = \sum_m (E' G_k' H)(H' H \times E' E)^{-1}$$

In dem hieraus abgeleiteten Algorithmus werden folgende Schritte iterativ durchlaufen, wobei  $l$  die Anzahl der durchgeführten Iterationen bezeichnet:

(i) "H-Teilschritt":

$$U_{A1} = \sum_k D_{k1} E_1' G_k'$$

$$V_{A1} = \sum_k D_{k1} E_1' E_1 D_{k1}$$

$$Q = U_{A1}' V_{A1}^{-1}$$

$$\Psi = \begin{bmatrix} \psi_1 & & \\ & \ddots & \\ & & \psi_m \end{bmatrix} \quad \text{mit } \psi_{m'} = \left( \sum_m q_{m',m} \right)^{1/2},$$

$$m' = 1, \dots, m$$

$$H_{1+1} = Q \Psi^{-1}$$

(ii) "E-Teilschritt":

$$U_{B1} = \sum_k D_{k1} H_{1+1}' G_k$$

$$V_{B1} = \sum_k D_{k1} H_{1+1}' H_{1+1} D_{k1}$$

$$Q = U_{B1}' V_{B1}^{-1}$$

$$\Psi = \begin{bmatrix} \psi_1 & & \\ & \ddots & \\ & & \psi_m \end{bmatrix} \quad \text{mit } \psi_{m'} = \left( \sum_m q_{m',m} \right)^{1/2},$$

$$m' = 1, \dots, m$$

$$E_{1+1} = Q \Psi^{-1}$$

(iii) "D-Teilschritt":

aus den zuletzt errechneten Transformationsmatrizen H und E wird  $D_{k,1+1}$  bestimmt.

Dieser Algorithmus entspricht in leicht modifizierter Form der Anwendung des CANDECOMP/PARAFAC-Verfahrens auf die Tuckersche Kernmatrix, die zuvor mit der Faktormatrix C aus Modus k verknüpft wurde. Der Diagonalisierungsalgorithmus eignet sich in gleicher Weise auch zur Transformation der Kernmatrix G ohne vorhergehende Verknüpfung mit der Faktormatrix C. Seine Anwendung erfolgt unter einer inhaltlichen

und einer modelltheoretischen Perspektive; einmal dient er der Vereinfachung der Struktur der Kernmatrix, um einen leichteren Zugang zu der Interpretation der Kernmatrix des Tuckerschen Ansatzes zu finden. Zum zweiten wird hauptsächlich die Möglichkeit der Einbettung einer Lösung nach dem Tuckerschen Modell in die Klasse der trilinearen Zerlegungen untersucht. Kann die Kernmatrix einer spezifischen Lösung nicht erfolgreich in eine Diagonalform überführt werden, so wird hieraus die Schlußfolgerung gezogen (vgl. MacCallum, 1976; Harshman & Lundy, 1984), daß allein unterschiedliche Gewichtungen der extrahierten Faktoren zur Erfassung inter-individueller Unterschiede der Komplexität der zugrunde liegenden Daten nicht gerecht werden und zusätzlich multiple Interaktionen der Faktoren in Betracht gezogen werden müssen. In diesem Fall wird eine mangelhafte Diagonalstruktur als Indikator der Unangemessenheit einer CANDECOMP/PARAFAC-Lösung gewertet.

Kroonenberg (1983b, S.116ff) weist daraufhin, daß der beschriebene Algorithmus im Grunde kein Diagonalisierungsverfahren darstellt, denn es wird mit

$$(92) f(H,E,D) = \sum_k \text{Spur}(G_k - H D_k E')'(G_k - H D_k E')$$

ein Ansatz spezifiziert, der selbst keine Funktion der Elemente außerhalb der Diagonalen der Kernmatrix ist. Letzteres wäre durch eine Funktion des Typs

$$(96) g(H,E,D) = \sum_k \text{Spur}(D_k - H G_k E')'(D_k - H G_k E')$$

gewährleistet, in welche die Abweichungen der Kernrelationen von null außerhalb der Diagonalen eingehen. Zur Konstruktion eines adäquaten Verfahrens ist hier zusätzlich zu fordern, daß H und E aus Gleichung (96) eine Determinante vom Betrage eins haben oder orthonormale Transformationsmatrizen sind (Kroonenberg, 1983b). Während für die erst genannte Restriktion bisher keine Algorithmen existieren, sind zur Bestimmung orthonormaler Matrizen H und E entsprechende Verfahren

wiederum von Kroonenberg konzipiert worden. Unter Berücksichtigung der Eigenschaften trilinearere Ansätze, die zu schiefwinkligen und linear abhängigen Faktoren führen können, umfassen orthonormale Transformationen der Kernmatrix ihrerseits nur spezifische Sonderfälle, in denen auch die CANDECOMP/PARAFAC-Faktoren wechselseitig orthogonal sein müssen. Dies ist aber in der Regel nicht erfüllt (vgl. Abschnitt 4).

Neben den vorgenannten Unzulänglichkeiten ist das beschriebene Diagonalisierungsverfahren auch technisch problematisch. Ebenso wie der CANDECOMP/PARAFAC-Algorithmus ist die Prozedur mit dem Problem behaftet, daß möglicherweise nur ein lokales Optimum aufgefunden wird, das mit dem globalen Optimum nicht übereinstimmt. Aufgrund der Konvergenzeigenschaften ist daher nicht gesichert, daß eine Lösung dem eingangs geforderten Minimum auch entspricht. Wenn keine optimale Approximation für die Teilmatrizen  $G_k$  durch  $H D_k E'$  aufgefunden wird, dann weichen die transformierten Matrizen  $H^{-1}G_k(E')^{-1}$  erheblich von einer Diagonalstruktur ab; dies konnte in einer Reihe empirischer Datenanalysen gezeigt werden (Kroonenberg, 1983b). Darüber hinaus konvergierte das Verfahren in einigen Anwendungen gegen eine Lösung, deren Transformationsmatrizen  $H$  und  $E$  die Bedingung der Regularität nicht erfüllten.

Obwohl sich der Diagonalisierungsansatz inzwischen als ein Verfahren zur Herleitung des Zusammenhangs zwischen dem Tuckerschen Modell und den trilinearen Verfahren durchgesetzt hat (vgl. Abschnitt 4), ist er insgesamt durch folgende Restriktionen limitiert:

- a) exakte Lösungen existieren nur für Sonderfälle mit höchstens zwei Faktoren in jedem Modus
- b) das zu bestimmende Optimum ist keine Funktion der Kernrelationen außerhalb der Diagonalen
- c) in der Prozedur wird möglicherweise nur ein lokales Optimum aufgefunden, und die resultierenden Transformationsmatrizen  $H$  und  $E$  sind nicht notwendig regulär.

Dennoch stellt er bei günstiger Datenkonfiguration ein Verfahren dar, welches einen verbesserten Zugang zu der Interpretation der Tuckerschen Kernmatrix liefert.

Das Diagonalisierungsverfahren wurde im wesentlichen konzipiert, um Rückschlüsse vom Tuckerschen Modell auf CANDECOMP/PARAFAC zu ermöglichen. Dagegen enthält der Ansatz der vorliegenden Arbeit einen allgemeinen Zusammenhang zwischen den Modellen, in dem umgekehrt von der trilinearen Klasse auf die quadrilinearen Zerlegungen geschlossen wurde. Dabei wurde vorausgesetzt, daß die CANDECOMP-Repräsentation einen minimalen Rang  $r$  besitzt, so daß  $r = \text{Rang}(X)$  und die Tuckerschen Faktormatrizen den Rang der jeweils zugeordneten Skalarproduktmatrix aufweisen. Die Annahme hinsichtlich der Tuckerschen Faktormatrizen ist bei vollständiger Extraktion der Eigenvektoren aus den Skalarproduktmatrizen, deren Eigenwerte größer als null sind, immer erfüllt. Die Annahme über den Rang der CANDECOMP-Repräsentation impliziert, daß eine genügend hohe Anzahl an CANDECOMP/PARAFAC-Faktoren vorliegt, die keine degenerierte Lösung konstituieren. Auf der Datenebene wurde eine dreimodale Matrix vorausgesetzt, die nicht mit der Nullmatrix identisch ist und deren Teilmatrizen  $X_i$  nicht alle linear abhängig sind. Im Rahmen empirischer Datenanalysen kann davon ausgegangen werden, daß beide Annahmen erfüllt sind; die Einschränkungen des in dieser Arbeit vorgestellten Transformationsansatzes zur Überführung der Konfigurationen aus den einzelnen Verfahren ineinander sind daher wesentlich geringer als in den Diagonalisierungsansätzen.

Innerhalb der trilinearen Klasse konnte eine generelle Äquivalenz des SUMMAX- und des CANDECOMP/PARAFAC-Verfahrens gezeigt werden; d.h. die Faktormatrizen eines jeden Modus aus beiden Ansätzen haben die gleichen Spalten bis auf eine Permutation und Multiplikation mit Skalaren ungleich null, wobei sich die Wirkungen der Permutationen der Faktoren und ihre Reskalierungen wechselseitig aufheben (Numerierung der Gleichungen wie in Abschnitt 5):

$$(57a) \quad \bar{A}_r = A_r P_A D$$

$$(57b) \quad \bar{B}_r = B_r P_B M$$

$$(57c) \quad \bar{C}_r = C_r P_C N \quad \text{mit}$$

$$(58) \quad [P_A D, P_B M, P_C N] = I \quad .$$

Der Übergang zum Tuckerschen Modell erfolgt auf der Grundlage des quadrilinearen SUMMAX-Verfahrens, welches durch Hauptachsentransformation der Faktormatrizen aus der trilinearen Grundgleichung hervorgeht. Eine Reskalierung der SUMMAX-Konfiguration nach dem Rationale des Tuckerschen Ansatzes und anschließende Extraktion der Eigenvektoren der SUMMAX-Kernmatrix liefert die orthogonalen Transformationen, welche die normierten Hauptkomponenten aus SUMMAX in die Faktoren des Tuckerschen Modells überführen:

$$(74a) \quad \bar{A} = A \Delta_A^{-1/2} U_A$$

$$(74b) \quad \bar{B} = B \Delta_B^{-1/2} U_B$$

$$(74c) \quad \bar{C} = C \Delta_C^{-1/2} U_C \quad .$$

Damit sind die betrachteten Verfahren der dreimodalen Faktorenanalyse aus SUMMAX durch einfache Transformationen analytisch ableitbar unter Bedingungen, die zu keiner wesentlichen Einschränkung der Aussagen führen. Unter Spezifikation von drei Transformationen, Reskalierungen, Permutationen und orthogonale Rotationen, ergeben sich Zusammenhänge, die hinsichtlich ihrer Allgemeingültigkeit über die Diagonalisierungsansätze hinausgehen.

Einige Autoren (MacCallum, 1976; Harshman & Lundy, 1984; Kroonenberg, 1983b) sehen einen prinzipiellen Unterschied in der angenommenen Struktur individueller Differenzen zwischen CANDECOMP/PARAFAC und dem Tuckerschen Ansatz. MacCallum kontrastiert das Konzept der Dimensionsgewichtung in den

trilinearen Verfahren mit der ungleich komplexeren Abbildung individueller Differenzen in der Tuckerschen Kernmatrix und zieht die Schlußfolgerung, daß nicht diagonalisierbare Kernmatrizen als Indikatoren der Unangemessenheit der trilinearen Modelle aufzufassen sind. In ähnlicher Weise bemerken Harshman und Lundy (1984, S.182): "If ... there are a few large off-diagonal cells in the core matrix, even after transformation to achieve closest approximation to diagonal form, then by studying the core, the analyst may be able to gain further insight into the pattern of individual difference in the data. Current research ... suggests that strict minimization of all off-diagonal core cells may not lead to the most interpretable rotation when certain kinds of T3 structure are present". Damit wird angenommen, daß das Tuckersche Modell, hier mit "T3" bezeichnet, in Abhängigkeit vom untersuchten Gegenstandsbereich gewisse Strukturaspekte des Datenkörpers abbildet, die durch reine Dimensionsgewichtungen nicht erfaßt werden können. In diesen Fällen sind die Lösungen aus den betrachteten Modellen strukturell nicht vergleichbar, und es ist eine Entscheidung zugunsten eines Modells zu treffen. Die Befunde einer von Rösler (1979) durchgeführten Modelluntersuchung liefern zusätzliche Belege für diese Auffassung. Nach Ansicht des Autors führt die dreimodale Faktorenanalyse nach Tucker zu Konfigurationen, die im Vergleich zu den Parametern trilinearere Verfahren, die individuelle Differenzen repräsentieren, detailliertere Aussagen enthalten. Daher stellt nach Rösler (1979, S.163) das Tuckersche Verfahren aufgrund des Konzeptes der Kernmatrix ein angemesseneres Instrumentarium bereit, wenn Unterschiede zwischen den Urteilssystemen verschiedener Personen durch einen dimensionsanalytischen Ansatz beschrieben werden sollen. Die in der vorliegenden Arbeit diskutierten Bezüge der Modelle zeigen aber, daß der Übergang von dem Konzept differentieller Dimensionsgewichtungen zu dem Konstrukt "multipler Interaktionen" der Faktoren oder Dimensionen bedingt ist durch eine Sequenz von Rotationen. Die Entstehung eines komplexen Faktorengflechts in Form der Kernmatrix ist daher lediglich ein Effekt der Transformationen

der Faktormatrizen aus den trilinearen Verfahren. Die Informationen, die zu individuellen Differenzen in einer trilinearen Konfiguration enthalten sind, werden bei der Einführung von Rotationstransformationen in der Kernmatrix gebündelt. Damit können die Informationen zu interindividuellen Differenzen in beiden Modellklassen aufeinander abgebildet werden. Im Gegensatz zu der Auffassung der genannten Autoren implizieren daher die verschiedenen Ansätze unter theoretischen Gesichtspunkten keine grundlegende Inkompatibilität. Diese ist möglicherweise eher in den zugeordneten Algorithmen und der hohen Anzahl an Faktortripeln begründet, die zur Ausschöpfung relevanter Strukturmerkmale der Daten bei dem Einsatz trilinearere Verfahren erforderlich ist.

#### 11.2. Algorithmen- und Prozessor-bedingte Einschränkungen der empirischen Befunde:

Im Rahmen empirischer Datenanalysen erscheint die Auswahl der Algorithmen und deren Konvergenzeigenschaften kritisch. In SUMMAX und den ALS-Verfahren wird jeweils eine Funktion definiert, deren globales Minimum oder Maximum durch eine iterative Prozedur bestimmt wird. Wie Press, Flannery, Teukolsky und Vetterling (1986) bemerken, existieren bisher keine Algorithmen, die allgemein zu einer globalen Optimierung bei multidimensionalen Problemen führen, und daher werden statt analytischer Verfahren Heuristiken verwendet, die zumeist die Definition einer rationalen Startkonfiguration erfordern. Der SUMMAX-Algorithmus wird initialisiert durch Auswahl einer Startkonfiguration für die Eigenpattern in Modus  $i$  und  $j$  nach dem in Abschnitt 3 dargestellten Vorgehen (vgl. Kohler, 1980). Nach Orlik (1980) entsprechen die Konvergenzeigenschaften des Verfahrens im wesentlichen denen des klassischen Centroid-Ansatzes.

Der ALS-Algorithmus für das INDSCAL- und das CANDECOMP-Verfahren wird mittels einer Zufallskonfiguration initialisiert. Carroll, De Soete und Pruzansky (1989) haben in einer

Monte Carlo-Studie das Konvergenzverhalten des Verfahrens spezifisch für den SINDSCAL-Ansatz untersucht, der eine Analyse symmetrischer "two mode three-way"-Daten gestattet. Zur Ermittlung einer Startkonfiguration wurde neben dem Einsatz von Zufallszahlen u.a. die CANDELINC-Methode (Carroll, Pruzansky & Kruskal, 1980) und die COSPA-Methode von Schönemann (vgl. Carroll, De Soete & Pruzansky, 1989, S.108) verwendet. Bei Anwendung auf 75 Datensätze, die nach der Anzahl der angenommenen Stimuli, Personen, Dimensionen und Höhe des Fehleranteils variierten, ergaben sich Unterschiede in der Zahl der benötigten Iterationen, während die erzeugten Konfigurationen jeweils übereinstimmten. Carroll, De Soete und Pruzansky werten diesen Befund als Indikator für die Qualität des Algorithmus hinsichtlich der Bestimmung des globalen Optimums. Für die Initialisierung des Algorithmus im Rahmen des CANDECOMP-Verfahrens schlagen die Autoren die Verwendung von Eigenvektormatrizen nach dem Tuckerschen Ansatz vor; die Effizienz dieses Vorgehens gegenüber dem Einsatz einer Zufallskonfiguration ist allerdings bisher nicht untersucht worden. Ferner existieren zahlreiche Programmversionen, die spezifische Modifikationen des ALS-Verfahrens umfassen (vgl. Schiffman, Reynolds & Young, 1981). Wie in Abschnitt 6 ausgeführt, wurden im Rahmen der empirischen Datenanalysen der vorliegenden Arbeit die Prozeduren "Candecomp" und "Indscal" des MDS(X)-Programmpakets verwendet. Mögliche Abweichungen der Lösungen infolge unterschiedlicher Algorithmus-Versionen und Startkonfigurationen können für die Ergebnisse nach INDSCAL, CANDECOMP wie auch nach dem Tuckerschen Verfahren nicht abgeschätzt werden, da jeweils nur eine Version zur Verfügung stand.

Für die Konfigurationen nach dem SUMMAX-Algorithmus wurde die verwendete Pascal-Version mit dem Fortran-Programm von Kohler (vgl. Kohler, 1980), implementiert auf dem Siemens-Großrechner der Universität des Saarlandes, anhand des in Abschnitt 10 dargestellten Datensatzes verglichen. Die in den Tabellen von Abschnitt 10 angegebenen Werte stimmen zwischen beiden Programmen überein; allerdings wurden nach

Extraktion des neunten Faktortripels Eigenpattern gefunden, die in einzelnen Komponenten voneinander abweichen. Dies ist bedingt durch die unterschiedliche Rechengenauigkeit der beiden eingesetzten Rechner; wird die Extraktion einer hohen Anzahl an Faktortripeln gewünscht, ist daher die Verwendung des auf dem Großrechner implementierten Programms zu empfehlen.

Der von Kroonenberg und De Leeuw (1980) entwickelte Algorithmus setzt metrische Daten voraus und dient spezifisch der Parameterschätzung des Tuckerschen Modells. In einer Reihe von Anwendungen konnte Kroonenberg (1983b; 1984; Kroonenberg & Ten Berge, 1987; Van Der Kloot & Kroonenberg, 1985) die Güte der Lösungen hinsichtlich ihrer Interpretierbarkeit und des faktoriell aufgeklärten Varianzanteils bei geringer Anzahl an extrahierten Faktoren aufzeigen. Ein entsprechender Befund wurde auch in der in Abschnitt 10 dargestellten Reanalyse des Datensatzes von Osgood und Luria deutlich, in der zusätzlich ein Vergleich der ALS-Lösung von Kroonenberg mit der Konfiguration nach dem Tuckerschen Verfahren vorgenommen wurde. Aus der Interpretation der ALS-Lösung folgten Aussagen, die die grundlegenden, auf der Datenebene aufzeigbaren Befunde reproduzierten, während die Konfiguration nach dem Tuckerschen Ansatz in einem wesentlichen Aspekt keinen Bezug zu den Ausgangsdaten aufwies. Auf das ALS-Verfahren von Kroonenberg und De Leeuw mußte in den weiteren Untersuchungen verzichtet werden, da der im Quelltext zur Verfügung gestellte Algorithmus nicht erfolgreich kompiliert werden konnte.

### 11.3. Alternative non-metrische Algorithmen:

Neben den in der vorliegenden Arbeit eingesetzten Verfahren der dreimodalen Datenanalyse, die metrische Ausgangsdaten voraussetzen, sind Algorithmen entwickelt worden, die auch auf nominal- oder ordinal-skalierte Datenkörper anwendbar sind. Hierzu zählt die ALSCAL-Methode von Takane, Young und

De Leeuw (1977), ALSCOMP von Sands und Young (1980) sowie GEMSCAL, MULTISCALE (vgl. Schiffman, Reynolds & Young, 1981) und PINDIS (Lingoes & Borg, 1984). Die meisten der genannten Entwicklungen mit Ausnahme von ALSCOMP können den Ansätzen der Multidimensionalen Skalierung individueller Differenzen zugeordnet werden und sind daher primär nicht Gegenstand einer Diskussion dreimodaler faktorenanalytischer Verfahren. Ein Überblick und Vergleich der MDS-Verfahren im Rahmen empirischer Datenanalysen findet sich in der Monographie von Schiffman, Reynolds und Young (1980) sowie in den Untersuchungen von Easterling (1987), Jones und MacCallum (1987), Takane, Young und De Leeuw (1977), Takane (1984), Young (1984), Lingoes und Borg (1984).

Sands und Young konzipierten den ALSCOMP-Algorithmus zur faktorenanalytischen Dekomposition nicht-metrischer Daten. ALSCOMP stellt kein neues strukturelles Modell dar, sondern es wird der Tuckersche Ansatz zugrunde gelegt, wobei der dritte Modus nicht faktoriell reduziert wird, wie auch die trilineare Grundgleichung von CANDECOMP/PARAFAC. Während Sands und Young das geforderte Skalenniveau der Daten in den metrischen Verfahren problematisieren und aus dieser Perspektive die allgemeine Verwendbarkeit von ALSCOMP in den Vordergrund stellen, sehen Harshman und Lundy (1984, S.188ff) in der Prozedur keine essentielle Neuerung. Harshman und Lundy führen an, daß nach bisherigen Erfahrungen die non-metrischen Verfahren in den erzeugten Konfigurationen nicht wesentlich von den metrischen Varianten abweichen, aber in höherem Maße mit dem Problem der Auffindung lokaler Optima behaftet sind und als weniger robust zu beurteilen sind.

#### 11.4. Befunde aus den empirischen Datenanalysen:

Es konnten folgende Verfahren in den empirischen Teil der vorliegenden Arbeit einbezogen werden: CANDECOMP/PARAFAC, INDSCAL, SUMMAX, die Tuckerschen Methoden der dreimodalen Faktorenanalyse und Multidimensionalen Skalierung indivi-

dueller Differenzen.

Das Ziel der empirischen Datenanalysen bestand in einer Modelluntersuchung der betrachteten Verfahren der dreimodalen Faktorenanalyse. Dazu wurden Datensätze aus vier Methodenexperimenten ausgewählt, deren zugrunde liegender Versuchsaufbau jeweils eine Vorhersage der relevanten faktoranalytischen Befunde gestattete. Die hauptsächlichsten Struktureigenschaften der Daten waren bereits vor der Analyse bekannt, und daher konnten die Hypothesen, die aus den Konfigurationen der verschiedenen Verfahren resultierten, hinsichtlich der Reproduktion wesentlicher Befunde auf der Ebene der Ausgangsdaten überprüft werden. Die Ergebnisse stehen nur teilweise in Einklang mit den in der Literatur diskutierten Befunden aus Untersuchungen, in denen die Verfahren als heuristische Ansätze eingesetzt wurden.

Obwohl das Tuckersche Modell hinsichtlich der Interpretierbarkeit der Kernmatrix als problematischer Ansatz dargestellt wird (vgl. Bartussek, 1973; Snyder, Law & Hattie, 1984), erwies sich dieser Ansatz in einer Reihe von empirischen Datenanalysen als ein praktikables heuristisches Verfahren. Aus diesen Untersuchungen wurde die Schlußfolgerung gezogen, daß das Modell in hohem Maße geeignet ist, komplexe Datenstrukturen in befriedigender Weise aufzuschlüsseln (Vavra, 1973; Zenisek, 1980; Levin, 1965; Frederiksen, 1972; Cornelius, Hakel & Sackett, 1979; Imada & London, 1979; Kjerulff & Wiggins, 1976; Muthen & Olsson, 1974; Wiggins & Blackburn, 1976; Huysamen, 1971) und differenziertere Aussagen bereitstellen kann als trilineare Ansätze (MacCallum, 1974; Rösler, 1979). Hierbei diente das Tuckersche Verfahren der Auffindung von Strukturhypothesen in teilweise noch wenig explorierten Untersuchungsfeldern. Auch in den Analysen der vorliegenden Arbeit erbrachten die Tuckerschen Konfigurationen nach Durchführung von Faktorrotationen entsprechend dem Varimax-Kriterium bzw. Procrustischen Verfahren interpretierbare, detaillierte Aussagen zu dem zugrunde liegenden Datenmaterial. Allerdings wiesen in der Mehrzahl die Hypothesen keine Bezüge zu den eingeführten

experimentellen Versuchsbedingungen und den Strukturmerkmalen der Daten auf. Lediglich in der Reanalyse der Sortierstudie von Rosenberg und Kim wurden die Befunde auf der Ebene der Ausgangsdaten in der Lösung nach dem Tuckerschen Modell reproduziert. Aufgrund der kontroversen Diskussion um die Angemessenheit von Faktorrotationen wurden zusätzlich die unrotierten Konfigurationen von drei Auswertungen berichtet. Diese waren entweder nicht inhaltlich interpretierbar oder führten zu widersprüchlichen Hypothesen.

Während sich das INDSCAL-Verfahren im Rahmen empirischer Datenanalysen bewährt hat, war nach Carroll und Pruzansky (1984, S.380) der übergeordnete CANDECOMP-Ansatz nur in wenigen Anwendungen erfolgreich. Dagegen konnte Harshman (vgl. Harshman, Ladefoged & Goldstein, 1977; Ladefoged, Harshman, Goldstein & Rice, 1978) die Praktikabilität und Brauchbarkeit des Verfahrens bei angemessenen Normierungen der Daten vor der Analyse demonstrieren. So bemerkt Kruskal (1984, S.50): "... the method has changed in my opinion from an interesting idea to a major advance in geometrical data analysis".

Die Reanalyse des Modellexperimentes von Orlik aus dem Bereich der optischen Wahrnehmung erbrachte Aufschlüsse über die Struktur der Farb- und Größenbeurteilungen, die in hohem Maße mit den aus dem Versuchsaufbau erwarteten Aussagen übereinstimmten. Dagegen ist eine inadäquate Zentrierung des von Osgood und Luria erhobenen Datensatzes wohl nicht die Ursache für die widersprüchlichen Befunde der CANDECOMP/PARAFAC-Konfiguration unter Beachtung der Ergebnisse, die die Anwendung von SUMMAX und des ALS-Verfahrens von Kroonenberg und De Leeuw erbrachte. Nach den von Harshman diskutierten "diagnostics" kann nicht ausgeschlossen werden, daß hier eine nach dem Konvergenzverhalten des Algorithmus degenerierte Lösung interpretiert wurde. Zur Auswertung der beiden übrigen Datensätze wurde das INDSCAL-Verfahren eingesetzt. Entgegen den oben genannten Befunden von Rösler und MacCallum ergaben sich in beiden Fällen interpretierbare "saliences", die die wesentlichen Unterschiede zwischen den

Sortiervarianten in der Studie von Rosenberg und Kim bzw. zwischen den Probanden im Helmschen Experiment widerspiegeln.

Die Lösungen nach dem trilinearen SUMMAX-Modell zur Multidimensionalen Skalierung individueller Differenzen erwiesen sich in diesen beiden Datensätzen als gleichwertig zu den Konfigurationen des eingeführten und häufig angewendeten INDESCAL-Verfahrens; INDESCAL und die unrotierte SUMMAX-Lösung führten zu identischen Aussagen. Ebenso wurde eine hohe Übereinstimmung zwischen der CANDECOMP- und der trilinearen SUMMAX-Konfiguration im Zusammenhang mit der faktorenanalytischen Untersuchung des Modellexperimentes von Orlik deutlich; allerdings erbrachte die Extraktion weiterer SUMMAX-Faktortripel mit anschließenden Hauptachsentransformationen zusätzliche Aufschlüsse über die Struktur der Ratingdaten, die durch die zugeordneten Ergebnisse aus der Auswertung der Ausgangsdaten gesichert werden konnten. Schließlich erwies sich der Datensatz aus der Studie von Osgood und Luria als kritisch in der Gegenüberstellung der dreimodalen faktorenanalytischen Verfahren insofern, als sich erhebliche Unterschiede zwischen den Ansätzen ergaben; im Vergleich mit CANDECOMP/PARAFAC und der Tuckerschen Methode standen nur die Konfigurationen nach dem SUMMAX-Verfahren in Einklang mit den relevanten Befunden auf der Datenebene.

Insgesamt zeigen die Reanalysen der vier Methodenexperimente vor allem wesentliche Unterschiede zwischen dem SUMMAX-Verfahren und dem Tuckerschen Ansatz in der Reproduktion der bekannten Struktureigenschaften der Daten. Hinsichtlich der Generalisierbarkeit dieser Schlußfolgerung müssen die Kriterien, die zur Auswahl der Datensätze verwendet wurden, und das Untersuchungsziel in Betracht gezogen werden. In die Analysen wurden Daten aus Untersuchungen einbezogen, deren Variationsquellen hauptsächlich auf die eingeführten experimentellen Bedingungen zurückführbar waren mit dem Ziel einer Modelluntersuchung. Diese Überprüfbarkeit ist in den Studien nicht gegeben, in denen die Modelle der dreimodalen

Faktorenanalyse als hypothesen-generierende Ansätze, als Heuristiken im Rahmen einer noch zu formulierenden strukturellen Theorie eingesetzt wurden. Nur in wenigen der genannten Untersuchungen wurde ein direkter Methodenvergleich vorgenommen (vgl. MacCallum, 1974; Rösler, 1979). In der Mehrzahl wurde ein Verfahren verwendet, dessen Brauchbarkeit anhand der Interpretierbarkeit der Ergebnisse und deren Kompatibilität mit eventuell schon bestehenden Theorien des untersuchten Gegenstandsbereiches beurteilt wurde. Dagegen orientierte sich die Bewertung der verschiedenen Verfahren in der vorliegenden Arbeit an der Struktur der analysierten Ausgangsdaten. Dies war möglich durch die Variation unabhängiger Variablen in den Versuchsplänen der zugrunde liegenden Untersuchungen und durch die überschaubare Struktur und Größe der einbezogenen Datensätze; die resultierenden Befunde können nicht ohne Beachtung dieser Evaluationskriterien und Untersuchungsziele auf eine Exploration umfangreicheren Materials übertragen werden, in der die Gewinnung von Hypothesen über das Ordnungssystem von abhängigen Variablen im Vordergrund steht.

In den Analysen der vorliegenden Arbeit erwiesen sich die trilinearen Modelle als Ansätze mit hoher Aussagekraft. In der überwiegenden Anzahl der Anwendungen der Verfahren wurde ein hoher Prozentsatz an faktoriell erklärter Varianz bereits nach Extraktion von wenigen Faktoren erreicht, welches einen leichten interpretativen Zugang gestattete. Ist hingegen die Extraktion einer hohen Anzahl an Faktoren bei weniger strukturierten Ausgangsdaten erforderlich, so liefert möglicherweise ein Verfahren eine sparsamere Beschreibung, welches durch Hauptachsentransformationen bestehende Redundanzen (d.h. linear abhängige Faktoren) in den Lösungen trilinearere Ansätze eliminiert und relevante Strukturinformationen durch das Konzept der Kernmatrix erfaßt. Im Vergleich der empirischen Befunde der vorliegenden Arbeit mit den Ergebnissen der in der Literatur diskutierten Anwendungen scheint die Brauchbarkeit trilinearere oder quadrilinearere Modelle wesentlich von der strukturellen Komplexität

der betrachteten Daten abzuhängen. Eine differenzierte Beurteilung der Angemessenheit der Verfahren bei unterschiedlich strukturierten Gegenstandsbereichen und Untersuchungszielen wird künftigen Studien offengehalten bleiben.



12.

Zusammenfassung

Zu den Modellen der dreimodalen Faktorenanalyse, die in empirischen Untersuchungen hauptsächlich eingesetzt wurden und als Grundkonzeptionen in der Literatur diskutiert wurden, gehören der Tuckersche Ansatz, CANDECOMP, PARAFAC und das SUMMAX-Verfahren. In der vorliegenden Arbeit werden die theoretisch-algebraischen Zusammenhänge dieser Modelle sowie Gemeinsamkeiten oder Unterschiede in der Anwendung der Verfahren untersucht.

Die Modelle der dreimodalen Faktorenanalyse werden klassifiziert nach der Anzahl der postulierten Konfigurationen und ihrer Verknüpfungen, welches zu einer Unterteilung in eine trilineare und eine quadrilineare Modellklasse führt. Der trilinearen Klasse ist CANDECOMP, PARAFAC und die Grundform von SUMMAX zugeordnet, zu den quadrilinearen Verfahren gehören der Tuckersche Ansatz und das erweiterte SUMMAX-Verfahren. Während aus dem CANDECOMP-Modell der PARAFAC-Ansatz direkt ableitbar ist, wurde bisher davon ausgegangen, daß sich die Verfahren der beiden Klassen in ihren Eigenschaften so grundlegend unterscheiden, daß keine allgemeingültigen Bezüge zwischen den Modellklassen bestehen. Tatsächlich kann gezeigt werden, daß sich die aus den verschiedenen Modellen resultierenden Faktoren über das SUMMAX-Verfahren ineinander überführen lassen. Dazu wird ein Transformationsansatz vorgestellt, der die analytische Ableitung der quadrilinearen Ansätze aus den trilinearen Modellen ohne einschränkende Voraussetzungen gestattet. Innerhalb der trilinearen Klasse sind die Konfigurationen nach CANDECOMP und PARAFAC durch einfache Reskalierungen und Permutationen der SUMMAX-Faktoren rekonstruierbar; werden die nach der Grundform des SUMMAX-Modells extrahierten Faktoren entsprechend dem CANDECOMP/PARAFAC-Rationale skaliert und angeordnet, so resultieren identische Konfigurationen. Damit stellen die trilinearen Verfahren eine konzeptuell und formal einheitliche Klasse dar. Werden diese Transformationen auf orthogonale Rotationen der Faktoren erweitert, so ist die Tuckersche dreimodale Hauptkomponentenanalyse aus der quadrilinearen Zerlegung des erweiterten SUMMAX-Verfahrens formal ableit-

bar. Aus der Kenntnis der trilinearen SUMMAX-Konfiguration und deren Transformation auf Hauptachsen sind die Faktor- und Kernmatrizen des von Tucker entwickelten Ansatzes rekonstruierbar. Dabei werden den dreimodalen Konfigurationen keine spezifischen Bedingungen unterlegt, so daß die aus dem Transformationsansatz aufzeigbaren modelltheoretischen Zusammenhänge allgemeingültig sind.

Allerdings implizieren die engen algebraischen Bezüge der verschiedenen dreimodalen Verfahren aufgrund der notwendigen Transformationen keine Identität der entsprechenden Konfigurationen. Anhand von vier ausgewählten Datensätzen wird exemplarisch gezeigt, welche divergierenden oder übereinstimmenden Strukturen die Verfahren in empirischen Datenanalysen erzeugen. Dabei stammen die gewählten Datensätze aus Methodenexperimenten, deren zugrunde liegender Versuchsaufbau jeweils eine Vorhersage der relevanten faktorenanalytischen Befunde ermöglicht und deren Struktur bereits vor der Analyse bekannt ist. Wesentliche Unterschiede zeigen sich vor allem zwischen dem SUMMAX-Verfahren und dem Tuckerschen Ansatz in der Reproduktion der bekannten Struktureigenschaften der Daten. Während die Interpretationen der SUMMAX-Konfigurationen in Einklang mit den entsprechenden Ergebnissen auf der Ebene der Ausgangsdaten stehen, weisen die aus den Tuckerschen Konfigurationen ableitbaren Hypothesen in der Mehrzahl keine Bezüge zu den eingeführten experimentellen Versuchsbedingungen auf. Schließlich erweisen sich die CANDECOMP/PARAFAC-Lösungen von drei Datensätzen als gleichwertig zu den entsprechenden SUMMAX-Konfigurationen, während die strukturelle Interpretation der CANDECOMP/PARAFAC-Faktoren aus der Reanalyse des vierten Methodenexperimentes widersprüchliche Hypothesen impliziert. Die in den Modelluntersuchungen gegebene Überprüfbarkeit der Strukturhypothesen aus den verschiedenen Verfahren führt zu einer Differenzierung der Ansätze, die über eine Beurteilung ihrer Brauchbarkeit anhand der reinen Interpretierbarkeit der Ergebnisse hinausgeht.



## Literatur:

- Abelson, R.P. (1958). A discriminant approach to factoring three-way tables (Abstract). *American Psychologist*, 13, 375. ?
- Abelson, R.P. (1960). Scales derived by consideration of variance components in multi-way tables. In: Gulliksen, H. & Messick, S. (Eds.). *Psychological Scaling: theory and applications*. New York: Wiley. ?
- Ahrens, H.J. (1974). Multidimensionale Skalierung. Weinheim: Beltz.
- Arabie, P., Carroll, J.D. & De Sarbo, W. (1987). *Three-way scaling and clustering*. Newbury Park: Sage.
- Bartussek, D. (1973). Zur Interpretation der Kernmatrix in der dreimodalen Faktorenanalyse von L.R. Tucker. *Psychologische Beiträge*, 15, 169-184.
- Bentler, P.M. & Lee, S.Y. (1978). Statistical aspects of a three-mode factor analysis model. *Psychometrika*, 43, 343-352.
- Bentler, P.M. & Lee, S.Y. (1979). A statistical development of three-mode factor analysis. *British Journal of Mathematical and Statistical Psychology*, 32, 87-104.
- Betten, J. (1977). *Elementare Tensorrechnung für Ingenieure*. Braunschweig: Vieweg.
- Bloxom, B. (1968). A note on invariance in three-mode factor analysis. *Psychometrika*, 33, 347-350.
- Borg, I. (1981). *Anwendungsorientierte Multidimensionale Skalierung*. Berlin: Springer.
- Borg, I. (1986). *Facettentheorie: Prinzipien und Beispiele*. *Psychologische Rundschau*, 37, 121-137.
- Bortz, J. (1979). *Lehrbuch der Statistik*. Berlin: Springer.
- Burt, C. (1955). Test reliability estimated by analysis of variance. *British Journal of Statistical Psychology*, 8, 103-118. (with Cattell's data box?)
- Carroll, J.D. (1976). Spatial, non-spatial and hybrid models for scaling. *Psychometrika*, 41, 439-463.
- Carroll, J.D. & Arabie, P. (1980). Multidimensional scaling. *Annual Review of Psychology*, 31, 607-649.
- Carroll, J.D. & Arabie, P. (1983). INDCLUS: an individual differences generalization of the ADCLUS model and the MAPCLUS algorithm. *Psychometrika*, 48, 157-169.

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

Carroll, J.D., De Soete, G. & Pruzansky, S. (1989). An evaluation of five algorithms for generating an initial configuration for SINDSCAL. *Journal of Classification*, 6, 105-119.

Carroll, J.D. & Pruzansky, S. (1980). Discrete and hybrid scaling methods. In: Lantermann, E.D. & Feger, H. (Eds.). *Similarity and choice*. Bern: Huber.

Carroll, J.D. & Pruzansky, S. (1983). Representing proximities data by discrete, continuous or "hybrid" models. In: Felsenstein, J. (Ed.). *Numerical taxonomy*. Berlin: Springer.

Carroll, J.D. & Pruzansky, S. (1984). The CANDECOMP-CANDELINC family of models and methods for multidimensional data analysis. In: Law, H.G., Snyder, C.W., Hattie, J.A. & McDonald, R.P. (Eds.). *Research methods for multimode data analysis*. New York: Praeger.

Carroll, J.D. & Pruzansky, S. (1986). Discrete and hybrid models for proximity data. In: Gaul, W. & Schader, M. (Eds.). *Classification as a tool of research*. Amsterdam: North-Holland.

Carroll, J.D., Pruzansky, S. & Kruskal, J.B. (1980). CANDELINC: a general approach to multidimensional analysis of many-way arrays with linear constraints on parameters. *Psychometrika*, 45, 3-24.

Carroll, J.D. & Wish, M. (1974). Models and methods for three-way multidimensional scaling. In: Krantz, D.H., Atkinson, R.C., Luce, R.D. & Suppes, P. (Eds.). *Contemporary developments in mathematical psychology*, Vol. 2.

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

Cattell, R.B. (1950). *Personality. A systematic theoretical and factual study*. New York: McGraw-Hill.

Cattell, R.B. (1952a). The three basic factor-analytic research designs - their interrelations and derivatives. *Psychological Bulletin*, 49, 499-520.

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

Cattell, R.B. (1966). The data box. Its ordering of total resources in terms of possible relational systems. In: Cattell, R.B. (Ed.). *Handbook of multivariate experimental psychology*, Chicago: Rand McNally.

- Cattell, R.B. (1980a). Guest Editorial: two basic models for personality-environment interaction and the need for their substantive investigation. *Multivariate Behavioral Research*, 15, 243-247.
- Cattell, R.B. (1980b). The separation and evaluation of personal and environmental contributions to behavior by the person-centered model (PCER). *Multivariate Behavioral Research*, 15, 371-402.
- Cattell, R.B., Blaine, D.D. & Brennan, J.M. (1984). N-way factor analysis for obtaining personality-environment-test contribution systems to any response: a plasmode illustration. In: Law, H.G., Snyder, C.W., Hattie, J.A. & McDonald, R.P. (Eds.). *Research methods for multimode data analysis*. New York: Praeger.
- Cornelius, E.T., Hakel, M.D. & Sackett, P.R. (1979). A methodological approach to job classification for performance appraisal purposes. *Personnel Psychology*, 32, 283-297.
- Coxon, A.P.M. (1981). *The MDS(X) series of multidimensional scaling programs*. Program Library Unit, University of Edinburgh, Edinburgh.
- Dunn-Ranking, P. (1983). *Scaling methods*. Hillsdale, N.J.: Erlbaum.
- Easterling, D.V. (1987). Political science: using the generalized euclidean model to study ideological shifts in the U.S. senate. In: Young, F.W. *Multidimensional scaling. History, theory and applications*. Hillsdale, N.J.: Erlbaum.
- Eckart, C. & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1, 211-218.
- Engeln-Müllges, G. & Reutter, F. (1985). *Formelsammlung zur numerischen Mathematik mit Pascal-Programmen*. Mannheim: Bibliographisches Institut.
- Finkbeiner, C.T. (1984). Book review: P.M. Kroonenberg, three-mode principal component analysis. *Journal of Classification*, 1, 133-138.
- Frederikson, N. (1972). Towards a taxonomy of situations. *American Psychologist*, 27, 114-123.
- Gröbner, W. (1966). *Matrizenrechnung*. Mannheim: Bibliographisches Institut.
- Guilford, J.P. (1954). *Psychometric methods*. New York: McGraw-Hill.
- Guttman, L. (1958). What lies ahead for factor analysis? *Educational and Psychological Measurement*, 18, 497-515.
- Harman, H.H. (1960). *Modern factor analysis*. Chicago: University of Chicago Press.

- Harshman, R.A. (1976). PARAFAC: methods of three-way factor analysis and multidimensional scaling according to the principle of proportional profiles. *Dissertation Abstracts International*, 1976, 37, 2478-2479.
- Harshman, R.A. (1984). "How can I know if it's 'real' ?". In: Law, H.G., Snyder, C.W., Hattie, J.A. & McDonald, R.P. (Eds.). *Research methods for multimode data analysis*. New York: Praeger.
- Harshman, R.A. & Berenbaum, S.A. (1981). Basic concepts underlying the PARAFAC-CANDECOMP three-way factor analysis model and its application to longitudinal data. In: Eichorn, D.H., Clausen, J.A., Haan, N., Honzik, M.P. & Mussen, P.H. (Eds.). *Present and past in middle life*. New York: Academic Press.
- Harshman, R.A. & De Sarbo, W. (1984). An application of PARAFAC to a small sample problem, demonstrating preprocessing, orthogonality constraints and split-half diagnostic techniques. In: Law, H.G., Snyder, C.W., Hattie, J.A. & McDonald, R.P. (Eds.). *Research methods for multimode data analysis*. New York: Praeger.
- Harshman, R.A., Ladefoged, P. & Goldstein, L. (1977). Factor analysis of tongue shapes. *Journal of the Acoustical Society of America*, 62, 693-707.
- Harshman, R.A. & Lundy, M.E. (1984). The PARAFAC model for three-way factor analysis and multidimensional scaling. In: Law, H.G., Snyder, C.W., Hattie, J.A. & McDonald, R.P. (Eds.). *Research methods for multimode data analysis*. New York: Praeger.
- Helm, C.E. (1959). Multidimensional ratio scaling analysis of color relations. Technical Report, Princeton University, Princeton, N.J.
- Helm, C.E. (1964). Multidimensional ratio scaling analysis of perceived color relations. *Journal of the Optical Society of America*, 54, 256-262.
- Helm, C.E. & Tucker, L.R. (1962). Individual differences in the structure of color-perceptions. *The American Journal of Psychology*, 75, 437-444.
- Hofstätter, P.R. (1977). *Persönlichkeitsforschung*. Stuttgart: Kröner.
- Holzinger, K.J. & Harman, H.H. (1941). *Factor analysis. A synthesis of factorial methods*. Chicago: University of Chicago Press.
- Horan, C.B. (1969). Multidimensional scaling: combining observations when individuals have different perceptual structures. *Psychometrika*, 34, 139-165.

- Horst, P. (1965). Factor analysis of data matrices. New York: Holt, Rinehart and Winston.
- Huysamen, G.K. (1971). Models for factor analyzing data in a completely crossed three-way design. *Dissertation Abstracts International*, 1971, 32, 4891.
- Imada, A.S. & London, M. (1979). Relationships between subjects, scales and stimuli in research on social perception. *Perceptual and Motor Skills*, 48, 691-697.
- Jolliffe, I.T. (1986). Principal component analysis. New York: Springer.
- Jones, M.R. & MacCallum, R. (1987). Psychology: an application of principal directions scaling to auditory pattern perception. In: Young, F.W. *Multidimensional scaling. History, theory and applications*. Hillsdale, N.J.: Erlbaum.
- Jöreskog, K.G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36, 109-133.
- Kapteyn, A., Neudecker, H. & Wansbeek, T. (1986). An approach to n-mode component analysis. *Psychometrika*, 51, 269-275.
- Kjerulff, K. & Wiggins, N.H. (1976). Graduate student styles for coping with stressful situations. *Journal of Educational Psychology*, 68, 247-254.
- Koecher, M. (1985). *Lineare Algebra und analytische Geometrie*. Berlin: Springer.
- Kohler, A. (1980). Das Trimod-Programm-System (TRIPSY) zur Berechnung der dreimodalen Faktorenanalyse nach Orlik. Manuskript, Fachrichtung Psychologie der Universität Saarbrücken, Saarbrücken.
- Kowalsky, H.J. (1970). *Lineare Algebra*. Berlin: De Gruyter.
- Kroonenberg, P.M. (1983a). Annotated bibliography of three-mode factor analysis. *British Journal of Mathematical and Statistical Psychology*, 36, 81-113.
- Kroonenberg, P.M. (1983b). *Three-mode principal component analysis*. Leiden: DSWO Press.
- Kroonenberg, P.M. (1984). Three-mode principal component analysis: illustrated with an example from attachment theory. In: Law, H.G., Snyder, C.W., Hattie, J.A. & McDonald, R.P. (Eds.). *Research methods for multimode data analysis*. New York: Praeger.
- Kroonenberg, P.M. (1985). User's guide to TUCKALS3 (Version 4.0). WEP-reeks, WR 85-09-RP, Vakgroep W.E.P., University of Leiden, Leiden.

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.

Kroonenberg, P.M. & Ten Berge, J.M.F. (1987). Cross-validation of the WISC-R factorial structure using three-mode principal components analysis and perfect congruence analysis. *Applied Psychological Measurement*, 11, 195-210.

Kruskal, J.B. (1976). More factors than subjects, tests and treatments: an indeterminacy theorem for canonical decomposition and individual differences scaling. *Psychometrika*, 41, 281-293.

Kruskal, J.B. (1977). Three-way arrays: rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics. *Linear Algebra and its Applications*, 18, 95-138.

Kruskal, J.B. (1981). Multilinear models for data analysis. *Behaviormetrika*, 10, 1-20.

} { Kruskal, J.B. (1983). Multilinear methods. *Proceedings of Symposia in Applied Mathematics*, 28, 75-104.

Kruskal, J.B. (1984). Multilinear methods. In: Law, H.G., Snyder, C.W., Hattie, J.A. & McDonald, R.P. (Eds.). *Research methods for multimode data analysis*. New York: Praeger.

Ladefoged, P., Harshman, R., Goldstein, L. & Rice, L. (1978). Generating vocal tract shapes from formant frequencies. *Journal of the Acoustical Society of America*, 64, 1027-1035.

Lamprecht, E. (1978). *Einführung in die Algebra*. Basel: Birkhäuser.

Lamprecht, E. (1980). *Lineare Algebra 1*. Basel: Birkhäuser.

Lamprecht, E. (1983). *Lineare Algebra 2*. Basel: Birkhäuser.

Lastovicka, J.L. (1981). The extension of component analysis to four-mode matrices. *Psychometrika*, 46, 47-57.

Law, H.G. & Snyder, C.W. (1979). Three-mode models for the analysis of psychological data (Abstract). *American Psychologist*, 14, 214.

Levin, J. (1963). Three-mode factor analysis. *Dissertation Abstracts International*, 1963, 24, 5530-5531.

Levin, J. (1965). Three-mode factor analysis. *Psychological Bulletin*, 64, 442-452.

Lilly, R.S. (1965). A developmental study of the semantic differential. *Dissertation Abstracts International*, 1966, 26, 4063-4064.

- Lingoes, J.C. & Borg, I. (1984). A direct approach to individual differences scaling using increasingly complex transformations. In: Law, H.G., Snyder, C.W., Hattie, J.A. & McDonald, R.P. (Eds.). Research methods for multimode data analysis. New York: Praeger.
- Lohmöller, J.B. (1979). Die trimodale Faktorenanalyse von Tucker: Skalierungen, Rotationen, andere Modelle. Archiv für Psychologie, 131, 137-166.
- Lyttkens, E. (1966). On the fix-point property of Wold's iterative estimation method for principal components. In: Krishnaiah, P.R. (Ed.). Multivariate analysis. New York: Academic Press.
- MacCallum, R.C. (1974). A comparison of two individual differences models for multidimensional scaling: Carroll and Chang's INDSCAL and Tucker's three-mode factor analysis. Dissertation Abstracts International, 1975, 35, 3619.
- MacCallum, R.C. (1976). Transformations of a three-mode multidimensional scaling solution to INDSCAL form. Psychometrika, 41, 385-400.
- Mahmoud, A.F. (1955). Test reliability in terms of factor theory. British Journal of Statistical Psychology, 8, 119-135. *(vgl. Burt)*
- Miller, G. (1969). A psychological method to investigate verbal concepts. Journal of Mathematical Psychology, 6, 169-191.
- Muthen, B. & Olsson, U. (1974). Tucker's three-mode factor model as applied to religious semantic differential data (Abstract). Multivariate Behavioral Research, 9, 498-499.
- Orlik, P. (1965). Eine Modellstudie zur Psychophysik des Polaritätsprofils. Zeitschrift für experimentelle und angewandte Psychologie, 12, 614-647.
- Orlik, P. (1967). Das Dilemma der Faktorenanalyse - Zeichen einer Aufbaukrise in der modernen Psychologie. In: Diemer, A. (Ed.). Geschichte und Zukunft. Meisenheim: Hain.
- Orlik, P. (1973). The SUMMAX method of factor analysis - a new basic linear data model? Arbeiten der Fachrichtung Psychologie, Universität des Saarlandes, Saarbrücken.
- Orlik, P. (1976). The SUMMAX model of three-mode factor analysis. Arbeiten der Fachrichtung Psychologie, Universität des Saarlandes, Saarbrücken.
- Orlik, P. (1980). Das SUMMAX-Modell der dreimodalen Faktorenanalyse mit interpretierbarer Kernmatrix. Archiv für Psychologie, 133, 189-218.

Orlik, P. (1981). Das SUMMAX-Modell der dreimodalen multiplen Prädiktion, gezeigt an einem Demonstrationsexperiment. In: Janke, W. (Ed.). Beiträge zur Methodik in der differentiellen, diagnostischen und klinischen Psychologie. Königstein/Ts.: Hain.

Osgood, C.E. & Luria, Z. (1954). A blind analysis of a case of multiple personality using the semantic differential. *Journal of Abnormal and Social Psychology*, 49, 579-591.

Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. (1957). *The measurement of meaning*. Urbana, Ill.: University of Illinois Press.

Pawlik, K. (1968). *Dimensionen des Verhaltens*. Bern: Huber.

Pohl, R.W. (1967). *Optik und Atomphysik*. Berlin: Springer.

Press, W.H., Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T. (1986). *Numerical recipes. The art of scientific computing*. Cambridge: Cambridge University Press.

Rao, C.R. (1964). The use and interpretation of principal component analysis in applied research. *Sankhya A*, 26, 329-358.

Rao, C.R. & Mitra, S.K. (1971). *Generalized inverse of matrices and its applications*. New York: Wiley.

Revenstorf, D. (1980). *Faktorenanalyse*. Stuttgart: Kohlhammer.

Rösler, F. (1979). Identifying interindividual judgment differences: INDSCAL or three-mode factor analysis? *Multivariate Behavioral Research*, 14, 145-167.

Rosenberg, S. & Kim, M.P. (1975). The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, 489-502.

Rowe, H.A.H. (1979). Three-mode factor analysis: problems of interpretation and possible solutions (Abstract). *American Psychologist*, 14, 222-223. *Australian*

Sands, R. & Young, F.W. (1980). Component models for three-way data: an alternating least squares algorithm with optimal scaling features. *Psychometrika*, 45, 39-67.

Schiffman, S.S., Reynolds, M.L. & Young, F.W. (1981). *Introduction to multidimensional scaling. Theory, methods and applications*. New York: Academic Press.

Seber, G.A.F. (1984). *Multivariate observations*. New York: Wiley.

Soli, S.D. & Arable, P. (1979). Auditory versus phonetic accounts of observed confusions between consonant phonemes. *Journal of the Acoustical Society of America*, 66, 46-59.

- Snyder, F.W. (1969). A unique variance model for three-mode factor analysis. *Dissertation Abstracts International*, 1969, 30, 1349.
- Snyder, C.W., Law, H.G. & Hattie, J.A. (1984). Overview of multimode analytic methods. In: Law, H.G., Snyder, C.W., Hattie, J.A. & McDonald, R.P. (Eds.). *Research methods for multimode data analysis*. New York: Praeger.
- Strube, G. (1984). Techniken der empirischen Wortfeldanalyse. In: von Eye, A. & Marx, W. (Eds.). *Semantische Dimensionen. Verhaltenstheoretische Konzepte einer psychologischen Semantik*. Göttingen: Hogrefe.
- Takane, Y. (1984). The weighted additive model. In: Law, H.G., Snyder, C.W., Hattie, J.A. & McDonald, R.P. (Eds.). *Research methods for multimode data analysis*. New York: Praeger.
- Takane, Y., Young, F.W. & De Leeuw, J. (1977). Nonmetric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features. *Psychometrika*, 42, 7-67.
- Thigpen, C.H. & Cleckley, H. (1954). A case of multiple personality. *Journal of Abnormal and Social Psychology*, 49, 135-151.
- Thurstone, L.L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.
- Torgerson, W.S. (1958). *Theory and methods of scaling*. New York: Wiley.
- Torgerson, W.S. (1986). Scaling and Psychometrika: spatial and alternative representations of similarity data. *Psychometrika*, 51, 57-63. kritisch  
INSICAL  
etc.
- Tucker, L.R. (1964a). The extension of factor analysis to three-dimensional matrices. In: Frederiksen, N. & Gulliksen, H. (Eds.). *Contributions to mathematical psychology*. New York: Holt, Rinehart and Winston.
- Tucker, L.R. (1964b). Experiments in multimode factor analysis. In: Anastasi, A. (Ed.). *Testing problems in perspective*. Princeton, N.J.: Educational Testing Service.
- Tucker, L.R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, 31, 279-311.
- Tucker, L.R. (1967). Three-mode factor analysis of Parker-Fleishman complex tracking behavior data. *Multivariate Behavioral Research*, 2, 139-151.
- Tucker, L.R. (1972). Relations between multidimensional scaling and three-mode factor analysis. *Psychometrika*, 37, 3-27.

- Tucker, L.R. & Messick, S. (1963). An individual differences model for multidimensional scaling. *Psychometrika*, 28, 333-367.
- Turbo Lader Statistics. Wiesbaden: Lauer & Wallwitz (1986).
- Überla, K. (1968). *Faktorenanalyse*. Berlin: Springer.
- Van Der Kloot, W.A. & Kroonenberg, P.M. (1985). External analysis with three-mode principal component models. *Psychometrika*, 50, 479-494.
- Vavra, T.G. (1974). A three-mode factor analytic investigation into the effectiveness of advertising. *Dissertation Abstracts International*, 1974, 34, 7802-7803.
- Wallace, A.F.C. & Atkins, J. (1960). The meaning of kinship terms. *American Anthropologist*, 62, 58-80.
- Weber, E. (1974). *Einführung in die Faktorenanalyse*. Stuttgart: Fischer.
- Wiggins, N.H. & Blackburn, M.C. (1976). Implicit theories of personality: an individual differences approach. *Multivariate Behavioral Research*, 11, 267-285.
- Wold, H. (1966). Estimation of principal components and related models by iterative least squares. In: Krishnaiah, P.R. (Ed.). *Multivariate analysis*. New York: Academic Press.
- Young, F.W. (1984). The general euclidean model. In: Law, H.G., Snyder, C.W., Hattie, J.A. & McDonald, R.P. (Eds.). *Research methods for multimode data analysis*. New York: Praeger.
- Young, F.W. (1987). *Multidimensional scaling: history, theory and applications*. Hillsdale, N.J.: Erlbaum.
- Zenisek, T.J. (1980). The measurement of job satisfaction: a three-mode factor analysis. *Dissertation Abstracts International*, 1980, 41, 75.

Thomas Marx

**Die multivariaten statistischen Analysemethoden  
und deren Anwendung auf den Weinbau**

Frankfurt/M., Bern, New York, 1986. 194 S.

Europäische Hochschulschriften: Reihe 5, Volks- und Betriebswirtschaft. Bd. 688

ISBN 3-8204-8385-3

br. DM 52.--/sFr. 43.--

Die vorliegende Untersuchung hat das Ziel, Zusammenhänge zwischen den in der empirischen Wirtschafts- und Sozialforschung gängigen statistischen univariaten und multivariaten Analysemethoden herzustellen und zu erklären. Als einheitliches theoretisches Modell erweist sich ein ausführlich bewiesenes lineares Regressionsmodell, auf dessen Grundlage Gemeinsamkeiten der statistischen Verfahren in allen Einzelheiten dargestellt werden. Das Modell erfährt eine empirische Anwendung im Rahmen einer Analyse des Badischen Weinsektors mittels Daten aus der chemischen und sensorischen Qualitätsweinprüfung.



**Verlag Peter Lang Frankfurt a.M. · Bern · New York · Paris**

Auslieferung: Verlag Peter Lang AG, Jupiterstr. 15, CH-3000 Bern 15  
Telefon (004131) 321122, Telex pela ch 912 651, Telefax (004131) 321131

- Preisänderungen vorbehalten -

Michael Röchner

**Personspezifische Aspekte und Determinanten  
der Weiterbildungsteilnahme**  
Eine empirische Analyse eines multivariaten Modells

Frankfurt/M., Bern, New York, 1987. VIII, 226 S.  
Europäische Hochschulschriften: Reihe 11, Pädagogik. Bd. 324  
ISBN 3-8204-0215-2 br./lam. DM 60.--/sFr. 50.--

Die Frage, warum Erwachsene Weiterbildungsveranstaltungen besuchen, berührt die Organisatoren solcher Veranstaltungen in vielfältiger Weise. Konkrete Antworten scheitern jedoch häufig am Fehlen theoretischer und empirischer Grundlagen als Folge einer bisher nur in Ansätzen formulierten Weiterbildungstheorie.

Über die kritische Aufarbeitung deutschsprachiger und anglo-amerikanischer Partizipationsmodelle hinaus wird hier ein multivariates Modell personspezifischer Bedingungen zur Weiterbildungsteilnahme skizziert und in einer empirisch-kausalanalytischen Vorgehensweise überprüft.

*Aus dem Inhalt:* Der Stand der Bedeutungsentwicklung der Erwachsenenbildung – Adressaten und Teilnehmerforschung im Rahmen der Weiterbildungsforschung – Theoretische Ansätze zur Weiterbildungsforschung – Kritik und Weiterentwicklung des Motivationskonstrukts in der Partizipationsforschung – Entwurf eines multivariaten Modells zu Determinanten der Weiterbildung.



**Verlag Peter Lang Frankfurt a.M. · Bern · New York · Paris**  
Auslieferung: Verlag Peter Lang AG, Jupiterstr. 15, CH-3000 Bern 15  
Telefon (004131) 321122, Telex pela ch 912 651, Telefax (004131) 321131

- Preisänderungen vorbehalten -



Faktorenanalyse bezeichnet eine Klasse von multivariaten Verfahren, welche die Rückführung einer Vielzahl von Variablen auf wenige Grundvariablen - die Faktoren - gestattet. In den traditionellen Ansätzen wird davon ausgegangen, daß die Variablen an den Individuen einer Stichprobe unter einer einzigen Untersuchungsbedingung gemessen oder beobachtet wurden. Liegen Beobachtungen aus unterschiedlichen Untersuchungsbedingungen vor, so kommen die Verfahren der dreimodalen Faktorenanalyse zum Einsatz, mit dem Ziel, zusätzlich Unterschiede zwischen den Individuen und den situativen Bedingungen in Form von Faktoren darzustellen. In der vorliegenden Arbeit werden die Verfahren der dreimodalen Faktorenanalyse diskutiert. Allgemeingültige algebraische Zusammenhänge zwischen den Verfahren, die bisher unverbunden nebeneinander stehen, werden abgeleitet. Eine vergleichende Modelluntersuchung wird im empirischen Teil anhand der Daten aus Experimenten vorgenommen, die aufgrund des zugrundeliegenden Versuchsaufbaus eine Bewertung wesentlicher Hypothesen aus der Anwendung der verschiedenen Verfahren gestatten.

Sabine Krolak-Schwerdt wurde 1958 in Düsseldorf geboren. Sie studierte Psychologie an der Universität des Saarlandes in Saarbrücken und schloß das Studium 1986 mit dem Diplom ab. Nach Teilnahme an einem Forschungsprojekt von 1986 bis 1987 ist sie seit 1988 als wissenschaftliche Mitarbeiterin des Psychologischen Instituts in Saarbrücken tätig.