

## Einführung in die Logik der Skalierungsverfahren

Bacher, Johann

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

### Empfohlene Zitierung / Suggested Citation:

Bacher, J. (1990). Einführung in die Logik der Skalierungsverfahren. *Historical Social Research*, 15(3), 4-170. <https://doi.org/10.12759/hsr.15.1990.3.4-170>

### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/deed.de>

### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:

<https://creativecommons.org/licenses/by/4.0>

## Einführung in die Logik der Skalierungsverfahren

*Johann Bacher\**

Abstract: This article gives an introduction to the logic of scaling and to scaling techniques. It starts with a definition of scaling. From this starting point the steps necessary to solve a scaling task are deduced. These steps are:

- Specification of a scaling model, that fits to theory and data.
- Estimation of the parameters of the scaling model.
- Validation of the scaling model.
- Error analysis if the validation test fails.
- Estimation of scores for the scaled units if the models fit to data.

The solutions of these steps are discussed in detail for real-life data sets from sociology and historical sciences. In addition it is shown how the solutions can be realized within statistical software packages, like SPSS-X or LISREL. It is important not to confound these steps not with scaling techniques. Scaling techniques are instruments, that can help to solve these steps. In order to use these instruments, we must know their assumptions and their properties.

The following scaling techniques are described: Guttman Scaling, Mokken Scaling, Rasch Scaling, Scaling according to Proctor, Guttman Scaling for polytomous data, Latent Class Analysis and Factor Analysis.

The statistical software packages used are: SPSS-PC+, PANMARK and LISREL. An introduction to PANMARK and LISREL is given.

---

\* Address all communications to: Johann Bacher, Johannes-Kepler-Universität, Institut für Soziologie, A-4040 Linz-Auhof, Austria.

## Vorbemerkungen

Diese Einführung ist auf der Grundlage eines Workshops entstanden, den der Autor 1985 und 1986 während des »Quantkurses«, einer jährlich in Salzburg stattfindenden postgradualen Ausbildung in den Methoden der historischen Sozialforschung (BOTZ u.a. 1988), abhielt. Aufgrund der Erfahrungen dieser beiden Workshops sowie den Erfahrungen in anderen Lehrveranstaltungen wurde das Skript vollkommen überarbeitet, als der Herausgeber der Zeitschrift »Historical Social Research/Historische Sozialforschung« eine Publikation der Arbeit anregte.

Die Probleme, die beim Erlernen von Skalierungsverfahren auftreten, sind:

- Bücher über Skalierungsverfahren (auch einführende) sind oft mit mathematischen Formeln überfrachtet. Dies führt u.a. dazu, daß die Grundlogik von Skalierungsverfahren sowie die Beziehung zwischen inhaltlichen Theorien, Skalierungsverfahren und Daten nicht erkannt wird.
- Weitere Probleme treten bei der Anwendung der Skalierungsverfahren in der Umsetzung von einzelnen Verfahrensschritten in die Befehlssprache eines Statistikprogramms, wie z.B. BMDP, SAS oder SPSS-X, auf.

Aufgrund dieser Probleme verfolgt die Einführung zwei Zielsetzungen:

- Vermittlung der Grundlogik der Skalierung sowie von Skalierungsverfahren.

Die Kenntnis der Grundlogik der Skalierung, des Ablaufs der Lösung einer Skalierungsaufgabe, ist zur Vermittlung der Grundlogik von Skalierungsverfahren sowie zur Darstellung der Beziehung zwischen Theorie, Skalierungsverfahren und empirischen Daten unerlässlich.

- Darstellung der Umsetzung der Schritte der Lösung einer Skalierungsaufgabe in die Befehlssprache eines Statistikprogramms, nämlich in SPSS-PC +.

So wird beispielsweise gezeigt, wie - auch ohne Vorliegen einer entsprechenden Prozedur - eine Guttman- oder Raschskalierung durchgeführt werden kann.

Diese beiden Zielsetzungen machten eine Einschränkung auf wenige Skalierungsverfahren erforderlich, die im Detail besprochen werden. Die vorliegende Arbeit ist mithin kein Überblick über Skalierungsverfahren, sondern eine Einführung in die Skalierung. Die vorgenommene Einschränkung ist tatsächlich aber nicht besonders schwerwiegend, da alle Skalierungsverfahren dieselbe Grundlogik besitzen. Bei den behandelten Verfahren wurde ferner versucht, die neuesten Entwicklungen in die Darstel-

lung miteinzubeziehen. Dies soll u.a. dazu beitragen, daß der Leser seine Kenntnisse durch Nachschlagen in der betreffenden Spezialliteratur vertiefen kann.

Für die kritische Durchsicht des Skripts und für wertvolle Anregungen sei Herrn Dr. Thome gedankt. Ferner möchte ich danken: dem ZHSF für die redaktionellen Arbeiten, den Student/innen und Teilnehmer/innen von Workshops für kritische Anregungen.

# 1. Aufgaben und Grundlogik der Skalierung

## 1.1 Aufgaben

Sozialwissenschaftliche Theorien bestehen in der Regel aus abstrakten theoretischen Begriffen, wie z.B. Industrialisierung, Urbanisierung, soziale Ungleichheit, soziale Schichtung, latente interne Konflikte, Revolutionen usw. Kennzeichen dieser theoretischen Begriffe ist u.a., daß sie nicht direkt empirisch beobachtbar sind. Zur empirischen Überprüfung der Theorien ist es erforderlich, die theoretischen Begriffe mit empirisch beobachtbaren Phänomenen zu verknüpfen. Dies kann auf unterschiedlicher Weise erfolgen. Ein Zugang besteht darin, für die theoretischen Begriffe eine Menge von Indikatoren (direkt empirisch beobachtbare Variablen oder aus diesen berechenbare Variablen) zu definieren. So können z.B. für den Begriff »Industrialisierung« folgende Indikatoren (DENZ 1989:33-35, 120-22) verwendet werden:

- Anteil der Zuwanderer an der Gesamtbevölkerung
- Anteil der Abwanderer an der Gesamtbevölkerung
- Anteil der in der Land- und Forstwirtschaft Erwerbstätigen
- Anteil der im Bergbau und Hüttenwesen Erwerbstätigen
- Anteil der in der Textilindustrie Erwerbstätigen
- Anteil der Tagelöhner an den in der Land- und Forstwirtschaft Erwerbstätigen
- Anteil der Selbständigen an den in der Land- und Forstwirtschaft Erwerbstätigen
- Anteil der Selbständigen an den in Industrie und Gewerbe Erwerbstätigen

Oder um ein weiteres Beispiel zu geben: Für den Begriff »latente interne Konflikte« können folgende Indikatoren ausgewählt werden:

- Durchschnittliche Zahl der Eigentumsdelikte je Einwohner
- Anteil der Personen in Armenhäusern an der Gesamtbevölkerung
- Anteil der politisch Verfolgten an der Gesamtbevölkerung
- Anzahl der Fälle ohne Kompromiß zwischen den politischen Eliten an der Gesamtzahl der Treffen der politischen Eliten

Entscheidend bei diesem Vorgehen ist folgendes:

1. Durch die Definition von Indikatoren wird (implizit oder explizit) eine (kausale oder korrelative) Beziehung zwischen den theoretischen Begriffen und Indikatoren angenommen, wie z.B.

- »Je höher der Grad der Industrialisierung, desto höher der Anteil der Zuwanderer« (positive Beziehung)
- »Je höher der Grad der Industrialisierung, desto höher der Anteil der Abwanderer.« (positive Beziehung)
- »Je höher der Grad der Industrialisierung, desto geringer der Anteil der selbständigen Landwirte.« (negative Beziehung) usw.

Diese Beziehungen lassen sich graphisch wie in der Abbildung 1-1 dargestellt verdeutlichen.

Die graphische Darstellung der Beziehung zwischen theoretischen Begriffen und Indikatoren kann noch weiter spezifiziert werden. So z.B. kann festgelegt werden, ob eine kausale oder korrelative Beziehung zwischen den theoretischen Begriffen und Indikatoren besteht (BAGOZZI/FORNELL 1982). Die Richtung des Zusammenhangs kann durch ein Vorzeichen symbolisiert werden. Bei einem positiven Zusammenhang kann ein »+« verwendet werden, bei einem negativen ein »-«.

2. Die zwischen Indikatoren und theoretischen Begriffen spezifizierten Beziehungen sind Hypothesen. Dies impliziert:

- a) Die Definition von Indikatoren ist noch keine Garantie dafür, daß sie den theoretischen Begriff (bzw. deren Subdimension) tatsächlich messen (Gültigkeitsproblem).
- b) Die angenommenen Beziehungen müssen zumindest prinzipiell empirisch überprüfbar sein.
- c) Sie können mehr oder weniger präzise sein. Ein Beispiel für eine sehr präzise Spezifikation wäre:

»Zwischen dem Indikator X und dem theoretischen Begriff T soll folgende mathematische Beziehung bestehen:  $X = T + E$ , wobei E ein zufälliger Meßfehler ist.«

Eine sehr vage Spezifikation würde vorliegen, wenn nur angegeben werden kann:

»Der Indikator X mißt in einer noch nicht näher bekannten Weise den theoretischen Begriff T.«

Auf der Grundlage der oben durchgeführten Explikation kann bereits eine erste Aufgabe der Skalierung formuliert werden:

Eine Aufgabe der Skalierung ist, empirisch zu prüfen, ob die angenommenen Beziehungen zwischen Indikatoren und theoretischen Begriffen zutreffen. Bei der Lösung dieser Aufgaben werden Skalierungsverfahren (formale, mathematische, statistische Modelle) eingesetzt, die (statistische) Beurteilungskriterien für obige Aufgabe bereitstellen.

Abhängig davon, wie präzise die Beziehungen zwischen Indikatoren und theoretischen Begriffen definiert sind, werden bei sehr vagen Vorstellungen explorative Skalierungsverfahren angewendet, bei sehr präzisen

Vorstellungen konfirmatorische Skalierungsverfahren (siehe dazu insbesondere Abschnitt 5.3.1). Allgemein gilt: Bei konfirmatorischen Skalierungsverfahren bestehen mehr Prüfmöglichkeiten, da die theoretischen Vorstellungen präziser sind. Andererseits bieten explorative Skalierungsverfahren die Möglichkeit, nicht vermutete Strukturen zu erkennen.

Nach der Prüfung der angenommenen Beziehungen zwischen Indikatoren und theoretischen Begriffen besteht eine zweite Aufgabe der Skalierung in der Schätzung von Skalenwerten (Ausprägungen) der Untersuchungseinheiten in den theoretischen Begriffen. Dies setzt voraus, daß die angenommenen Beziehungen nicht verworfen werden. Ist dies dagegen der Fall, ist eine Fehleranalyse erforderlich.

Die einzelnen Schritte der Lösung dieser Aufgaben werden in Abschnitt 1.2 dargestellt. Zuvor soll aber noch die Bedeutung der bisher sehr salopp verwendeten Begriffe präzisiert werden. Die wichtigsten Begriffe sowie deren Bedeutung wurden dazu in der Abbildung 1-2 zusammengefaßt.

Zu beachten ist dabei, daß die Begriffe »theoretischer Begriff«, »manifest«, »latent« und »Variable« sehr allgemeine Konzepte sind. Der Begriff »Indikator« dagegen ist ein untergeordneter Begriff, da er sich auf eine spezielle Klasse von manifesten Variablen bezieht.

## 1.2 Schritte der Lösung einer Skalierungsaufgabe

Im vorausgehenden Abschnitt wurden Skalierungsverfahren als formale (Meß-)Modelle definiert. Sie bestehen als solche aus:

- inhaltsleeren latenten und manifesten Variablen,
- bestimmten Parametern und
- formalen Beziehungen zwischen latenten und manifesten Variablen (i.d.R. mathematische Gleichungen).

Eine inhaltliche Interpretation der Variablen ergibt sich erst, wenn die Skalierungsverfahren konkret zur Lösung einer Skalierungsaufgabe eingesetzt werden. Allgemein verläuft die Lösung einer Skalierungsaufgabe in den in der Abbildung 1-3 dargestellten Schritten.

Die Schritte sind:

- Spezifikation eines inhaltlich-theoretischen Modells der angenommenen Beziehung zwischen Indikatoren und theoretischen Begriffen.
- Auswahl der Daten.
- Auswahl des Skalierungsverfahrens. Die Wahl des Skalierungsverfahrens hängt von den vorhandenen empirischen Daten, den theoretischen Vorstellungen über die Indikatoren sowie deren Beziehung (theoretisches Modell) zu den theoretischen Begriffen ab. So z.B. set-

zen bestimmte Skalierungsverfahren intervall- oder ratioskalierte Indikatoren voraus (wie z.B. die Faktorenanalyse), während andere von nominalskalierten Indikatoren ausgehen (wie z.B. die Analyse latenter Klassen).

- Zwischen diesen drei Schritten bestehen gegenseitige Abhängigkeiten. So z.B. kann die Kenntnis eines Skalierungsverfahrens die Auswahl der Daten und/oder die theoretischen Vorstellungen beeinflussen. Dies bedeutet, daß die Anordnung der Schritte vom konkreten Forschungsprojekt abhängt.
- Übersetzung in ein Schätzmodell. Die Entscheidung für ein bestimmtes Skalierungsverfahren führt dazu, daß das theoretische Modell in ein statistisch schätzbares Modell übersetzt werden kann. Voraussetzung für eine Schätzung ist das Vorliegen eines identifizierten Modells.
- Datentransformationen. Die Auswahl eines Skalierungsverfahrens kann bestimmte Datentransformationen erforderlich machen, um die Anwendungsvoraussetzungen des Skalierungsverfahrens zu erfüllen. Datentransformationen können sich aber aufgrund des theoretischen Modells sowie der empirischen Daten ergeben.
- Prüfung der Identifikation. Die Prüfung der Identifikation gewährleistet, daß die berechneten Parameter inhaltlich interpretierbar sind.
- Schätzung der Modellparameter. Durch die Anwendung des Schätzmodells auf die transformierten Daten werden bestimmte Parameter des Modells geschätzt.

Mit den bisherigen Schritten sind die »Vorarbeiten« zur Lösung der ersten Aufgabe von Skalierungsverfahren, die Überprüfung der angenommenen inhaltlichen, theoretischen Vorstellungen abgeschlossen. Aus der Aufzählung der Schritte ist ersichtlich, daß bereits bis zu diesem Punkt eine Reihe von Entscheidungen getroffen werden müssen, die fehlerhaft sein können. Die weiteren Schritte sind:

- Prüfung der Modellanpassung. Die Schätzung der Modellparameter ermöglicht die Berechnung von theoretischen Daten (erwartete Modelldaten), also Daten, die sich ergeben würden, wenn das spezifizierete Modell zutrifft. Diese können mit den empirischen Daten und/oder mit den transformierten Daten verglichen werden. Liegt eine Übereinstimmung vor, ist das Modell geeignet. Bei explorativen Verfahren besteht die Modellüberprüfung auch darin, ob die Ergebnisse inhaltlich interpretierbar sind. Bei konfirmatorischen Skalierungsverfahren ist dies nicht erforderlich, da ja präzise theoretische Vorstellungen gegeben sind.
- Schätzung von Skalenwerten der Untersuchungseinheiten in den latenten Variablen, sofern das Modell geeignet ist.



- Fehleranalyse. Bei einer schlechten Modellanpassungen werden mögliche Fehlerquellen untersucht, um ein besseres Modell zu finden.

Mögliche Fehlerquellen können sein:

- Die durchgeführten Datentransformationen sind unbrauchbar.
- Die vorhandenen Daten sind unbrauchbar.
- Das ausgewählte Skalierungsverfahren ist unbrauchbar.
- Die angenommenen theoretischen Vorstellungen sind falsch.

Wie die beispielhafte Aufzählung der Ursachen für eine schlechte Modellanpassung zeigt, kann aus einer schlechten Modellanpassung beispielsweise nicht eindeutig abgeleitet werden, daß die theoretischen Vorstellungen falsch sind, da andere Fehlerquellen vorliegen können. Diese Tatsache ist außerordentlich wichtig, um zu vorschnelle Schlußfolgerungen zu vermeiden. Auf der anderen Seite sollte diese Tatsache aber nicht zur Immunisierung theoretischer Vorstellungen mißbraucht werden.

Abschließend ist anzumerken, daß das Schema eine »idealtypische Rekonstruktion« des Vorgehens bei der Lösung einer Skalierungsaufgabe darstellt. Vielfach werden einzelne Schritte nur implizit oder/und aufgrund von etablierten Regeln gelöst. Dieses Vorgehen ist nicht unproblematisch: Mögliche Fehlerquellen werden dadurch vielfach eliminiert. Die Tatsache, daß der Anwendung von Skalierungsverfahren inhaltliche Vorstellungen über die Beziehung von theoretischen Begriffen und Indikatoren zugrunde liegen, wird übersehen usw.

### 1.3 Kommentierte Inhaltsübersicht

Aufgrund der in der Vorbemerkung erwähnten Zielsetzungen, zum einen die Grundlogik der Skalierung und von Skalierungsverfahren sowie deren Annahmen zu verdeutlichen und zum anderen die konkrete Anwendung von Skalierungsverfahren darzustellen, werden in den folgenden Kapiteln nur wenige Skalierungsverfahren ausführlich behandelt. Drei »Grundtypen« werden besprochen:

- Die Guttmanskala (Kapitel 2).
- Die Analyse latenter Klassen (Kapitel 4).
- Die Faktorenanalyse (Kapitel 5).

Für jeden dieser Grundtypen werden anhand eines Beispiels aus der Forschungspraxis die einzelnen, in dem vorausgehenden Abschnitt dargestellten Schritte zur Lösung einer Skalierungsaufgabe behandelt (siehe Abbildung 1-4).

Die Gründe für die Auswahl dieser drei Verfahren waren:

- Die Guttmanskala besitzt eine sehr einfache Logik. Sie eignet sich deshalb als Einstieg in die Skalierung.
- Zahlreiche Skalierungsverfahren können als Weiterentwicklung bzw. Modifikation der Guttmanskala gedacht werden.
- Die Analyse latenter Klassen stellt zum einen ein eigenständiges Skalierungsverfahren für nominalskalierte Indikatoren dar. Zum anderen kann sie zur Schätzung der Guttmanskala sowie von Skalierungsmodellen, in denen die Guttmanskala weiterentwickelt wurde, eingesetzt werden.
- Die Faktorenanalyse schließlich stellt das in der Forschungspraxis am häufigsten verwendete Skalierungsverfahren dar.

Neben diesen drei Verfahren gibt es noch zwei »Grundtypen«, die nicht behandelt werden:

- Mehrdimensionale Skalierungsverfahren.
- Unfoldingverfahren.

Zur Einführung in diese beiden Verfahren empfehlen wir:

- Mehrdimensionale Skalierung: BORG (1981), BORG/STAUFGENBIEL (1989:82-120), COXON (1982).
- Unfolding: COOMBS (1964: 80-180), McIVER/C ARMINES (1981: 71-85), SIXTL (1982: 414-430).

Neben den drei behandelten Grundtypen werden in Kapitel 3 noch Verfahren behandelt, die als Weiterentwicklung der Guttmanskala gedacht werden können. Hier wird auf eine Darstellung der einzelnen Schritte verzichtet, da sie wie bei der Guttmanskala durchgeführt werden können. Die einzelnen Kapitel sind zum Teil unterschiedlich didaktisch aufgebaut. Im Kapitel 2 beispielsweise wird die Anwendung der Guttmanskalierung zur Lösung einer Skalierungsaufgabe ohne ein EDV-Programm dargestellt. Die Umsetzung in ein Programm wird in einem Exkurs behandelt. Im Kapitel 3 dagegen steht die technische Umsetzung in Form eines Programms im Vordergrund. Allgemein ist zu beachten, daß die Kapitel schrittweise methodisch anspruchsvoller werden. Das mathematisch anspruchsvollste Kapitel ist mithin das Kapitel 5 über die Faktorenanalyse.

Mit den behandelten Verfahren können die in der Abbildung 1-5 dargestellten Indikatoren untersucht werden. Als Einteilungskriterium wurde das Meßniveau der Variablen behandelt.

## 1.4 Vorkenntnisse

Obwohl die Arbeit als Einführung konzipiert ist, müssen aus Platzgründen bestimmte Vorkenntnisse vorausgesetzt werden. Diese sind u.a.:

methodologische Vorkenntnisse:

- Gültigkeit (Validität)
- Zuverlässigkeit (Reliabilität)
- Operationalisierung
- Meßniveau

statistische Vorkenntnisse:

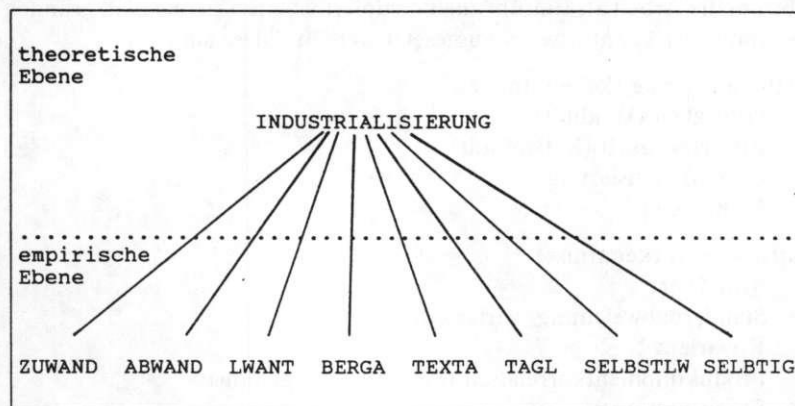
- Mittelwert
- Standardabweichung, Varianz
- Kovarianz
- Produktmomentkorrelation (Pearson-Korrelation)
- Erwartungswert
- Rechnen mit Erwartungswerten
- statistische Unabhängigkeit

EDV-Vorkenntnisse:

- Grundzüge in SPSS-PC+ bzw. SPSS-X

Diese Vorkenntnisse können in den meisten einführenden Lehrbüchern der empirischen Sozialforschung und der Statistik nachgelesen werden, wie z.B. in dem bereits zitierten Buch von DENZ (1989) oder bei BENNINGHAUS (1990). Sie sind für das Verständnis der nachfolgenden Abschnitte unbedingte Voraussetzung.

Abbildung 1-1: Graphische Darstellung der Beziehung zwischen theoretischen Begriffen und Indikatoren



verwendete Abkürzungen:

ZUWAND	Anteil der Zuwanderer
ABWAND	Anteil der Abwanderer
LWANT	Anteil der in der Land- und Forstwirtschaft Erwerbstätigen
BERGA	Anteil der im Bergbau und Hüttenwesen Erwerbstätigen
TEXTA	Anteil der in der Textilindustrie Erwerbstätigen
TAGL	Anteil der Tagelöhner an den in der Land- und Forstwirtschaft Erwerbstätigen
SELBSTLW	Anteil der Selbständigen an den in der Land- und Forstwirtschaft Erwerbstätigen
SELBTIG	Anteil der Selbständigen an den in Industrie und Gewerbe Erwerbstätigen

Abbildung 1-2: Definition der wichtigsten Begriffe der Skalierungsliteratur

---

Begriff	Definition/Synonyme Bezeichnungen
theoretischer Begriff	Element einer wissenschaftlichen Theorie, der bestimmte Phänomene abbildet. Synonyme Bezeichnungen: theoretisches Konzept, theoretisches Konstrukt.
manifest	Diese Bezeichnung wird für direkt empirisch beobachtbare Phänomene verwendet. Synonyme Bezeichnungen: empirisch, direkt beobachtbar.
latent	Diese Bezeichnung wird für nicht direkt empirisch beobachtbare Phänomene verwendet. Synonyme Bezeichnungen: theoretisch, nicht direkt beobachtbar.
Variable	Eine direkt oder indirekt meßbare Eigenschaft mit bestimmten Ausprägungen.
Indikator	Manifeste Variable, von der angenommen wird, daß sie eine oder mehrere latente Variable mißt.
Item	Spezielle Form von Indikatoren, nämlich die Statements in einer Fragebatterie zur Messung von Einstellungen.
Meßdimension	Latente Variable, die von Indikatoren gemessen wird. Synonyme Bezeichnungen: latente Dimension, Faktor (Faktorenanalyse), Zieldimension.

---

Abbildung 1-3: Schritte der Lösung einer Skalierungsaufgabe

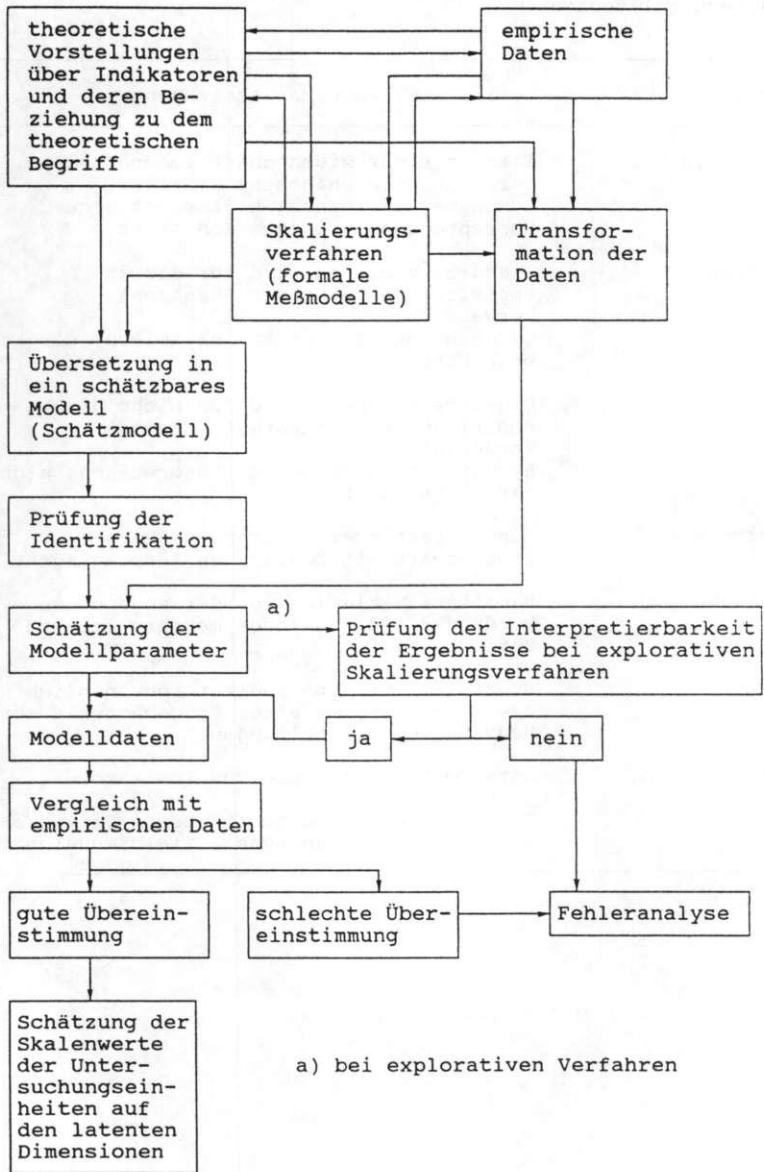


Abbildung 1-4: Kapitelnummer, in denen die einzelnen Schritte der Lösung einer Skalierungsaufgabe behandelt werden.

Schritt	GUTTMAN	ANALYSE LATENTER KLASSEN	FAKTOREN- ANALYSE
- Spezifikation eines theoretischen Modells	2.1	4.1	5.1, (1.1)
- Beschreibung der empirische Daten	2.2	(2.2)	5.1, (1.1)
- Auswahl des Skalierungsverfahrens	2.3	(2.7)	5.1
- Beschreibung der Grundlogik des Verfahrens	2.4	4.2	5.2, 5.3, 5.8.2
- Datentransformationen	2.3	(2.3)	5.1, 5.8.1
- Identifikationsprüfung	-	4.3	5.2, 5.4
- Schätzung der Modellparameter	2.5	4.4	5.5, 5.6, 5.8.2
- Modellprüfung	2.6	4.4	5.7.1
- Fehleranalyse	2.7	4.5	5.7.2, 5.8
- Schätzung der Skalenergebnisse	2.8	4.6	5.8.2.3
Anmerkungen:	Analyse latenter Klassen wird aus den Ergebnissen der Fehleranalyse der Guttmankala entwickelt.		
Anwendungsbeispiel:	aus der Soziologie "politisches Aktivitätsprofil" von Personen	aus der historischen Sozialfor. "Industrialisierung" (Abschnitt 1.1)	

Abbildung 1-5: Erforderliches Meßniveau der Indikatoren zur Anwendung der behandelten Skalierungsverfahren

Verfahren	erforderliches Meßniveau	Anmerkung zum Meßniveau
Guttmanskala (Kapitel 2)	dichotome, ordinalskalierte Variablen	Bei polytomen, ordinalskalierten Variablen sowie bei intervall- oder ratioskalierten Variablen ist eine Dichotomisierung erforderlich (Infor- mationsverlust).
Weiterentwicklung der Guttmanskala (Kapitel 3)	dichotome oder polytome, ordinalskalierte Variablen	Bei intervall- oder ratioskalierten Variablen ist eine Zusammenfassung in Kategorien erforder- lich (Informations- verlust).
Analyse latenter Klassen (Kapitel 4)	dichotome oder polytome, nominalskalierte Variablen	Wegen der Hierarchie des Meßniveaus können auch ordinalskalierte Variablen untersucht werden. Bei inter- vall oder ratioska- lierten Variablen ist eine Zusammenfassung in Kategorien erfor- derlich (Informa- tionsverlust).
Faktorenanalyse (Kapitel 5)	intervall- oder ratioskalierte Variablen	Hat sich in der For- schungspraxis auch für ordinalskalierte Variablen als brauchbar erwiesen (siehe Abschnitt 5.9).



## 2. Guttmanskalierung

### 2.1 Inhaltlich-theoretische Vorstellungen

Eine in den Sozial- und Politikwissenschaften häufig untersuchte Forschungsfrage ist die Bereitschaft von Personen, an bestimmten politischen Aktivitäten teilzunehmen (Aktivitätsbereitschaft), oder die tatsächliche Beteiligung von Personen an bestimmten politischen Aktivitäten (Aktivitätsprofil). Eine Skala zur Messung des Aktivitätsprofils einer Person könnte wie in der Abbildung 2-1 dargestellt aussehen. Die Skala wurde leicht modifiziert aus dem Sozialen Survey Österreichs (HALLER/HOLM 1988) entnommen (siehe dazu Abschnitt 2-2).

Bei der Entwicklung der Skala wurde u.a. davon ausgegangen, daß die einzelnen Aktivitäten unterschiedliche Grade der politischen Beteiligung messen. Als ordnende Dimension kann man sich beispielsweise die Zeit vorstellen, die für die einzelnen Aktivitäten aufgewendet werden muß. Entsprechend dieser Dimension könnte sich folgende Ordnung der Items ergeben:

geringstes zeitliches  
Engagement

- B - Unterstützung eines Volksbegehrens
- A - Teilnahme an einer Kundgebung
- C - Kontakt mit Politikern
- D - Mitwirken in einer Bürgerinitiative

höchstes zeitliches  
Engagement

- E - aktive Mitarbeit in einer politischen Partei

Bezüglich des Verhaltens der Personen bestehen die theoretischen Vorstellungen aus zwei Annahmen:

- (a) Jede Person besitzt einen maximalen zeitlichen Engagementpunkt, z.B. die Aktivität C.
- (b) Sie nimmt an Aktivitäten teil, die ein geringeres zeitliches Engagement erfordern, nicht aber an Aktivitäten mit einem höheren zeitlichen Engagement.

Die Annahmen bedeuten z.B. für eine Person mit einem maximalen zeitlichen Engagementpunkt bei der Aktivität C, daß aufgrund der inhaltlichen Vorstellungen eine Teilnahme an den Aktivitäten B und A erwartet wird, nicht aber an den Aktivitäten D und E. Spielt man alle theoretischen Möglichkeiten durch, würden in einer empirischen Erhebung die in der Abbildung 2-2 dargestellten Antwortmuster erwartet werden.

Ein Antwortmuster der Art

B	A	C	D	E
0	0	1	0	0

würde dagegen nicht erwartet werden (obwohl es empirisch auftreten kann), da entsprechend den theoretischen Vorstellungen eine Person mit einem Engagement bei der Aktivität C auch an den Aktivitäten mit geringerem zeitlichen Aufwand (B und A) teilnimmt.

Zusammenfassend liegen bezüglich der Skala (des Verhaltens der Personen bei Beantwortung der Fragen) sehr einfache, aber präzise theoretische Vorstellungen vor. Da diese Vorstellungen Hypothesen darstellen, können sie auch falsch sein. So z.B. könnte ein Teil der Personen die Items nach einer anderen Dimension (z.B. befürchtete politische Sanktionen oder nach der Wirksamkeit) ordnen oder den erforderlichen Zeitaufwand unterschiedlich einschätzen. Es könnten aber auch andere Verhaltensstrategien vorliegen. Entscheidend ist nun, daß die Hypothesen empirisch überprüfbar sind und es sich um inhaltliche Hypothesen handelt.

An dieser Stelle sei eine Warnung vor zu großen Erwartungen an Skalierungsverfahren ausgesprochen. Mit ihnen kann nur überprüft werden, ob die Items eine für alle Personen identische Anordnung besitzen und sich die Personen entsprechend den beiden Annahmen (a) und (b) verhalten. Ob die Items aber nach dem zeitlichen Aufwand, den Sanktionen, die mit ihnen verbunden sind, oder ihrer Wirksamkeit von den Personen geordnet werden, kann mit der Hilfe von Skalierungsverfahren nicht nachgewiesen werden. Dies ist Aufgabe einer inhaltlichen Interpretation sowie von externen Gültigkeitsprüfungen.

## 2.2 Empirische Daten

Zur Überprüfung der theoretischen Vorstellungen der entwickelten Skala sollen die Daten des Sozialen Survey Österreichs untersucht werden. Befragt wurden in dieser Erhebung 2014 erwachsene Österreicher (16- bis 69-Jährige). Die in dieser Befragung verwendete Skala zur Messung des politischen Aktivitätsprofils unterscheidet sich von der Skala der Abbildung 2-1 dadurch, daß anstelle von zwei Antwortkategorien drei verwendet wurden. Diese waren:

- 1 habe ich bereits getan bzw. tue ich
- 2 noch nicht getan, würde es aber tun, wenn es an mich herantritt
- 3 noch nicht getan, würde es auf keinen Fall tun

Mit den Antwortkategorien wurde versucht, sowohl die *politische* Beteiligung der Personen (Kategorie 1) als auch die Bereitschaft zur Beteiligung

(Kategorien 1 und 2) zu erheben. Implizit wird dabei bezüglich der Messung der politischen Aktivitätsbereitschaft angenommen, daß Personen, die derzeit an einer bestimmten Aktivität teilnehmen bzw. an einer solchen teilgenommen haben, auch bereit sind, sich an einer solchen erneut zu beteiligen. Sollen diese Daten zur Überprüfung der in Abschnitt 2.1 entwickelten Vorstellungen verwendet werden, ist eine Datentransformation erforderlich. Da das Aktivitätsprofil erfaßt werden soll, können die Ausprägung »2« und »3« zusammengefaßt und auf die Ausprägung »0« umkodiert werden. Durch diese Operation werden die Antwortkategorien in Richtung der Meßdimension verkodet. Eine »1« bedeutet eine Teilnahme, eine »0« keine Teilnahme.

### 2.3 Auswahl des Skalierungsverfahrens

Nach der Auswahl der empirischen Daten besteht der nächste Schritt der Überprüfung der in Abschnitt 2.1 entwickelten Vorstellungen in der Auswahl eines geeigneten Skalierungsverfahrens.

Ein Skalierungsverfahren, das den theoretischen Vorstellungen des Abschnitts 2-1 und den Daten entspricht, ist die Guttmanskala. Das bedeutet: Die Modellannahmen der Guttmanskala stimmen mit den inhaltlichen Vorstellungen sowie der Datengrundlage überein. Um auch in anderen Forschungsfragen prüfen zu können, ob sich die Guttmanskalierung eignet, ist es wichtig, die Modellannahmen und Eigenschaften der Guttmanskala zu kennen. Unter einer Guttmanskala kann man sich dabei ein »Meßinstrument« mit bestimmten Eigenschaften vorstellen. Aufgabe der Guttmanskalierung ist zu prüfen, inwieweit dieses »Meßinstrument« vorliegt.

### 2.4 Modellannahmen und Eigenschaften der Guttmanskala

Analog den inhaltlich-theoretischen Vorstellungen unseres Beispiels werden bei einer Guttmanskala allgemein folgende Annahmen getroffen:

- Es liegt eine Menge von Indikatoren vor, die eine gemeinsame Dimension messen. In unserem Beispiel ist dies das tatsächliche Aktivitätsprofil einer Person.
- Die Indikatoren besitzen nur zwei Ausprägungen und sind in Richtung der Meßdimension kodiert. Im folgenden soll eine 0-1-Codierung angenommen werden. Eine »0« bedeutet dann, daß keine Teilnahme vorliegt, eine »1« dagegen das Vorliegen einer Teilnahme.
- Jeder Indikator und jede Person bzw. allgemein jede Untersuchungseinheit besitzen Skalenwerte auf der Meßdimension.

- Bezüglich des Auftretens der Ausprägung »0« oder »1« in einem Indikator bei einer Person  $g$  wird angenommen, daß in allen Indikatoren, die links des Skalenwerts der Person  $g$  liegen, eine »1« auftritt. In allen Indikatoren, die »rechts« des Skalenwerts der Person  $g$  liegen, wird dagegen das Auftreten einer »0« angenommen.

Überträgt man die Modellannahmen auf unser Beispiel, würden bei Vorliegen einer perfekten Guttmanskala empirisch nur die in der Abbildung 2-2 spezifizierten Antwortmuster auftreten. Diese können selbstverständlich unterschiedliche Auftrittshäufigkeiten besitzen. Das Antwortmuster »00000« kann beispielsweise 15-mal auftreten, das Antwortmuster »10000« 20-mal, usw..

Zur Verdeutlichung der Eigenschaften der Guttmanskala soll aber zunächst von der Annahme ausgegangen werden, daß jedes Antwortmuster nur einmal auftritt (siehe Abbildung 2-3).

In der Abbildung 2-3 wurde für jede Person und jedes Item ein Summenwert berechnet. Diese haben folgende Bedeutung:

- Summe einer Person über die Items (Gesamtpunktwert): Der Gesamtpunktwert gibt an, an wievielen Aktivitäten eine Person teilnimmt. Personen, die an mehr Aktivitäten teilnehmen, haben einen höheren Gesamtpunktwert. Entsprechend den theoretischen Vorstellungen besitzen sie einen höheren maximalen zeitlichen Engagementpunkt.
- Summe eines Items über die Personen (Auftrittshäufigkeit der Ausprägung »1«): Ein höherer Summenwert bedeutet, daß die Aktivität von mehr Personen ausgeübt wird. Da angenommen wird, daß die Items auf der Meßdimension geordnet sind, wird allgemein von der Schwierigkeit bzw. Leichtigkeit von Items gesprochen. Ein Item  $F$  ist schwieriger als ein Item  $E$ , wenn es einen höheren Wert auf der Meßdimension besitzt. Dies bedeutet umgekehrt, daß es einen geringen Summenwert besitzt. In der Abbildung 2-3 ist das Item  $E$  das schwierigste, da es den geringsten Summenwert besitzt.

Bei Vorliegen einer perfekten Guttmanskala reicht die Kenntnis der Schwierigkeit der Items und der Gesamtpunktwerte der Personen aus, um die Datenmatrix zu reproduzieren: Aufgrund der Auftrittshäufigkeiten der Items kann die Anordnung der Items festgelegt werden. Daran anschließend kann aus den Gesamtpunktwerten jeder Personen das Antwortmuster abgeleitet werden. Die Guttmanskala besitzt diese Eigenschaft, da sie spezifisch objektiv mißt und einen deterministischen »Antwortmechanismus« annimmt.

Eine Skala wird dann als spezifisch objektiv bezeichnet, wenn die Itemparameter (bei der perfekten Guttmanskala: Anordnung der Items nach

dem Schwierigkeitsgrad) unabhängig von den Personenparametern (bei der perfekten Guttmanskala: Gesamtpunktwerte der Personen) sind. Diese Unabhängigkeitseigenschaft heißt spezifisch, da sie nur unter bestimmten Bedingungen, die theoretisch und empirisch zu bestimmen sind, gilt. Aus der spezifischen Objektivität folgt eine Reihe von weiteren Eigenschaften der (perfekten) Guttmanskala:

Die perfekte Guttmanskala ist stichprobenunabhängig. Unabhängig von der Personengruppe, die das Aktivitätsprofil beantwortet, bleibt die Anordnung der Items nach der Schwierigkeit gleich. Im ungünstigsten Fall kann die Situation eintreten, daß zwei Items gleich schwer werden.

Die perfekte Guttmanskala ist invariant. Werden andere Items, die ebenfalls eine perfekte Guttmanskala bilden und dieselbe Meßdimension messen, verwendet, ändert sich die Rangfolge zwischen den Personen nicht. Personen mit einem größeren Gesamtpunktwert werden auch bei Verwendung anderer Items einen größeren/gleichen Gesamtpunktwert haben als Personen mit einem kleineren Gesamtpunktwert in der ursprünglichen Skala.

Bei einer perfekten Guttmanskala ist eine unabhängige Schätzung der Itemparameter (Anordnung der Schwierigkeitsgrade) und der Personenparameter (Gesamtpunktwerte) möglich (Eigenschaft der separierten Schätzbarkeit).

Neben diesen Eigenschaften besitzt die perfekte Guttmanskala wegen des deterministischen »Mechanismus« noch folgende Monotonieeigenschaften (Konsistenzigenschaften):

Tritt ein bestimmtes Item auf, treten auch alle leichteren Items auf.  
Bei Personen mit einem größeren Gesamtpunktwert tritt die Ausprägung »1« in denselben Items wie bei Personen mit einem geringeren Gesamtpunktwert auf.

Zu beachten ist, daß sich die angegebenen Eigenschaften nur auf das Modell der perfekten Guttmanskala beziehen und selbst für den Fall, daß die empirischen Daten exakt einer perfekten Guttmanskala genügen, der spezifische Geltungsbereich geklärt werden muß.

Aufgabe der Skalierung ist es nun zu überprüfen, inwieweit die empirischen Daten eine perfekte Guttmanskala bilden. Diese Aufgabe wird - entsprechend dem Abschnitt 1.2 - in folgenden Schritten durchgeführt:

Schätzung der Modellparameter aus den empirischen Daten.

Berechnen einer theoretischen Datenmatrix. Diese wird unter der Annahme berechnet, daß eine perfekte Guttmanskala vorliegt.

Vergleich der theoretischen und empirischen Datenmatrix.

- Entscheidung, ob die Ergebnisse (Annahme einer perfekten Gutt-manskala) akzeptiert werden können.
- Fehlersuche, wenn die Ergebnisse nicht akzeptiert werden können.
- Berechnung der »Skalenwerte« der Personen, wenn die Ergebnisse akzeptiert werden können.

## 2.5 Schätzung der Modellparameter

Die Gutt-manskala besitzt - wie im vorausgehenden Abschnitt gezeigt - eine sehr einfache Logik. Bei der Anwendung auf (transformierte) empirische Daten ist es nur erforderlich, die Anordnung der Items als Modellparameter zu schätzen, wenn diese theoretisch nicht bekannt ist, wenn also nur vage Vorstellungen dahingehend bestehen, daß die Items unterschiedlich schwer sind. Bestehen dagegen genauere Vorstellungen, können diese durch einen Vergleich mit der empirisch berechneten Anordnung »geprüft« werden. Die letzte Situation trifft für unser Beispiel zu.

Beispiel: Für 19 zufällig ausgewählte Personen aus dem Sozialen Survey Österreichs (siehe Abbildung 2-4) ergeben sich für die Items A bis E folgende Auftrittshäufigkeiten der Ausprägung »1« (Teilnahme):

Item A: 5  
Item B: 8  
Item C: 4  
Item D: 3  
Item E: 2

Item B ist in dieser Stichprobe das leichteste, Item A das zweit leichteste, usw. Die Rangfolge der Items ist B, A, C, D und E und entspricht somit den theoretischen Vorstellungen. In dieser Hinsicht stimmen die theoretischen Vorstellungen mit den bisher berechneten Ergebnissen überein.

## 2.5 Berechnung der theoretischen Datenmatrix und Modellprüfung

Nach der Schätzung der Rangordnung der Items (sofern eine solche erforderlich ist) kann eine theoretische Datenmatrix erzeugt werden (siehe Abbildung 2-4). Dazu ist es erforderlich, für jede Person den Gesamtpunktwert zu berechnen. Die Berechnung der theoretischen Datenmatrix erfolgt unter der Annahme, daß eine perfekte Gutt-manskala vorliegt. Hat also beispielsweise eine Person einen Gesamtpunktwert von 1,0, wird angenommen, daß eine »1« im leichtesten Item auftritt, usw.

Wie aus der Abbildung ersichtlich, weichen für die Personen 11, 12, 14, 15 und 17 die empirischen Antwortmuster von den theoretischen ab. Diese Personen lassen sich nicht durch die angenommene Guttmanskala rekonstruieren.

Zur Messung der Reproduzierbarkeit der empirischen Daten durch die sich bei einer perfekten Guttmanskala ergebenden theoretischen Datenmatrix wurden unterschiedliche Maßzahlen entwickelt. Das gebräuchlichste Vorgehen wurde von GOODENOUGH (1944) vorgeschlagen. Den Ausgangspunkt bildet dabei die Berechnung einer Fehlermatrix. Diese Berechnung kann man sich folgendermaßen vorstellen: die empirische und theoretische Datenmatrix werden übereinander gelegt. Bei einer Übereinstimmung in einer Zelle wird in die Fehlermatrix eine »0« eingetragen, bei einer Nichtübereinstimmung eine »1« (Fehler). Für die Person 11 ergeben sich beispielsweise folgende Eintragungen in die Fehlermatrix:

	B	A	C	D	E	
theoret. Antw.	1	0	0	0	0	
empir. Antw.	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>0</u>	
Fehlermuster	1	1	0	0	0	— Eintragungen in die Fehlermatrix

Es ergeben sich somit zwei Fehler, ein Fehler im Item B und ein Fehler im Item A, da bei diesen beiden items die beiden Antwortmuster nicht übereinstimmen.

Die Fehlermatrix für alle (n= 19) Personen enthält die Abbildung 2-5. In ihr wurden wiederum Zeilen- und Spaltensummen berechnet. Die Zeilensummen enthalten die Fehler je Person, die Spaltensummen die Fehler je Item. Addition der Zeilen- oder Spaltensummen ergibt die Gesamtfehlersumme E (E = Errors) von 10. Diese hängt u.a. von der Zahl der Personen und der Zahl der items ab: Bei gegebener Reproduzierbarkeit einer Skala werden für mehr Personen/mehr items absolut betrachtet mehr Fehler auftreten. Um diesen Effekt zu eliminieren, normierte GOODENOUGH (1944) und GUTTMAN (1950) die Gesamtfehlersumme mit der Zahl der Zellen der Datenmatrix. Diese ist gleich  $n \cdot m$  (n = Zahl der Personen, m = Zahl der items). Die Größe  $E/(n \cdot m)$  mißt also den durchschnittlichen Fehler je Zelle. Da man aber nicht an einem durchschnittlichen Fehlermaß interessiert ist, sondern an einem Maß für die Reproduzierbarkeit wird der durchschnittliche Fehler von 1 subtrahiert. Die so entstehende Maßzahl wird als Reproduktionskoeffizient bezeichnet:

$$\text{Rep} = 1 - \frac{\text{Gesamtfehlerzahl (E)}}{\text{Zahl der Zellen (n \cdot m)}}$$

Der Reproduktionskoeffizient variiert zwischen 0 (keine Reproduzierbarkeit der empirischen Daten) und 1 (perfekte Reproduzierbarkeit der empirischen Daten).

Für die 19 zufällig ausgewählten Personen ergeben sich folgende Kennwerte:

Gesamtfehlersumme E =	10
Zellen der Datenmatrix (n-m): 5-19 =	95
durchschnittl. Fehler je Zelle (E/(n-m)): 10/95 =	0.105
Reproduktionskoeffizient Rep (1 - E/(n-m)): 1 - (10/95) =	0.895

Als Schwellenwert, ab dem die Reproduzierbarkeit noch akzeptiert werden kann, wird 0.85 (GUTTMAN 1950) oder 0.90 (Mc IVER/CARMINES 1983: 48) angegeben. In unserem Beispiel könnte somit die Annahme, daß die empirischen Daten einer Guttmanskala genügen, bei einem Schwellenwert von 0.85 akzeptiert werden, bei einem Schwellenwert von 0.90 dagegen nicht.

Der Reproduktionskoeffizient wurde unmittelbar nach seiner Publikation kritisiert (siehe dazu CLIFF 1983). Die Kritik bestand darin, daß die Reproduzierbarkeit überschätzt wird(1), da die Gesamtfehlerzahl statt mit der Zahl der Zellen (n-m) mit der Zahl der maximal möglichen Fehler (maxE) bei gegebener Anordnung der Items und gegebenen Gesamtpunktwerten normiert werden sollte. Die Zahl der maximal möglichen Fehler ist aber immer kleiner als die Zahl der Zellen. Es findet somit eine Überschätzung an Reproduktion statt. Man kann sich diese Tatsache leicht verdeutlichen: Bei Personen, die in allen Items eine »1« oder »0« besitzen, können keine Fehler auftreten, unabhängig davon, wie die Items angeordnet sind. Sie eignen sich somit für jede beliebige Guttmanskala. Ähnliches gilt für Items: Bei Items mit einem sehr hohen/geringen Anteilswert in der Ausprägung »1«, können - wenn ein rein zufälliges Antwortverhalten angenommen wird - weniger Fehler auftreten.

Von dieser letzten Überlegung geht die Berechnung der minimalen marginalen Reproduktivität aus. Diese versucht zu bestimmen, wieviele Fehler gemacht werden, wenn die empirischen Antworten alleine von den Randverteilungen (margins, Auftrittshäufigkeiten der Ausprägung »1«) der Items bestmöglichst, also mit einem minimalen Fehler reproduziert werden. Eine beste Reproduktion der empirischen Antworten eines Items i ist dann gegeben, wenn die Ausprägung (»0« oder »1«) mit der größten Auftrittshäufigkeit gewählt wird. Bei einem Anteilswert von 0.3 für die Ausprägung »1« würde man die Ausprägung »0« wählen, da  $1-0.3 = 0.7$  größer ist als 0.3, bei einem Anteilswert von 0.6 für die Ausprägung »1« dagegen die Ausprägung »1«, da  $0.6$  größer ist als  $1-0.6 = 0.4$ . Für unsere Beispieldaten sind die Anteilswerte und die entsprechenden Reproduktionen aufgrund der Randverteilungen:



Anteilswerte  
der Auspr. »1«

- Item A:  $P(A) = 0.421$  -> beste Reproduktion »0«  
 Item B:  $P(B) = 0.316$  -> beste Reproduktion »0«  
 Item C:  $P(C) = 0.263$  -> beste Reproduktion »0«  
 Item D:  $P(D) = 0.211$  -> beste Reproduktion »0«  
 Item E:  $P(E) = 0.105$  -> beste Reproduktion »0«

Die beste Reproduktion ist also gegeben, wenn für jede Person ein Antwortmuster, das in allen Items die Ausprägung »0« enthält, angenommen wird. Die so entstehende theoretische Datenmatrix sowie die Fehlermatrix enthält die Abbildung 2-6.

Bei einer Reproduktion aufgrund der Randverteilungen der Items treten insgesamt 25 Fehler auf. Analoge Normierung wie beim Reproduktionskoeffizienten ergibt den minimalen marginalen Reproduktionskoeffizienten (MMR):

$$\text{MMR} = 1 - \frac{\text{Fehler bei Reproduktion durch Randverteilungen}}{\text{Zahl der Zellen (=n} \cdot \text{m)}}$$

$$\text{MMR} = 1 - \frac{25}{95} = 0.737$$

Da bei der Guttmanskala neben den Randverteilungen der Items noch ein bestimmter Mechanismus des Antwortverhaltens angenommen wird, kann die Differenz

$$\text{PI} - \text{Rep} - \text{MMR} = 0.895 - 0.737 - 0.158$$

als prozentuelle Verbesserung der Reproduktion der empirischen Daten durch die Guttmanskala interpretiert werden. In unserem Beispiel wird die Reproduktion der Daten durch die Guttmanskala um 16% verbessert. Neben dieser absoluten Maßzahl kann auch ein PRE-Koeffizient berechnet werden, der die relative prozentuelle Verbesserung der Reproduktion durch die Guttmanskala gegenüber der Reproduktion aus den Randverteilungen der Items mißt. Überträgt man die allgemeine Definition des PRE-Koeffizienten auf die Guttmanskala ergibt sich folgende Formel:

$$\text{PRE} = \frac{\text{Fehler bei Reproduktion durch die Randverteilungen} - \text{Fehler bei Reproduktion durch die Guttmanskala}}{\text{Fehler bei Reproduktion durch die Randverteilungen}}$$

$$= \frac{25 - 10}{25} = 0.6 \text{ (für Beispielsdaten)}$$

Für unsere Beispielsdaten wird die Reproduktion durch die Guttmanskala relativ um 60% verbessert. Theoretische und empirische Arbeiten, ab welcher Höhe von PI oder PRE eine Guttmanskala akzeptiert werden kann, fehlen. Nach MENZEL (zit. in Mc IVER/CARMINES 1983: 50) sollte der PRE-Koeffizient größer/gleich 0.6 sein.

Alle bisher behandelten Maßzahlen machen nur in einem geringen Umfang von den Monotonieeigenschaften der Guttmanskala Gebrauch. Entsprechend der Logik der Guttmanskala setzt eine 1-Anzeige in einem schweren Item eine 1-Anzeige in allen leichteren Items voraus. Umgekehrt ist eine 1-Anzeige in einem schweren Item und eine 0-Anzeige in einem leichten Item ein Hinweis auf einen Fehler. Eine 1-Anzeige in einem leichten Item und eine 0-Anzeige in einem schweren Item ist dagegen mit der Guttmanskala verträglich. Aufgrund dieser Überlegung entwickelte CLIFF (1983) einen Konsistenzkoeffizienten. Dabei werden die Items/Personen paarweise miteinander verglichen. Das Ergebnis des Vergleichs der Itempaare für eine Person kann in drei Ausgängen bestehen:

- konsistentes Paar: Das leichtere Item besitzt die Ausprägung »1« und das schwere die Ausprägung »0«.
- inkonsistentes Paar: Das leichtere Item besitzt die Ausprägung »0« und das schwere die Ausprägung »1«.
- gebundenes Paar: Das leichtere und das schwerere Item besitzen die Ausprägungen »1« oder »0«.

Zur Verdeutlichung der Logik betrachten wir als Beispiel das empirische Antwortmuster der Person 11:

	B	A	C	D	E
Person 11:	0	1	0	0	0

Itempaare:	B-A	inkons. Paar, da B leichter als A
	B-C	gebundenes Paar
	B-D	gebundenes Paar
	B-E	gebundenes Paar
	A-C	kons. Paar, da A leichter als C
	A-D	kons. Paar, da A leichter als D
	A-E	kons. Paar, da A leichter als E
	C-D	gebundenes Paar
	C-E	gebundenes Paar
	D-E	gebundenes Paar

Für die weiteren Berechnungen werden die gebundenen Paare eliminiert, da sie keine Anhaltspunkte über die Konsistenz geben. Dies führt u.a. dazu, daß Personen, die in allen Items die Ausprägung »1« oder »0« be-

sitzen und sich deshalb für beliebige Guttman-skalen eignen, eliminiert werden. Die Informationen über die konsistenten und inkonsistenten Itempaare können nun über alle Personen aufsummiert (dadurch ergeben sich die Besetzungszahlen der Vierfeldertafeln der Items außerhalb der Diagonalen) und übersichtlich in einer sogenannten Dominanzmatrix dargestellt werden (siehe Abbildung 2-7). Das (i,j)-te Element der Dominanzmatrix ist wie folgt zu interpretieren: Das Zeilenelement i (Item i) besitzt die Ausprägung »1«, das Spaltenelement j (Item j) besitzt die Ausprägung »0«. In der Abbildung 2-7 wurde die Dominanzmatrix für die theoretische und empirische Datenmatrix der Abbildung 2-5 berechnet.

Wie der Abbildung 2-7 zu entnehmen ist, bildet die theoretische Dominanzmatrix eine obere Dreiecksmatrix (Voraussetzung: Die Items sind nach dem Schwierigkeitsgrad geordnet). Sind die Zellen der unteren empirischen Dominanzmatrix besetzt, ist dies ein Hinweis auf Fehler (Inkonsistenzen). CLIFF (1983) hat zur Messung der Konsistenz u.a. folgende Maßzahl vorgeschlagen:

$$C = \text{empC}/\text{maxC}$$

mit

empC = empirisch auftretende Konsistenzen

maxC = maximal mögliche Konsistenzen

Die Zahl der maximal möglichen Konsistenzen ist gleich der Summe der Besetzungszahlen der theoretischen Dominanzmatrix, in dem Beispiel der Abbildung 2-7 ist maxC

$$\text{maxC} = 4 + 7 + 9 + 10 + 3 + \dots = 50$$

Die empirisch auftretenden Konsistenzen ergeben sich dadurch, daß von der Zahl der konsistenten Paare (obere Dreiecksmatrix) die Zahl der inkonsistenten Paare (untere Dreiecksmatrix) subtrahiert wird:

Zahl der Konsistenzen in der Abbildung 2-7:

$$C(+) = 4 + 4 - 5 + 5 + 5 + 6 + 2 + \dots = 39$$

Zahl der Inkonsistenzen in der Abbildung 2-7:

$$C(-) = 2 + 2 + 1 + 1 + \dots = 11$$

Die Zahl der empirischen Konsistenzen ist folglich

$$\text{empC} = C(+) - C(-) = 39 - 11 = 28$$

Einsetzen in die Formel des Konsistenzkoeffizienten ergibt:

$$C = \frac{28}{50} = 0.560$$

Wie beim minimalen Reproduktionskoeffizient kann man nun fragen, ob man die Zahl konsistenter oder inkonsistenter Paare nicht durch ein einfacheres Modell reproduzieren kann. Ein solches einfaches Modell ist gegeben, wenn statistische Unabhängigkeit, also ein zufälliges Antwortverhalten, angenommen wird. In diesem Fall würde man für ein Itempaar (j,i)  $p(j) \cdot (1-p(i)) \cdot n$  inkonsistente Paare erwarten ( $p(j)$  = Anteilswert der Ausprägung »1« des schwereren Items,  $p(i)$  = Anteilswert der Ausprägung »1« des leichteren Items,  $n$  = Zahl der Personen) erwarten. Die sich für unsere Beispieldaten bei statistischer Unabhängigkeit ergebende Zahl inkonsistenter und konsistenter Paare enthält Abbildung 2-8.

Da bei Vorliegen einer perfekten Guttmanskala die Zahl der Inkonsistenzen eines Paares gleich 0 ist, ist zu fordern, daß die empirischen Inkonsistenzen deutlich kleiner der Inkonsistenzen bei statistischer Unabhängigkeit ist. Von dieser Überlegung wird bei der Berechnung des Homogenitätskoeffizienten, der von LOEVINGER (1948) aufgrund anderer Überlegungen entwickelt wurde (CLIFF 1983), ausgegangen. Der Homogenitätskoeffizient ist definiert als (KINGMA/TAERUM 1989):

$$H = \frac{\text{Inkonsistenzen bei statistischer Unabhängigkeit} - \text{empirische Inkonsistenzen}}{\text{Inkonsistenzen bei stat. Unabhängigkeit}}$$

Entsprechend der Abbildung 2-8 ist die Zahl inkonsistenter Paare bei statistischer Unabhängigkeit gleich 23.368, die Zahl der empirisch inkonsistenten Paare gleich 11. Für den Homogenitätskoeffizienten ergibt sich daher ein Wert von:

$$H = \frac{23.368 - 11.000}{23.368} = 0.529$$

Nach MOKKEN (1971) können aufgrund von H folgende Skalen typen unterschieden werden:

$H \geq 0.500$	Die Items bilden eine starke Skala.
$0.500 > H \geq 0.400$	Die Items bilden eine mittlere Skala.
$0.400 > H \geq 0.300$	Die Items bilden eine schwache Skala.
$0.300 > H$	Die Items bilden keine Skala.

In unserem Beispiel liegt somit eine starke Skala nach MOKKEN vor. Die theoretischen Vorstellungen können akzeptiert werden.

Zu beachten ist, daß auf dem ersten Blick der Homogenitätskoeffizient analog dem PRE-Koeffizienten definiert ist. Tatsächlich wird beim Homogenitätskoeffizienten aber nicht auf die Reproduktion der empirischen Daten durch eine perfekte Guttmanskala zurückgegriffen, sondern nur

auf die Anordnung nach der Schwierigkeit. MOKKEN (1971) hat gezeigt, daß damit ein sehr viel allgemeineres Skalierungsmodell, das die Guttman-skala als Submodell enthält, überprüft wird. Dieses Modell wird als Mokkenskala bezeichnet und in Kapitel 3 behandelt.

Für alle Befragten (n = 2014) ergeben sich folgende Werte:

Zahl der Befragten	= 2014.
Wegen fehlenden Werten eliminierte Befragte	= 36.
Befragte in der Analyse	= 1978.
Gesamtfehler (nach GOODENOUGH)	= 1188.
Zahl der Zellen (m-n)	— 9890.
Reproduktionskoeff. REP	= 0.880
MMR-Koeffizient	= 0.765
PRE-Koeffizient	= 0.489
Analyse nach CLIFF:	
maximale Zahl konsistenter Paare (maxC)	= 5432.
Differenz zw. konsistenten und inkonsistenten Paare (empC)	= 3036.
Konsistenzkoeffizient (C)	— 0.559
Analyse nach MOKKEN:	
emp. Zahl inkonsistenter Paare	= 1198.000
Zahl inkonsistenter Paare bei statistischer Unabhängigkeit	— 2111.713
Homogenitätskoeffizient	= 0.433

Die Skala für alle Befragten würde entsprechend den Kriterien von MOKKEN als mittlere Skala bezeichnet werden. Entsprechend dem von MENZEL für den PRE-Koeffizienten angegebenen Schwellenwert würde keine brauchbare Guttman-skala vorliegen. Unter der Annahme eines Schwellenwertes von 0.85 für den Reproduktionskoeffizienten würde eine brauchbare Skala vorliegen, bei einem Schwellenwert von 0.90 dagegen nicht. Ob man die vorliegende Skala akzeptiert, kann somit nicht eindeutig entschieden werden. Diese Entscheidung kann in diesem Fall vom nächsten Schritt der Lösung einer Skalierungsaufgabe, der Fehleranalyse, abhängig gemacht werden.

## 2.7 Fehleranalyse

Für eine schlechte Reproduktion der empirischen Daten durch das gewählte Skalierungsverfahren sind eine Reihe von Ursachen denkbar, die bereits in Abschnitt 1-2 aufgezählt wurden. Mit Hilfe der Guttman-skala können zwei Fehlerursachen untersucht werden:

- a) Einzelne Items lassen sich schlecht durch die Guttmanskala reproduzieren. Eine mögliche Ursache dafür kann beispielsweise Mehrdimensionalität sein.
- b) Einzelne Personen lassen sich schlecht durch die Guttmanskala reproduzieren. Eine mögliche Ursache dafür können beispielsweise individuelle Unterschiede in der Anordnung der Items sein. Dabei ist es möglich, daß die empirische Datenmatrix eine Mischung von zwei oder mehr Guttmanskalen darstellt.

Um nun Anhaltspunkte für diese beiden Fehlerquellen zu finden, können die im vorausgehenden Abschnitt dargestellten Koeffizienten für Personen oder Items berechnet werden. So kann z.B. für jede Person und/oder für jedes Item ein Reproduktionskoeffizient berechnet werden. Die entsprechenden Formeln sind:

$$\text{Rep (Item } i) = 1 - \frac{\text{Gesamtfehler Item } i}{\text{Zahl der Zellen in denen Item } i \text{ vorkommt}} \\ (= \text{Zahl der Personen})$$

$$\text{Rep (Person } g) = 1 - \frac{\text{Gesamtfehler bei der Person } g}{\text{Zahl der Zellen, in denen Person } g \text{ vorkommt}} \\ (= \text{Zahl der Items})$$

Betrachten wir zur Verdeutlichung den Datensatz der 19 Personen (Abbildung 2-4 und 2-5). In Item B treten 3 Fehler auf. Der Reproduktionskoeffizient dieses Items ist  $1 - 3/19 = 1 - 0.158 = 0.842$ , da 19 Personen in die Analyse einbezogen wurden. Für die Person 11 beispielsweise, würde sich ein Reproduktionskoeffizient von  $1 - 2/5 = 1 - 0.4 = 0.6$  ergeben, da bei dieser Person 2 Fehler auftreten und 5 Items in die Analyse einbezogen wurden. Im Prinzip ist es möglich, alle behandelten Koeffizienten sowohl für Personen als auch für die einzelnen Items zu berechnen. Für Personen entsteht dadurch aber ein zum Teil sehr beträchtlicher Rechenaufwand, da beispielsweise bei 2000 Personen eine  $2000 \times 2000$  Dominanzmatrix berechnet werden müßte, die die Grundlage der Berechnung des Konsistenzindex bilden würde. Auch für den Fall, daß nur alle empirisch auftretenden Antwortmuster (bei fünf Items maximal  $2^5 = 32$ ) untersucht werden, ist noch ein beträchtlicher Rechenaufwand gegeben. Es wurden deshalb weniger rechenintensive Verfahren vorgeschlagen, bei denen nicht jede Person mit jeder Person verglichen werden muß (siehe dazu die Literaturverweise in CLIFF 1983). Für die Items ist dagegen i.d.R. ein geringerer Rechenaufwand erforderlich.

Für unser Beispiel ( $n = 2014$  Befragte) ergeben sich für die einzelnen Items folgende Kennwerte:

Item	Fehler	Rep.	MMR	PI(a)	PRE	C(b)	H(c)
A	307.	0.845	0.758	0.087	0.358	0.454	0.416
B	221.	0.888	0.537	0.351	0.758	0.782	0.587
C	225.	0.886	0.807	0.079	0.409	0.427	0.399
D	241.	0.878	0.841	0.037	0.232	0.461	0.384
E	194.	0.902	0.880	0.022	0.185	0.572	0.394

(a) prozentuelle Verbesserung

(b) Konsistenzkoeffizient

(c) Homogenitätskoeffizient

Betrachtet man den Reproduktionskoeffizient der Personen, so ergeben sich fünf Antwortmuster mit einem sehr niedrigen Reproduktionskoeffizienten von 0.2:

B	A	C	D	E	Häufigkeit	
0	0	0	1	0	4	Mitwirken in Bürgerinitiativen.
0	0	0	1	1	5	Mitwirken in Bürgerinitiativen und Mitarbeit in Parteien.
0	0	1	1	0	12	Mitwirken in Bürgerinitiativen und Kontakt mit Politikern
0	1	0	1	1	3	Teilnahme an Kundgebungen, Mitwirken in Bürgerinitiativen und Mitarbeit in Parteien
1	0	0	1	0	1	Unterstützung von Volksbegehren und Mitwirken in Bürgerinitiativen

Nach dieser Berechnung können schlechte Items oder nicht skalierbare Personen eliminiert werden. Dabei ist ein iteratives Vorgehen vorstellbar:

Schritt 1: Eliminiere schlechtestes Item oder die am wenigsten skalierbare Person.

Schritt 2: Für die so entstandene neue Datenmatrix wird erneut eine Analyse gerechnet.

Schritt 3: Kann die Skala nicht akzeptiert werden, wird wiederum mit Schritt 1 fortgefahren. Kann die Skala akzeptiert werden, wird das Verfahren abgebrochen.

Welche der beiden Strategien (Elimination von Personen oder von Items) beim Schritt 1 verfolgt wird, ist letzten Endes eine Entscheidung des Forschers, obwohl üblicherweise von einer homogenen Population (keine individuellen Unterschiede) ausgegangen und eine Verbesserung der Reproduktion durch die Elimination von Items versucht wird. In Abhängigkeit von der Wahl des Koeffizienten würden für unsere Beispielsdaten in einem ersten Schritt folgende Items eliminiert werden:

Reproduktionskoeffizient Rep	→ Elimination von Item A
prozentuelle Verbesserung PI	→ Elimination von Item E
PRE-Koeffizient	→ Elimination von Item E
Konsistenzkoeffizient C	→ Elimination von Item C
Homogenitätskoeffizient H	→ Elimination von Item D

Dieses Beispiel dürfte deutlich machen, daß bei einer rein automatischen Vorgehensweise Vorsicht angebracht ist, da unterschiedliche Items eliminiert werden. Jede Entscheidung ist inhaltlich zu begründen.

Eine andere Vorgehensweise bei der Fehleranalyse geht auf GOODMAN (1975) zurück. GOODMAN schlägt vor, die nicht skalierbaren Personen zu einer Gruppe zusammenzufassen. Bezüglich dieser Gruppe wird angenommen, daß ein rein zufälliges Antwortverhalten vorliegt. Die Überprüfung dieses erweiterten Modells ist mit Hilfe der Analyse latenter Klassen möglich (s. Kapitel 4).

Die Annahme, daß die nicht skalierbaren Personen eine Gruppe von Personen mit zufälligem Antwortverhalten bilden, kann aber auch mit Hilfe der Guttmanskala überprüft werden. Dazu wird für die nicht skalierbaren Personen eine Guttmanskala gerechnet. Das Ergebnis der Skalierung kann in folgenden Ausgängen bestehen:

- a) Die nicht skalierbaren Personen bilden eine eigene Guttmanskala. In diesem Fall können die empirischen Daten als eine Mischung von zwei Guttmanskalen interpretiert werden. Zu beachten ist dabei, daß die Personen, die in allen Items die Ausprägung »1« oder »0« besitzen, nicht mehr eindeutig einer der beiden Skalen zugeordnet werden können. Zur graphischen Darstellung derartiger gemischter Skalen siehe BORG/ STAUFENBIEL (1989: 136-146).
- b) Die Annahme des zufälligen Antwortverhaltens kann nicht verworfen werden. Dies ist dann der Fall, wenn der PRE-Koeffizient und der Homogenitätskoeffizient klein sind (2).

Führt man für die nicht skalierbaren Personen unseres Datensatzes eine Guttmanskalierung durch, ergibt sich ein PRE-Koeffizient von 0.081. Der Homogenitätskoeffizient beträgt 0.153. Es kann somit von einer Gruppe nicht skalierbarer Personen mit zufälligem Antwortverhalten ausgegangen werden.

Auf der Grundlage dieses Ergebnisses kann in einem nächsten Schritt gefragt werden, ob das Modell zufälliger Fehler nicht auf alle Personen verallgemeinert werden kann. Dies hätte den Vorteil, daß für alle Personen - unter Berücksichtigung von zufälligen Fehlern - ein Skalenwert auf der latenten Dimension berechnet werden kann. Dies bedeutet, daß als Skalierungsverfahren ein probabilistisches Skalierungsverfahren verwendet werden muß. Einige Verfahren, die dieser Gruppe von Skalierungsverfahren angehören, werden im nächsten Kapitel behandelt.



## 2.8 Berechnen von Skalenwerten für Personen

Erweist sich die Guttmanskala als brauchbar, besteht der letzte Schritt der Lösung einer Skalierungsaufgabe in der Berechnung von Skalenwerten für die Personen. Dabei müssen zwei Entscheidungen getroffen werden:

- Behandlung fehlender Werte. Soll für Personen mit fehlenden Werten ein Skalenwert geschätzt werden oder sollen diese Personen aus der Berechnung eliminiert werden?
- Behandlung nicht skalierbarer Personen. Soll für nicht skalierbare Personen, also für Personen mit einer Fehlersumme größer 0, ein Skalenwert geschätzt werden, oder sollen diese als eine eigene Gruppe nicht skalierbarer Personen behandelt werden?

Anzumerken ist, daß die Guttmanskalierung selbst, also die Prüfung der Skala, immer ohne fehlende Werte durchgeführt wird. Personen mit einem oder mehreren fehlenden Werten aus der Skalierung eliminiert.

Eine sehr allgemeine Lösung würde darin bestehen, für jede Person folgende Informationen zur Verfügung zu haben:

- Einen »Skalenwert«
- Die Zahl der fehlenden Werte
- Die Zahl aller Fehler

Diese Informationen würden es beispielsweise ermöglichen, in bestimmten Anwendungsfällen Personen mit fehlenden Werten zu eliminieren oder nicht-skalierbare Personen als eigenständige Gruppe zu behandeln. Die Berechnung ist in der Abbildung 2-9 verdeutlicht. Sie besteht darin, für jede Person einen »Quasi-Skalenwert« zu berechnen. Liegen keine fehlenden Werte vor, ist dieser identisch mit dem Gesamtpunktwert. Bei fehlenden Werten wird der Anteil der »1«-Anzeigen an den gültigen Ausprägungen berechnet und anschließend mit der Zahl der Items multipliziert. Die so entstandene Zahl wird gerundet und als Skalenwert interpretiert.

Ein verfeinertes Vorgehen würde in der Berechnung eines gewichteten Anteilwertes bestehen. Als Gewichte werden dabei die Schwierigkeitsgrade der Items verwendet. Diesem Vorgehen liegt die Annahme zugrunde, daß eine »1«-Anzeige in einem schweren Item einem »größeren« Skalenwert auf der Meßdimension entspricht als eine »1«-Anzeige in einem leichten Item. Zu beachten ist auf jeden Fall, daß die berechneten Skalenwerte ein ordinales Meßniveau besitzen. Das bedeutet, daß die Abstände zwischen den Skalenwerten nicht definiert sind. Eine ausführliche Diskussion der Berechnung von Skalenwerten sowie der Behandlung fehlender Werte geben McIVER/CARMINES (1983: S. 1-53, 64-65).

## 2.9 Anwendungsfelder in der historischen Sozialforschung

Entsprechend dem formalen Modell der Guttmanskala kann die Guttmanskala für Skalierungsaufgaben eingesetzt werden, wo (a) ein eindimensionaler Begriff, (b) eine Anordnung von Items (Indikatoren) nach ihrer Schwierigkeit, (c) ein deterministischer Mechanismus angenommen werden kann und (d) dichotome Items vorliegen. Die Annahmen (a) bis (c) hängen von den theoretischen Vorstellungen ab. Die Voraussetzung (d) bezieht sich auf das Datenmaterial. Bei statistischen Datenquellen (Volkszählungen, u.ä.) werden i.d.R. quantitative Indikatoren vorliegen. Eine Anwendung der Guttmanskala wäre nicht sinnvoll, da die Indikatoren dichotomisiert werden müßten. Dies würde einen erheblichen Informationsverlust bedeuten. Ein derartiger Informationsverlust wäre nur vertretbar, wenn durch die Dichotomisierung die Zuverlässigkeit der Daten erhöht werden könnte. Bei inhaltsanalytischen Auswertungen nicht statistischer Datenquellen, wie z.B. von Zeitungen, Berichte u.ä., dagegen werden vielfach keine quantitativen Indikatoren gewonnen werden können. Hier ist eine Anwendung der Guttmanskala möglich. So kann beispielsweise eine auf der Grundlage der Guttmanskala entwickeltes Kategorienschema angewendet werden (siehe Abbildung 2-10).

In der Abbildung 2-10 wurde ein zur Abbildung 2-1 analoges Schema zur Messung des politischen Aktivitätsprofil von politischen Gruppen entwickelt. Zum Unterschied von Personen werden politische Gruppen skaliert. Das Schema kann beispielsweise zur Analyse eines Konfliktes eingesetzt werden. Es ist aber auch denkbar, daß das Schema zur Skalierung politischer Konflikte selbst angewendet wird. Dabei wird dann von folgendem Modell politischer Konflikte ausgegangen:

Kundgebungen -> Warnstreiks -> Streiks -> Revolten

Hat man sich aufgrund von theoretischen und/oder empirischen Überlegungen für die Anwendung der Guttmanskala entschieden, tritt neben den bereits behandelten Problemen u.a. das praktische Problem auf, daß in den Standardsoftwarepaketen BMDP, SAS und SPSS-X (3) die Guttmanskalierung nicht als Programmmodul enthalten ist. In dem Programmsystem ALMO (HOLM 1989) steht die Guttmanskala als Submodell des latenten Distanzmodells der Analyse latenter Strukturen zur Verfügung (DENZ 1982, 1989a). Dabei wird aber nur ein pauschale Chi-Quadrat-Testgröße berechnet, die für diagnostische Zwecke (Fehlersuche) wenig brauchbar ist.

Da die Guttmanskala aber eine sehr einfache Logik besitzt, kann dieses Problem dadurch gelöst werden, daß ein eigenständiges Programm (4) geschrieben oder die Guttmanskala innerhalb eines Standardstatistiksystem programmiert wird. Diese letzte Möglichkeit wird im nachfolgenden Ab-

schnitt für das Programmsystem dargestellt.

Eine andere Möglichkeit besteht darin, ein Programm zur Analyse latenter Klassen zu verwenden (siehe dazu Abschnitt 4.7). Die Einarbeitung in ein solches Programm ist zu empfehlen, da mit ihm auch zahlreiche der nachfolgend besprochenen Skalierungsverfahren gerechnet werden können.

## 2.10 Exkurs: Durchführung einer Guttmanskalierung mit SPSS-PC +

Die Durchführung einer Guttmanskalierung mit SPSS-X und mit SPSS-PC+ soll am Beispiel der Daten des Sozialen Surveys demonstriert werden. Da die PC-Version von SPSS über einen geringeren Umfang an Datenmanipulationsmöglichkeiten verfügt, werden die Programme in der PC-Version wiedergegeben (5). Für SPSS-X existieren zum Teil über DO REPEAT-Schleifen elegantere Lösungswege. Mit den folgenden zwei Programmen kann in SPSS-PC-f (SPSS-X) eine Guttmanskalierung durchgeführt werden:

- Programm 1: Berechnen der **Anordnung** der Items.
- Programm 2: Berechnen des Reproduktions-, des **minimalen marginalen** Reproduktions- und des PRE-Koeffizienten.

Entsprechend der Logik des Programms 2 kann ein Programm zur Berechnung des Homogenitäts- und Konsistenzkoeffizienten geschrieben werden.

Ein Programm zur Berechnung von Skalenwerten bei fehlenden Werten sowie für nicht skalierbare Personen kann nach der im Abschnitt 2.8 beschriebenen Logik ebenfalls leicht programmiert werden.

### Programm 1:

```
1      data list file = 'pol.dat' free/a,b,c,d,e.
2      missing value a to e (-1).
3      count ig = a to e(missing).
4      freqvar=ig.
5      select if (ig eq 0).
6      descriptives var = a to e.
```

Die Programmzeilen bedeuten:

- 1 Definition der Daten. In dem Beispiel werden die Daten als Rohdatenfile mit dem Namen 'pol.dat' definiert. Sie sind im freien Format (free) gespeichert. Die Variablen a,b,c... sind dabei wie folgt verkodet:
  - 1 = fehlender Wert
  - 0 = keine **Teilnahme**
  - 1 = Teilnahme

- Liegt diese Verkodung nicht vor, müssen die Daten mit recode umkodiert werden.
- In SPSS-X ist die Definition des Datenfiles vom Betriebssystem abhängig. I.d.R. wird eine FILE HANDLE-Anweisung benötigt.
- 2 Definition von fehlenden Werten. Die Ausprägung -1 wird als Ausprägung für fehlende Daten definiert.
- 3 Die Zahl der fehlenden Werte wird für jeden Datensatz (Person) berechnet und in die Variable ig geschrieben.
- 4 Die Variable ig wird ausgezählt. Dadurch ergeben sich Anhaltspunkte, wieviele Datensätze wegen fehlender Werte für die Skalierung eliminiert werden müssen.
- 5 Diese Anweisung bewirkt, daß nur Datensätze ohne fehlende Werte ( $ig = 0$ ) in die Analyse einbezogen werden.
- 6 Die Prozedur deskriptives berechnet die Anteilswerte der Items. Aufgrund dieser kann die Anordnung der Items festgelegt werden. Eine Alternative wäre die Verwendung der Prozedur FREQUENCIES.

#### Programm 2:

```
1 data list file - 'pol.dat' free/a,b,c,d,e.
2 missing value a to e (-1).
3 count ig = a to e(missing).
4 select if (ig eq 0).
5 compute ig = b-fa + c + d+e.
6 compute aa = 0.
7 compute bb = 0.
8 compute cc = 0.
9 compute dd = 0.
10 compute ee = 0.
11 if (ig ge 1) bb=1.
12 if (ig ge 2) aa = 1.
13 if (ig ge 3) cc = 1.
14 if (ig ge 4) dd= 1.
15 if (ig ge 5) ee= 1.
16 compute fa = 0.
17 compute fb = 0.
18 compute fc = 0.
19 compute fd = 0.
20 compute fe = 0.
21 if (aa ne a) fa= 1.
22 if (bb ne b) fb= 1.
23 if (cc ne c) fc= 1.
```

```

24   if (dd ne d) fd=1.
25   if (ee ne e) fe = 1.
26   compute gg= 1.
27   aggregate outfile = *
28       /break = gg
29       /pa,pb,pc,pd,pe = mean(a,b,c,d,e)
30       /ea,eb,ec,ed,ee = sum(fa,fb,fc,fd,fe)
31       /nn = sum(gg).
32   if (pa gt 0.5) pa = 1.-pa.
33   if (pb gt 0.5) pb = 1.-pb.
34   if (pc gt 0.5) pc = 1.-pc.
35   if (pdgt0.5)pd=1.-pd.
36   if (pe gt 0.5) pe = 1.-pe.
37   compute pa = pa*nn.
38   compute pb = pb*nn.
39   compute pc = pc*nn.
40   compute pd = pd*nn.
41   compute pe = pe*nn.
42   compute item1 = 'A\
43   compute item2 = 'B\
44   compute item3 = 'C\
45   compute item4 = 'D\
46   compute item 5 = 'E\
47   set results = 'poi.res'.
48   write var=item1,nn,ea,pa,
49       item2,nn,eb,pb,
50       em3,nn,ec,pc;
51       item4,nn,ed,pd,
52       em5,nn,ee,pe.
53   data list file='poi.res' free/item (A),nn,feh1,mfehl.
54   list variables = item,nn,feh1,mfehl.
55   if (mfehl gt 0) pre = (mfehl-fehl)/mfehl.
56   compute rep= 1.-feh1/nn.
57   compute mnr = 1.-mfeh1/nn.
58   compute pi = rep-mmr.
59   compute rep=rep* 100.
60   compute pi = pi* 100.
61   compute mmr = mnr* 100.
62   compute pre = pre* 100.
63   report format = list
64       /variables = item,feh1,mfeh1,rep,mmr,pi,pre
65       /break = (nobreak)
66       /summary=mean.

```

Die Zeilen des Programms bedeuten:

- 1 - 4 Diese Zeilen entsprechen den Zeilen 1,2,3 und 5 des Programms 1.
- 5 Für jede Person wird der Gesamtpunktwert berechnet.
- 6-10 Durch diese Anweisungen werden die Variablen der theoretischen Datenmatrix definiert und gleich Null gesetzt. Die Variable aa der theoretischen Datenmatrix entspricht der Variablen a der empirischen Datenmatrix, usw...
- 11-15 Durch diese Anweisungen wird für jede Person das theoretische Antwortmuster berechnet. Die erste if-Anweisung bewirkt, daß alle Personen mit einem Gesamtpunktwert größer/gleich 1 (ig ge l) in der Variablen bb den Wert 1 enthalten, da das Item B das leichteste ist, usw. Zu beachten ist, daß die Variablen nach dem Schwierigkeitsgrad angeordnet werden müssen.
- 16-20 Durch diese Anweisungen werden die Variablen der Fehlermatrix definiert und gleich Null gesetzt.
- 21-25 Für jede Person wird geprüft, ob das theoretische mit dem empirischen Antwortmuster übereinstimmt. Ist dies bei einer oder mehreren Variablen nicht der Fall, erhält (erhalten) die entsprechende^ Fehlervariable(n) den Wert 1. Besteht also beispielsweise eine Abweichung in der Variablen a (aa ne a) erhält die Fehlervariable fa den Wert 1.
- 26 Durch diese Anweisung wird eine fiktive Variable gebildet, die für die anschließende Aggregation benötigt wird. Sie besitzt für alle Personen den Wert 1.
- 28-31 Durch diese Anweisungen werden Spaltensummen bzw. Spaltenmittelwerte der Variablen über alle in die Analyse einbezogenen Personen berechnet, da die Aggregatsvariable gg (Variable nach break =) für alle Personen die Ausprägung 1 besitzt. Die Variablen der Personen werden aggregiert. Durch die Definition outfile = \* bilden die aggregierten Variable die neue SPSS-PC+-internen Datenmatrix, auf die sich die folgenden Befehle beziehen. Sie besteht nur aus einem Datensatz (einer Zeile), der folgende Variablen enthält:
- pa = Anteilswert der Ausprägung »1« in Item A  
pb = Anteilswert der Ausprägung »1« in Item B  
pc = Anteilswert der Ausprägung »1« in Item C  
pd = Anteilswert der Ausprägung »1« in Item D  
pe = Anteilswert der Ausprägung »1« in Item E  
fa = Zahl der Fehler in Item A  
fb = Zahl der Fehler in Item B  
fc = Zahl der Fehler in Item C  
fd = Zahl der Fehler in Item D

- fe = Zahl der Fehler in Item E  
nn = Zahl der in die Skalierung einbezogenen Personen.
- 32 - 33 Durch diese Anweisungen werden die Anteilswerte der Variablen bestimmt, die zu minimalen Reproduktionsfehler bei einer Reproduktion aus den Randverteilungen der Items führen. Ist der Anteilswert der Ausprägung »1« eines Items größer 0.5 (z.B. 0.6), beträgt der kleinst mögliche Fehler 0.4. Es muß also mit »1 - Anteilswert« gerechnet werden.
- 34 - 41 Durch diese Anweisungen wird für jedes Item die absolute Höhe des minimalen Fehlers bei einer Reproduktion aus der Randverteilung berechnet.
- 42 - 46 Durch diese Anweisungen werden die Namen der Items für die Ausgabe bei report festgelegt.
- 47 Der Name der Datei für die nachfolgende write-Anweisung wird definiert (Voreinstellung: spss.prc als Dateiname).
- 48 - 52 Der Name, die Zahl der in die Skalierung einbezogenen Personen, die Zahl der Fehler unter der Annahme einer Guttmanskala sowie unter der Annahme einer Reproduktion aus der Randverteilung für jedes Item werden in die Datei pol.res geschrieben.
- 53 Die mit write hinausgeschriebenen Informationen werden eingelesen. Dabei enthalten die Variablen:
- item: Name des Items (A,B,C,..)
  - nn: Zahl der in die Skalierung einbezogenen Personen
  - fehl: Zahl der Fehler des Items bei einer Reproduktion unter der Annahme einer Guttmanskala
  - mfehl: Zahl der Fehler des Items bei einer Reproduktion aus der Randverteilung.
- 54 Kontrollausgabe.
- 55 Der PRE-Koeffizient wird für jedes Item berechnet.
- 56 Der Reproduktionskoeffizient wird für jedes Item berechnet.
- 57 Der minimale Reproduktionskoeffizient wird für jedes Item berechnet.
- 58 Die prozentuelle Verbesserung wird für jedes Item berechnet.
- 59 - 62 Die berechneten Koeffizienten werden mit 100 multipliziert. Sie werden also in % ausgegeben.
- 63 - 66 Die berechneten Koeffizienten werden mit Hilfe der Prozedur report ausgegeben.

Das Ergebnis sieht wie folgt aus:

ITEM	FEHL	MFEHL	REP	MMR	PI	PRE
A	307.00	478.00	84.48	75.83	8.65	35.77
B	221.00	915.00	88.83	53.74	35.09	75.85
C	225.00	381.00	88.62	80.74	7.89	40.94
D	241.00	314.00	87.82	84.13	3.69	23.25
E	194.00	238.00	90.19	87.97	2.22	18.49
MEAN	237.6000	465.2000	387.9879	76.4813	11.5066	38.8603

Vergleicht man die Ergebnisse mit den Ergebnissen aus den Abschnitten 2-7 und 2-8 so zeigt sich, daß nur der PRE-Koeffizient für die Gesamtskala falsch ist. Er muß aus den Mittelwerten der Variablen FEHL und MFEHL berechnet werden:

$$\text{PRE für Gesamtskala} = \frac{465.2 - 237.6}{465.2} = .489$$



Abbildung 2-1: Skala zur Messung des Aktivitätsprofil

---

Es gibt unterschiedliche Möglichkeiten, sich politisch zu betätigen. Die nachfolgende Liste enthält verschiedene Aktivitäten. Geben Sie bitte bei jeder Tätigkeit an, ob Sie sie schon gemacht haben oder noch nicht?

	habe ich schon ge- macht	habe ich noch nicht gemacht
A - An einer Kundgebung teil- nehmen.	<input type="radio"/>	<input type="radio"/>
B - Ein Volksbegehren unter- stützen.	<input type="radio"/>	<input type="radio"/>
C - Kontakt mit einem Politiker suchen, z.B. einen Brief schreiben, anrufen, ihn in den Amtsstunden aufsuchen.	<input type="radio"/>	<input type="radio"/>
D - Bei einer Bürgerinitiative mitmachen, z.B. gegen neue Straßenprojekte, für einen Kinderspielplatz.	<input type="radio"/>	<input type="radio"/>
E - In einer politischen Partei aktiv mitarbeiten, z.B. ein Amt übernehmen, beim Wahl- kampf helfen.	<input type="radio"/>	<input type="radio"/>

---

Abbildung 2-2: Theoretisch erwartete Antwortmuster auf die Skala der Abbildung 2-1

		nach zeitlichem Engagement ge- ordnete Items					
		B	A	C	D	E	
Personen mit geringer							
Beteiligung ----->		0	0	0	0	0	"0" = keine Teil-
		1	0	0	0	0	nahme
		1	1	0	0	0	"1" = Teilnahme
		1	1	1	0	0	
		1	1	1	1	0	
Personen mit hoher----->		1	1	1	1	1	
Beteiligung							

Abbildung 2-3: (Fiktive) Datenmatrix bei Vorliegen einer perfekten Guttman-skala

Personen	nach Schwierig- keit geordnete Items					Summe
	B	A	C	D	E	
1	0	0	0	0	0	0
2	1	0	0	0	0	1
3	1	1	0	0	0	2
4	1	1	1	0	0	3
5	1	1	1	1	0	4
6	1	1	1	1	1	5
Summe	5	4	3	2	1	

Abbildung 2-4: Theoretische und empirische Datenmatrix für n=19 Personen für die politische Aktivitätsskala

Person	Gesamtpunktwert	theoretische Antworten					empirische Antworten					Gesamtpunktwert
		B	A	C	D	E	B	A	C	D	E	
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0
9	1	1	0	0	0	0	1	0	0	0	0	1
10	1	1	0	0	0	0	1	0	0	0	0	1
11	1	1	0	0	0	0	0	1	0	0	0	1
12	1	1	0	0	0	0	0	0	1	0	0	1
13	2	1	1	0	0	0	1	1	0	0	0	2
14	2	1	1	0	0	0	1	0	0	1	0	2
15	2	1	1	0	0	0	1	0	0	0	1	2
16	3	1	1	1	0	0	1	1	1	0	0	3
17	3	1	1	1	0	0	0	1	1	1	0	3
18	4	1	1	1	1	0	1	1	1	1	0	4
19	5	1	1	1	1	1	1	1	1	1	1	5
Summe		11	7	4	2	1	8	5	4	3	2	

theoretische Datenmatrix

empirische Datenmatrix

empirische Informationen, die in die Konstruktion der theoretischen Datenmatrix eingehen.

Abbildung 2-5: Fehlermatrix für die Datenmatrix der Abbildung 2-4 unter der Annahme einer perfekten Guttman skala

Person	theoretische Antworten					empirische Antworten					Fehlermuster					Summe
	B	A	C	D	E	B	A	C	D	E	B	A	C	D	E	
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
10	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
11	1	0	0	0	0	0	1	0	0	0	1	1	0	0	0	2
12	1	0	0	0	0	0	0	1	0	0	1	0	1	0	0	2
13	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0
14	1	1	0	0	0	1	0	0	1	0	0	1	0	1	0	2
15	1	1	0	0	0	1	0	0	0	1	0	1	0	0	1	2
16	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0
17	1	1	1	0	0	0	1	1	1	1	0	1	0	0	1	2
18	1	1	1	1	0	1	1	1	1	0	0	0	0	0	0	0
19	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
						Summe	3	3	1	2	1					10

Fehlermatrix — Zahl der Fehler bei Item A  
 Gesamtfehlerzahl — Zahl der Fehler bei Person 17

Abbildung 2-6: Fehlermatrix für die Datenmatrix der Abbildung 2-4 bei einer Reproduktion aus den Randverteilungen der Items.

	Person theoretische Antworten					empirische Antworten					Fehler- muster					Summe
	B	A	C	D	E	B	A	C	D	E	B	A	C	D	E	
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
10	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
11	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1
12	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1
13	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	2
14	0	0	0	0	0	1	0	0	1	0	1	0	0	1	0	2
15	0	0	0	0	0	1	0	0	0	1	1	0	0	0	1	2
16	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	3
17	0	0	0	0	0	0	1	1	1	0	0	1	1	1	0	3
18	0	0	0	0	0	1	1	1	1	0	1	1	1	1	0	4
19	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	5
						Summe					8	6	5	4	2	25

theoretische Matrix, wenn die Antworten aufgrund der Randverteilungen der Items reproduziert werden.

Fehlermatrix

Gesamtfehler

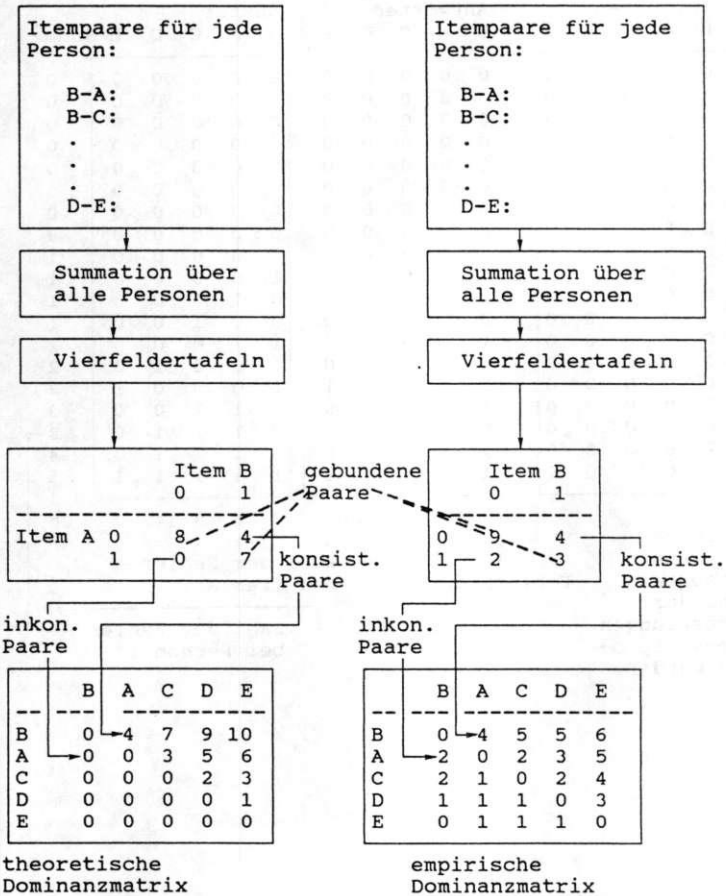
Zahl der Fehler bei Item A

Zahl der Fehler bei Person 17

Abbildung 2-7: Berechnung der theoretischen und empirischen Dominanzmatrix

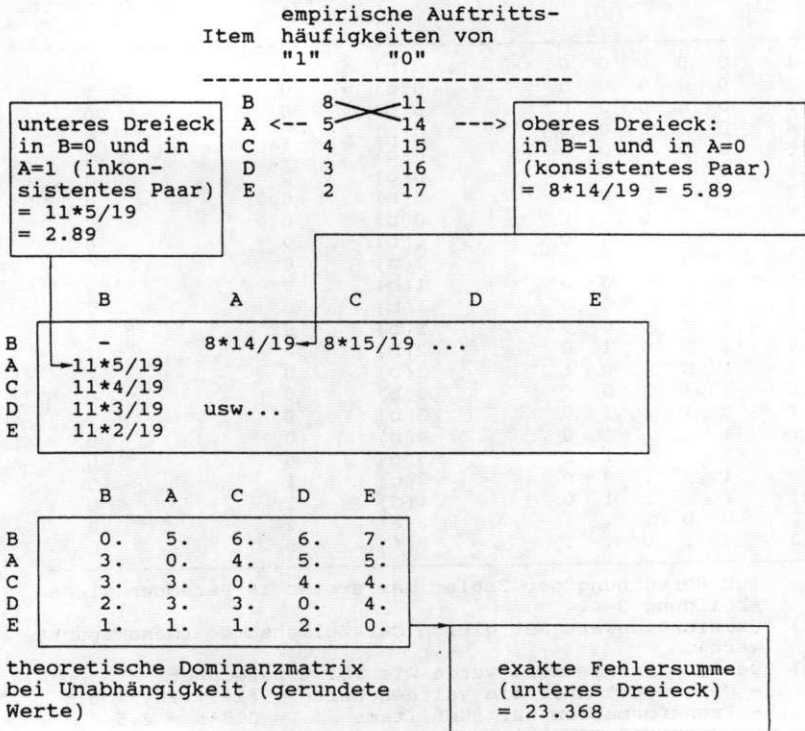
theoretische Daten:

empirische Daten:



Anmerkung: Item A ist schwieriger als Item B.

Abbildung 2-8: Dominanzarotrix bei statistischer Unabhängigkeit



Anmerkung: Item A ist schwieriger als Item B.

Abbildung 2-9: Berechnen von Quasiskalenwerten

Person	empirische Antworten					Quasi- Skalenwert	Zahl fehlender Werte	Zahl der Fehler im Antwortm. (a)
	B	A	C	D	E			
1	0	0	0	0	0	0(b)	0	0
2	0	0	0	0	0	0(b)	0	0
3	0	0	0	0	0	0(b)	0	0
4	0	0	0	0	0	0(b)	0	0
5	0	0	0	0	0	0(b)	0	0
6	0	0	0	0	0	0(b)	0	0
7	0	0	0	0	0	0(b)	0	0
8	0	0	0	0	0	0(b)	0	0
9	1	0	0	0	0	1(b)	0	0
10	1	0	0	0	0	1(b)	0	0
11	0	1	0	0	0	1(b)	0	2
12	0	0	1	0	0	1(b)	0	2
13	1	1	0	0	0	2(b)	0	0
14	1	0	0	1	0	2(b)	0	2
15	1	0	0	0	1	2(b)	0	2
16	1	1	1	0	0	3(b)	0	0
17	0	1	1	1	0	3(b)	0	2
18	1	1	1	1	0	4(b)	0	0
19	1	1	1	1	1	5(b)	0	0
20	1	1	.	0	0	3(c)	1	0
21	1	.	1	1	0	4(d)	1	0
22	0	0	1	.	.	2(e)	2	3
23	.	.	.	.	.	.(f)	5	.

(a) Zur Berechnung der Fehler der ersten 19 Personen siehe Abbildung 2-4.

(b) Quasiskalenwert ist gleich der Zeilensumme (Gesamtpunkt- wert).

(c) Der Quasiskalenwert wurde wie folgt berechnet:  
 - Mittelwert über die validen Werte =  $(1+1+0+0)/4=0.5$   
 - Transformation auf fünf Items =  $0.5*5 = 2.5$   
 - Rundung des Wertes = 3

Bei einem Gesamtpunktwert von drei sieht das theoretische Antwortmuster wie folgt aus: (1,1,1,0,0).

Vergleicht man dieses mit dem empirischen Antwortmuster, so treten bei den validen Werten keine Fehler auf.

(d) - (e) Es wurde analog zu (c) vorgegangen.

(f) Berechnung nicht möglich, da kein valider Wert.



Abbildung 2-10: Kategorienschema für eine Inhaltsanalyse zur Messung des politischen Aktivitätsprofils von politischen Gruppen/Gruppierungen in Konflikten

Name der politischen Gruppe	Teilnahme an....							
	Kundgebung		Warnstreik		Streik		Revolte	
	ja	nein	ja	nein	ja	nein	ja	nein
.....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
usw..								

### 3. Skalierungsverfahren, die als Weiterentwicklung der Guttmanskala gedacht werden können

#### 3.1 Skalierungsverfahren für dichotome Items

##### 3.1.1 Grundlogik und Probleme

Die Guttmanskala geht von einem deterministischen Mechanismus aus: Tritt eine »1« in einem Item auf, muß auch in allen leichteren Items eine »1« auftreten. Eine Reihe von Skalierungsverfahren versucht nun die Guttmanskala dahingehend zu modifizieren, daß Zufallsfehler auftreten können. Sehr einfach ausgedrückt wird dabei die empirische Datenmatrix als eine mit Zufallsfehlern überlagerte perfekte Guttmanskala gedacht:

$$\begin{array}{l} \text{empirische} \quad \quad \text{perfekte} \quad \quad \text{Zufalls-} \\ \text{Datenmatrix} = \text{Guttman-} + \text{fehler} \\ \quad \quad \quad \quad \text{skala} \end{array}$$

Durch die Einführung von Zufallsfehlern geht der deterministische Mechanismus verloren: Tritt eine »1« in einem Item auf, muß nicht notwendigerweise in allen leichteren Items eine »1« auftreten, sondern nur mit einer bestimmten Wahrscheinlichkeit. Durch diesen Modellcharakter gestaltet sich die Lösung einer Skalierungsaufgabe schwieriger, da beispielsweise - durch Zufallsfehler bedingt - das Antwortmuster (0,0,0,0,0) empirisch auch bei Personen mit einem theoretischen Antwortmuster (1,1,1,1,1) auftreten kann. Dies bedeutet, daß die Auftrittshäufigkeit des theoretischen Antwortmusters (1,1,1,1,1) der perfekten Guttmanskala nicht gleich der Auftrittshäufigkeit des empirischen Antwortmusters (1,1,1,1,1) ist. Allgemein muß somit geschätzt werden, mit welcher Wahrscheinlichkeit die theoretischen Antwortmuster einer perfekten Guttmanskala auftreten. Zum zweiten muß die Höhe der Zufallsfehler bestimmt werden. Bei fünf Items müßten somit insgesamt 35 Parameter geschätzt werden:

- Die Wahrscheinlichkeit des Auftretens der theoretischen Antwortmuster der perfekten Guttmanskala. Bei fünf Items können bei Vorliegen einer perfekten Guttmanskala sechs Antwortmuster auftreten. Da die Summe dieser Auftrittswahrscheinlichkeiten 1.0 ergeben muß, sind 5 Parameter zu schätzen.
- Die Höhe der Zufallsfehler. Da bei jedem theoretischen Antwortmuster in jedem Item ein Zufallsfehler auftreten kann, sind bei fünf

Items 30 Parameter (6 theoretische Antwortmuster mal 5 Items) zu schätzen.

Den 35 zu schätzenden Parameter stehen  $2^{**}5 = 32$  empirisch mögliche Antwortmuster gegenüber, von denen - bei gegebenem Stichprobenumfang - 31 frei variieren können. Eine Schätzung ist unmöglich, da aus 31 bekannten Größen nicht 35 unbekannte Parameter berechnet werden können. Dieses Problem wird in der Literatur als Identifikationsproblem bezeichnet. Sehr salopp ausgedrückt heißt ein Skalierungsmodell dann identifiziert, wenn die unbekannt Parameter als Funktionen der bekannten Parameter und der auf Grund des spezifizierten Modells vorhandenen empirischen Information (= theoretisch vorhandenen empirischen Information) ausgedrückt werden können (siehe dazu ausführlicher Abschnitt 4.1.3). Eine notwendige - aber keinesfalls hinreichende Bedingung der Identifikation ist, daß die Zahl der unbekannt Parameter kleiner/gleich der Zahl der theoretisch vorhandenen empirischen Information ist.

Um die notwendige Bedingung der Identifikation zu erfüllen, müssen bestimmte Restriktionen bezüglich der Parameter getroffen werden. Nach der Art der Restriktionen lassen sich drei Klassen (Typen) von Skalierungsverfahren unterscheiden:

Typ I: Skalierungsverfahren mit parameterfreien Restriktionen. Bezüglich der Höhe der Zufallsfehler wird nur angenommen, daß sie nicht die Monotonieeigenschaften der Guttmanskala zerstören. Beispiel: Mokka-Skala (MOKKEN 1971; siehe Abschnitt 3.1.2).

Typ II: Skalierungsverfahren mit Parameterrestriktionen aber ohne Spezifikation einer mathematischen Funktion. Es werden bestimmte Annahmen bezüglich der Fehlerhöhe getroffen, wie z.B.: Die Fehlerhöhe ist für alle Items und theoretischen Antwortmuster gleich; die Fehlerhöhe hängt nur von den Items ab usw...

Beispiele: Skalierung nach PROCTOR (1970; siehe Abschnitt 3.1.3), Strukturanalyse von DAYTON/MACREADY (1976).

Typ III: Skalierungsverfahren mit einer Spezifikation einer mathematischen Funktion. Hier wird die Höhe der Zufallsfehler als eine mathematische Funktion von Item- und Personenparametern definiert.

Beispiele: Raschskalierung (RASCH 1966; siehe Abschnitt 3.1.4), Analyse latenter Dimensionen (LAZARFELD/HENRY 1968 sowie DENZ 1982).

Bei der praktischen Anwendung tritt - wie bei der Guttmanskala - das Problem auf, daß die hier nur beispielhaft angeführten Skalierungsverfahren in den Standardstatistikprogrammen (BMDP, SAS und SPSS-X) nicht als Prozeduren zur Verfügung stehen. ALMO enthält dagegen ein Modul

zur Analyse latenter Dimensionen von LAZARFELD/HENRY (DENZ 1989a) sowie ein Programm zur Raschskalierung (POTUSCHAK 1989). Für SPSS-X bieten KINGMA/TAERUM (1989) ein Zusatzmodul zur Mokkenskalierung an.

In den Abschnitten 3.1.2 und 3.1.4 wird deshalb am Beispiel der Daten des Kapitels 2 gezeigt, wie in SPSS-PC+ eine Mokken- und Raschskalierung ohne Zusatzmodule durchgeführt werden kann.

### 3.1.2 Ein Skalierungsverfahren des Typs I: Mokkenskalierung

Das von MOKKEN (1971) entwickelte Skalierungsverfahren unterscheidet sich von der Guttmanskala dadurch, daß Zufallsfehler zugelassen werden. Diese dürfen aber die Monotonieeigenschaften der Guttmanskala nicht zerstören. MOKKEN unterscheidet zwei Monotoniebedingungen (zit. nach KINGMA/TAERUM 1989):

- Monotonie der Itemschwierigkeiten. Diese Forderung bedeutet, daß für jedes theoretische Antwortmuster der perfekten Guttmanskala die Auftretswahrscheinlichkeit der Ausprägung »1« eines schwereren Items kleiner/gleich der Auftretswahrscheinlichkeit der Ausprägung »1« eines leichteren Items sein muß.
- Monotonie der Auftretswahrscheinlichkeit der Ausprägung »1« (monotone homogeneity). Diese Forderung bedeutet, daß bei einem theoretischen Antwortmuster mit einem bestimmten Gesamtpunktwert die Auftretswahrscheinlichkeit der Ausprägung »1« in jedem Item größer/gleich sein muß als bei theoretischen Antwortmustern mit kleineren Gesamtpunktwerten.

Beide Bedingungen sind in der Abbildung 3-1 veranschaulicht. Sie garantieren, daß sich die Kurven der Items, die als Itemcharakteristiken bezeichnet werden, nicht überschneiden und mit dem theoretischen Gesamtpunktwert ansteigen.

Zur Überprüfung der Monotoniebedingungen werden in der Mokkenskalierung folgende Tests durchgeführt:

- Untersuchung des Homogenitätskoeffizienten und daraus abgeleiteter Teststatistiken.
- Untersuchung der sogenannten P- und PO-Matrix.

Die Logik des Homogenitätskoeffizienten wurde bereits ausführlich behandelt (siehe Abschnitt 2.6). Der Homogenitätskoeffizient überprüft, inwieweit - bei gegebener Anordnung der Items - die Zahl der empirischen Inkonsistenzen kleiner der Zahl der Inkonsistenzen bei statistischer Unabhängigkeit ist. Bei der Berechnung des Homogenitätskoeffizienten wird also von der Annahme einer für alle (theoretischen) Antwortmuster kon-

stanten Anordnung der Items (Annahme der Monotonie der Itemschwierigkeiten) ausgegangen. Diese Annahme kann empirisch durch eine Untersuchung der P- und PO-Matrix überprüft werden.

Die P-Matrix enthält als Elemente die Auftrittswahrscheinlichkeiten der Ausprägung (1,1) für alle Itempaare, die PO-Matrix die Auftrittswahrscheinlichkeiten der Ausprägung (0,0) für alle Itempaare. Ist die Bedingung der Monotonie der Itemschwierigkeiten erfüllt, muß die P-Matrix folgende Struktur besitzen, wenn die Items nach dem Schwierigkeitsgrad geordnet sind:

- Wird der Zeilenindex nach unten bewegt, nehmen im unteren Dreieck der P-Matrix die Auftrittswahrscheinlichkeiten ab.
- Wird der Spaltenindex nach rechts bewegt, nehmen im unteren Dreieck der P-Matrix die Auftrittswahrscheinlichkeiten ab.

Die PO-Matrix muß ein umgekehrtes Verhalten zeigen. Der Prüfungsvorgang soll anhand der Beispieldaten aus Kapitel 2 verdeutlicht werden. Die P- und PO-Matrix sowie die Grundlogik der Prüfung sind in der Abbildung 3-2 dargestellt.

Wie der Abbildung 3-2 zu entnehmen, weicht in der P-Matrix in einem Fall das empirische Muster von der erwarteten Struktur ab (Zeilenvergleich von D mit E im Item A). Ob diese eine Abweichung signifikant ist, kann durch nichtparametrische Tests geprüft werden (KINGMA/TAERUM 1989).

Technisch lassen sich die P- und PO-Matrix durch Kreuztabellierungen aller Items untereinander gewinnen (Abbildung 3-4). Die entsprechenden SPSS-PC+ (Version 2.0) Anweisungen sind:

```
1      data list file = 'pol.dat' free/a,b,c,d,e.
2      missing values a to e (-1).
3      count ig = a to e(missing).
4      select if (ig eq 0).
5      crosstabs tables = a,c,d,e by b/
6                          c,d,e by a/
7                          d,e by c/
8                          d by e/
9      options = 5,14,15,16,17.
```

Die Zeilen bedeuten:

- 1 Definition der Daten (siehe Programm 1 des Abschnitts 2.10).
- 2 Die Ausprägung -1 wird als Ausprägung für fehlende Werte definiert.
- 3 Die Zahl der fehlenden Werte wird für jeden Datensatz (Person) berechnet und in die Variable ig geschrieben.
- 4 Diese Anweisung bewirkt, daß nur Datensätze ohne fehlende

Werte (ig = 0) in die Analyse einbezogen werden.

5 - 9 Kreuztabellierung aller Items.

Aus den Vierfeldertafeln aller Items können zunächst die Elemente der P- und PO-Matrix herausgeschrieben werden (siehe Abbildung 3-3). Darüber hinaus kann für jedes Itempaar der Homogenitätskoeffizient H und die von MOKKEN entwickelte Teststatistik DELTA berechnet werden (siehe Abbildung 3-3). Dazu wird in der Vierfeldertafel die inkonsistente Zelle betrachtet, also jene Zelle, wo im schwierigeren Item eine »1« auftritt und im leichteren Item ein »0«. Der Homogenitätsindex eines Items ist analog dem Homogenitätsindex der Gesamtskala definiert mit (siehe Abschnitt 2.6):

$$H = \frac{\text{Inkonsistenzen bei statistischer Unabhängigkeit} - \text{empirische Inkonsistenzen}}{\text{Inkonsistenzen bei statistischer Unabhängigkeit}}$$

$$= 1 - \frac{\text{empirische Inkonsistenz}}{\text{Inkonsistenzen bei statistischer Unabhängigkeit}}$$

In dem Beispiel der Abbildung 3-3 beträgt die Zahl der empirischen Inkonsistenz 110 (= erste Zeile der Zelle), die bei statistischer Unabhängigkeit erwartete Zahl ist gleich 256.9. H ist folglich

$$H = \frac{256.9 - 110}{256.9} = 1 - \frac{110}{256.9} = 0.57$$

Mit Hilfe der Teststatistik DELTA kann geprüft werden, ob die Zahl der empirischen Inkonsistenzen signifikant kleiner der Zahl der bei statistischer Unabhängigkeit erwarteten Zahl von Inkonsistenzen ist. Ihre Berechnung und Anwendung wird im nachfolgenden SPSS-PC + Programm dargestellt.

Für die Gesamtskala ergibt sich bei diesem Vorgehen (Berechnung aus den Kreuztabellen) ein beträchtlicher Rechenaufwand, da die Koeffizienten der Gesamtskala nicht gewichtete Summen der Koeffizienten der Itempaare sind. Es kann deshalb ein zum Programm 2 der Guttmankala (siehe Abschnitt 2.10) leicht modifiziertes Programm geschrieben werden, in dem anstelle der Items die Itempaare untersucht werden. Das nachfolgende Programm berechnet für jedes Itempaar sowie für die Gesamtskala den Homogenitätskoeffizienten und die Teststatistik DELTA:

```

1      data list file = 'pol.dat' free/a,b,c,d,e.
2      missing values a to e (-1).
3      count ig = a to e(missing).
4      select if (ig eq 0).
5      compute ba = 0.
6      compute bc = 0.
7      compute bd = 0.
    
```

```

8      compute be = 0.
9      compute ac = 0.
10     compute ad = 0.
11     compute ae = 0.
12     compute cd = 0.
13     compute ce = 0.
14     compute de = 0.
15     if (b eq 0 and a ge 1) ba = 1.
16     if (b eq 0 and c ge 1) be = 1.
17     if (b eq 0 and d ge 1) bd = 1.
18     if (b eq 0 and e ge 1) be = 1.
19     if (a eq 0 and c ge 1) ac = 1.
20     if (a eq 0 and d ge 1) ad = 1.
21     if (a eq 0 and e ge 1) ae = 1.
22     if (c eq 0 and d ge 1) cd = 1.
23     if (c eq 0 and e ge 1) ce = 1.
24     if (d eq 0 and e ge 1) de = 1.
25     compute gg = 1.
26     aggregate outfile = *
27         /break = gg
28         /pa,pb,pc,pd,pe = mean(a,b,c,d,e)
29         /fba,fb,fbd,fbce,fac,fad,fae,fce,fde =
30             sum(ba,bc,bd,be,ac,ad,ae,cd,ce,de)
31         /nn = sum(gg).
32     compute zba = pa*(1-pb)*nn.
33     compute zbc = pc*(1-pb)*nn.
34     compute zbd = pd*(1-pb)*nn.
35     compute zbe = pe*(1-pb)*nn.
36     compute zac = pc*(1-pa)*nn.
37     compute zad = pd*(1-pa)*nn.
38     compute zae = pe*(1-pa)*nn.
39     compute zed = pd*(1-pe)*nn.
40     compute zce = pe*(1-pe)*nn.
41     compute zde = pe*(1-pd)*nn.
42     compute vbb = pb*(1-pb).
43     compute vaa = pa*(1-pa).
44     compute vec = pc*(1-pc).
45     compute vdd = pd*(1-pd).
46     compute vee = pe*(1-pe).
47     compute sba = vbb*vaa* 100.
48     compute sbc = vbb*vec* 100.
49     compute sbd = vbb*vdd* 100.
50     compute sbe = vbb*vee* 100.

```

```

51     compute sac = vaa*vcc* 100.
52     compute sad  = vaa*vdd*100.
53     compute sae  = vaa*vee*100.
54     compute scd  = vcc*vdd*100.
55     compute see  = vcc*vee*100.
56     compute sde  = vdd*vee*100.
57     compute item1 = 'B-A'.
58     compute item2 = 'B-C'.
59     compute item3 = 'B-D'.
60     compute item4 = 'B-B'.
61     compute item5 = 'A-C'.
62     compute item6 = 'A-D'.
63     compute item7 = 'A-E'.
64 compute item8 = 'C-D'.
65 compute item9 = 'C-E'.
66     compute item10 = 'D-E'.
67     set results = 'mokken.dat'.
68     write var = item1,nn,fba,zba,sba,
69             item2,nn,fbz,zbc,sbc,
70             item3,nn,fbd,zbd,sbd,
71             item4,nn,fbe,zbe,sbe,
72             item5,nn,fac,zac,sac,
73             item6,nn,fad,zad,sad,
74             item7,nn,fae,zae,sae,
75             item8,nn,fcd,zcd,scd,
76             item9,nn,fce,zce,sce,
77             item10,nn,fde,zde,sde.
78     data list file = 'mokken.dat' free/
79     itemp (A) nn,empink,zufink,varink.
80     compute hh = 0.
81     if (zufink gt 0) hh = 1.-empink/zufink.
82     compute sa_delta = sqrt(varink/100).
83     compute delta = (sqrt(nn-1)*((zufink-empink)/nn))/
84     (sa_delta).
85     report format = list
86             /variables = itemp,empink,zufink,sa_delta,hh,delta
87             /break = (nobreak)
88             /summary = sum.
89     compute gg = 1.
90     aggregate outfile = *
91             /break = gg
92             /emp.zuf ,var = sum(empink,zufink,varink)
93             /nnn = mean(nn).

```



```
94 compute hh = 0.
95 compute itemp = 'Gesamt'.
96 if (zuf gt 0) hh = 1.-emp/zuf.
97 compute sa_delta = sqrt(var/100).
98 compute delta = (sqrt(nnn-1)) * ((zuf-emp)/nnn)/sa_delta.
99 report format = list
100 /variables = itemp,emp,zuf,sa_delta,hh,delta
101 /break = (nobreak)
102 /summary = sum.
```

Die Programmzeilen bedeuten:

- 1 Definition der Daten (siehe Programm 1 des Abschnitts 2.10).
- 2 Die Ausprägung -1 wird als Ausprägung für fehlende Werte definiert.
- 3 Die Zahl der fehlenden Werte wird für jeden Datensatz (Person) berechnet und in die Variable ig geschrieben.
- 4 Diese Anweisung bewirkt, daß nur Datensätze ohne fehlende Werte (ig = 0) in die Analyse einbezogen werden.
- 5-14 Durch diese Anweisungen werden die Variablen für alle Itempaare definiert, ba steht als für das Itempaar (B,A), bc für das Itempaar (B,C), usw..
- 15-24 Für jede Person wird in jedem Itempaar geprüft, ob eine Inkonsistenz vorliegt. Ist dies der Fall erhält die entsprechende Variable den Wert 1. Tritt also z.B. im Itempaar (B,A) eine Inkonsistenz auf, wird ba gleich 1 gesetzt.
- 25 Durch diese Anweisung wird eine fiktive Variable für die anschließende Aggregation gebildet. Sie besitzt für alle Personen den Wert 1.
- 26-31 Durch diese Anweisungen werden die Variablen über alle in die Analyse einbezogenen Personen aggregiert, da die Aggregatsvariable gg (Variable nach break =) für alle Personen die Ausprägung 1 besitzt. Durch die Definition outfile = \* bilden die aggregierten Variablen die neue SPSS-PC + -interne Datenmatrix:
  - pa = Anteilswert der Ausprägung »1« im Item A
  - pb = Anteilswert der Ausprägung »1« im Item B
  - pc = Anteilswert der Ausprägung »1« im Item C
  - pd = Anteilswert der Ausprägung »1« im Item D
  - pe = Anteilswert der Ausprägung »1« im Item E
  - fba = Zahl der Inkonsistenzen im Itempaar(B,A)
  - fbc = Zahl der Inkonsistenzen im Itempaar(B,C)
  - .
  - .
  - nn = Zahl der in die Skalierung einbezogenen Personen.
- 32-41 Durch diese Anweisungen wird für jedes Itempaar die Zahl der

Inkonsistenzen bei statistischer Unabhängigkeit berechnet. Für das Itempaar (B,A) (B leichter als A) ist die Zahl der Inkonsistenzen zba bei statistischer Unabhängigkeit gleich dem Erwartungswert der Zelle (0,1) =

$$zba = (1-pb)-pa \gg nn$$

- 42 - 46 Durch diese Anweisungen wird für jedes Item die Varianz nach der Formel  $p*(1-p)$  berechnet, wobei p der Anteilswert der Ausprägung »1« ist. Für Item B beispielsweise ist die Varianz vbb =

$$vbb = pb*(1-pb)$$

- 47 - 56 Durch diese Anweisungen wird für jedes Itempaar die Varianz des Anteilswertes der Inkonsistenzen bei statistischer Unabhängigkeit berechnet. Für das Itempaar (B,A) ist die Varianz sba definiert als

$$sba = vbb*vaa$$

mit  $vbb = pb*(1-pb)$  = Varianz von B

$$vaa = pa*(1-pa)$$
 = Varianz von A

Die anschließende Multiplikation mit 100 dient der Erhöhung der Rechengenauigkeit.

- 57 - 66 Durch diese Anweisungen werden die Namen der Itempaare für die Ausgabe bei report festgelegt.

- 67 Der Name der Datei für die nachfolgende write-Anweisung wird definiert.

- 68 - 77 Für jedes Itempaar werden folgende Werte in die Datei \*moken.dat' geschrieben:

- der Name
- die Zahl der in die Skalierung einbezogenen Personen
- die Zahl der empirischen Inkonsistenzen
- die Zahl der Inkonsistenzen bei statistischer Unabhängigkeit
- die Varianz des Anteils der Inkonsistenzen bei statistischer Unabhängigkeit

- 78 - 79 Die für jedes Itempaar gespeicherten Informationen werden eingelesen:

itemp der Name

nn die Zahl der in die Skalierung einbezogenen Personen

empink die Zahl der empirischen Inkonsistenzen

zufink die Zahl der Inkonsistenzen bei statistischer Unabhängigkeit

varink die Varianz des Anteils der Inkonsistenzen bei statistischer Unabhängigkeit

80 - 84 Für jedes Itempaar wird der Homogenitätskoeffizient hh und die Teststatistik delta berechnet.

$$hh = \frac{zufink - empink}{zufink} = 1 - \frac{empink}{zufink}$$

$$sa\_delta = \sqrt{(\text{varink}/100)}$$

Standardabweichung des Anteilswerts der Inkonsistenzen bei stat. Unabhängigkeit

$$delta = \sqrt{(nn-1) \cdot \left(\frac{zufink - empink}{nn}\right) \cdot \left(\frac{1}{sa\_delta}\right)}$$

85 - 88 Ausgabe der Koeffizienten für jedes Itempaar.

89 - 102 Diese Anweisungen dienen der Berechnung des Homogenitätskoeffizienten hh und der Teststatistik delta für die Gesamtskala. Dazu werden die Itempaare zunächst aggregiert. Die neue SPSS-PC-f Arbeitsdatei besteht dann aus einem Datensatz mit folgenden Variablen:

itemp der Name

nnn die Zahl der in die Skalierung einbezogenen Personen

emp die Zahl der empirischen Inkonsistenzen in allen Itempaaren

zuf die Zahl der Inkonsistenzen bei statistischer Unabhängigkeit in allen Itempaaren

var die Varianz des Anteils der Inkonsistenzen bei statistischer Unabhängigkeit in allen Itempaaren

Anschließend werden dieselben Berechnungsformeln wie für die Itempaare angewendet. Die Ausgabe erfolgt über die Prozedur report.

Das Programm berechnet folgende Ergebnisse:

ITEMP	EMPINK	ZUFINK	SA_DELTA	HH	DELTA
B-A	110.00	256.88	.21	.57	15.46
B-C	89.00	204.75	.20	.57	13.23
B-D	53.00	168.75	.18	.69	14.28
B-E	61.00	127.90	.16	.52	9.27
A-C	201.00	288.93	.17	.30	11.71
A-D	167.00	238.12	.16	.30	10.21
A-E	85.00	180.49	.14	.53	15.41
C-D	153.00	253.52	.14	.40	15.67
C-E	122.00	192.16	.13	.37	12.28
D-E	157.00	200.22	.12	.22	8.18
Gesamt	1198.00	2111.72	.52	.43	39.71

Aus den Ergebnissen ist ersichtlich, daß der Homogenitätskoeffizient für das Itempaar (D,E) sehr klein ist. Die Größe DELTA besitzt bei großen  $n$  annähernd eine Standardnormalverteilung. Es kann folglich - wie bereits erwähnt - geprüft werden, ob DELTA signifikant größer 0 ist. Ist dies der Fall, kann die HO-Hypothese der statistischen Unabhängigkeit verworfen werden: Die Zahl der empirischen Inkonsistenzen ist signifikant kleiner als bei einem zufälligen Antwortverhalten (statistische Unabhängigkeit). Bei einem Fehlerniveau von 5% beträgt der kritische Schwellenwert 1.96. Die Testgröße der Gesamtskala liegt deutlich über diesem Schwellenwert. Wird jedes Item getrennt getestet, empfiehlt sich anstelle des Fehlerniveaus von 5% das Fehlerniveau mit  $0.05 / (m(m-1)/2)$  mit  $m =$  Zahl der Items festzulegen, da der Test mehrfach angewandt wird. In unserem Beispiel ist das korrigierte Fehlerniveau  $0.05/10 = 0.005$ . Der Schwellenwert beträgt 2.58. Alle Testgrößen liegen über diesem Schwellenwert. Zu beachten ist, daß bei dem Test nicht geprüft wird, ob  $H$  größer einem bestimmten Schwellenwert (z.B. 0.5 siehe Abschnitt 2.6) ist. Dies bedeutet, daß ein signifikanter DELTA-Wert nur eine notwendige Bedingung für das Vorliegen einer Mokkenskala ist. Zum genaueren Vorgehen siehe MOKKEN (1971) oder KINGMA/TAERUM (1984).

Neben der Prüfung einer theoretisch angenommenen Skala eignet sich die Mokkenskalisierung insbesondere zur Konstruktion einer oder mehrerer Skalen aus einer Menge von dichotomen Items (explorative Skalierung). Die Grundlogik der Skalenkonstruktion besteht darin, daß aus der Menge der Items ein »bestes« Itempaar ausgewählt wird. Diesem werden schrittweise neue Items hinzugefügt, solange bestimmte Kriterien (z.B. signifikanter Homogenitätskoeffizient) erfüllt sind. Für Items, die nicht der so gebildeten Skala zugeordnet werden konnten, wird versucht, eine neue Skala zu bilden, usw.. (siehe z.B. KINGMA/TAERUM 1989). Steht kein Programm zur Mokkenskalisierung zur Verfügung, kann aushilfsweise die Verschmelzung der Items zu einer oder mehreren Skalen mit Hilfe einer hierarchischen Clusteranalyse durchgeführt werden. Dabei wird die Matrix der Homogenitätskoeffizienten als Ähnlichkeitsmatrix definiert, da ein größerer Homogenitätskoeffizient eine größere »Ähnlichkeit« hinsichtlich der Messung des theoretischen Begriffs bedeutet. Zu beachten ist aber, daß dieses Verfahren nur eine sehr »grobe« Annäherung darstellt und die Ergebnisse mit Mokkenskalisierung nicht übereinstimmen müssen.

### 3.13 Ein Skalierungsverfahren des Typs II: Skalierung nach PROCTOR

Bei der Mokkenskalisierung werden nur bestimmte parameterfreie Annahmen bezüglich der Zufallsfehler getroffen. Gegenüber der Guttmanskala müssen keine weiteren Modellparameter geschätzt werden. Ein zweiter Typ von Skalierungsverfahren ist dadurch gekennzeichnet, daß durch

Parameterrestriktionen die notwendige Identifikationsbedingung (mehr theoretisch vorhandene empirische Information als unbekannte Parameter) erreicht wird. Ein sehr einfaches und restriktives Modell dieses Typ hat PROCTOR (1970) spezifiziert: Die Höhe der Zufallsfehler (= Fehler-rate) ist für alle Items und theoretischen Antwortmuster der perfekten Guttmanskala konstant. Unter dieser Annahme sind folgende Parameter zu schätzen.

- Die Wahrscheinlichkeiten  $\pi(v)$  für das Auftreten des theoretischen Antwortmusters  $v$  der perfekten Guttmanskala. Da jedes theoretische Antwortmuster eindeutig (erschöpfend) durch den Gesamtpunktwert bestimmt ist, kann für  $\pi(v)$  geschrieben werden:  $\pi(x=v)$  mit  $x$  = theoretischer Gesamtpunktwert.

$$\begin{array}{ccccc} \text{Beispiel: } & \pi(0,0,0) & = & \pi(x=0) & = & \pi(0) \\ & \downarrow & & \downarrow & & \downarrow \\ & \text{theoretisches} & & \text{Gesamt-} & & \text{Kurzschreibweise} \\ & \text{Antwortmuster} & & \text{punktwert} & & \end{array}$$

- Die Fehlerrate  $\alpha$ .

Da zufällige Fehler angenommen werden, ist das Auftreten des Fehlers in einem Item vom Auftreten eines Fehlers in einem anderen Item statistisch unabhängig. Wegen der statistischen Unabhängigkeit kann die bedingte Wahrscheinlichkeit  $p(g/v)$  für das Auftreten des empirischen Antwortmusters  $g$  (z.B. 0,0,0), gegeben dem theoretischen Antwortmuster  $v$  (z.B. 1,0,0), leicht berechnet werden. Betrachten wir dazu ein Beispiel: Drei Items A, B und C bilden theoretisch eine perfekte Guttmanskala, wobei A leichter als B und B leichter als C ist. Durch Zufallsfehler bedingt, können nun für Personen mit dem theoretischen Antwortmuster (0,0,0) empirisch alle acht möglichen Antwortmuster (0,0,0), (1,0,0), (0,1,0) usw. auftreten (siehe Abbildung 3-4).

Die bedingte Wahrscheinlichkeit für das Auftreten des empirischen Antwortmusters (1,0,0) gegeben dem theoretischen Antwortmuster (0,0,0) (dem Gesamtpunktwert 0) ist beispielsweise:

$$p(1,0,0/0,0,0) - p(1,0,0/x=0) - p(\text{Fehler in A}) \cdot p(\text{kein Fehler in B}) \cdot p(\text{kein Fehler in C})$$

wobei  $p(\dots)$  für Wahrscheinlichkeit steht.

Da

$$\begin{aligned} p(\text{kein Fehler in B}) &= 1 - p(\text{Fehler in B}) \\ p(\text{kein Fehler in C}) &= 1 - p(\text{Fehler in C}) \end{aligned}$$

ist, ergibt sich bei Einsetzen der Fehlerrate  $\alpha$  ( $= p(\text{Fehler in A}) = p(\text{Fehler in B}) - p(\text{Fehler in C})$ ):

$$p(1,0,0/0,0,0) - \alpha \cdot (1-\alpha) \cdot (1-\alpha)$$

Für  $\alpha = 0.2$  beispielsweise, ist  $p(1,0,0/0,0,0) = 0.2 \cdot (1-0.2) \cdot (1-0.2) = 0.128$ . Die bedingten Wahrscheinlichkeiten  $p(g/v)$  für alle empirischen und theoretischen Antwortmuster bei einer Fehlerrate von  $\alpha = 0.2$  sind in der Abbildung 3-4 dargestellt.

Aus der Abbildung 3-4 ist folgendes ersichtlich:

- Das bei einer perfekten Guttmankala erwartete Antwortmuster besitzt die größte bedingte Auftrittswahrscheinlichkeit. (Dies gilt, solange  $\alpha < 0.5$  ist.)
- Da die Fehlerrate für alle Items und für alle theoretischen Antwortmuster konstant ist, hängen die bedingten Auftrittswahrscheinlichkeiten  $p(g/v)$  nur von der Zahl der Fehler ab.

Um in der Abbildung 3-4 die Besetzungszahlen  $n(g,v)$  der einzelnen Zellen zu erhalten, muß die bedingte Wahrscheinlichkeit  $p(g/v)$  mit der Auftrittswahrscheinlichkeit  $\pi(v)$  des theoretischen Antwortmuster  $v$  und der Stichprobengröße ( $= n$ ) multipliziert werden:

$$n(g,v) = p(g/v) \cdot n(v) \cdot n$$

Für

$$\pi(0,0,0) = \pi(x=0) = \pi(0) = .30$$

$$\pi(1,0,0) = \pi(x=1) = \pi(1) = .40$$

$$\pi(1,1,0) = \pi(x=2) = \pi(2) = .20$$

$$\pi(1,1,1) = \pi(x=3) = \pi(3) = .10$$

und

$$n = 10000$$

ergeben sich die in der Abbildung 3-4 angegebenen Werte für das empirische Antwortmuster (0,0,0) und das theoretische Antwortmuster (0,0,0) also beispielsweise der Wert  $0.512 \cdot 0.30 \cdot 10000 = 1536$ , da  $p(0,0,0/0) = 0.512$  und  $n(0) = 0.30$  und  $n = 10000$  angenommen wurde. Wäre die Tabelle der  $n(g,v)$  der Abbildung 3-4 bekannt, könnten die unbekannt Parameter leicht geschätzt werden. Die Schätzwerte (Maximum Likelihood) wären:

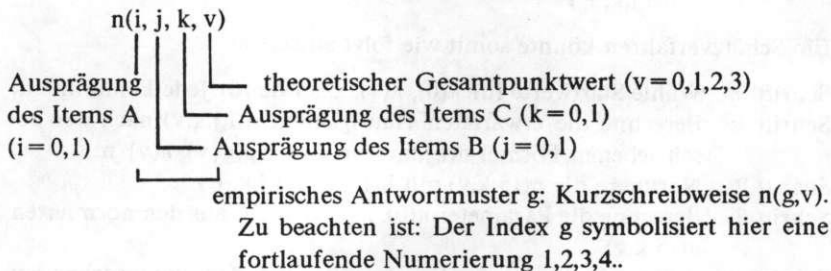
$$\pi(0) = \frac{n(+,+,+,0)}{n(+,+,+,+)} = \frac{3000}{10000} = 0.300$$

$$\pi(1) = \frac{n(+,+,+,1)}{n(+,+,+,+)} = \frac{4000}{10000} = 0.400$$

$$\pi(2) = \frac{n(+,+,+,2)}{n(+,+,+,+)} = \frac{2000}{10000} = 0.200$$

$$\pi(3) = \frac{n(+, +, +, 3)}{n(+, +, +, +)} = \frac{1000}{10000} = 0.100$$

Bei den Formeln wurde folgende Notation verwendet:



Wird über ein Item summiert, wird dies durch ein »+« symbolisiert.  $n(+, j, k, v)$  bedeutet also, daß über die Ausprägungen des Items A summiert wird usw.

Zur Schätzung der Fehlerrate muß zunächst die Zahl der Fehler in den Items berechnet werden:

$$f(A) - 384 + 96 + 96 + 24 + 512 + \dots - 2000$$

$$f(B) = 384 + 96 + \dots - 2000$$

$$f(C) = 384 + 96 + \dots - 2000$$

Die Fehlerrate kann nun mit

$$\alpha = \frac{f(A) + f(B) + f(C)}{n(A) + n(B) + n(C)} = \frac{2000 + 2000 + 2000}{10000 + 10000 + 10000} = 0.2$$

mit  $n(A) \ll n(B) = n(C) = n(+, +, +, +) = n$  geschätzt werden. Damit sind alle Modellparameter geschätzt.

Bei empirischen Daten tritt das Problem auf, daß die Tabelle der Abbildung 3-4 nicht bekannt ist. Gegeben ist nur eine Tabelle mit den empirischen Auftretshäufigkeiten der Antwortmuster (0,0,0), (1,0,0) usw. Diese sollen im folgenden mit  $h(i, j, k)$  bezeichnet werden.

Um die eben beschriebene einfache Schätzmethode verwenden zu können, kann man sich überlegen, ob die Tabelle der Abbildung 3-4 nicht aus der vorhandenen empirischen Information, den  $h(i, j, k)$ , und den Modellparametern  $\pi(0)$ ,  $\pi(1)$ , ... und  $\alpha$  geschätzt werden kann. Die aufgrund der Modellparameter erwarteten Häufigkeiten sind:  $n(i, j, k, v)$  und  $n(i, j, k) = n(i, j, k, +) = \sum n(i, j, k, v)$  wobei über  $v$  summiert wird.

Da durch das Modell die empirischen Häufigkeiten  $h(i, j, k)$  möglichst gut reproduziert werden sollen, ist es naheliegend, die  $n(i, j, k, v)$  so zu normieren daß

$$n(i, j, k, +) \approx h(i, j, k)$$

$$n(i, j, k, +) \approx h(i, j, k)$$

gilt. Dies wird dadurch erreicht, daß die  $n(i,j,k,v)$  wie folgt normiert werden:

$$n(i,j,k,v) = \frac{h(i,j,k)}{n(i,j,k,+)} \cdot n(i,j,k,v)$$

Ein Schätzverfahren könnte somit wie folgt aussehen:

Schritt 0: Wähle Startwerte für  $\pi(0)$ ,  $\pi(1)$ ,... sowie für jede Fehlerrate  $a$ .

Schritt 1: Berechne die erwarteten Häufigkeiten  $n(i,j,k,v)$  nach der beschriebenen Formel  $n(i,j,k,v) = r(g,v) = p(g|v) \cdot \pi(v) \cdot n$

Schritt 2: Normiere die  $n(i,j,k,v)$  mit  $h(i,j,k)/n(i,j,k,+)$ .

Schritt 3: Berechne die Parameter  $\pi(0)$ ,  $\pi(1)$ ,... und  $a$  aus den normierten  $n(i,j,k,v)$ .

Schritt 4: Prüfe, ob die neuen Schätzwerte von den vorausgehenden Schätzwerten abweichen. Wenn ja, gehe zu Schritt 1. Wenn nein, beende das Verfahren.

Das eben beschriebene Verfahren wird als EM-Algorithmus bezeichnet (VAN DE POL/DE LEEUW, 1986). Zu den Eigenschaften des EM-Algorithmus siehe DEMPSTER (1977). Die Schritte 1 und 2 werden als E-Schritte (»expectation«) bezeichnet, der Schritt 3 als M-Schritt (»maximum likelihood«). Im Prinzip ist es wiederum möglich, den EM-Algorithmus in SPSS-X bzw. SPSS-PC+ zu programmieren. Dazu werden zwei Programme geschrieben. Ein Programm für den E-Schritt und ein Programm für den M-Schritt. Auf eine ausführliche Beschreibung der beiden Programme soll aber verzichtet werden, da in Kapitel 4 das Programm PANMARK (VAN DE POL/LANGEWEINE/DE JONG, 1989) behandelt wird, mit dem sich eine Reihe von Skalierungsaufgaben lösen lassen.

Für die Beispieldaten des vorausgehenden Kapitels berechnet PANMARK folgende Schätzwerte:

$$\begin{aligned} \pi(0,0,0,0,0) &= \pi(0) = 0.552 (\pm 0.014) \\ \pi(1,0,0,0,0) &= \pi(1) = 0.238 (\pm 0.014) \\ \pi(1,1,0,0,0) &= \pi(2) = 0.091 (\pm 0.010) \\ \pi(1,1,1,0,0) &= \pi(3) = 0.029 (\pm 0.007) \\ \pi(1,1,1,1,0) &= \pi(4) = 0.042 (\pm 0.007) \\ \pi(1,1,1,1,1) &= \pi(5) = 0.048 (\pm 0.006) \\ \alpha &= 0.097 (\pm 0.004) \end{aligned}$$

In der Untersuchung besitzen ungefähr 55% der Personen das theoretische Antwortmuster (0,0,0,0,0) (Standardabweichung  $\pm 1.4\%$ ), 24% das theoretische Antwortmuster (1,0,0,0,0) usw. Die Fehlerrate beträgt 9,7%.

Die Reproduktion der empirischen Häufigkeiten  $h(i,j,k,l,m)$  ( $i$  = Ausprägung des Item A,  $j$  — Ausprägung des Item B,  $m$  = Ausprägung des Item E;  $i,j,k,l,m = 0,1$ ) durch das angenommene Skalierungsmodell kann durch eine Chiquadrat-verteilte Testgröße CHI



$$CHI = \sum \frac{(h(i,j,k,l,m) - n(i,j,k,l,m,+))^2}{n(i,j,k,l,m,+)}$$

mit

$h(i,j,k,l,m)$  = empirische Häufigkeit des Antwortmusters (i,j,k,l,m)

$n(i,j,k,l,m,+)$  = erwartete Häufigkeit des Antwortmusters (i,j,k,l,m)

geprüft werden. Diese Testgröße besitzt bei großem Stichprobenumfang eine Chiquadrat-Verteilung mit

$$df = 2 \cdot r - 2$$

Freiheitsgraden ( $r$  — Zahl der Items). Für unsere Daten ist

$$CHI = 458.7$$

$$df = 2 \cdot 5 - 2 = 32 - 2 = 25$$

Der kritische Tabellenwert bei einem Signifikanzniveau von 95% beträgt 37.7. Die Testgröße liegt deutlich über diesem Wert, die Hypothese, daß die Daten durch das Skalierungsmodell reproduziert werden, muß verworfen werden. Untersucht man die Abweichungen im Detail (Fehleranalyse), so zeigt sich, daß die größte Abweichung beim Antwortmuster (0,1,1,1,0) auftritt, also bei Personen, die in Bürgerinitiativen mitwirken (Item B), Kontakt mit Politikern suchen (Item C) und ein Volksbegehren unterstützen (Item B).

### 3.1.4 Ein Skalierungsverfahren des Typs III: Raschskalierung

Kennzeichen der Modelle des Typs III ist die Vorstellung, daß durch Zufallsfehler Itemcharakteristiken entstehen, die sich durch eine mathematische Funktion beschreiben lassen. Dabei wird angenommen, daß die Meßdimension eine kontinuierliche latente Variable ist, auf der jede Person und jedes Item einen Skalenwert besitzen. Bei den anderen bisher besprochenen Verfahren (Guttman-, Mokkenskalisierung und Skalierung nach PROCTOR) kann zwar ebenfalls von diesen Vorstellungen ausgegangen werden, sie gehen aber nicht in die Entwicklung des formalen Modells ein. Ein Skalierungsverfahren, das diesem Typ angehört, wurde von RASCH (1966) entwickelt. Bei der Entwicklung wurde von der Vorstellung ausgegangen, daß der Gesamtpunktwert der empirischen Antwortmuster eine erschöpfende Statistik für die Skalenwerte der Personen ist. Jedes empirische Antwortmuster soll also eindeutig einem theoretischen Antwortmuster (Skalenwert) zugeordnet sein (ANDRICH 1985). Um dies zu erreichen, wird beim Raschmodell angenommen, daß Personen mit einem theoretischen Gesamtpunktwert gleich  $v$  (= eine ganze Zahl zwischen 0 und  $r$ ;  $r$  = Zahl der Items) genau  $v$  Items mit »1« beantworten (siehe Abbildung 3-5 und ANDRICH 1985).

Bezüglich der bedingten Wahrscheinlichkeit  $p(g/v)$  für das Auftreten der empirischen Antwortmuster  $g$  wird angenommen, daß sie nur von den Skalenwerten  $\delta$  (Schwierigkeitsgrade) der Items abhängen. In der Abbildung 3-5 wurde für das Item A ein Skalenwert  $\delta(A)$  von -1.0 angenommen, für das Item B ein Skalenwert  $\delta(B)$  von 0.0 und für das Item C ein Skalenwert  $\delta(C)$  von 1.5. Item A ist also leichter als B und Item B leichter als C. Gegeben einem empirischen Gesamtpunktwert (= theoretischer Gesamtpunktwert) von 1.0 betragen die bedingten Wahrscheinlichkeiten  $p(g/v)$

$$p(1,0/1) = \exp(-\delta(A)) = \exp(+1) = 2.718$$

$$p(0,1/1) = \exp(-\delta(B)) = \exp(0) = 1.000$$

$$p(0,0,1/1) = \exp(-\delta(C)) = \exp(-1.5) = 0.223.$$

Die Schreibweise  $\exp(x)$  steht dabei für die Funktion  $e^x$ . Die  $e$ -Funktion gewährleistet, daß das leichtere Item (in unserem Beispiel das Item A) den größten Auftrittswert besitzt, das zweit leichteste (in unserem Beispiel das Item B) den zweitgrößten Auftrittswert usw.

Da die Summe der bedingten Wahrscheinlichkeiten gleich 1.0 sein muß, ist eine Normierung mit

$$\tau(1) = \tau(x=1) = \exp(+1) + \exp(0) + \exp(-1.5) = 3.941$$

erforderlich. Die entsprechenden Werte enthält die Abbildung 3-5. Aus ihr ist ersichtlich, daß das bei einer perfekten Guttmanskala erwartete Antwortmuster die größte bedingte Auftrittswahrscheinlichkeit besitzt.

Aus den Modellgleichungen für die bedingten Wahrscheinlichkeiten können unmittelbar Schätzgleichungen für die Skalenwerte der Items abgeleitet werden. Wird die ML-Schätzmethode (Maximum-Likelihood) verwendet, werden die entsprechenden Schätzgleichungen als bedingte ML-Schätzgleichungen bezeichnet. Zur bedingten ML-Schätzung informiert man sich beispielsweise bei ROST (1988:198-253). Der Vorteil der bedingten ML-Schätzung liegt u.a. darin, daß die Skalenwerte der Items und die Güte der Modellanpassung mit Hilfe eines loglinearen Modells berechnet werden können (siehe dazu z.B. CLOGG 1988). Dazu können die in den Standardstatistiksoftwarepaketen enthaltenen loglinearen Modelle verwendet werden.

Ein Nachteil der bedingten ML-Schätzung ist dadurch gegeben, daß die Skalenwerte der Personen erst in einem zweiten Schritt geschätzt werden können. Dieses Problem läßt sich umgehen, indem die bedingte Wahrscheinlichkeit für das Auftreten der Ausprägung »1« in einem Item bei einem Gesamtpunktwert von  $v$  untersucht wird. Dabei wird angenommen, daß die noch nicht bekannten Skalenwerte  $\theta(v)$  sowie die Schwierigkeitsgrade  $\delta(i)$  gegeben sind. Diese Wahrscheinlichkeit ist gegeben durch

$$p(x(i,v) = 1/\delta(i), \vartheta(v)) = \frac{\exp(\vartheta(v) - \delta(i))}{1 + \exp(\vartheta(v) - \delta(i))}$$

mit

$x(i,v)$  = Ausprägung des theoretischen Antwortmusters  $v$  im Item  $i$ .

$\delta(i)$  = Skalenwert des Items  $i$

$\vartheta(v)$  = Skalenwert des theoretischen Antwortmusters  $v$

$\exp(x)$  = e-Funktion

Der Zusammenhang zwischen den bedingten Wahrscheinlichkeiten  $p(g/x=v)$  und  $p(x(i,v)=1/\delta(i), \vartheta(v))$  ist in ROST (1988: 110-111) dargestellt. Anwendung der ML-Schätzmethode auf die Wahrscheinlichkeiten  $p(x(i,v) = 1/\vartheta(v), \delta(i))$  ergibt die unbedingte ML-Schätzmethode (ROST 1988: 264-272; SIXTL 1982: 474-484). Das in SIXTL (1982) beschriebene Schätzverfahren ist im Programmsystem ALMO (POTUSCHAK 1989) implementiert und gehört dieser Klasse von Schätzverfahren an.

Da die Standardsoftwareprogramme (BMDP, SAS und SPSS-X) keine Prozedur zur Raschskalierung enthalten, soll nachfolgend am Beispiel der Daten des Kapitels 2 dargestellt werden, wie sich die Parameter des Raschmodells direkt (ohne Iterationen) schätzen lassen, wenn kein Programm zur Verfügung steht. Das Vorgehen ist ausführlich in SIXTL (1982: 400-410) beschrieben und besteht aus folgenden Schritten:

Schritt 1: Die Wahrscheinlichkeiten  $p(x(i,g)= 1/\vartheta(v), \delta(i))$  werden für jedes Item  $i$  und jeden Gesamtpunktwert  $x = 1,2,3,4$  geschätzt. Für unsere Daten ergeben sich folgende Schätzwerte

emp.	Item				
Gesamt- punktwert	B	A	C	D	E
1	.68	.12	.10	.05	.05
2	.82	.50	.27	.25	.16
3	.93	.65	.62	.52	.28
4	.98	.88	.85	.65	.65

Personen mit einem Gesamtpunktwert von 0 oder 5 müssen eliminiert werden, da ansonsten in den nachfolgenden Berechnungen nicht definierte Rechenergebnisse auftreten würden. Der erste Zahlenwert von 0.68 bedeutet beispielsweise, daß bei Personen mit einem Gesamtpunktwert von 1 im Item B die Ausprägung »1« in 68% aller Fälle dieser Gruppe auftritt.

Schritt 2: Auf die geschätzten bedingten Wahrscheinlichkeiten wird die Logittransformation angewendet. Diese ist definiert als:

$$\ln \left( \frac{p(x(i,v) = 1/9(v), \delta(i))}{1 - p(x(i,v) = 1/9(v), \delta(i))} \right)$$

Für die logittransformierten Wahrscheinlichkeiten gilt bei Vorliegen des Raschmodells:

$$\ln \left( \frac{p(x(i,v) = 1/9(v), \delta(i))}{1 - p(x(i,v) = 1/9(v), \delta(i))} \right) = 9(v) - \delta(i)$$

Aus den logittransformierten Wahrscheinlichkeiten können also Schätzwerte für die gesuchten Skalenwerte berechnet werden.

**Schritt 3:** Dazu werden in der Matrix der logittransformierten Wahrscheinlichkeiten Spalten- und Zeilenmittelwerte berechnet. Diese stellen Schätzwerte für die Skalenwerte der Personen und der Items dar, die noch normiert werden müssen.

Spaltenmittelwerte \* (- 1) = Schätzwerte für die Skalenwerte der Items.

Zeilenmittelwerte = Schätzwerte für die Skalenwerte der Personen.

Matrix der logittransformierten  $p(x(i,v) = 1/9(v), \delta(i))$

IG	LLB	LLA	LLC	LLD	LLE	Zeilen mittel
1.00	.7753	-1.9835	-2.2044	-2.9980	-2.9980	-1.4408
2.00	1.5440	-.0176	-1.0106	-1.1025	-1.6271	-.2406
3.00	2.6168	.6156	.4802	.0943	-.9570	.4740
4.00	3.8607	1.9577	1.6987	.6168	.6168	1.4104
Spalten- mittel						
	2.1992	.1431	-.2590	-.8474	-1.2413	.0507

**Schritt 4:** Die Schätzwerte können dadurch normiert werden, daß die Summe der Zeilenmittelwerte gleich Null gesetzt wird. Technisch wird dies dadurch erreicht, daß von den Spalten- und Zeilenmittelwerten der Mittelwert der Zeilenmittelwerte subtrahiert wird, in unserem Beispiel also der Wert 0.0507. Die geschätzten Skalenwerte sind dann:

Skalenwerte der Items:

$$\delta(B) = (-1) \cdot (2.1992 - 0.0507) = -2.1485$$

$$\delta(A) = (-1) \cdot (0.1431 - 0.0507) = -0.0924$$

$$\delta(C) = (-1) \cdot (-0.2590 - 0.0507) = 0.3097$$

$$\delta(D) = (-1) \cdot (-0.8474 - 0.0507) = 0.8908$$

$$\delta(E) = (-1) \cdot (-1.2413 - 0.0507) = 1.2920$$

Skalenwerte der Personen:

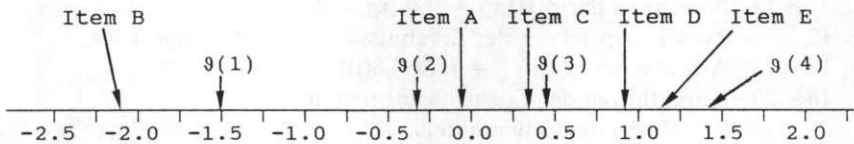
$$\vartheta(1) = (-1.4408 - 0.0507) = -1.4915$$

$$\vartheta(2) = (-0.2406 - 0.0507) = -0.2913$$

$$\vartheta(3) = (0.4740 - 0.0507) = 0.4233$$

$$\vartheta(4) = (1.4104 - 0.0507) = 1.3597$$

Die Skalenwerte  $\vartheta$  und  $\delta$  können zur besseren Interpretation auf einer Zahlengerade dargestellt werden:



Aus der Darstellung ist ersichtlich, daß die Items A (Kundgebung) und C (Kontakt mit Politikern) sowie die Items D (Mitwirken in einer Bürgerinitiative) und E (Mitarbeit in einer Partei) sehr eng zusammenliegen und  $\vartheta(2)$  von  $\vartheta(3)$  durch die Items A und C getrennt werden und  $\vartheta(3)$  von  $\vartheta(4)$  durch die Items E und D.

Das SPSS-PC-f, mit dem die Ergebnisse der Schritte 1 bis 4 berechnet wurden, ist:

```

1      data list file = 'pol.dat' free/a,b,c,d,e.
2      missing values a to e (-1).
3      count ig = a to e(missing).
4      select if (ig eq 0).
5      compute ig = a + b + c + d + e.
6      select if (ig gt 0 and ig lt 5).
7      compute gg = 1.
8      aggregate outfile = *
9          /break = ig
10         /pb,pa,pc,pd,pe = mean(b,a,c,d,e)
11         /mm = sum(gg).
12      save outf ile = 'rasch.dat'.
13      report format = list
14         /variables = pb,pa,pc,pd,pe
15         /break = ig.
16      compute llb = ln(pb/(1-pb)).
17      compute lla = ln(pa/(1-pa)).
18      compute llc = ln(pc/(1-pc)).
19      compute lld = ln(pd/(1-pd)).
    
```

- 20 compute lle = ln(pe/(1-pe)).
- 21 compute llf = (llb + lla + lld + lle)/5.
- 22 report format = nolist
- 23 /variables = llb,lla,llc,lld,lle,llf
- 24 /break = ig (total)
- 25 /summary = mean.

Die Zeilen bedeuten:

- 4 **Elimination von Personen mit fehlenden Werten.**
- 5 **Elimination** von Personen mit einem Gesamtpunktwert von 0 oder 5.
- 7 - 11 Berechnen der  $p(x(i,v) = 1/9(v), \delta(i))$ .
- 12 Zwischenspeichern der Ergebnisse für die Modellprüfung.
- 13 - 15 **Ausgabe** der  $p(x(i,v) = 1/9(v), \delta(i))$ .
- 16 - 20 Durchführen der Logittransformation.
- 21 **Berechnen des Zeilenmittels.**
- 22 - 25 **Ausgabe** der logittransformierten  $p(x(i,v) = 1/9(v), \delta(i))$  sowie der Zeilen- und Spaltenmittel.

Zur Überprüfung des Modells sind folgende weitere Schritte erforderlich:

Schritt 5: Es werden aus den geschätzten Parametern die bei einer Raschskala erwarteten Häufigkeiten  $n(i,v) = n(v) \cdot p(x(i,v) = 1/9(v), \delta(i))$  berechnet ( $n(v)$  = Zahl der Personen mit Gesamtpunktwert  $v$ ).

Schritt 6: Die erwarteten Häufigkeiten  $n(i,v)$  werden mit den empirischen Häufigkeiten  $h(i,v)$  verglichen. Zur statistischen Beurteilung der Abweichungen kann eine Chiquadrat-verteilte Testgröße berechnet werden. Für die Daten ergeben sich folgende Testgrößen.

emp. Gesamt- punktw.	Chiquadrat-Beiträge						Zeilen- summe
	Items						
IG	CHIB	CHIA	CHIC	CHID	CHIE	CHI	
1.00	.03	24.07	10.78	2.11	11.01	47.99	
2.00	8.62	.54	17.80	9.29	13.94	50.20	
3.00	.04	.00	3.19	37.61	13.41	54.26	
4.00	.12	1.74	3.89	4.50	.06	10.32	
Summe	8.81	26.36	35.66	53.52	38.42	162.77	

Die Zahl der Freiheitsgrade für jedes Item ist 3 (allgemein  $r-2$  mit  $r$  = Zahl der Items) und für die Gesamtskala 12 (allgemein  $(r-2)(r-1)$ ). Die kritischen Schwellenwerte für ein Signifikanzniveau von 95% sind für die Items 7.82 und für die Gesamtskala 21.00. Alle Chiquadrat-Werte liegen über diesen Schwellenwerten. Das Raschmodell ist also für eine Repro-

duktion der empirischen Daten nicht geeignet. Das SPSS-PC+ Programm, das die Chi-Quadrat-Prüfgrößen berechnet, ist:

```

1      get file = 'rasch.dat'.
2      list variables = all.
3      compute db = -2.1992.
4      compute da = -0.1431.
5      compute dc = 0.2590.
6      compute de = 0.8474.
7      compute dd = 1.2413.
8      compute m1 = -1.4408.
9      compute m2 = -0.2406.
10     compute m3 = 0.4740.
11     compute m4 = 1.4104.
12     if (ig eq 1) mm = m1.
13     if (ig eq 2) mm = m2.
14     if (ig eq 3) mm = m3.
15     if (ig eq 4) mm = m4.
16     compute ppb = exp(mm-db).
17     compute ppb = ppb/( 1 + ppb).
18     compute ppa = exp(mm-da).
19     compute ppa = ppa/( 1 + ppa).
20     compute ppc = exp(mm-dc).
21     compute ppc = ppc/( 1 + ppc).
22     compute ppd = exp(mm-dd).
23     compute ppd = ppd/( 1 + ppd).
24     compute ppe = exp(mm-de).
25     compute ppe = ppe/( 1 + ppe).
26     compute chib = nn*(pb-ppb)*(pb-ppb).
27     compute chia = nn*(pa-ppa)*(pa-ppa).
28     compute chic = nn*(pc-ppc)*(pc-ppc).
29     compute chid = nn*(pd-ppd)*(pd-ppd).
30     compute chie = nn*(pe-ppe)*(pe-ppe).
31     compute vb = ppb*(1-ppb).
32     compute va = ppa*(1-ppa).
33     compute vc = ppc*(1-ppc).
34     compute vd = ppd*(1-ppd).
35     compute ve = ppe*(1-ppe).
36     compute chib = chib/vb.
37     compute chia = chia/va.
38     compute chic = chic/vc.
39     compute chid = chid/vd.
40     compute chie = chie/ve.
41     compute chi = chib + chia + chic + chid + chie.
42

```

43        report format = nolist  
44            /variables = chib,chia,chic,chid,chie,chi  
45            /break = ig (total)  
46            /summary = sum.

Die Zeilen bedeuten:

- 3-11    Eingabe der Skalenwerte. Es können die normierten oder nicht normierten Skalenwerte verwendet werden.  
12-15   Berechnen des Skalenwertes für einen bestimmten Gesamtpunktwert. Ist ig beispielsweise gleich 1, dann wird mm gleich ml (= -1.4408).  
16-25   Berechnen der vom Modell erwarteten Wahrscheinlichkeiten  $p(x(i,v) - 1/\vartheta(v), \delta(i))$ .  
26 - 42   Berechnen der Chiquadrat-Prüfgrößen.  
43 - 46   Ausgabe der Chiquadrat-Prüfgrößen.

Abschließend ist anzumerken, daß bei einer iterativen ML-Schätzung die Modellanpassung entscheidend verbessert wird. Führt man beispielsweise eine Schätzung mit ALMO durch, so liegen die Prüfgrößen für die Items C und E unter dem 95%-Schwellenwerten. Sie eignen sich folglich für eine Raschskala. Zu beachten ist aber, daß zwei Items keine Raschskala bilden können. Neben der Berechnung einer Chiquadrat-verteilten Testgröße wurden für die Raschskala eine Reihe von Teststatistiken entwickelt (siehe dazu zusammenfassend VAN DEN WOLLENBERG 1988; ANDERSEN 1983). DENZ (1982) bringt einen Vergleich der Raschskala mit Modellen der Analyse latenter Dimensionen von LAZ ARS FELD/HENRY.

### 3.2 Erweiterungen: Polytome Guttmanskalierung

Die Weiterentwicklung der Guttmanskala kann in zwei Phasen gedacht werden:

Phase 1: Erweiterung um Zufallsfehler.

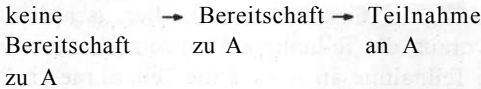
Phase 2: Erweiterung auf polytome Items und auf mehrdimensionale Begriffe.

Die Phase 1 wurde bereits in Abschnitt 3.1 behandelt. Im folgenden soll als Beispiel der Phase 2 die Erweiterung der Guttmanskala auf ordinale polytome Items beschrieben werden. (Zur mehrdimensionalen Modellierung siehe BORG/STAUFBENBIEL 1984: 134-146, SAMEJIMA 1974; SIXTL 1982: 378-282; zur Verallgemeinerung der Raschskalierung auf nominale und ordinale polytome Items siehe z.B. ANDRICH 1985; ROST 1988; CLOGG 1988). Betrachten wir dazu die Beispieldaten des Kapitels 2. In der ursprünglichen Skala wurden drei Antwortkategorien verwendet:



- 1 = habe ich bereits getan bzw. tue ich
- 2= noch nicht getan, würde es aber tun, wenn es an mich herantritt
- 3= noch nicht getan, würde es auf keinen Fall tun

In den bisherigen Beispielen wurden die Ausprägungen »2« und »3« zusammengefaßt und als »Nichtteilnahme« bezeichnet, da nur die tatsächlichen politischen Aktivitäten gemessen werden sollten. Man kann sich nun fragen, ob sich die politischen Aktivitäten durch folgendes Modell erklären lassen:



Entsprechend der Logik der Guttmanskala könnte angenommen werden, daß eine Teilnahme an der Aktivität A eine Bereitschaft zur Aktivität A voraussetzt. Würden die Bereitschaft und die Aktivität getrennt gemessen werden, müßte die Vierfeldertafel folgende Struktur besitzen:

		Aktivität A		
		nein	ja	
Bereitschaft zu A	nein	x	0	<-- nicht erlaubt (inkonsistent)
	ja	x	x	

Da keine getrennten Messungen vorliegen, können die Inkonsistenzen innerhalb einer Aktivität nicht berechnet werden. Dies ist aber zwischen zwei oder mehr »Aktivitäten« (Items) möglich. Betrachten wir dazu die Items A (Kundgebung) und B (Unterstützung eines Volksbegehrens). Diese besitzen folgende Randverteilungen:

Ausprägung	Item B (Kundgebung)	Item A (Volksbegehren)
1	46,3%	24,4%
2	36,0%	30,0%
3	17,7%	45,9%
	100,0%	100,0%
	(n = 1978)	(n = 1978)

} 82,3%      } 54,2%

Nimmt man an, daß die theoretischen Vorstellungen zutreffen, können die Ausprägungen »1« und »2« als »Bereitschaft« interpretiert werden, die Ausprägung »1« als »Teilnahme«. Es ergibt sich somit folgende Anordnung der Items nach dem Schwierigkeitsgrad:

leichtestes Item	Bereitschaft zu B (Ausprägung 1 + 2) Bereitschaft zu A (Ausprägung 1 + 2) Teilnahme an B (Ausprägung 1)
schwerstes Item	Teilnahme an A (Ausprägung 1)

Bei Vorliegen einer Guttmanskala würden folgende Antwortmuster erwartet werden:

Item B	Item A	
0	0	»0« keine Bereitschaft
1	0	»1« Bereitschaft
1	1	»2« Teilnahme
2	1	
2	2	

Inhaltlich ist dieses Modell wie folgt zu interpretieren: Die Bereitschaft zu A setzt die Bereitschaft an B voraus, die Teilnahme an B setzt die Bereitschaft zu A und B voraus, die Teilnahme an A setzt die Teilnahme an B sowie die Bereitschaft zu B und A voraus. Entsprechend der Logik der Guttmanskala sind folgende Antwortmuster nicht erlaubt:

Item B	Item A	
0	1	»0« keine Bereitschaft
0	2	»1« Bereitschaft
1	2	»2« Teilnahme
2	0	

Tabelliert man beide Items, müssen die in der Abbildung 3-6 dargestellten Zellen leer sein.

Für das Itempaar A und B können nun alle in Kapitel 2 und Abschnitt 3.1.1 behandelten Maßzahlen aus den Kreuztabellierungen der Items berechnet werden (siehe Abbildung 3-7). Damit ist bereits eine Prüfmöglichkeit angedeutet.

Eine andere besteht darin, die Programme des Kapitels 2 zu modifizieren.

Programm 1:

```

1      data list file = 'poll.dat' free/a,b,c,d,e.
2      missing value a to e(0).
3      count ig = a to e(missing).
4      select if (ig eq 0).
5      freq var = a to e.
```

Aus den Ergebnissen des Programms kann die Anordnung nach dem Schwierigkeitsgrad berechnet werden. (Anstelle der Prozedur descriptives wurde die Prozedur frequencies verwendet). Für alle Items ergibt sich folgende Anordnung:

leichtestes Item	Bereitschaft zu B (0.82)
	Bereitschaft zu D(0.71)
	Bereitschaft zu C (0.63)
	Bereitschaft zu A (0.54)

	Teilnahme an B	(0.46)
	Bereitschaft zu E	(0.28)
	Teilnahme an A	(0.24)
	Teilnahme an C	(0.19)
	Teilnahme an D	(0.16)
schwierigstes Item	Teilnahme an E	(0.12)

Inhaltlich interpretiert würde bei Vorliegen einer perfekten Guttmanskala beispielsweise eine Bereitschaft zu E (Mitarbeit in einer Partei) eine Teilnahme an B (Unterstützung eines Volksbegehrens) sowie eine Bereitschaft zu B, D, C und A voraussetzen, usw....

Programm 2:

```
1      data list file = 'poll.dat' free/a,b,c,d,e.
2      recode a to e(0 = -1).
3      missing value a to e (-1).
4      count ig = a to e(missing).
5      select if (ig eq 0).
6      recode a to e (3 = 0) (2 = 1) (1 = 2).
7      compute ig = a + b + c + d + e.
8      compute aa 1_2 = 0.
9      compute aa 1 = 0.
10     compute bbl_2 = 0.
11     compute bbl = 0.
12     compute ccl_2 = 0.
13     compute ccl = 0.
14     compute dd 1_2 = 0.
15     compute dd 1 = 0.
16     compute ee 1_2 = 0.
17     compute ee 1 = 0.
18     if (ig ge 1) bbl_2 = 1.
19     if (ig ge 2) dd 1_2 = 1.
20     if (ig ge 3) ccl_2 = 1.
21     if (ig ge 4) aa 1_2 = 1.
22     if (ig ge 5) bbl = 1.
23     if (ig ge 6) ee 1_2 = 1.
24     if (ig ge 7) aa 1 = 1.
25     if (ig ge 8) ccl = 1.
26     if (ig ge 9) dd 1 = 1.
27     if (ig ge 10) ee 1 = 1.
28     compute aa = aa 1_2 + aa 1.
29     compute bb = bbl_2 + bbl.
30     compute ce = ccl_2 + ccl.
31     compute dd = dd 1_2 + dd 1.
```

```

32     compute ee = ee 1_2 + ee 1.
33     compute fa = 0.
34     compute fb = 0.
35     compute fc = 0.
36     compute fd = 0.
37     compute fe = 0.
38     if (aa ne a) fa = 1.
39     if (bb ne b) fb = 1.
40     if (cc ne c) fc = 1.
41     if (dd ne d) fd = 1.
42     if (ee ne e) fe = 1.
43     compute gg = 1.
44     aggregate outfile = *
45         /break = gg
46         /pa0,pb0,pc0,pd0,pe0 = pin(a,b,c,d,e,0,0)
47         /pa1,pb1,pc1,pd1,pe1 = pin(a,b,c,d,e,1,1)
48         /pa2,pb2,pc2,pd2,pe2 = pin(a,b,c,d,e,2,2)
49         /ea,eb,ec,ed,ee = sum(fa,fb,fc,fd,fe)
50         /nn = sum(gg).
51     compute pa = pa0.
52     compute pb = pb0.
53     compute pc = pc0.
54     compute pd = pd0.
55     compute pe = pe0.
56     if (pa1 gt pa) pa = pa1.
57     if (pb1 gt pb) pb = pb1.
58     if (pc1 gt pc) pc = pc1.
59     if (pd1 gt pd) pd = pd1.
60     if (pe1 gt pe) pe = pe1.
61     if (pa2 gt pa) pa = pa2.
62     if (pb2 gt pb) pb = pb2.
63     if (pc2 gt pc) pc = pc2.
64     if (pd2 gt pd) pd = pd2.
65     if (pe2 gt pe) pe = pe2.
66     compute pa = (100-pa)*nn/100.
67     compute pb = (100-pb)*nn/100.
68     compute pc = (100-pc)*nn/100.
69     compute pd = (100-pd)*nn/100.
70     compute pe = (100-pe)*nn/100.
71     compute item1 = 'A'.
72     compute item2 = 'B'.
73     compute item3 = 'C'.
74     compute item4 = 'D'.

```

```

75     compute item5 = 'E'.
76     set results = 'pol.res'.
77     write var= item1,nn,ea,pa,
78         item2,nn,eb,pb,
79         item3,nn,ec,pc,
80         item4,nn,ed,pd,
81         item5,nn,ee,pe.
82     data list file = 'pol.res' free/item (A),nn,fehl,mfehl.
83     if (mfehl gt 0) pre = (mfehl-fehl)/mfehl.
84     compute rep = 1.-fehl/nn.
85     compute mmr = 1.-mfehl/nn.
86     compute pi = rep-mmr.
87     compute rep = rep* 100.
88     compute pi = pi* 100.
89     compute mmr = mmr* 100.
90     compute pre = pre* 100.
91     report format = list
92         /variables = item,fehl,mfehl,rep,mmr,pi,pre
93         /break = (nobreak)
94         /summary = mean.
95     compute gg = 1.
96     aggregate outfile = *
97         /break = gg
98         /ffehl,mmfehl,nnn = sum(fehl,mfehl,nn).
99     if (mmfehl gt 0) pre = (mmfehl-ffehl)/mmfehl.
100    compute rep = 1.-ffehl/nnn.
101    compute mmr = 1.-mmfehl/nnn.
102    compute pi = rep-mmr.
103    compute rep = rep* 100.
104    compute pi = pi* 100.
105    compute mmr = mmr* 100.
106    compute pre = pre* 100.
107    compute item = 'Gesamt'.
108    report format = list
109        /variables = item,ffehl,mmfehl,rep,mmr,pi,pre
110        /break = (nobreak)
111        /summary = mean.

```

Das Programm 2 ist analog dem Programm 2 des Kapitels 2 aufgebaut. Die Zeilen bedeuten:

- 1 Definition des Datenfiles. Die Variablen A bis E sind im Unterschied zum Kapitel 2 wie folgt kodiert:
  - 0 - fehlender Wert

- 1 = Teilnahme (Aktivität)
- 2 = Bereitschaft
- 3 = keine Bereitschaft
- 2 Umkodieren der Ausprägung für fehlende Werte.
- 3 Definition der Ausprägung für fehlende Werte.
- 4 Berechnen der Zahl fehlender Werte.
- 5 Elimination der Personen mit einem oder mehreren fehlenden Werten.
- 6 Umkodierung der Ausprägungen. Nach der Umkodierung haben die Ausprägungen folgende Bedeutung
  - 0 = keine Bereitschaft
  - 1 = Bereitschaft
  - 2 = Teilnahme
- 7 Berechnen des Gesamtpunktwertes.
- 8 - 17 Definition von (fiktiven) dichotomen Items. Die Variable aal\_2 entspricht der Bereitschaft für A, die Variable aal der Teilnahme zu A usw.
- 18-27 Berechnen des aufgrund des Gesamtpunktwertes erwarteten Antwortmusters. Bei der Eingabe ist die Anordnung der Variablen aal\_2, aal, ... nach dem Schwierigkeitsgrad zu beachten. In Abhängigkeit vom Gesamtpunkt看wert nehmen die Variablen folgende Werte an:
 

ig	bbl_2	ddl_2	ccl_2	aal_2	bbl	eel_2	aal	ccl	ddl	eel
0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0
2	1	1	0	0	0	0	0	0	0	0
3	1	1	1	0	0	0	0	0	0	0
4	1	1	1	1	0	0	0	0	0	0
5	1	1	1	1	1	0	0	0	0	0
6	1	1	1	1	1	1	0	0	0	0
7	1	1	1	1	1	1	1	0	0	0
8	1	1	1	1	1	1	1	1	0	0
9	1	1	1	1	1	1	1	1	1	0
10	1	1	1	1	1	1	1	1	1	1
- 28 - 32 Berechnen der erwarteten Antworten in den Items A bis E. Dazu werden die (fiktiven) dichotomen Items eines Items aufsummiert. Die erwartete Antwort im Item A ist  $\frac{a a l_2 + a a l}{3}$ , bei einem Gesamtpunkt看wert von 3 beispielsweise also 0, bei einem Gesamtpunkt看wert von 4 gleich 1, bei einem Gesamtpunkt看wert von 7 gleich 2, usw.. (siehe vorausgehenden Tabelle).
- 33 - 37 Definition der Fehlervariablen.
- 38 - 42 Berechnen der Werte der Fehlervariablen.
- 43 Definition einer fiktiven Variablen für die nachfolgende Aggre-

gation.

- 44- 50 Aggregieren der Daten. Die neue SPSS-PC+ interne Arbeitsdatei besteht dann aus einem Datensatz. Die Variablen paO bis peO enthalten den Anteil der Ausprägung »0« (in %) der Items A bis E, die Variablen pal bis pbl den Anteil der Ausprägung »1« der Items A bis E, die Variablen pa2 bis pb2 den Anteil der Ausprägung 2 der Items A bis E. In den Variablen ea bis ee stehen die Fehlersummen in den Items, in der Variable nn steht die Zahl der in die Analyse einbezogenen Personen.
- 51-65 Diese Anweisungen dienen der Berechnung der Ausprägung mit der besten Reproduktion der empirischen Antworten durch die Randverteilungen der Items. Für ein Item ist dies die Ausprägung mit dem größten Anteilswert. Für Item A also das Maximum von paO, pal und pa2 usw....
- 66 - 70 Die Zahl der minimalen Fehler bei einer Reproduktion aus den Randverteilungen wird berechnet. Die Verwendung von 100 ist erforderlich, da pa bis pe in % gemessen wird.
- 71-75 Definition von Namen für die nachfolgende Ausgabe.
- 76 Definition einer Datei für die nachfolgende write-Anweisung.
- 77-81 Hinausschreiben von Variablen in die Datei pol.res.
- 82 - 94 Diese Anweisungen entsprechen dem Programmzeilen 55 bis 66 des Programms 2 des Kapitels 2. Es werden die Koeffizienten für die einzelnen Items berechnet.
- 95 - 111 Berechnen der Koeffizienten für die Gesamtskala.

Das Programm erbringt folgende Ergebnisse:

ITEM	FEHL	MFEHL	REP	MMR	PI	PRE
A	611.00	1071.00	69.11	45.85	23.26	42.95
B	560.00	1063.00	71.69	46.26	25.43	47.32
C	611.00	1113.00	69.11	43.73	25.38	45.10
D	584.00	891.00	70.48	54.95	15.52	34.46
E	558.00	556.00	71.79	71.89	-1.10	-3.36
Gesamt	2924.00	4694.00	70.43	52.54	17.90	37.71

FEHL = Zahl der Fehler bei der Guttmanskala.

MFEHL = minimale Fehlerzahl bei Reproduktion aus den Randverteilungen.

REP = Reproduktionskoeffizient (in %).

MMR = minimale marginale Reproduktivität (in %).

PI = prozentuelle Verbesserung (in %).

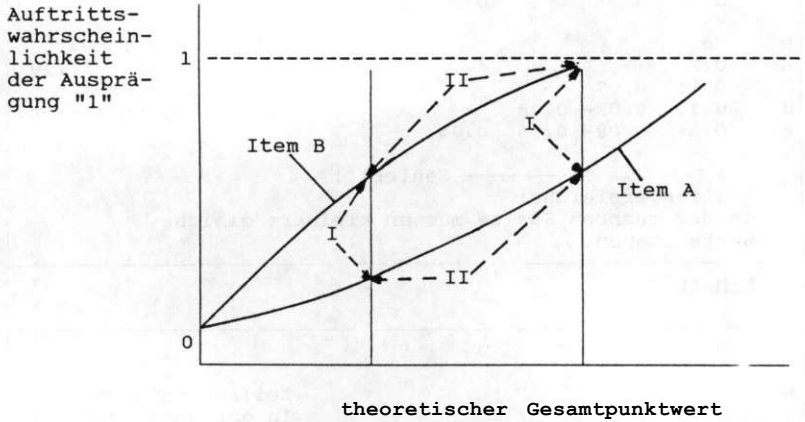
PRE « PRE-Koeffizient (in %).

Aus den Ergebnissen ist ersichtlich, daß sich die empirischen Daten nicht durch eine Guttmanskala reproduzieren lassen. Der Reproduktionskoeffizient ist kleiner 85% (= von Guttman angegebener Schwellenwert). Der

PRE-Koeffizient liegt deutlich unter dem Schwellwert von 60%. Bei der Fehlersuche kann nun wie in Abschnitt 2.8 beschrieben vorgegangen werden. Zu beachten ist, daß im Unterschied zur Guttman-skalierung bei polytomen ordinalskalierten Variablen negative PRE- und PI- Koeffizienten auftreten können.



Abbildung 3-1: Bedingung der doppelten Monotonie der Mokkenskala



I = Monotonie der Itemschwierigkeiten  
II = Monotonie der Auftrittswahrscheinlichkeiten der Ausprägung "1"

Abbildung 3-2: P- und PO-Matrix der Items A bis E des Kapitels

2

	B	A	C	D	E	
B	-					Zeilenvergleiche: In der unteren Zeile müssen kleinere/ gleiche Werte stehen.
A	0.19	-				
C	0.15	0.09	-			
D	0.13	0.07	0.08	-		
E	0.09	0.08	0.06	0.04	-	

Fehler!!!

Spaltenvergleiche:  
 In der rechten Spalte müssen kleinere/gleiche  
 Werte stehen.

P-Matrix

	B	A	C	D	E	
B	-					Zeilenvergleiche: In der unteren Zeile müssen größere/ gleiche Werte stehen.
A	0.48	-				
C	0.49	0.66	-			
D	0.51	0.67	0.73	-		
E	0.51	0.72	0.75	0.76	-	

Spaltenvergleiche:  
 In der rechten Spalte müssen größere/gleiche  
 Werte stehen.

PO-Matrix

Abbildung 3-3: Berechnung der Elemente der P- und PO-Matrix sowie des Homogenitätskoeffizienten und der Teststatistik DELTA aus den Vierfeldertafeln der Items

Crosstabulation: A (schwierigeres Item)  
By B (leichteres Item)

		Count	Exp Val	Tot Pct	Residual	Std Res	Adj Res	Row Total
A	0.0	953	806.1	48.2%	146.9	5.2	15.5	1500 75.8%
	1.0	110	256.9	5.6%	-146.9	-9.2	-15.5	478 24.2%
		1063	1063	53.7%	0	0	0	1978 100.0%

Element der PO-Matrix: 953/1978 = 0.48 = = Tot Pct	
--	--

Element der P-Matrix 368/1978 = 0.19 = TotPct	
--	--

Homogenitätskoeffizient B-A:  
 $1 - (\text{Count}/\text{expVal}) = 1 - (110/256.9) = 0.57$

Teststatistik DELTA für B-A:  
 $(\sqrt{(n-1)/n}) * (-1) * \text{AdjRes} = (\sqrt{1977}/\sqrt{1978}) * 15.5 = 0.9997 * 15.5 = 15.5$

Abbildung 3-4: Die Beziehung zwischen theoretischen und empirischen Antwortmustern in der Skalierung nach PROCTOR

theoret. Antwortmuster v	empir. Antwortmuster g	Zahl der Fehler	bedingte Wahrscheinlichkeit p(g/v)	Auftrittshäufigkeit n(g,v)	
0 0 0	0 0 0	0	.8*.8*.8=.512(a)	1536	
	<u>1</u> 0 0	1	.2*.8*.8=.128	384	
	0 <u>1</u> 0	1	.8*.2*.8=.128	384	
	0 0 <u>1</u>	1	.8*.8*.2=.128	384	
	<u>1</u> <u>1</u> 0	2	.2*.2*.8=.032	96	
	<u>1</u> 0 <u>1</u>	2	.2*.8*.2=.032	96	
	0 <u>1</u> <u>1</u>	2	.8*.2*.2=.032	96	
	<u>1</u> <u>1</u> <u>1</u>	3	.2*.2*.2=.008	24	
	-----				
	(Σ= 3000)				
1 0 0	<u>0</u> 0 0	1	.2*.8*.8=.128	512	
	1 0 0	0	.8*.8*.8=.512(a)	2048	
	<u>0</u> <u>1</u> 0	2	.2*.2*.8=.032	128	
	<u>0</u> 0 <u>1</u>	2	.2*.8*.2=.032	128	
	1 <u>1</u> 0	1	.8*.2*.8=.128	512	
	1 0 <u>1</u>	1	.8*.8*.2=.128	512	
	<u>0</u> <u>1</u> <u>1</u>	3	.2*.2*.2=.008	32	
	1 <u>1</u> <u>1</u>	2	.8*.2*.2=.032	128	
-----					
(Σ= 4000)					
1 1 0	<u>0</u> <u>0</u> 0	2	.2*.2*.8=.032	64	
	1 <u>0</u> 0	1	.8*.2*.8=.128	256	
	<u>0</u> 1 0	1	.2*.8*.8=.128	256	
	<u>0</u> <u>0</u> <u>1</u>	3	.2*.2*.2=.008	16	
	1 1 0	0	.8*.8*.8=.512(a)	1024	
	1 <u>0</u> <u>1</u>	2	.8*.2*.2=.032	64	
	<u>0</u> 1 <u>1</u>	2	.2*.8*.2=.032	64	
	1 1 <u>1</u>	1	.8*.8*.2=.128	256	
-----					
(Σ= 2000)					
1 1 1	<u>0</u> <u>0</u> <u>0</u>	3	.2*.2*.2=.008	8	
	1 <u>0</u> <u>0</u>	2	.8*.2*.2=.032	32	
	<u>0</u> 1 <u>0</u>	2	.2*.8*.2=.032	32	
	<u>0</u> <u>0</u> 1	2	.2*.2*.8=.032	32	
	1 1 <u>0</u>	1	.8*.8*.2=.128	128	
	1 <u>0</u> 1	1	.8*.2*.8=.128	128	
	<u>0</u> 1 1	1	.2*.8*.8=.128	128	
	1 1 1	0	.8*.8*.8=.512(a)	512	
-----					
(Σ= 1000)					

Fehlerrate  $\alpha=0.2$ .

Auftrittswahrscheinlichkeiten der theoretischen Antwortmuster:  $\pi(0)=0.3$ ,  $\pi(1)=0.4$ ,  $\pi(2)=0.2$  und  $\pi(3)=0.1$ .  $n=10000$ .

Unterstrichene Werte = Fehler.

(a) bei einer perfekten Guttmanskala erwartetes Antwortmuster

Abbildung 3-5: Die Beziehung zwischen theoretischen und empirischen Antwortmustern in der Raschskalierung

theoret. Antwortmuster v	empir. Antwortmuster g	emp. und theoret. Gesamtpunktw.	bedingte Wahrscheinlichkeit $p(g/v)$
0 0 0	0 0 0	0	1.000(a)
1 0 0	1 0 0	1	$\exp(-\delta(A))/\tau(1) = 0.690(a)$
	0 <u>1</u> 0	1	$\exp(-\delta(B))/\tau(1) = 0.254$
	0 0 <u>1</u>	1	$\exp(-\delta(C))/\tau(1) = 0.056$
1 1 0	1 1 0	2	$\exp(-\delta(A)-\delta(B))/\tau(2) = 0.776(a)$
	1 <u>0</u> <u>1</u>	2	$\exp(-\delta(A)-\delta(C))/\tau(2) = 0.171$
	<u>0</u> <u>1</u> <u>1</u>	2	$\exp(-\delta(B)-\delta(C))/\tau(2) = 0.063$
1 1 1	1 1 1	3	1.000(a)

Skalenwerte der Items:  $\delta(A) = -1.0$ ,  $\delta(B) = 0.0$ ,  $\delta(C) = 1.5$

Unterstrichene Werte = Fehler bei perfekter Guttmanskala.

(a) bei einer perfekten Guttmanskala erwartetes Antwortmuster

Abbildung 3-6: Erwartete leere Zellen bei einer perfekten Guttmanskala für die Items A und B

		Item A		
		0	1	2
Item B	0	x	-	-
	1	x	x	-
	2	-	x	x

"-" - erwartete leere Zellen

Abbildung 3-7: Berechnung der Koeffizienten der Guttmanskala für Itempaare

Crosstabulation: B (Unterstützung Volksbegehren)  
By A (Kundgebung)

Count Exp Val					Row Total
		0.0	1.00	2.00	
A->	B				
	0.0	287 160.5 x geb.P.	36 104.9 - ink.P.	27 84.6 - ink.P.	350 17.7%
	1.00	344 326.9 x kon.P.	286 213.8 x geb.P.	83 172.3 - ink.P.	713 36.0%
	2.00	276 419.6 - ink.P.	271 274.3 x kon.P.	368 221.1 x geb.P.	915 46.3%
Column Total		907 45.9%	593 30.0%	478 24.2%	1978 100.0%

Ausprägung mit  
der besten marg.  
Reproduktion in B

Ausprägung mit der besten  
marginalen Reproduktion in A

Homogenitätskoeffizient des Itempaars:

$$H = 1 - \frac{\text{emp. Ink.}}{\text{zuf. Ink.}} = 1 - \frac{36 + 27 + 83 + 276}{104.9 + 84.6 + 172.3 + 419.3} = 1 - \frac{422.0}{788.1} = 0.46$$

Konsistenzkoeffizient des Itempaars:

$$C = \frac{\text{kon.P.}-\text{ink.P.}}{\text{kon.P.}+\text{ink.P.}} = \frac{615 - 422}{615 + 422} = 0.19$$

## 4. Analyse latenter Klassen

### 4.1 Modellspezifikation

In den beiden vorausgehenden Kapiteln (Abschnitt 2.8 und Abschnitt 3.1.3) wurde bereits erwähnt, daß die gesuchten Modellparameter mit Hilfe eines Programms zur Analyse latenter Klassen geschätzt werden können. Die Grundlogik der Analyse latenter Klassen soll in diesem Abschnitt anhand des bei der Fehlersuche der Guttmanskala angeführten Modells von GOODMAN (1975) verdeutlicht werden. Das Modell von GOODMAN geht von folgenden Annahmen aus:

- Es gibt zwei Gruppen von Personen: Eine Gruppe von Personen, die nach einer perfekten Guttmanskala skalierbar ist, und eine zweite Gruppe von nichtskalierbaren Personen.
- Bezüglich der Gruppe der nichtskalierbaren Personen wird angenommen, daß für jedes Item bestimmte Auftrittswahrscheinlichkeiten in den Ausprägungen vorliegen und die Antwort auf ein Item von den Antworten in den anderen Items (statistisch) unabhängig ist.

Für die fünf Items A bis E unserer Beispieldaten (1) gibt es somit 7 Klassen:

- Sechs Klassen für die perfekte Guttmanskala mit den theoretischen Antwortmustern:

Items	A	B	C	D	E	latenter Klassenindex
	0	0	0	0	0	0
	0	1	0	0	0	1
	1	1	0	0	0	2
	1	1	1	0	0	3
	1	1	1	1	0	4
	1	1	1	1	1	5

Diese Klassen gehören der Gruppe der skalierbaren Personen an.

Die Antwortmuster ergeben sich durch den Schwierigkeitsgrad der Items (B ist leichter als A, A leichter als C, C leichter als D und D leichter als E). Die durch die theoretischen Antwortmuster gebildeten latenten Klassen wurden mit 0,1,...,5 indiziert. Die latente Klasse 0 ist also durch das Antwortmuster (0,0,0,0,0) gekennzeichnet (in allen Items tritt eine »0« auf). In der latenten Klasse 1 tritt im leichtesten Item B ein »1« auf, usw.

- Eine Klasse nichtskalierbarer Personen. In dieser Klasse, die im folgenden mit »6« indiziert wird, können alle  $2^{*5} = 32$  möglichen Antwortmuster auftreten. Wegen der statistischen Unabhängigkeit der Antworten ist die Auftrittswahrscheinlichkeit für ein Antwortmuster  $(i,j,k,l,m)$  (z.B. 0,0,0,0,0) bei gegebener latenter Klasse »6« (Gruppe der nichtskalierbaren Personen) gleich

$$p(i,j,k,l,m/6) = p(x(A) = i/6) * p(x(B) = j/6) * p(x(C) = k/6) * p(x(D) = l/6) * p(x(E) = m/6)$$

mit

$p(x(A) = i/6)$  = bedingte Wahrscheinlichkeit des Auftretens der Ausprägung  $i$  im Item A ( $i = 0,1$ ) gegeben der Gruppe der nichtskalierbaren Personen.

$p(x(B) = j/6)$  = bedingte Wahrscheinlichkeit des Auftretens der Ausprägung  $j$  im Item B ( $j = 0,1$ ) gegeben der Gruppe der nichtskalierbaren Personen.

$p(x(C) = k/6)$  = bedingte Wahrscheinlichkeit des Auftretens der Ausprägung  $k$  im Item C ( $k = 0,1$ ) gegeben der Gruppe der nichtskalierbaren Personen.

$p(x(D) = l/6)$  = bedingte Wahrscheinlichkeit des Auftretens der Ausprägung  $l$  im Item D ( $l = 0,1$ ) gegeben der Gruppe der nichtskalierbaren Personen.

$p(x(E) = m/6)$  = bedingte Wahrscheinlichkeit des Auftretens der Ausprägung  $m$  im Item E ( $m = 0,1$ ) gegeben der Gruppe der nichtskalierbaren Personen.

Für das Antwortmuster (0,0,0,0,0) soll also gelten:

$$p(0,0,0,0,0/6) = (p(x(A) = 0/6) * \{p(x(B) = 0/6) * \{p(x(C) = 0/6) * \{p(x(D) = 0/6) * \{p(x(E) = 0/6)\}\}\}\})$$

Analog soll für das Antwortmuster (1,0,0,0,0) (Auftreten einer »1« im Item A) gelten:

$$p(1,0,0,0,0/6) = (p(x(A) = 1/6) * \{p(x(B) = 0/6) * \{p(x(C) = 0/6) * \{p(x(D) = 0/6) * \{p(x(E) = 0/6)\}\}\}\})$$

Da in der Gruppe der nichtskalierbaren Personen auch die bei einer Guttmanskala erwarteten Antwortmuster auftreten können, sind die Auftrittswahrscheinlichkeiten der theoretischen Antwortmuster der Guttmanskala (die latenten Klassen 0,1,...,5) nicht bekannt. Dies bedeutet, daß neben den bedingten Antwortwahrscheinlichkeiten  $p(x(A) = i/6)$ ,  $p(x(B) = j/6)$ ,... der latenten Klasse der nichtskalierbaren Personen die Auftrittswahrscheinlichkeiten der latenten Klassen 0,1,2,...,6 (latente Klassenwahrscheinlichkeiten) geschätzt werden müssen. Die zu schätzenden Parameter sind in der Abbildung 4-1 dargestellt.



## 4.2 Modellannahmen der Analyse latenter Klassen

Bevor das Modell der Abbildung 4-1 weiter diskutiert wird, soll anhand der Abbildung 4-1 die allgemeine Grundlogik der Analyse latenter Klassen und deren Annahmen dargestellt werden (siehe dazu z.B. LAZARUS-FELD/HENRY 1968: 46-156). Diese sind:

- Eine bestimmte Zahl latenter Klassen wird angenommen.
- Zwischen den latenten Klassen kann eine Ordnungsrelation bestehen. In der Abbildung 4-1 ist dies für die latenten Klassen **0,1,...,5**, die die bei einer perfekten Gutmanskala erwarteten Antwortmuster abbilden, der Fall. Alle Klassen zusammen, also einschließlich der Gruppe der nicht skalierbaren Personen, besitzen aber nur nominales Meßniveau. Die Analyse latenter Klassen geht somit grundsätzlich nur von der Vorstellung einer nominalskalierten latenten Dimension aus.
- Jede latente Klasse  $g$  ist durch bedingte Antwortwahrscheinlichkeiten  $p(x(A)=i/g)$ ,  $p(x(B)=j/g)$ , ... in den Items gekennzeichnet.
- Innerhalb jeder latenten Klasse wird (statistische) Unabhängigkeit des Auftretens einer Ausprägung in einem Item von dem Auftreten einer Ausprägung in einem anderen Item angenommen. Diese Annahme wird als lokale Unabhängigkeit bezeichnet. Formal ausgedrückt soll also gelten:

$$p(i,j,k,l,m/g) = p(x(A)=i/g)*p(x(B)=j/g)*p(x(C)=k/g)*p(x(D)=l/g)*p(x(E)=m/g)$$

Man kann sich leicht überzeugen, daß in der Abbildung 4-1 die Annahme der lokalen Unabhängigkeit auch für die theoretischen Antwortmuster der Guttmanskala (latente Klassen **0,1,...,5**) erfüllt ist. Betrachten wir dazu beispielsweise die latente Klasse 2 (theoretisches Antwortmuster **(1,1,0,0,0)**). Die bedingten Wahrscheinlichkeiten  $p(x(A)=i/2)$ , ... sind:

$$\begin{aligned} p(x(A) = 0/2) &= 0 & p(x(A) = 1/2) &= 1 \\ p(x(B) = 0/2) &= 0 & p(x(B) = 1/2) &= 1 \\ p(x(C) = 0/2) &= 1 & p(x(C) = 1/2) &= 0 \\ p(x(D) = 0/2) &= 1 & p(x(D) = 1/2) &= 0 \\ p(x(E) = 0/2) &= 1 & p(x(E) = 1/2) &= 0 \end{aligned}$$

Die bedingte Wahrscheinlichkeit für das erwartete Antwortmuster **(1,1,0,0,0)** bei Anwendung der Formel für die statistische Unabhängigkeit ist:

$$p(1,1,0,0,0/2) = p(x(A) = 1/2) * p(x(B) = 1/2) * p(x(C) = 0/2) * \\ p(x(D) = 0/2) * p(x(E) = 0/2) = 1.0$$

Für alle anderen bedingten Wahrscheinlichkeiten ergibt sich ein Wert 0. Die Annahme der lokalen Unabhängigkeit ist erfüllt.

### 4.3 Überprüfung der Identifikation

In Abschnitt 3.1.1 wurde bereits sehr allgemein das Identifikationsproblem andiskutiert (siehe dazu allgemein LAZARSELD/HENRY 1968: 59-68). Ein Modell wurde als identifiziert bezeichnet, wenn die unbekannt (zu schätzenden) Parameter als Funktionen der theoretisch vorhandenen empirischen Information und der bekannten Parameter ausgedrückt werden können. Dadurch ist garantiert, daß die unbekannt Parameter nicht beliebig wählbar und somit inhaltlich interpretierbar sind. Das Modell der Abbildung 4-1 ist durch folgende Größen gekennzeichnet:

- theoretisch vorhandene empirische Information:
  - $n(i,j,k,l,m)$  = durch das Modell erwartete Häufigkeiten
- unbekannte Parameter:
  - $\pi(0)$  bis  $\pi(6)$  = latente Klassenwahrscheinlichkeiten.
  - Bedingte Antwortwahrscheinlichkeiten  $p(x(A)=i/6)$ ,  $p(x(B)=j/6)$ , ... der latenten Klasse 6 (nichtskalierbare Personen)
- bekannte Parameter:
  - $n$  = die Stichprobengröße.
  - Bedingte Antwortwahrscheinlichkeiten  $p(x(A)=i/g)$ ,  $p(x(B)=j/g)$ ,  $p(x(E)=m/g)$  für die latenten Klassen 0 bis 5 (skalierbare Personen).

Bei der Identifikationsprüfung wird nun allgemein schrittweise wie folgt vorgegangen:

- Es wird versucht, zunächst einen Parameter zu identifizieren.
- Würde der Parameter identifiziert, ist er »bekannt« und kann wie die anderen bekannten Parameter zur Identifikation der noch nicht identifizierten Parameter verwendet werden.
- Anschließend wird versucht, den nächsten Parameter zu identifizieren usw...

Zur Lösung des Identifikationsproblems des Modells der Abbildung 4-1 betrachten wir in einem ersten Schritt die beiden fehlerhaften Antwortmuster (1,0,0,0,0) und (1,0,1,0,0). Das Antwortmuster (1,0,0,0,0) ist fehlerhaft, da im Item A eine »1« auftritt, im leichtesten Item B dagegen ein »0«. Das Antwortmuster (1,0,1,0,0) ist fehlerhaft, da in den schweren Items A

und C eine »1« auftritt, in dem leichtesten Item B dagegen eine »0«. Die erwarteten Häufigkeiten besitzen folgende Darstellung, da sie nur von der latenten Klasse 6, der Gruppe der nichtskalierbaren Personen, abhängen:

$$n(1,0,0,0,0) = p(A/6) * (1-p(B/6)) * (1-p(C/6)) * (1-p(D/6)) * (1-p(E/6)) * \pi(6) * n$$

$$n(1,0,1,0,0) = p(A/6) * (1-p(B/6)) * p(C/6) * (1-p(D/6)) * (1-p(E/6)) * \pi(6) * n$$

In den beiden Gleichungen wurden folgende Kurzbezeichnungen verwendet

$$p(A/6) \text{ für } p(x(A)=1/6); \quad 1-p(A/6) \text{ für } p(x(A) = 0/6)$$

$$p(B/6) \text{ für } p(x(B)=1/6); \quad 1-p(B/6) \text{ für } p(x(B) = 0/6)$$

usw. . . .

Division von  $n(1,0,0,0,0)$  mit  $n(1,0,1,0,0)$  ergibt

$$\frac{n(1,0,0,0,0)}{n(1,0,1,0,0)} = \frac{1 - p(C/6)}{p(C/6)}$$

da sich die anderen Parameter wegekürzen. Umformen der Gleichung ergibt für  $p(C/6)$  die Darstellung

$$p(C/6) = \frac{n(1,0,1,0,0)}{n(1,0,1,0,0) + n(1,0,0,0,0)}$$

Der unbekannte Parameter  $p(C/6)$  ist somit als Funktion der theoretisch bekannten empirischen Information, den erwarteten Häufigkeiten  $n(i,j,k,l,m)$ , dargestellt und identifiziert. Ähnliche Überlegungen ergeben für  $p(D/6)$ ,  $p(E/6)$ ,  $p(A/6)$  und  $p(C/6)$  folgende Darstellungen:

$$p(D/6) = \frac{n(0,0,1,1,0)}{n(0,0,1,1,0) + n(0,0,1,0,0)}$$

$$p(E/6) = \frac{n(0,0,0,1,1)}{n(0,0,0,1,1) + n(0,0,0,1,0)}$$

$$p(A/6) = \frac{n(1,0,1,0,0)}{n(1,0,1,0,0) + n(0,0,1,0,0)}$$

$$p(B/6) = \frac{n(0,1,0,0,1)}{n(0,1,0,0,1) + n(0,0,0,0,1)}$$

Damit sind die bedingten Antwortwahrscheinlichkeiten  $p(A/6)$ ,  $p(B/6)$ , ... der latenten Klassen der nichtskalierbaren Personen identifiziert. Sie können in der Folge zur Identifikation der latenten Klassenwahrscheinlichkeiten  $\pi(0)$ ,  $\pi(1)$ , ... verwendet werden. Betrachten wir dazu die Darstellung der erwarteten Häufigkeit des Fehlerantwortmusters  $(0,0,1,1,1)$ :

$$n(0,0,1,1,1) = (1-p(A/6)) * (1-p(B/6)) * p(C/6) * p(D/6) * p(E/6) * \pi(6) * n$$

Außer  $\pi(6)$  sind alle Parameter bekannt. Für  $\pi(6)$  ergibt sich somit folgende Darstellung:

$$\pi(6) = \frac{n(0,0,1,1,1)}{(1-p(A/6)) \cdot (1-p(B/6)) \cdot p(C/6) \cdot p(D/6) \cdot p(E/6) \cdot n}$$

Zur Berechnung der latenten Klassenwahrscheinlichkeiten  $\pi(0)$ ,  $\pi(1)$ , ...,  $\pi(5)$  können nun die Gleichungen für die erwarteten Häufigkeiten der Antwortmuster der perfekten Guttmankala verwendet werden:

$$n(0,0,0,0,0) = \pi(0) \cdot n + (1-p(A/6)) \cdot (1-p(B/6)) \cdot (1-p(C/6)) \cdot (1-p(D/6)) \cdot (1-p(E/6)) \cdot \pi(6) \cdot n$$

$$n(0,1,0,0,0) = \pi(1) \cdot n + (1-p(A/6)) \cdot p(B/6) \cdot (1-p(C/6)) \cdot (1-p(D/6)) \cdot (1-p(E/6)) \cdot \pi(6) \cdot n$$

$$n(1,1,0,0,0) = \pi(2) \cdot n + p(A/6) \cdot p(B/6) \cdot (1-p(C/6)) \cdot (1-p(D/6)) \cdot (1-p(E/6)) \cdot \pi(6) \cdot n$$

$$n(1,1,1,0,0) = \pi(3) \cdot n + p(A/6) \cdot p(B/6) \cdot p(C/6) \cdot (1-p(D/6)) \cdot (1-p(E/6)) \cdot \pi(6) \cdot n$$

$$n(1,1,1,1,0) = \pi(4) \cdot n + p(A/6) \cdot p(B/6) \cdot p(C/6) \cdot p(D/6) \cdot (1-p(E/6)) \cdot \pi(6) \cdot n$$

$$n(1,1,1,1,1) = \pi(5) \cdot n + p(A/6) \cdot p(B/6) \cdot p(C/6) \cdot p(D/6) \cdot p(E/6) \cdot \pi(6) \cdot n$$

In diesen Gleichungen sind alle Größen mit Ausnahme der latenten Klassenwahrscheinlichkeiten  $\pi(0)$ ,  $\pi(1)$ ,  $\pi(2)$ ,  $\pi(3)$ ,  $\pi(4)$  und  $\pi(5)$  bekannt. Durch einfaches Umformen können Darstellungen für diese unbekannt Parameter gefunden werden.

Das Modell ist somit identifiziert, da alle Parameter identifiziert sind. Mitunter ist das eben beschriebene Vorgehen der Überprüfung der Identifikation sehr aufwendig und erfordert komplexe mathematische Operationen (siehe z.B. LAZARSFELD/HENRY 1968: 59-68). Deshalb wurden Verfahren zur computerunterstützten Identifikationsprüfung entwickelt (VAN DE POL/LANGEHEINE/DE JONG 1989: 29). Dabei wird i.d.R. untersucht, ob die Korrelationsmatrix der geschätzten Parameter bzw. ihre Inverse vollen Rang besitzt. Ist dies nicht der Fall, liegt eine lineare Abhängigkeit der geschätzten Parameter vor. Das heißt, daß zumindest ein Parameter als lineare Funktion der anderen Parameter ausgedrückt werden kann, z.B. kann gelten:

$$\begin{aligned} \Omega(1) &= \text{Konstante} + \Omega(2) \\ \text{mit } \Omega(i) &= \text{unbekannter Parameter} \end{aligned}$$

In dem obigen Beispiel kann ein Parameter ( $\Omega(1)$  oder  $\Omega(2)$ ) beliebig gewählt werden, da gilt:

$$\text{Konstante} - \Omega(2) = \Omega(1).$$

Das Vorliegen einer linearen Abhängigkeit ist somit ein Hinweis auf die Nichtidentifikation eines oder mehrerer Parameter. Zu beachten ist, daß beim Vorliegen keiner linearen Abhängigkeit aber noch nicht die Identifikation nachgewiesen ist.

Abschließend ist zur Identifikation folgendes anzumerken:

- Die Identifikation ist eine Eigenschaft des spezifizierten Modells.
- Sie hat somit nichts mit der Schätzung der Parameter zu tun. Das bedeutet, daß auch nicht identifizierte Parameter geschätzt werden können, die aber dann nicht sinnvoll inhaltlich interpretierbar sind.
- Bei der Schätzung kann ferner das Problem auftreten, daß theoretisch identifizierte Parameter empirisch nicht identifiziert sind. In unserem Beispiel z.B. für  $p(C/6)$  dann, wenn  $n(1,0,0,0,0)$  und  $n(1,0,1,0,0)$  Null wären. Diese würde zu einer Division durch Null führen.
- Einzelne Parameter können überidentifiziert sein. Das heißt, daß für sie mehrere Darstellungsweisen gefunden werden können. Für  $p(C/6)$  beispielsweise gibt es folgende Darstellungsmöglichkeiten:

$$\begin{aligned}
 p(C/6) &= \frac{n(1,0,1,0,0,0)}{n(1,0,1,0,0) + n(1,0,0,0,0)} \\
 &= \frac{n(1,0,1,1,0)}{n(1,0,1,1,0) + n(1,0,0,1,0)} \\
 &= \frac{n(1,0,1,1,1)}{n(1,0,1,1,1) + n(1,0,0,1,1)} \\
 &= \dots
 \end{aligned}$$

- Sind alle Parameter eines Modells identifiziert, wird das Modell als identifiziert bezeichnet.
- Ist mindestens ein Parameter eines Modells überidentifiziert, wird das Modell als überidentifiziert bezeichnet.

#### 4.4 Schätzung der Modellparameter und Modellprüfung

Die unbekannt Parameter des Modells der Abbildung 4-1 können mit den bereits beschriebenen EM-Algorithmus (siehe Abschnitt 3.1.3) geschätzt werden. Für die Beispieldaten ergeben sich folgende Schätzwerte:

$\pi(0)$	=	.396 ( $\pm 0.012$ )	}	.525
$\pi(1)$	=	.090 ( $\pm 0.012$ )		
$\pi(2)$	=	.013 ( $\pm 0.008$ )		
$\pi(3)$	=	.000 ( $\pm 0.000$ )		
$\pi(4)$	=	.006 ( $\pm 0.019$ )		
$\pi(5)$	=	.019 ( $\pm 0.003$ )		

$\pi(6)$	= .475 ( $\pm 0.021$ )
$p(A/6)$	= .427 ( $\pm 0.021$ )
$p(B/6)$	= .702 ( $\pm 0.018$ )
$p(C/6)$	= .352 ( $\pm 0.021$ )
$p(D/6)$	= .281 ( $\pm 0.019$ )
$p(E/6)$	= .213 ( $\pm 0.016$ )

Die Schätzung wurde mit dem Programm PANMARK (VAN DE POL/LANGEHEINE/DE JONG 1989) durchgeführt (siehe Abschnitt 4.7). In Klammern stehen die Standardabweichungen der geschätzten Parameter. Inhaltlich sind die Ergebnisse wie folgt zu interpretieren: 52.5% der Personen gehören den skalierbaren Personen an, 47.5% der Gruppe der nicht skalierbaren Personen. In der Gruppe der skalierbaren Personen dominiert das theoretische Antwortmuster (0,0,0,0,0) ( $\pi(0) = 39.6\%$ ). Alle anderen Muster besitzen nur eine geringe Auftrittswahrscheinlichkeit. Die Gruppe der nicht skalierbaren Personen ist durch folgende Antwortwahrscheinlichkeiten gekennzeichnet: Die bedingte Auftrittswahrscheinlichkeit der Ausprägung »1« im Item A beträgt 42.7%, die im Item B 70.2% usw. Ordnet man die bedingten Auftrittswahrscheinlichkeiten nach ihrer Größe, ergibt sich die bei der Guttmankala angenommene Anordnung der Items nach dem Schwierigkeitsgrad. Die Gruppe der nichtskalierbaren Personen unterscheidet sich folglich von der Gruppe der skalierbaren Personen nicht in der Anordnung der Items, sondern nur darin, daß kein deterministischer Mechanismus (beim Auftreten einer »1« in einem schweren Item muß in allen leichteren Items ebenfalls eine »1« auftreten) vorliegt.

Zur Überprüfung der Modellanpassung (Reproduktion der empirischen Häufigkeiten  $h(i,j,k,l,m)$ ) berechnet PANMARK u.a. eine Chi-Quadrat-verteilte Testgröße, die definiert ist als:

$$CHI = \sum_i \sum_j \sum_k \sum_l \sum_m \frac{(h(i,j,k,l,m) - n(i,j,k,l,m))^2}{n(i,j,k,l,m)} = 103.9$$

mit

$h(i,j,k,l,m)$  = empirische Häufigkeit des Antwortmusters (i,j,k,l,m)

$n(i,j,k,l,m)$  = erwartete Häufigkeit des Antwortmusters (i,j,k,l,m)

Die Zahl der Freiheitsgrade ergibt sich mit

2**5 = 32	empirische Antwortmuster
- 1	da die Stichprobe fixiert ist
- 7	zu schätzende Parameter $\pi(0), \dots, \pi(6)$
+ 1	da $\pi(0) + \pi(1) + \dots + \pi(6) = 1.0$ ist
- 5	zu schätzende Parameter $p(A/6), \dots, p(E/6)$
<hr/>	
20	frei variierbare Größen (= Freiheitsgrade)

Der kritische Wert der Chiquadrat-Verteilung bei einem 95%-Signifikanzniveau und bei 20 Freiheitsgraden beträgt 31.4. Die berechnete Prüfgröße liegt über diesem Wert. Die HO-Hypothese, daß die empirischen Häufigkeiten durch das Modell reproduziert werden, muß somit verworfen werden.

#### 4.5 Fehleranalyse

Wegen der schlechten Modellanpassung tritt das Problem der Fehlersuche auf. Dazu bieten sich mehrere Strategien an. Eine Strategie besteht in einem explorativen Vorgehen. Dabei werden bezüglich der bedingten Antwortwahrscheinlichkeiten  $p(x(A)=i/g)$ ,  $p(x(B)=j/g)$ ,... der latenten Klassen  $g$  keine Restriktionen getroffen. Die Zahl der latenten Klassen wird systematisch variiert, z.B. von 2 bis 5 latenten Klassen, um Anhaltspunkte über die den Daten zugrundeliegende Zahl latenter Klassen zu gewinnen. In Abhängigkeit von der Zahl der latenten Klassen ergeben sich für die Beispielsdaten folgende Prüf werte:

Zahl der latenten Klassen	Chiquadrat- Prüfgröße	Freiheitsgrade	Differenz der Chiquadrat- Prüfgrößen	Freiheitsgrade der Differenz
2	134.6	20		
3	66.1	14	68.5	6
4	12.4	8	54.1	6
5	4.9	2	7.5	6

Um zu überprüfen, ob eine größere Klassenzahl zu einer signifikant besseren Modellanpassung führt, können zwischen den Prüfgrößen Differenzen berechnet werden. Für die Differenz zwischen 3 und 2 latenten Klassen ergeben sich folgende Werte:

$$\text{CHI}(3-2) = \text{CHI}(2) - \text{CHI}(3) = 134.6 - 66.1 = 68.5$$

$$\text{df}(3-2) = \text{df}(2) - \text{df}(3) = 20 - 14 = 6$$

Die Differenzen zwischen den Chiquadrat-Prüfgrößen besitzen wiederum eine Chiquadrat-Verteilung mit 6 Freiheitsgraden. Es kann somit geprüft werden, ob die Differenzen signifikant von Null verschieden sind. Der 95%-Schwellwert der Chiquadrat-Verteilung bei 6 Freiheitsgraden beträgt 12.6. Die 3-Klassenlösung verbessert mithin die Modellanpassung signifikant gegenüber der 2-Klassenlösung. Gleiches gilt für die 4-Klassenlösung. Die 5-Klassenlösung dagegen bringt keinen signifikanten Zuwachs in der Modellanpassung. In diesem Fall wird man sich also für die 4-Klassenlösung entscheiden. Anzumerken ist, daß mit der beschriebenen Teststrategie die zusätzliche Erklärungskraft der 6-Klassenlösung nicht

geprüft werden kann, da eine Schätzung einer 7-Klassenlösung nicht möglich ist. Die notwendige Identifikationsbedingung, daß mehr bekannte als unbekannte Größen vorliegen (siehe Abschnitt 3.1.1), ist nicht erfüllt. Die vier Klassen lassen sich wie folgt interpretieren (siehe Abbildung 4-2):

latente Klasse 1: Personen mit geringem Aktivitätsniveau

latente Klasse 2: Personen mit hohem Aktivitätsniveau

latente Klasse 3: Personen mit mittlerem Aktivitätsniveau im Item A (Teilnahme an einer Kundgebung) und im Item B (Unterstützung eines Volksbegehrens)

latente Klasse 4: Personen mit mittlerem bis starkem Aktivitätsniveau an Aktivitäten der direkten Demokratie (Unterstützung eines Volksbegehrens, Kontakt mit Politikern, Mitwirken in einer Bürgerinitiative) und mit einem geringeren Aktivitätsniveau in dem Item »Teilnahme an Kundgebungen«

#### 4.6 Zuordnung der empirischen Antwortmuster zu den latenten Klassen

Die vorausgehenden Abschnitte 4.1 bis 4.5 haben sich mit der Spezifikation und Prüfung eines Skalierungsmodells beschäftigt. Wurde ein brauchbares Modell gefunden (in unserem Beispiel eine 4-Klassenlösung), tritt das Problem der Zuordnung der empirischen Antwortmuster zu den latenten Klassen auf. Dabei ist zu beachten, daß die empirischen Antwortmuster nicht eindeutig einer bestimmten Klasse zugeordnet werden können, sondern nur mit einer bestimmten Wahrscheinlichkeit. Gesucht ist die bedingte Wahrscheinlichkeit für das Auftreten der latenten Klasse  $g$  gegeben einem empirischen Antwortmuster  $k$ . Diese bedingten Wahrscheinlichkeiten werden als Zuordnungswahrscheinlichkeiten bezeichnet. Ihre Berechnung soll am Beispiel des empirischen Antwortmusters (0,0,0,0,0) für die 4-Klassenlösung verdeutlicht werden.

Die bedingten Wahrscheinlichkeiten für das Auftreten des empirischen Antwortmusters  $k$  bei gegebener latenter Klasse  $g$  können aus der nachfolgenden Gleichung berechnet werden mit:

$$p(0,0,0,0,0/g) = p(x(A)=0/g) * p(x(B)=0/g) * p(x(C)=0/g) * p(x(D)=0/g) = p(x(C)=0/g)$$

Einsetzen der Werte der berechneten bedingten Wahrscheinlichkeiten aus der Abbildung 4-2 ergibt folgende Werte:

$$p(0,0,0,0,0/1) = (1-0.000) * (1-0.194) * (1-0.037) * (1-0.013) * (1-0.009) = .759$$



$$p(0,0,0,0,0/2) = .000$$

$$p(0,0,0,0,0/3) = .101$$

$$p(0,0,0,0,0/4) = .020$$

Im nächsten Schritt der Berechnung der Zuordnungswahrscheinlichkeiten werden die Wahrscheinlichkeiten des gemeinsamen Auftretens des empirischen Antwortmusters und der latenten Klassen  $g$  berechnet. Diese sind definiert als:

$$p(0,0,0,0,0,g) = p(0,0,0,0,0/g) \cdot \pi(g).$$

Für die errechneten latenten Klassenwahrscheinlichkeiten ergeben sich folgende Werte:

$$p(0,0,0,0,0,1) = 0.76 \cdot 0.523 = 0.397$$

$$p(0,0,0,0,0,2) = 0.00 \cdot 0.064 = 0.000$$

$$p(0,0,0,0,0,3) = 0.10 \cdot 0.253 = 0.025$$

$$p(0,0,0,0,0,4) = 0.02 \cdot 0.159 = 0.003$$

$$p(0,0,0,0,0,+) = \frac{\quad}{0.425}$$

Die gesuchten Zuordnungswahrscheinlichkeiten lassen sich nun nach der Formel berechnen mit:

$$p(g/0,0,0,0,0) = p(0,0,0,0,0,g)/p(0,0,0,0,0,+)$$

Einsetzen ergibt folgende Zuordnungswahrscheinlichkeiten:

$$p(1/0,0,0,0,0) = 0.307/0.425 = 0.93$$

$$p(2/0,0,0,0,0) = 0.000/0.425 = 0.00$$

$$p(3/0,0,0,0,0) = 0.025/0.425 = 0.06$$

$$p(4/0,0,0,0,0) = 0.003/0.425 = 0.01$$

Für das empirische Antwortmuster (0,0,0,0,0) besitzt die latente Klasse 1 (Personen mit geringem Aktivitätsniveau) - wie erwartet - die größte bedingte Zuordnungswahrscheinlichkeit (= 0.93). Die bedingten Zuordnungswahrscheinlichkeiten für alle möglichen Antwortmuster enthält Abbildung 4-3.

Die so für jede Person berechneten Zuordnungswahrscheinlichkeiten können in den weiteren Analysen als unabhängige oder abhängige Variablen verwendet werden. Werden sie auf Seite der unabhängigen Variablen in eine Analyse einbezogen, muß eine latente Klasse eliminiert werden, da ansonsten lineare Abhängigkeiten (Multikollinearitäten) vorliegen, da die Summe der Zuordnungswahrscheinlichkeiten gleich 1 ist.

Nachfolgendes SPSS-PC+ Programm berechnet Zuordnungswahrscheinlichkeiten für die 4-Klassenlösung. Es werden sowohl für Personen ohne fehlenden Werte als auch für Personen mit fehlenden Werten Zuordnungswahrscheinlichkeiten berechnet.

```
1      data list file = 'pol.dat' free/a,b,c,d,e.
2      compute pa1 = 0.000.
3      compute pb1 = 0.194.
4      compute pc1 = 0.037.
5      compute pd1 = 0.013.
6      compute pe1 = 0.009.
7      compute pa2 = 0.812.
8      compute pb2 = 0.950.
9      compute pc2 = 0.815.
10     compute pd2 = 0.620.
11     compute pe2 = 0.905.
12     compute pa3 = 0.529.
13     compute pb3 = 0.639.
14     compute pc3 = 0.171.
15     compute pd3 = 0.078.
16     compute pe3 = 0.227.
17     compute pa4 = 0.350.
18     compute pb4 = 0.865.
19     compute pc4 = 0.488.
20     compute pd4 = 0.582.
21     compute pe4 = 0.000.
22     compute pi1 = 0.523.
23     compute pi2 = 0.064.
24     compute pi3 = 0.253.
25     compute pi4 = 0.151.
26     count mm = a to e(-1).
27     compute pa = pa1.
28     compute pb = pb1.
29     compute pc = pc1.
30     compute pd = pd1.
31     compute pe = pe1.
32     if (a eq 0) pa = 1-pa.
33     if (b eq 0) pb = 1-pb.
34     if (c eq 0) pc = 1-pc.
35     if (d eq 0) pd = 1-pd.
36     if (e eq 0) pe = 1-pe.
37     if (a eq -1) pa = 1.
38     if (b eq -1) pb = 1.
39     if (c eq -1) pc = 1.
40     if (d eq -1) pd = 1.
41     if (e eq -1) pe = 1.
42     compute ppl = pa*pb*pc*pd*pe.
43     if (mm ge 3) ppl = -1.
```

```

44     compute pa = pa2.
45     compute pb = pb2.
46     compute pc = pc2.
47     compute pd = pd2.
48     compute pe = pe2.
49     if (a eq 0) pa = 1-pa.
50     if (b eq 0) pb = 1-pb.
51     if (c eq 0) pc = 1-pc.
52     if (d eq 0) pd = 1-pd.
53     if (e eq 0) pe = 1-pe.
54     if (a eq -1) pa = 1.
55     if (b eq -1) pb = 1.
56     if (c eq -1) pc = 1.
57     if (d eq -1) pd = 1.
58     if (e eq -1) pe = 1.
59     compute pp2 = pa*pb*pc*pd*pe.
60     if (mm ge 3) pp2 = -1.
61     compute pa = pa3.
62     compute pb = pb3.
63     compute pc = pc3.
64     compute pd = pd3.
65     compute pe = pe3.
66     if (a eq 0) pa = 1-pa.
67     if (b eq 0) pb = 1-pb.
68     if (c eq 0) pc = 1-pc.
69     if (d eq 0) pd = 1-pd.
70     if (e eq 0) pe = 1-pe.
71     if (a eq -1) pa = 1.
72     if (b eq -1) pb = 1.
73     if (c eq -1) pc = 1.
74     if (d eq -1) pd = 1.
75     if (e eq -1) pe = 1.
76     compute pp3 = pa*pb*pc*pd*pe.
77     if (mm ge 3) pp3 = -1.
78     compute pa = pa4.
79     compute pb = pb4.
80     compute pc = pc4.
81     compute pd = pd4.
82     compute pe = pe4.
83     if (a eq 0) pa = 1-pa.
84     if (b eq 0) pb = 1-pb.
85     if (c eq 0) pc = 1-pc.
86     if (d eq 0) pd = 1-pd.

```

```

87     if (e eq 0) pe = 1-pe.
88     if (a eq -1) pa = 1.
89     if (b eq -1) pb = 1.
90     if (c eq -1) pc = 1.
91     if (d eq -1) pd = 1.
92     if (e eq -1) pe = 1.
93     compute pp4 = pa*pb*pc*pd*pe.
94     if (mm ge 3) pp4 = -1.
95     missing value pp1,pp2,pp3,pp4(-1).
96     list variables = a to e,pp1 to pp4/cases from 1 to 50.
97     compute pp1 = pp1*pi1.
98     compute pp2 = pp2*pi2.
99     compute pp3 = pp3*pi3.
100    compute pp4 = pp4*pi4.
101    compute pp = pp1 + pp2 + pp3 + pp4.
102    list variables = a to e,pp1 to pp4,pp/cases from 1 to 50.
103    compute pp1 = pp1/pp.
104    compute pp2 = pp2/pp.
105    compute pp3 = pp3/pp.
106    compute pp4 = pp4/pp.
107    list variables = a to e,pp1 to pp4/cases from 1 to 50.

```

Die Programmzeilen bedeuten:

- 1        Definition der Daten.
- 2 - 6    Eingabe der bedingten Wahrscheinlichkeiten  $p(x(A)=1/1)$ ,  $p(x(B) = 1/1)$ ,... für die erste latente Klasse. (SPSS-PC+ Variablen: pa1, pb1...)
- 7 - 11   Eingabe der bedingten Wahrscheinlichkeiten  $p(x(A)=1/2)$ ,  $p(x(B) = 1/2)$ , ... für die zweite latente Klasse. (SPSS-PC+ Variablen: pa2, pb2...)
- 12- 16   Eingabe der bedingten Wahrscheinlichkeiten  $p(x(A)=1/3)$ ,  $p(x(B) = 1/3)$ , ... für die dritte latente Klasse. (SPSS-PC+ Variablen: pa3, pb3...)
- 17-21   Eingabe der bedingten Wahrscheinlichkeiten  $p(x(A)=1/4)$ ,  $p(x(B) = 1/4)$ , ... für die vierte latente Klasse. (SPSS-PC+ Variablen: pa4, pb4...)
- 22 - 25   Eingabe der latenten Klassenwahrscheinlichkeiten  $n(1)$ ,  $n(2)$ ,... (SPSS-PC+Variablen: pi1, pi2...)
- 26       Berechnen der Zahl der fehlenden Werte.
- 27 - 43   Berechnen der bedingten Wahrscheinlichkeit für das Auftreten des empirischen Antwortmusters gegeben der latenten Klasse 1. Dabei wird wie folgt vorgegangen:
- 27-31   Die Variablen pa, pb, ..., erhalten zunächst den Wert der beding-

- ten Wahrscheinlichkeiten  $p(x(A)= 1/1), \dots$ . Es wird also angenommen, daß die Ausprägung »1« vorliegt.
- 32 - 36 Es wird geprüft, ob die Ausprägung »0« vorliegt. Ist dies beispielsweise im Item A der Fall, wird  $p_a$  gleich  $1-p_a$  gesetzt.
- 37 - 41 Es wird geprüft, ob in einem Item ein fehlender Wert auftritt. Ist dies beispielsweise im Item A der Fall, wird  $p_a$  gleich 1 gesetzt. Dies führt dazu, daß der fehlende Wert die nachfolgende Multiplikation nicht beeinflußt.
- 42 Die bedingte Wahrscheinlichkeit für das Auftreten des untersuchten Antwortmusters gegeben der latenten Klasse 1 wird berechnet. Tritt beispielsweise im Item A ein fehlender Wert auf, während alle anderen Items valide Werte besitzen, wird die bedingte Wahrscheinlichkeit nur aus den Variablen  $p_b, p_c, p_d$  und  $p_e$  berechnet, da  $p_a = 1.000$  gesetzt wurde.
- 43 Es wird geprüft, ob ausreichend valide Werte für die Berechnung der bedingten Wahrscheinlichkeit für das Auftreten des empirischen Antwortmusters gegeben der latenten Klasse 1 vorliegen. Ist die Zahl der fehlenden Werte größer/gleich 3, wird die bedingte Wahrscheinlichkeit auf -1 (= Kode für fehlenden Wert) gesetzt. Die maximale Zahl fehlender Werte muß vom Benutzer festgelegt werden.
- 44 - 60 Berechnen der bedingten Wahrscheinlichkeit für das Auftreten des untersuchten empirischen Antwortmusters gegeben der latenten Klasse 2. Es wird analog wie bei der latenten Klasse 1 vorgegangen.
- 61-77 Berechnen der bedingten Wahrscheinlichkeit für das Auftreten des untersuchten empirischen Antwortmusters gegeben der latenten Klasse 3.
- 78 - 94 Berechnen der bedingten Wahrscheinlichkeit für das Auftreten des untersuchten empirischen Antwortmusters gegeben der latenten Klasse 4.
- 85 Definition der Ausprägung für fehlenden Wert (= -1).
- 86 Kontrollausgabe.
- 97 -100 Berechnen der gemeinsamen Auftrittswahrscheinlichkeit des untersuchten empirischen Antwortmusters und der latenten Klasse.
- 101 Berechnen der Summe der gemeinsamen Auftrittswahrscheinlichkeiten (z.B.  $p(0,0,0,0,0,+)$ ).
- 102 Kontrollausgabe.
- 103-106 Berechnen der Zuordnungswahrscheinlichkeiten.
- 107 Kontrollausgabe.

Die Ergebnisse der von dem SPSS/PC+ Programm vorgenommenen Berechnung der Zuordnungswahrscheinlichkeiten bei Antwortmustern mit fehlenden Werten ist in der Abbildung 4-4 dargestellt.

## 4.7 Anwendungsmöglichkeiten in der historischen Sozialforschung

Ein Anwendungsbereich der Analyse latenter Klassen ist durch die in Abschnitt 2.9 dargestellten Anwendungsmöglichkeiten gegeben. Bei der Inhaltsanalyse besteht ferner die Möglichkeit die Übereinstimmung von »Ratern« zu prüfen (BACHER 1990, sowie den Überblicksaufsatz von LANGEHEINE 1988:89-91). Weitere Anwendungsfelder sind in folgenden Bereichen gegeben:

- Soziale Klassenanalyse. Mit Hilfe der Analyse latenter Klassen kann geprüft werden, ob eine theoretisch angenommene Zuordnung von Berufen zu sozialen Klassen oder sozialen Schichten empirisch haltbar ist. Dazu kann beispielsweise eine Mobilitätsmatrix untersucht werden, wobei angenommen wird, daß soziale Klassen sozial geschlossene statistische Gebilde sind (siehe LANGEHEINE 1988: 100-101).
- Mobilitätsanalysen. Bei der Untersuchung einer Mobilitätsmatrix mit einem latenten Klassenmodell werden konstante Klassen angenommen. Mitunter können aber die theoretischen Vorstellungen darin bestehen, daß Übergänge zwischen den latenten Klassen auftreten. In diesem Fall ist die Anwendung der Analyse latenter Klassen nicht angemessen. Zur Schätzung der Übergänge zwischen den latenten Klassen sowie der Zuordnung von Berufen zu latenten Klassen kann aber ein latentes Markovmodell verwendet werden (LANGEHEINE 1988:101-103; LAZARFELD/HENRY 1968: 245-263). Dieses geht von folgenden Annahmen aus:

- Es gibt eine bestimmte Anzahl latenter Klassen.
- Zwischen den latenten Klassen sind Übergänge möglich.

Voraussetzung für die Anwendung ist, daß Angaben für mindestens drei Zeitpunkte vorliegen. Das latente Markovmodell ist in dem Programmsystem PANMARK enthalten.

## 4.8 Exkurs: Das Programm PANMARK

Das Programm PANMARK (VAN DE POL/LANGEHEINE/DE JONG 1989) wurde ursprünglich zur Analyse latenter Markovketten entwickelt (VAN DE POL/DE LEEUW 1986) und in der Folge zu einem Modell gemischter latenter Markovketten erweitert. PANMARK enthält neben dem latenten Markovmodell und dem gemischten latenten Markovmodell als weiteres Submodell die Analyse latenter Klassen (in PANMARK: static factor model). Mit diesem lassen sich sowohl dichotome als auch polytome

nominalskalierte Variablen untersuchen. PANMARK wird über einen SE-TUP-File gesteuert. Dieser enthält u.a. den Namen der Datendatei, den Namen der Datei der Parameterrestriktionen und den Namen der Datei der Startwerte.

Datendatei: Sie wird für alle Analysen benötigt und enthält die empirischen Auftretishäufigkeiten der empirischen Antwortmuster, die wie folgt eingegeben werden müssen:

$h(0,0,0,0,0)$ ,  $h(0,0,0,0,1)$ ,  $h(0,0,0,1,0)$ ,  $h(0,0,0,1,1)$ ,  $h(0,0,1,0,0)$ ,  $h(0,0,1,0,1)$ ,  $h(0,0,1,1,0)$ ,  $h(0,0,1,1,1)$ ,  $h(0,1,0,0,0)$ ...

Zunächst werden also die Ausprägungen des letzten Items variiert, dann die des zweitletzten usw. Zu beachten ist, daß die Datendatei alle möglichen Antwortmuster enthalten muß, also auch Antwortmuster mit einer Häufigkeit gleich 0. Diese Tatsache macht die Nachbearbeitung einer mit SPSS-PC+ erzeugten Datendatei erforderlich, da diese nur Antwortmuster mit Häufigkeiten größer 0 erzeugt. Das entsprechende SPSS-PC+ Programm lautet:

```

1      data list file = 'pol.dat' free/a,b,c,d,e.
2      missing value a to e(-1).
3      count ig = a to e(missing).
4      select if (ig eq 0).
5      compute gg = 1.
6      aggregat outfile = *
7          /break = a,b,c,d,e
8          /nn = sum(gg).
9      list variables = all.
10     set results = 'goodman.dat'.
11     write var = nn.
```

Es erzeugt eine Datei mit 31 Datensätzen. Das empirisch nicht besetzte Antwortmuster (0,0,1,1,1) muß ergänzt werden (siehe Abbildung 4-5).

Datei der Parameterrestriktionen. PANMARK enthält drei Typen von Parametern:

- frei variierbare Parameter.
- auf einen bestimmten Wert gefixte Parameter, wie z.B.:  $p(x(A)=1/2) = 0$
- gleichgesetzte Parameter, wie z.B.:  $p(x(B)=1/2) = p(x(C)=1/2)$

Für diese drei Typen werden in PANMARK folgende Codes verwendet:

- 0 = frei variierbare Parameter
- 1 = auf einen bestimmten Wert gefixte Parameter
- 1 = erste Menge gleichgesetzter Parameter

- 2 = zweite Menge gleichgesetzter Parameter
- 3 = dritte Menge gleichgesetzter Parameter
- usw. ...

Für das zu Beginn dieses Kapitels behandelte Modell von GOODMAN sieht die Datei der Parameterrestriktionen wie in der Abbildung 4-6 dargestellt aus.

Die Datei der Parameterrestriktionen wird nicht benötigt, wenn keine Parameterrestriktionen vorliegen. Sie kann mit PANMARK wie folgt erzeugt werden (VAN DE POL/LANGEHEINE/DE JONG 1989: 25): Zunächst wird eine Schätzung ohne Parameterrestriktionen durchgeführt. Dabei wird eine Datei erzeugt, welche die Schätzwerte enthält und wie die gesuchte Datei der Parameterrestriktionen aufgebaut ist. Diese Datei besitzt die Erweiterung .EST. Mit einem Editor können die entsprechenden Änderungen vorgenommen werden.

Datei der Startwerte. Diese Datei wird nur benötigt, wenn bestimmte Parameter gleich einem bestimmten Wert gesetzt werden. Sie kann analog wie die Datei der Parameterrestriktionen erzeugt werden. Für das GOODMAN-Modell besitzt sie folgende Struktur (Abbildung 4-7).

Zu beachten ist, daß für die anderen Parameter Startwerte ungleich 0 oder 1 eingegeben werden, da sie ansonsten in der nachfolgenden Parameterschätzung nicht mehr geändert werden.

Die Parameterschätzung wird in PANMARK mit dem bereits beschriebenen EM-Algorithmus (siehe Abschnitt 3.1.3) durchgeführt. Zusätzlich berechnet PANMARK noch Standardabweichungen der geschätzten Parameter und führt nach der in Abschnitt 4.3 dargestellten Logik einen empirischen Identifikationstest durch.

Ausgegeben werden können u.a. die geschätzten Parameter, die Standardabweichungen der geschätzten Parameter, Chiquadrat-Prüfgrößen, deren Signifikanz sowie die empirischen und erwarteten Häufigkeiten der Antwortmuster zur Fehlerkontrolle.



Abbildung 4-1: Parameter des Modells von GOODMAN

	latente Gruppe (Klasse) g						
	0	1	2	3	4	5	6
$p(x(A)=1/g) =$	0	0	1	1	1	1	<u><math>p(x(A)=1/6)</math></u>
$p(x(B)=1/g) =$	0	1	1	1	1	1	<u><math>p(x(B)=1/6)</math></u>
$p(x(C)=1/g) =$	0	0	0	1	1	1	<u><math>p(x(C)=1/6)</math></u>
$p(x(D)=1/g) =$	0	0	0	0	1	1	<u><math>p(x(D)=1/6)</math></u>
$p(x(E)=1/g) =$	0	0	0	0	0	1	<u><math>p(x(E)=1/6)</math></u>

Auftrittswahrscheinlichkeiten  
der latenten  
Klassen

	<u><math>\pi(0)</math></u>	<u><math>\pi(1)</math></u>	<u><math>\pi(2)</math></u>	<u><math>\pi(3)</math></u>	<u><math>\pi(4)</math></u>	<u><math>\pi(5)</math></u>	<u><math>\pi(6)</math></u>
--	----------------------------	----------------------------	----------------------------	----------------------------	----------------------------	----------------------------	----------------------------

unterstrichene Werte = zu schätzende Parameter.  
Anordnung der Items A bis E nach der Schwierigkeit:  
 $B < A < C < D < E$  (siehe Kapitel 2)

Anmerkung zur Notation:

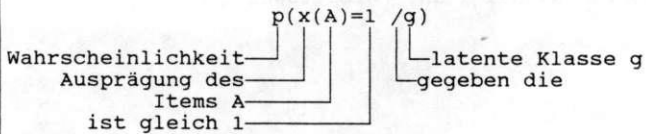


Abbildung 4-2: Ergebnisse der Analyse latenter Klassen bei 4 latenten Klassen

	latente Gruppe (Klasse) g			
	1	2	3	4
$p(x(A)=1/g) =$	0.000	0.812	0.529	0.350
$p(x(B)=1/g) =$	0.194*	0.950	0.639	0.865
$p(x(C)=1/g) =$	0.037	0.815	0.171	0.488
$p(x(D)=1/g) =$	0.013	0.620	0.078	0.582
$p(x(E)=1/g) =$	0.009	0.905	0.227	0.000
Auftrittswahrscheinlichkeiten der latenten Klassen	0.523**	0.064	0.253	0.159

\* Interpretation: Die bedingte Wahrscheinlichkeit des Auftretens einer "1" der latenten Klasse 1 ist 0.194.

\*\* Interpretation: 52,3% der untersuchten Personen gehören der latenten Klasse 1 an. (Arbeitswert: 0.523)

Abbildung 4-3: Zuordnungswahrscheinlichkeiten der empirischen Antwortmuster zu den latenten Klassen

Items					Zuordnungswahrscheinlichkeiten zu den latenten Klassen			
A	B	C	D	E	1	2	3	4
0	0	0	0	0	.93 *	.00	.06	.01
1	0	0	0	0	.00	.00	.95 *	.05
0	1	0	0	0	.60 *	.00	.28	.11
0	0	1	0	0	.66 *	.00	.23	.12
0	0	0	1	0	.46 *	.00	.19	.35
0	0	0	0	1	.32	.00	.67 *	.00
1	1	0	0	0	.00	.01	.83 *	.16
1	0	1	0	0	.00	.01	.79 *	.20
1	0	0	1	0	.00	.01	.53 *	.47
1	0	0	0	1	.00	.02	.98 *	.00
0	1	1	0	0	.12	.01	.30	.57
0	1	0	1	0	.04	.00	.12	.83 *
0	1	0	0	1	.06	.05	.89 *	.00
0	0	1	1	0	.05	.01	.10	.85 *
0	0	1	0	1	.08	.09	.83 *	.00
0	0	0	1	1	.06	.08	.85 *	.00
1	1	1	0	0	.00	.07	.49 *	.44
1	1	0	1	0	.00	.03	.23	.74 *
1	1	0	0	1	.00	.17	.83 *	.00
1	0	1	1	0	.00	.05	.19	.76 *
1	0	1	0	1	.00	.30	.70 *	.00 **
1	0	0	1	1	.00	.28	.72 *	.00
0	1	1	1	0	.00	.02	.03	.95 *
0	1	1	0	1	.01	.54 *	.46	.00
0	1	0	1	1	.00	.51 *	.48	.00
0	0	1	1	1	.00	.68 *	.32	.00
1	1	1	1	0	.00	.15	.05	.80 *
1	1	1	0	1	.00	.82 *	.18	.00
1	1	0	1	1	.00	.80 *	.20	.00
1	0	1	1	1	.00	.89 *	.11	.00
0	1	1	1	1	.00	.96 *	.04	.00
1	1	1	1	1	.00	.99 *	.01	.00

\* größte Zuordnungswahrscheinlichkeit.

\*\* Interpretation: Das Antwortmuster (1,0,1,0,1) gehört mit einer Wahrscheinlichkeit von 0 der latenten Klasse 1 an, mit einer Wahrscheinlichkeit von 0.3 der latenten Klasse 2, mit einer Wahrscheinlichkeit von 0.7 der latenten Klasse 3 und mit einer Wahrscheinlichkeit von 0.0 der latenten Klasse 4.

Abbildung 4-4: Berechnete Zuordnungswahrscheinlichkeiten für Antwortmuster mit fehlenden Werten

Items					Zuordnungswahrscheinlichkeiten zu den latenten Klassen			
A	B	C	D	E	1	2	3	4
0	1	0	0	-1	.56	.00	.34	.10
-1	1	0	0	0	.44	.00	.44	.13
-1	-1	-1	-1	-1	.	.	.	.
1	1	-1	0	0	.00	.02	.74	.23
1	0	1	0	-1	.00	.08	.77	.15
1	-1	1	0	0	.00	.05	.57	.38
-1	0	0	0	0	.87	.00	.12	.01
0	-1	1	0	1	.02	.43	.55	.00
-1	-1	-1	-1	-1	.	.	.	.
0	0	-1	0	0	.92	.00	.07	.01
-1	-1	0	0	0	.73	.00	.22	.05
0	0	0	-1	0	.92	.00	.06	.02
1	0	-1	0	0	.00	.00	.92	.08
-1	-1	-1	-1	-1	.	.	.	.
0	-1	-1	-1	-1	.	.	.	.
0	1	0	0	-1	.56	.00	.34	.10
0	1	0	0	-1	.56	.00	.34	.10

usw..

"-1" = fehlender Wert

"." = nicht berechnet, da mehr als zwei fehlende Werte

Abbildung 4-5: Erzeugen der Datendatei für PANMARK aus SPSS-PC+

-----  
mit SPSS-PC+ berechnet (Ergebnisse von AGGREGATE)

842.00  
22.00  
22.00  
4.00  
46.00  
5.00  
12.00  
317.00  
27.00  
62.00  
3.00  
62.00  
12.00  
52.00  
12.00  
56.00  
20.00  
7.00  
1.00  
12.00  
7.00  
5.00  
2.00  
130.00  
30.00  
39.00  
15.00  
42.00  
34.00  
34.00  
44.00

Nach "12.00" Einfügen einer Zeile mit der Besetzungszahl 0.00, da Antwortmuster (0,0,1,1,1) empirisch nicht auftritt. Einfügestelle ist durch vorausgehenden list-Befehl bekannt.

Abbildung 4-6: Datei der Parameterrestriktionen für das Modell von GOODMAN

---

goodman 77  0 0 0 0 0 0 0  -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 0 0	Name des SETUP-Files Zahl der Parameter insgesamt  Auftretswahrscheinlichkeiten der latenten Klassen (frei variierende Paramter)  bedingte Wahrscheinlichkeiten $p(x(A)=0/g)$ und $p(x(A)=1/g)$ für alle latenten Klassen:  Werden für die ersten 6 Klassen (skalierbare Personen) auf einen bestimmten Wert gesetzt.  Für die letzte Klasse (nicht skalierbare Personen) frei variierbar.
---	---

Für die Items B bis E folgt analoge Definition.

---

Abbildung 4-7: Datei der Startwerte für PANMARK

-----							
goodman							Name des SETUP-Files.
77							Zahl der Parameter insgesamt.
0.30	0.07	0.18	0.07	0.13	0.08	0.12	Startwerte für Auftretswahrscheinlichkeiten der latenten Klassen (= gerundete Werte aus der Schätzung ohne Restriktionen).
1.0	0.0						Durch Datei der Parameterrestriktionen festgelegte nicht mehr zuverändernde "Startwerte".
1.0	0.0						
0.0	1.0						
0.0	1.0						
0.0	1.0						
0.0	1.0						
0.7	0.3						Startwert für variierendes Parameter $p(x(A)=0/6)$ und $p(x(A)=1/6)$ (=gerundete Werte aus der Schätzung ohne Restriktionen).

Für die Items B bis E analoge Definitionen entsprechend Abbildung 4-1.

## 5. Faktorenanalyse

### 5.1 Theoretische Vorstellungen und Modellspezifikation

In den vorausgehenden drei Kapiteln (Kapitel 2 bis 4) wurden Skalierungsverfahren für dichotome und polytome ordinalskalierte Variablen (Guttman-, Mokken-, Raschskalierung und Skalierung nach Proctor) sowie für nominalskalierte Variablen (Analyse latenter Klassen) behandelt. Das Kapitel 5 behandelt die Faktorenanalyse als Skalierungsverfahren für ratio- und intervallskalierte Variablen. Intervall- und ratioskalierte Variable werden in der historischen Sozialforschung i.d.R. bei der Untersuchung von statistischem Zahlenmaterial (Volkszählungen, sozialstatistische Erhebungen u.ä.) auftreten. Als Beispiel zur Darstellung der Faktorenanalyse soll das Beispiel des Abschnitts 1.1 zur Messung der Industrialisierung dienen. Die Daten wurden DENZ (1989: 33-35) entnommen. Dabei bilden die politischen Bezirke und Städte des österreichischen Teils der Habsburgermonarchie zu Ende des vorigen Jahrhunderts (Volkszählung 1880) die Untersuchungseinheiten. Die zur Messung der Industrialisierung ausgewählten Indikatoren sind:

- Anteil der Zuwanderer an der Gesamtbevölkerung (=  $x(1)$ )
- Anteil der Abwanderer an der Gesamtbevölkerung (=  $x(2)$ )
- Anteil der in der Land- und Forstwirtschaft Erwerbstätigen (=  $x(3)$ )
- Anteil der im Bergbau und Hüttenwesen Erwerbstätigen (=  $x(4)$ )
- Anteil der in der Textilindustrie Erwerbstätigen (=  $x(5)$ )
- Anteil der Tagelöhner an den in der Land- und Forstwirtschaft Erwerbstätigen (=  $x(6)$ )
- Anteil der Selbständigen an den in der Land- und Forstwirtschaft Erwerbstätigen (=  $x(7)$ )
- Anteil der Selbständigen an den in Industrie und Gewerbe Erwerbstätigen (=  $x(8)$ )

Zwischen diesen Indikatoren und dem theoretischen Begriff der Industrialisierung werden die in Abschnitt 1.1 (siehe Abbildung 1-1) dargestellten Beziehungen angenommen. Diese können wie folgt weiter spezifiziert werden:

- Jede Untersuchungseinheit  $g$  besitzt auf der (latenten) Meßdimension »Industrialisierung« einen Skalenwert  $\mathfrak{I}(g)$ , der den »Grad« der Industrialisierung mißt.
- Die Messungen in den Indikatoren sind fehlerbehaftet. Die Ausprägung einer Untersuchungseinheit in einem Indikator  $i$  kann man sich wie folgt zusammengesetzt denken:



$$x(i,g) = \alpha(i) \cdot \vartheta(g) + \delta(i,g)$$

mit

$x(i,g)$  = Ausprägung der Untersuchungseinheit  $g$  im Indikator  $i$ .

$\alpha(i)$  = »Meßgenauigkeit« (Faktorladung) des Indikators  $i$

$\vartheta(g)$  = Skalenwert der Untersuchungseinheit  $g$  auf der (latenten) Meßdimension

$\delta(i,g)$  = Ausprägung des Meßfehlers in der Untersuchungseinheit  $g$  und im Indikator  $i$ .

Das fiktive Rechenbeispiel der Abbildung 5-1 dient der Verdeutlichung der Logik. Die Untersuchungseinheit 1 hat dabei einen Skalenwert von 1.0 auf der Meßdimension, bei der Messung des Indikators  $i$  (z.B. »Anteil Zuwanderer«) tritt ein Meßfehler in der Höhe von 0.1 auf. Der empirische Wert ist  $x(i,1) = 0.4 \cdot 1.0 + 0.1 = 0.5$ , da eine Faktorladung von 0.4 angenommen wurde.

Das eben spezifizizierte Modell entspricht dem Modell der Faktorenanalyse. Mit den bisher getroffenen Spezifikationen ist das Modell nicht identifiziert, da es mehr unbekannt Parameter als theoretisch vorhandene empirische Informationen enthält. (Die notwendige Identifikationsbedingung, daß weniger unbekannt Parameter als theoretisch vorhandene empirische Information vorliegt, ist nicht erfüllt.) Bei 300 Untersuchungseinheiten und 8 Indikatoren beispielsweise müßten folgende Parameter geschätzt werden:

300	Skalenwerte $\vartheta(g)$
8	Faktorladungen $\alpha(i)$
<u>2400</u>	<u>Ausprägungen der Meßfehler <math>\delta(i,g)</math></u>
2708	unbekannte Parameter

Diesen 2708 unbekannt Parameter stehen  $300 \cdot 8 = 2400$  empirische Beobachtungen gegenüber. Die Faktorenanalyse trifft deshalb bestimmte (allgemeine) Annahmen, die für alle faktoranalytischen Modelle gelten, um eine Identifikation zu gewährleisten. Bei der konfirmatorischen Faktorenanalyse besteht darüber hinaus die Möglichkeit für weitere Restriktionen, die auch zu einer Abschwächung der allgemeinen Annahmen führen können.

## 5.2 Modellannahmen der Faktorenanalyse

Um die Zahl der unbekannt Parameter einzuschränken, werden bei der Faktorenanalyse allgemein folgende Annahmen getroffen:

- Die Skalenwerte  $\mathfrak{g}$  der Untersuchungseinheiten  $g$  sind Ausprägungen (Realisierungen) einer Zufallsvariablen  $\mathfrak{g}$ . Diese soll einen Erwartungswert von 0 ( $E(\mathfrak{g}) = 0$ ) und eine zu schätzende Varianz  $\text{VAR}(\mathfrak{g})$  besitzen.
- Die Ausprägungen  $\delta(i,g)$  der Meßfehler der Untersuchungseinheiten  $g$  im Indikator  $i$  werden ebenfalls als Realisierungen einer Zufallsvariablen  $\delta(i)$  mit einem Erwartungswert von 0 und einer zu schätzenden Varianz  $\text{VAR}(\delta(i))$  gedacht.
- Die Meßfehler treten zufällig auf.

Aus diesen drei Annahmen folgt:

- Die empirischen Ausprägungen  $x(i,g)$  der Untersuchungseinheiten  $g$  in den Indikatoren  $i$  sind ebenfalls Realisierungen einer Zufallsvariablen  $x(i)$  mit einem Erwartungswert gleich 0. Formal ausgedrückt gilt (1):

$$E(x(i,g)) = E(x(i)) = E(\alpha(i) \cdot \mathfrak{g} + \delta(i,g)) = 0.$$

Für die Anwendung der Faktorenanalyse bedeutet dies, daß die Indikatoren mittelwertzentriert werden müssen.

- Da die Verteilungen der Untersuchungseinheiten  $g$  auf der Meßdimension  $\mathfrak{g}$  und in den Meßfehlern  $\delta(i)$  durch den Erwartungswert und die Varianz beschrieben werden, ist zu fordern, daß diese Kenngrößen »aussagefähig« sind. Dies ist u.a. dann der Fall, wenn die Verteilungen eingipfelig und um den jeweiligen Erwartungswert symmetrisch sind. Die Normalverteilung erfüllt diese beiden Bedingungen. Aus diesem Grund wird deshalb vielfach eine multidimensionale Normalverteilung der Zufallsvariablen  $\mathfrak{g}(i)$  und  $\mathfrak{g}$  angenommen (2). Aus der Annahme der multidimensionalen Normalverteilung der Zufallsvariablen  $\mathfrak{g}(i)$  und  $\mathfrak{g}$  folgt, daß die empirischen Ausprägungen  $x(i,g)$  ebenfalls multidimensional normalverteilt sind, da sie Summen von normalverteilten Zufallsvariablen sind.

Technisch bedeutet dies, daß mitunter eine Transformation der Indikatoren auf eine Normalverteilung bzw. zumindest auf eine symmetrische eingipfelige Verteilung erforderlich sein kann.

- Die Annahme der »Zufälligkeit« der Meßfehler bedeutet, daß das Auftreten eines Meßfehlers im Indikator  $i$  von dem Auftreten eines Meßfehlers im Indikator  $j$  unabhängig ist. Daraus folgt, daß die Kovarianz  $\text{KOV}(\delta(i)X\delta(j))$  der Meßfehler  $\delta(i)$  und  $\delta(j)$  gleich 0.0 ist. Ferner impliziert die Annahme der »Zufälligkeit« der Meßfehler, daß das Auftreten der Meßfehler nicht von den Skalenwerten  $\mathfrak{g}$  abhängt. Die Kovarianzen  $\text{KOV}(\mathfrak{g},\delta(i))$  der Meßfehler  $\delta(i)$  mit der Meßdimension  $\mathfrak{g}$  müssen somit für alle  $\delta(i)$  ebenfalls 0 sein. Wären diese beiden Annahmen nicht erfüllt, könnte nicht von zufälligen Meßfehlern gesprochen werden.

Aus den getroffenen Annahmen können aus dem Modell erwartete Varianzen und Kovarianzen der Indikatoren abgeleitet werden (3):

$$\begin{aligned} \text{VAR}(x(i)) &= \alpha(i)^2 \cdot \text{VAR}(\mathfrak{g}) + \text{VAR}(\delta(i)) \\ \text{KOV}(x(i), x(j)) &= \alpha(i) \cdot \alpha(j) \cdot \text{VAR}(\mathfrak{g}) \end{aligned}$$

Aus den Modellgleichungen ist ersichtlich, daß die erwarteten Kovarianzen der Indikatoren nur von den Faktorladungen  $\alpha(i)$  und von der Varianz  $\text{VAR}(\mathfrak{g})$  abhängen. Die Varianz eines Indikators hängt zudem noch von der Varianz des Zufallsfehlers ab. Zu beachten ist, daß in den Modellgleichungen von den Ausprägungen  $x(i, g)$  der Untersuchungseinheiten  $g$  in den Indikatoren  $i$  zu den Varianzen und Kovarianzen der Indikatoren übergegangen wurde. Die Skalenwerte  $\mathfrak{g}(g)$  und die Meßfehler  $\delta(i, g)$  sind in den Gleichungen nicht mehr enthalten. Dadurch wird die Zahl der zu schätzenden Parameter reduziert: Bei 8 Indikatoren können  $8 \cdot 7 = 56$  Kovarianzen und 8 Varianzen berechnet werden. Diesen stehen folgende unbekannte Parameter gegenüber:

1	Varianz von $\mathfrak{g}$ (= $\text{VAR}(\mathfrak{g})$ )
8	Faktorladungen $\alpha(i)$ ( $i = 1, 2, \dots, 8$ )
8	Meßfehlervarianzen (= $\text{VAR}(\delta(i))$ mit $i = 1, 2, \dots, 8$ )
17	unbekannte Parameter

Die notwendige Identifikationsbedingung (mehr theoretisch vorhandene empirische Informationen als unbekannte Parameter) ist somit erfüllt. Dennoch sind die Parameter nicht identifiziert (4). Um dies zu erreichen, muß entweder die Varianz  $\text{VAR}(\mathfrak{g})$  gleich 1.0 (oder einem anderen konstanten Wert) oder eine Faktorladung  $\alpha(i)$  gleich 1.0 (oder einem anderen konstanten Wert) gesetzt werden. Dadurch wird die Skaleneinheit der Zufallsvariablen  $\mathfrak{g}$  definiert und i.d.R. eine Identifikation erreicht. Welche dieser beiden Restriktionen gewählt wird, hängt von dem gewählten Verfahren der Faktorenanalyse ab. Bei explorativen Verfahren (siehe Abschnitt 8.2) besteht nur die Möglichkeit, die Varianz  $\text{VAR}(\mathfrak{g})$  gleich 1.0 zu setzen, da die Indikatoren i.d.R. auch schwach auf weiteren gemeinsamen Faktoren laden und deshalb durch die Restriktion  $\alpha(i) = 1.0$  keine Skalendefinition der Faktoren erreicht werden kann. Dies wäre nämlich nur dann der Fall, wenn der Indikator  $i$  nur auf einen Faktor lädt.

Bei explorativen Verfahren wird wegen der Normierung von  $\text{VAR}(\mathfrak{g}) = 1.0$  gefordert, daß die Indikatoren ebenfalls standardisiert sind, sodaß die Ergebnisse nicht von den Varianzen der Indikatoren abhängen. Technisch bedeutet dies, daß die Indikatoren mit

$$z(i,g) = \frac{x(i,g) - m(x(i))}{\sqrt{\text{var}(x(i))}}$$

mit

$x(i,g)$  = Ausprägung der Untersuchungseinheit  $g$  im Indikator  $i$

$m(x(i))$  = Mittelwert des Indikators  $i$

$\text{var}(x(i))$  = Varianz des Indikators  $i$

standardisiert werden. Die Modellgleichungen vereinfachen sich dann zu:

$$\text{VAR}(z(i)) = 1.0 = \alpha(i)^2 + \text{VAR}(\delta(i))$$

$$\text{KOV}(z(i),z(j)) = \text{KOR}(z(i),z(j)) = \alpha(i) \cdot \alpha(j)$$

Anstelle der Kovarianzen werden die Korrelationen zwischen den Indikatoren untersucht.

### 5.3 Das allgemeine faktorenanalytische Modell

Das für die Indikatoren der Industrialisierung spezifizierte Modell stellt nur ein Submodell innerhalb von faktorenanalytischen Modellen dar. Das allgemeine faktorenanalytische Modell besitzt folgende Struktur:

$$x(i,g) = \alpha(i,1) \cdot \vartheta(1,g) + \alpha(i,2) \cdot \vartheta(2,g) + \dots + \alpha(i,p) \cdot \vartheta(p,g) \\ + \beta(i,1) \cdot \Omega(1,g) + \beta(i,2) \cdot \Omega(2,g) + \dots + \beta(i,3) \cdot \Omega(q,g) \\ + \delta(i,g)$$

mit

$\alpha(i,k)$  = Faktorladung des Indikators  $i$  auf dem gemeinsamen Faktor  $k$

$\vartheta(k,g)$  = Skalenwert der Untersuchungseinheit  $g$  auf dem gemeinsamen Faktor  $k$

$\beta(i,1)$  = Faktorladung des Indikators  $i$  auf dem spezifischen Faktor 1

$\Omega(1,g)$  = Skalenwert der Untersuchungseinheit  $g$  auf dem spezifischen Faktor 1

$\delta(i,g)$  = Ausprägung des Meßfehlers in der Untersuchungseinheit  $g$  und im Indikator  $i$

Das allgemeine faktorenanalytische Modell besteht aus:

$p$  gemeinsamen Faktoren  $\vartheta(k)$ .

$q$  spezifischen Faktoren  $\Omega(l)$ . Ein spezifischer Faktor ist eine latente Dimension, auf der nur ein Indikator lädt. Er mißt also das »Spezifische« dieses Indikators,

$m$  Meßfehlern  $\delta(i)$ . ( $m$  = Zahl der Indikatoren)

Bezüglich der spezifischen Faktoren wird angenommen, daß sie von den gemeinsamen Faktoren, den Meßfehlern sowie untereinander unabhängig

sind. Andernfalls könnten sie nicht als »spezifisch« bezeichnet werden. Diese Annahme führt dazu, daß die spezifischen Faktoren eines Indikators nicht von dessen Meßfehler getrennt werden können. Die Modelle

$$x(i,g) = \alpha(i,1) \cdot \vartheta(1,g) + \alpha(i,2) \cdot \vartheta(2,g) + \dots + \alpha(i,p) \cdot \vartheta(p,g) \\ + \beta(i,1) \cdot \Omega(1,g) + \beta(i,2) \cdot \Omega(2,g) + \dots + \beta(i,3) \cdot \Omega(q,g) \\ + \delta(i,g)$$

und

$$x(i,g) = \alpha(i,1) \cdot \vartheta(1,g) + \alpha(i,2) \cdot \vartheta(2,g) + \dots + \alpha(i,p) \cdot \vartheta(p,g) \\ + \delta(i,g)$$

lassen sich also nicht unterscheiden. Es genügt also Modelle zu betrachten, die nur aus gemeinsamen Faktoren und Meßfehlern bestehen.

Ein Beispiel für ein faktorenanalytisches Modell mit zwei gemeinsamen Faktoren gibt die Abbildung 5-2. Dabei wurde die in der Faktorenanalyse übliche Matrixschreibweise verwendet. Das Modell besteht aus drei Matrizen:

- Faktorladungsmatrix. Diese enthält als Elemente die Faktorladungen der Indikatoren. In dem Beispiel laden die Indikatoren  $x(1)$  und  $x(2)$  nur auf dem gemeinsamen Faktor 1, die Indikatoren  $x(4)$  und  $x(5)$  nur auf dem gemeinsamen Faktor 2. Der Indikator  $x(3)$  dagegen lädt auf beiden Faktoren (Dimensionen).
- Kovarianzmatrix der latenten Dimensionen (gemeinsamen Faktoren). In dem Beispiel sind die beiden Dimensionen korreliert, d.h.  $KOV(\vartheta(1), \vartheta(2))$  ist ungleich 0. In der Diagonalen der Matrix stehen die Varianzen der latenten Dimensionen.
- Kovarianzmatrix der Meßfehler. Diese ist eine Diagonalmatrix und enthält als Elemente die Varianzen der Meßfehler.

Das Modell läßt sich graphisch wie in der Abbildung 5-2a dargestellt veranschaulichen. Zu beachten ist dabei, daß es sich um zwei vollkommen identische Darstellungsweisen handelt.

### 5.3.1 Explorative und konfirmatorische Faktorenanalyse

Abhängig von dem theoretischen Vorwissen werden zur Schätzung von faktorenanalytischen Modells entweder konfirmatorische oder explorative Verfahren eingesetzt werden (siehe Abbildung 5-3).

Die Vorteile der explorativen Faktorenanalyse und damit Nachteile der konfirmatorischen Faktorenanalyse bestehen darin, daß keine präzisen Vorstellungen bezüglich der Struktur der Daten bestehen müssen und somit das Vorgehen an den Daten orientiert ist. Es können dadurch im Unterschied zur konfirmatorischen Faktorenanalyse nicht erwartete Strukturen entdeckt werden. Die Nachteile der explorativen Faktorenanalyse

und damit Vorteile der konfirmatorischen Faktorenanalyse sind: Eine Testung eines theoretischen Modells sowie ein Vergleich unterschiedlicher faktorenanalytischer Modelle dahingehend, welches Modell die Daten besser erklärt, sind nicht möglich. Zur kritischen Diskussion dieser und anderer Vorteile der konfirmatorischen Faktorenanalyse siehe BACHER (1986: 51-62).

## 5.4 Überprüfung der Identifikation

Die Aufgabe der Identifikationsprüfung eines spezifizierten Modells stellt sich nur bei der konfirmatorischen Faktorenanalyse. Bei explorativen Verfahren wird sie durch die Wahl eines Verfahrens zur Bestimmung der Faktorladungen zu erreichen versucht (siehe dazu Abschnitt 5.8.2.2). Eine allgemeine Lösung des Identifikationsproblems existiert - wie für die Analyse latenter Klassen - nicht. Für einige spezielle faktorenanalytische Modelle wurden aber hinreichende Bedingungen entwickelt (LONG 1983: 34-55). In vielen praktischen Anwendungsfällen kann folgende Regel angewendet werden (5):

Regel (hinreichende Bedingung): Wird eine Meßdimension durch mindestens drei Indikatoren gemessen und laden diese drei Indikatoren nur auf dieser Meßdimension, dann sind die Faktorladungen und Meßfehlervarianzen der Indikatoren sowie die Varianz der Meßdimension identifiziert.

Bei der Regel werden die in Abschnitt 5.2 angegebenen Annahmen vorausgesetzt: Die Meßfehler sind voneinander und von der Meßdimension bzw. allgemein den gemeinsamen Faktoren unabhängig. Eine Faktorladung oder die Varianz der Meßdimension wird gleich 1.0 gesetzt.

Für unser Beispiel kann die Regel angewendet werden, da die Meßdimension »Industrialisierung« von 8 Indikatoren gemessen wird und die Indikatoren nur auf dieser Meßdimension laden. Ist die Regel nicht anwendbar, kann die Identifikation nach dem bei der Analyse latenter Klassen beschriebenen Vorgehen geprüft werden: Die unbekannt Parameter werden schrittweise als Funktionen der erwarteten Varianzen und Kovarianzen (= theoretisch empirisch bekannte Informationen) sowie der bekannten Parameter ausgedrückt.

## 5.5 Parameterschätzung

Im folgenden sollen die Grundzüge der Parameterschätzung bei der konfirmatorischen Faktorenanalyse beschrieben werden (zur Parameterschätzung bei der explorativen Faktorenanalyse siehe Abschnitt 5.8.2). Dabei werden nur die Schätzverfahren des Programmsystems LISREL (JÖRESKOG/SÖRBOM 1984) behandelt, das als eigenständiges Programm für den Großrechner sowie für PCs (Betriebssystem DOS) und als Zusatzmodell für SPSS-X verfügbar ist. LISREL (PC-Version 7.13) enthält drei Schätzfunktionen:

ULS (unweighted least Squares): Die unbekannt Parameter (Faktorladungen, Kovarianzmatrix der Faktoren und Kovarianzmatrix der Meßfehler) werden iterativ so bestimmt, daß die Funktion

$$F(\text{ULS}) = 1/2 \cdot \text{tr}(\mathbf{S} - \mathbf{\Sigma})^2$$

ein Minimum wird. Die Größen in der Gleichung bedeuten:

$\mathbf{S}$  = empirische Kovarianzmatrix der Indikatoren. Sie enthält als Elemente die berechneten (empirischen) Varianzen und Kovarianzen.

$\mathbf{\Sigma}$  = theoretische Kovarianzmatrix. Diese enthält als Elemente die aufgrund des Modells erwarteten Varianzen und Kovarianzen. Diese hängen von den zu schätzenden Parameter ab. In unserem Beispiel entsprechen die erwarteten Varianzen und Kovarianzen den abgeleiteten Modellgleichungen:

$$\begin{aligned} \text{VAR}(x(i)) &= \alpha(i)^2 \cdot \text{VAR}(\eta) + \text{VAR}(\delta(i)) \\ \text{KOV}(x(i), x(j)) &= \alpha(i) \cdot \alpha(j) \cdot \text{VAR}(\eta) \end{aligned}$$

tr = Spur einer Matrix (Summe der Diagonalelemente).

Ausführlich angeschrieben ist die Funktion

empirische Varianz von x(1)	empirische Kovarianz von x(1) und x(2)
$F(\text{ULS}) = 1/2 \cdot ([s(1,1) - \sigma(1,1)]^2 + [s(2,1) - \sigma(2,1)]^2 + \dots)$	
theoretische Varianz von x(1) (= VAR(X(1)))	theoretische Kovarianz von x(1) und x(2) (= KOV(X(1), X(2)))

Bei der ULS-Schätzung wird also versucht, die Parameter so zu schätzen, daß die quadratischen Abweichungen zwischen erwarteten und empirischen Varianzen und Kovarianzen ein Minimum wird.

GLS (general least Squares): Die unbekannt Parameter werden iterativ so bestimmt, daß die Funktion

$$F(\text{GLS}) = 1/2 \cdot \text{tr}[(S - \Sigma)S^{-1}]^2$$

ein Minimum wird. Im Unterschied zur ULS-Methode werden bei der GLS-Methode die Abweichungen  $(s(i,j) - \sigma(i,j))^2$  gewichtet (mit den Elementen der Inversen von  $S (= S^{-1})$ ). Diese Gewichtung führt dazu, daß die berechneten Ergebnisse skalunenabhängig sind (LONG 1983: 57-58). Das bedeutet, daß eine Änderung der Skaleneinheit eines Indikators den Wert des Minimums der Funktion  $F(\text{GLS})$  nicht ändert und sich die Parameter nur um die Skaleneinheit verändern. Wird also beispielsweise ein Indikator mit 2 multipliziert, verdoppelt sich die Faktorladung. Die Fehlervarianz und die empirische Varianz werden vervierfacht:

$$2 \cdot x(i,g) = 2 \cdot \alpha(i) \cdot \vartheta(g) + 2 \cdot \delta(i,g)$$

$$2^2 \cdot \text{VAR}(x(i)) = (2 \cdot \alpha(i))^2 \cdot \text{VAR}(\vartheta(g)) + 2^2 \cdot \text{VAR}(\delta(i))$$

Die Struktur des faktoranalytischen Modells bleibt unverändert.

ML (maximum likelihood): Die unbekannt Parameter werden iterativ so bestimmt, daß die Funktion

$$F(\text{ML}) = \text{tr}(\Sigma S^{-1}) + (\log(|\Sigma|) - \log|S|)$$

ein Minimum wird. »|..|« ist dabei die Determinante einer Matrix. Die Ergebnisse sind wie bei der GLS-Methode Skalen unabhängig.

Neben der Skaleninvarianz haben die GLS- und ML-Methode gegenüber der ULS-Methode den Vorteil, daß sie bestimmte wünschenswerte asymptotische Eigenschaften besitzen, die die Berechnung von Teststatistiken ermöglichen (LONG 1983:59; JÖRESKOG 1983; BROWNE 1982; 1984).

Neben den drei dargestellten Verfahren enthält LISREL noch zwei Verfahren zur Bestimmung von Startwerten. Die asymptotischen Eigenschaften dieser Verfahren werden in JÖRESKOG (1983) diskutiert.

## 5.6 Ergebnisse der Parameterschätzung

Führt man für die Beispieldaten der Industrialisierung eine Parameterschätzung durch, werden die in der Abbildung 5-4 dargestellten Ergebnisse berechnet. Aus der Abbildung 5-4 ist folgendes ersichtlich:

Es wurde die Korrelationsmatrix untersucht.

Die Faktorladung des ersten Indikators  $x(1)$  = Anteil der Zuwanderer) wurde gleich 1,0 gesetzt, um die Skaleneinheit der Meßdimension zu definieren.

Die Indikatoren »Anteil der in der Land- und Forstwirtschaft Erwerbstätigen«  $x(3)$ , »Anteil der Selbständigen an den in der Land-



und Forstwirtschaft Erwerbstätigen« (x(7)) und »Anteil der Selbständigen an den in Industrie und Gewerbe Erwerbstätigen« (x(8)) laden auf der Meßdimension negativ. Dies entspricht den Erwartungen, da ein hoher Wert in diesen Indikatoren eine geringe Industrialisierung (negative Beziehung zwischen Indikatoren und der Meßdimension) mißt.

- Eine im Absolutwert höhere Faktorladung bedeutet eine höhere Meßgenauigkeit. Der Indikator x(1) mißt also besser als der Indikator x(2), aber schlechter als der Indikator x(3) usw...
- Umgekehrt bedeutet eine höhere Meßfehlervarianz eine geringere Meßgenauigkeit. Diese »inverse« Beziehung zwischen Faktorladung und Meßfehlervarianz ergibt sich aus der Darstellung der Varianz der Indikatoren:

$$\text{VAR}(x(i)) = \alpha(i)^2 \cdot \text{VAR}(\theta) + \text{VAR}(\delta(i))$$

gegeben. Bei Verwendung der Korrelationsmatrix gleich 1.0      für alle Indikatoren gleich

hoher Wert	-->	geringer Wert
	<--	
geringer Wert	-->	hoher Wert
	<--	

- Aus der Gleichung für VAR(x(i)) kann ein weiteres Maß für die »Meßgenauigkeit« der Indikatoren abgeleitet werden, der sogenannte quadierte multiple Korrelationskoeffizient. Er ist definiert als

$$R^2(x(i)) = 1 - \frac{\text{VAR}(\delta(i))}{\text{VAR}(x(i))}$$

und variiert zwischen 0 und 1. Ein Wert von 0 bedeutet, daß die Varianz des Indikators nicht durch die Meßdimension (bzw. die Meßdimensionen) erklärt werden kann. Ein Wert von 1.0 dagegen, daß die gesamte Varianz des Indikators durch die Meßdimension (bzw. die Meßdimensionen) erklärt wird. Lädt ein Indikator auf mehreren gemeinsamen Faktoren (Dimensionen), mißt R<sup>2</sup> die durch diese Faktoren erklärte Varianz. Wurde mit standardisierten Indikatoren gerechnet, das heißt, die Korrelationsmatrix wurde untersucht, ist R<sup>2</sup>(x(i)) = 1 - VAR(δ(i)). Anzumerken ist, daß die Definition von R<sup>2</sup> der Definition der Zuverlässigkeit der klassischen Testtheorie entspricht und in der faktoranalytischen Literatur auch als Kommunalität des Indikators i bezeichnet wird.

- Betrachtet man die Fehlervarianzen, so zeigt sich, daß bei der GLS- und ML-Schätzung für den Indikator »Anteil der in der Land- und

Forstwirtschaft Erwerbstätigen« ( $x(3)$ ) eine negative Fehlervarianz berechnet wird. Dies ist ein nicht erlaubtes Ergebnis, da eine Varianz immer größer/gleich Null sein muß.

Mögliche Ursachen für das Auftreten dieses Ergebnisses können sein:

- Fehlspezifikation des Modells (siehe Abschnitt 5.7).
- Unbrauchbarkeit der ML- und GLS-Schätzung. Die Stichprobe ist zu klein, sodaß die asymptotischen Eigenschaften der Schätzer nicht gegeben sind (siehe ebenfalls Abschnitt 5.7).

## 5.7 Überprüfung der Modellanpassung und Fehleranalyse

### 5.7.1 Modellanpassung

LISREL berechnet eine Reihe von Testgrößen und Maßzahlen zur Überprüfung der Anpassung des gesamten Modells sowie zur Fehleranalyse (siehe dazu einfürend PFEIFER/SCHMIDT 1987: 34-40; JÖRESKOG/SÖRBOM 1984: I.38-I.42). Vor der Betrachtung dieser Testgrößen empfiehlt sich aber immer ein qualitativer Test der berechneten Ergebnisse dahingehend, ob

- die Faktorladungen das erwartete Vorzeichen besitzen,
- alle geschätzten Varianzen größer/gleich Null sind (ist dies nicht der Fall, bringt LISREL eine Warnung und bricht die Berechnung von Testgrößen zur Fehleranalyse ab) und ob
- die Faktorladungen relativ (zu den anderen Faktorladungen) hoch bzw. die Meßfehlervarianzen relativ gering sind.

Für unsere Beispieldaten wird eine negative Fehlervarianz berechnet. Ein deutlicher Hinweis, daß mit dem spezifizierten Modell etwas nicht stimmt. Zudem besitzen die Indikatoren »Anteil der im Bergbau und Hüttenwesen Erwerbstätigen« ( $x(4)$ ), »Anteil der in der Textilindustrie Erwerbstätigen« ( $x(5)$ ), »Anteil der Tagelöhner an den in der Land- und Forstwirtschaft Erwerbstätigen« ( $x(6)$ ) und »Anteil der Selbständigen an den in Industrie und Gewerbe Erwerbstätigen« ( $x(8)$ ) im Vergleich zu den anderen Indikatoren relativ geringe Faktorladungen bzw. relativ hohe Fehlervarianzen. Diese hohen Fehlervarianzen sind möglicherweise ein Hinweis, daß diese Indikatoren eine zweite Dimension messen.

Zur Messung der Gesamtanpassung des Modells berechnet LISREL u.a. folgende Testgröße (JÖRESKOG/SÖRBOM 1984:I.38-I.42):

- Eine Chiquadrat-verteilte Testgröße. Die Verwendung dieser Testgröße setzt die Annahme einer mehrdimensionalen Normalverteilung der Indikatoren voraus. Mit ihr kann getestet werden, ob die empirischen Daten durch das spezifizierte Modell reproduziert werden.

- Deskriptive Anpassungsindizes. Die Chiquadrat-verteilte Testgröße hat u.a. den Nachteil, daß sie von der Stichprobengröße abhängt (6). Deshalb berechnet LISREL den »goodness of fit index« (GFI) sowie den »adjusted goodness of fit index« (AGFI) als beschreibende Maßzahlen der Modellanpassung. Ein Wert von 1.0 bedeutet eine perfekte Modellanpassung, ein Wert nahe Null keine Modellanpassung. In unserem Beispiel nehmen die Testgrößen und Anpassungsindizes die in der Abbildung 5-5 dargestellten Werte an.

Aus inferenzstatistischen Gesichtspunkten muß das spezifizierte Modell verworfen werden, da die Fehlerwahrscheinlichkeit (die mit »P= ...« angegebene Wahrscheinlichkeit) für das Verwerfen der HO-Hypothese (das spezifizierte Modell ist angemessen) gleich Null ist. Die Anpassungsindizes führen dagegen zu unterschiedlichen Schlußfolgerungen. Zum einen sind sie relativ hoch (größer 0.80), zum anderen relativ niedrig (kleiner 0.70).

Zusammenfassend wird man den berechneten Ergebnissen kritisch gegenüberstehen, da sich »qualitative« Hinweise auf eine Fehlspezifikation ergeben, aus inferenzstatistischen Gesichtspunkten das Modell verworfen werden muß und die Anpassungsindizes kein eindeutiges Bild ergeben. Praktisch bedeutet dies, daß auf eine Fehleranalyse nicht verzichtet werden kann.

### 5.7.2 Fehleranalyse

Ziel der Fehleranalyse ist, mögliche Fehlerquellen zu erkennen und dadurch ein den Daten besser angepaßtes Modell zu finden. Eine allgemeine Strategie für das Auffinden von Fehlern liegt nicht vor. Grundsätzlich sind zwei Zugänge möglich, die sich gegenseitig aber nicht ausschließen, sondern ergänzen. Sie sind in der Abbildung 5-6 dargestellt.

Bei der als »empirisch« bezeichneten Vorgehensweise werden mit LISREL bestimmte Teststatistiken berechnet, die Hinweise darauf geben sollen, welche Parameter geändert werden sollen. Auf dieser Grundlage wird ein neues Modell spezifiziert, inhaltlich begründet und geschätzt. Bei der als »theoretisch« bezeichneten Vorgehensweise werden bestimmte Fehlerquellen spezifiziert und überprüft. Betrachten wir nochmals die bei der Guttman-skala (Abschnitt 2.8) und der Analyse latenter Klassen (Abschnitt 4.5) durchgeführten Fehleranalysen vor dem Hintergrund der in der Abbildung 5-6 dargestellten Strategien, so lassen sich diese wie folgt einordnen:

- Bei der Fehleranalyse der Guttman-skala (Abschnitt 2.8) wurden theoretisch spezifizierte Fehlerquellen, die sich mit Hilfe der Guttman-skala selbst überprüfen lassen, untersucht. Da von theoretisch spezi-

fizierten Fehlerquellen ausgegangen wurde, liegt gemäß Abbildung 5-6 eine »theoretische« Zugangsweise vor. Zu beachten ist, daß in Abschnitt 2.8 aber nicht alle theoretisch möglichen Fehlerquellen untersucht wurden.

Bei der Fehleranalyse des Modells von GOODMAN (Abschnitt 4.5) wurde eine explorative Strategie gewählt. Es wurde zunächst untersucht, wieviele latente Klassen den Daten zugrundeliegen. Daran anschließend wurden die bedingten Auftretswahrscheinlichkeiten der Items (Indikatoren) bei gegebener Zahl von latenten Klassen berechnet und inhaltlich interpretiert. Dieses Vorgehen ist nicht explizit in der Abbildung 5-6 dargestellt, könnte aber in diese dadurch integriert werden, daß eine Strategie eingezeichnet wird, die mit der Spezifikation eines neuen Modells (z.B. eines 4-Klassenmodells) beginnt, das anschließend geschätzt und interpretiert wird.

Im folgenden soll die in der Abbildung 5-6 als »theoretisch« bezeichnete Zugangsweise beschrieben werden, da die als »empirisch« bezeichnete Zugangsweise in den einführenden Lehrbüchern zu LISREL ausführlich behandelt wird (siehe z.B. PFEIFER/SCHMIDT 1987).

Der erste Schritt der »theoretischen« Zugangsweise zur Fehleranalyse besteht darin, daß mögliche Fehlerquellen spezifiziert werden. Diese können sehr allgemein in folgende Gruppen unterteilt werden:

- methodische Fehlerquellen: Das Schätzverfahren ist ungeeignet, da z.B. die Verteilungsannahmen (eingipfelige und symmetrische Verteilung der Indikatoren und bei ML-Schätzung mehrdimensionale Normalverteilung) nicht erfüllt sind.
- inhaltliche Fehlerquellen: Diese können sein:
  - Faktorenanalytisches Modell. (Z.B.: Die lineare Verknüpfung ist nicht geeignet.)
  - Heterogenität der Untersuchungseinheiten.
  - Mehrdimensionalität.

Zur Überprüfung dieser Fehlerquellen können in Abhängigkeit von der Präzision des theoretischen Vorwissens unterschiedliche Verfahren eingesetzt werden (siehe Abbildung 5-7).

Im folgenden soll kurz die Inspektion der eindimensionalen Verteilungen sowie die Anwendung der explorativen Faktorenanalyse zur Fehleranalyse beschrieben werden.

## 5.8 Überprüfung auf Fehlerquellen

### 5.8.1 Inspektion der eindimensionalen Verteilungen

Bei der Besprechung des faktoranalytischen Modells wurde abgeleitet, daß die empirischen Indikatoren eine symmetrische und eingipfelige Verteilung besitzen sollten, damit die verwendeten Maßzahlen zur Beschreibung der Verteilung (Mittelwert, Varianz) aussagekräftig sind. Bei der ML-Schätzung sowie bei der Berechnung von statistischen Testgrößen wird darüber hinaus eine mehrdimensionale Normalverteilung der Indikatoren angenommen (siehe Abschnitt 5.2). Diese Annahmen lassen sich durch eine graphische Inspektion der eindimensionalen Verteilungen der Indikatoren prüfen. In der Abbildung 5-8 ist die Verteilung des Indikators »Anteil der Zuwanderer« (= x(l)) dargestellt. Die Abbildung wurde mit folgendem SPSS-PC+ Programm berechnet:

```

1      data list file = 'hml880z.asc' free/nummer status(A)
2          bezirk(A) land(A) lqm25 lqm45 bevzahl sexprop
3          zuwand abwand lwanteil bergante textante taglohn1
4          selbstw selbstig lq25z,lq45z,region.
5      missing value zuwand to abwand(-1).
6      compute zuwand = 1-(zuwand/1000).
7      compute abwand = 1-(abwand/1000).
8      compute lwanteil = lwanteil/10000.
9      compute bergante = bergante/10000.
10     compute textante = textante/10000.
11     compute taglohn1 = taglohn1/1000.
12     compute selbstw = selbstw/1000.
13     compute selbstig = selbstig/1000.
14     freq var = zuwand/histogram = min(0) max( 1.0)
15     increment(0.05)/stats = all.
```

Die Zeilen bedeuten:

- 1 - 4 Definition der Eingabedaten. Zum Aufbau der Daten siehe DENZ (1989:33).
- 5 Definition von fehlenden Werten.
- 6-13 Durch diese Anweisungen werden die Indikatoren auf Anteilswerte transformiert.
- 14 Aufruf der Prozedur FREQUENCIES. Der Befehl »histogram =« bewirkt, daß für die untersuchte Variable ein Histogramm gezeichnet wird. Die Skala geht von 0 bis 1.00. Die Schrittweite beträgt 0.05. Durch »stat=all« wird dem Programm mitgeteilt, alle Statistiken zu berechnen.

In der Abbildung 5-8 wurden von den berechneten statistischen Maßzahlen nur die Schiefe und Steilheit der Verteilung eingetragen. Der Abbildung ist zu entnehmen, daß die Verteilung nicht eingipfelig ist. Ferner ist sie rechtsschief und spitz. Die Verteilung läßt zwei Interpretation zu:

- Es liegt eine heterogene Population mit drei Subpopulationen vor, die in der Abbildung mit POP I, POP II und POP III bezeichnet wurden. Die Verteilung wird also als ein Hinweis auf einen inhaltlichen Fehler (Heterogenität) interpretiert.
- Es liegt ein methodischer Fehler vor. Die Subpopulationen werden als »Ausreißer« interpretiert. In diesem Fall können folgende Verfahren angewendet werden:
  - Datentransformation. Durch Datentransformationen wird versucht eine symmetrische eingipfelige Verteilung zu erreichen.
  - Robuste Schätzung der Korrelationen (siehe dazu ARMINGER 1979:144-146).
  - Elimination des Indikators.

Für welche der beiden Interpretationen man sich entscheidet, ist letztlich eine Entscheidung, die vom Forscher getroffen werden muß. Als formales Kriterium kann der Anteil der »Ausreißer« verwendet werden (siehe dazu Abschnitt 5.8.2.4). Nachfolgend sollen einige Datentransformationstechniken behandelt werden:

#### Datentransformationen für Anteilswerte:

- Die Oddtransformation. Sie ist definiert als:

$$\text{odd}(p) = p/(1-p)$$

Sie führt dazu, daß die transformierten Anteilswerte zwischen 0 und  $+\infty$  variieren und hohe Anteilswerte auseinandergezogen werden. Für unser Beispiel ist dieser Effekt unerwünscht, da die kleinen Anteilswerte auseinandergezogen werden sollen. Man kann deshalb

- die Oddtransformation von  $q = 1-p$  verwenden. Diese ist definiert als:

$$\text{odd}(q) = \text{odd}(1-p) = (1-p)/p$$

Sie führt dazu, daß die transformierten Anteilswerte zwischen 0 und  $+\infty$  variieren und kleine Anteilswerte auseinandergezogen werden.

- Die Logittransformation schließlich führt dazu, daß die logittransformierten Anteilswerte zwischen  $-\infty$  und  $+\infty$  variieren (dem Variationsbereich der Normalverteilung) und daß kleine und große Anteilswerte stärker auseinandergezogen werden. Sie ist definiert als:  
$$\text{logit}(p) = \ln(p/(1-p))$$

Für die Verteilung des Indikators »Anteil der Zuwanderer« wird man aus den genannten Gründen die Oddtransformation für  $q$  ( $= 1-p$ ) oder die Logittransformation verwenden. Die bei Anwendung dieser Transformationen entstehenden Verteilungen enthält die Abbildung 5-9.

Der Abbildung 5-9 ist zu entnehmen, daß die Oddtransformation zu stark wirkt und Ausreißer entstehen. Die Logittransformation führt dagegen zu einer fast symmetrischen Verteilung, die Maßzahlen der Schiefe und Steilheit sind nahe bei Null.

Nach dem für den Indikator »Anteil der Zuwanderer« beschriebenen Vorgehen werden die Verteilungen der anderen Indikatoren geprüft und gegebenenfalls transformiert.

Liegen keine Anteilswerte vor, sind beispielsweise folgende Transformationen möglich:

- Logarithmische Transformation. Sie führt dazu, daß kleine Werte auseinandergezogen und große Werte zusammengestaucht werden. Vor Anwendung der logarithmischen Transformation muß sicher gestellt werden, daß alle Ausprägungen größer 0 sind, da ansonsten der Logarithmus nicht definiert ist. Dies gilt auch für die
- quadratische Transformation. Sie führt dazu, daß große Werte auseinandergezogen werden und kleine Werte zusammengestaucht werden.

## 5.8.2 Explorative Faktorenanalyse

Die explorative Faktorenanalyse kann zur Überprüfung auf Mehrdimensionalität angewendet werden, wenn keine genauen Vorstellungen über die Mehrdimensionalität bestehen. Für die explorative Faktorenanalyse wurde ein Vielzahl von Verfahren entwickelt, denen alle das folgende Vorgehen gemeinsam ist (siehe dazu z.B. einführend HOLM 1976):

- Bestimmung der Zahl der gemeinsamen Faktoren.
- Rotation der Faktoren, um eine inhaltlich interpretierbare Lösung zu erhalten.
- Berechnung von Faktorwerten (— Skalenwerte der Untersuchungseinheiten auf den gemeinsamen Faktoren)

### 5.8.2.1 Bestimmung der Zahl der Faktoren

Bei der Bestimmung der Zahl der Faktoren werden zunächst unabhängige Faktoren angenommen. Diese werden so berechnet, daß der erste Faktor die größte Erklärungskraft bezüglich der Streuung der empirischen Indikatoren besitzt, der zweite Faktor die zweitgrößte usw. Technisch wird die »Erklärungskraft« durch die Eigenwerte der Faktoren gemessen. Stellt man den Verlauf der Eigenwerte graphisch dar, ergeben sich Anhaltspunkte

te über die Zahl der Faktoren. Für unsere Beispieldaten ergibt sich bei Verwendung der Hauptkomponentenmethode (7) das in der Abbildung 5-11 dargestellte Muster.

Diese graphische Darstellung wurde durch folgendes SPSS-PC + Programm gewonnen:

```
1      data list file='hm1880z.asc' free/nummer status(A)
2          bezirk(A) land(A) lqm25 lqm45 bevzahl sexprop
3          zuwand abwand Iwanteil bergante textante taglohn1
4          selbstw selbstig lq25z,lq45z,region.
5      missing value zuwand to abwand(-1).
6      compute zuwand = 1-(zuwand/1000).
7      compute abwand = 1-(abwand/1000).
8      compute Iwanteil = Iwanteil/10000.
9      compute bergante = bergante/10000.
10     compute textante = textante/10000.
11     compute taglohn1 = taglohn1/1000.
12     compute selbstw = selbstw/1000.
13     compute selbstig = selbstig/1000.
14     factor var = zuwand to selbstig/plot = eigen.
```

Die Programmzeilen bedeuten:

- 1-13 siehe Programm im Abschnitt 5.8.1  
14 Durch die Anweisung factor wird die Faktorenanalyse aufgerufen. Die zu analysierenden Variablen werden durch die Anweisung var = zuwand to selbstig definiert. Die graphische Ausgabe der Eigenwerte wird durch den Befehl plot = eigen erreicht.

Die Zahl der Faktoren kann nun dort festgelegt werden, wo ein deutlicher Knick feststellbar ist. In dem Beispiel also zwischen einem und zwei Faktoren. Aufgrund dieses Kriteriums würde man sich also für eine eindimensionale Lösung (faktorenanalytisches Modell mit einem Faktor) entscheiden. Diese einfaktorielle Lösung erklärt ungefähr 43% der Gesamtvariation der Indikatoren (Summe der Varianzen der Indikatoren). Wie bei der konfirmatorischen Faktorenanalyse besitzen aber die Indikatoren »Anteil der im Bergau und Hüttenwesen Erwerbstätigen«, »Anteil der im Textilgewerbe Erwerbstätigen«, »Anteil der Tagelöhner an den in der Land- und Forstwirtschaft Erwerbstätigen« sowie »Anteil der in Industrie und Gewerbe selbständig Erwerbstätigen« eine geringe Faktorladung (siehe Abbildung 5-11).

In der Abbildung wurde auch die Berechnung der wichtigsten Maßzahlen der explorativen Faktorenanalyse eingetragen, nämlich die der Kommunalitäten sowie der Eigenwerte.

Neben dem Eigenwertabfall wurden noch weitere Kriterien zur Bestimmung der Zahl der Faktoren formuliert. In Abhängigkeit davon, ob bei



den geschätzten Kommunalitäten (Werte kleiner 1.0 oder 1.00) verwendet werden, werden alle Faktoren größer Null (bei Kommunalitätsschätzung) oder größer 1.00 (bei Verwendung von 1.00 für die Kommunalitäten) ausgewählt. (Zur Begründung dieses Kriteriums siehe ARMINGER 1979: 42). Für unsere Beispieldaten ergeben sich drei Faktoren, die 73% der Gesamtvariation erklären, da die Kommunalitäten bei Beginn der Schätzung auf 1.00 gesetzt wurden.

Den ersten Schritt der explorativen Faktorenanalyse zusammenfassend, kann somit ein faktoranalytisches Modell mit einem Faktor oder mit drei Faktoren akzeptiert werden. Für welches Modell man sich entscheidet, kann von den Ergebnissen des nächsten Schritts abhängig gemacht werden.

### 5.8.2.2 Rotation der Faktoren und inhaltliche Interpretation

Die im ersten Schritt berechneten Faktorladungen sind - mit Ausnahme des Modells mit einem Faktor - inhaltlich nicht interpretierbar, da sie nicht identifiziert sind. Bei der explorativen Faktorenanalyse wird die Identifikation durch eine Rotation der Faktoren zu erreichen versucht. Dabei werden die Faktoren so gedreht, daß die Indikatoren möglichst nur auf einem Faktor laden. Die Rotationsverfahren lassen sich in zwei Gruppen einteilen:

- Rechtwinkelige Rotationsverfahren. Es werden unabhängige Faktoren angenommen. Dies bedeutet, daß die Faktoren rechtwinkelig aufeinander stehen und unkorreliert sind.
- Schiefwinkelige Rotationsverfahren. Die Faktoren müssen nicht unabhängig sein. Der Winkel zwischen den Indikatoren muß nicht rechtwinkelig sein. Die Faktoren können also miteinander korrelieren.

Der Unterschied zwischen beiden Rotationsverfahren ist in der Abbildung 5-12 dargestellt. Die rechtwinkelige Rotation kann als Sonderfall der schiefwinkeligen Rotation betrachtet werden. Bei der schiefwinkeligen Rotation können die Indikatoren besser durch die Punktwolken gelegt werden.

In SPSS-PC+ stehen als Rotationsverfahren zur Verfügung:

- rechtwinkelige Rotation: VARIMAX, EQUAMAX und QUARTIMAX
- schiefwinkelige Rotation: OBLIMIN

Diese Verfahren sind ausführlich in ARMINGER (1979: 79-113) und NORUSIS (1985: 123-192) beschrieben. Die Ergebnisse der VARIMAX und OBLIMIN-Rotation enthält die Abbildung 5-13.

In der Abbildung 5-13 wurden nur Faktorladungen mit einem Absolutbetrag größer 0.50 eingetragen, um die Struktur der Faktorladungsmatrix

deutlicher zu machen. In SPSS-PC+ kann diese Darstellung durch nachfolgendes Programm erreicht werden:

```
1      data list file = 'hml880z.asc' free/nummer status(A)
2          bezirk(A) land(A) lqm25 lqm45 bevzahl sexprop
3          zuwand abwand Iwanteil bergante textante taglohn1 4
4          selbstlw selbstig lq25z,lq45z,region.
5      missing value zuwand to abwand(-1).
6      compute zuwand = 1-(zuwand/1000).
7      compute abwand = 1-(abwand/1000).
8      compute Iwanteil = Iwan teil/10000.
9      compute bergante = bergante/10000.
10     compute textante = textante/10000.
11     compute taglohn1 = taglohn1/1000.
12     compute selbstlw = selbstlw/1000.
13     compute selbstig = selbstig/1000.
14     factor var* zuwand to selbstig
15         /format = sort blank(.5)
16         /criteria = factors(3)
17         /extraction = pa 1
18         /rotation = varimax.
19     factor var = zuwand to selbstig
20         /format = sort blank(.5)
21         /criteria = factors(3)
22         /extraction = pa 1
23         /rotation = oblimin.
```

Die Zeilen bedeuten:

- 1-13 siehe Programm des Abschnitts 5.8.1
- 14 Aufruf der Prozedur FACTOR. Mit dem Befehl var= zuwand to selbstig werden die zu analysierenden Variablen definiert.
- 15 Durch die Anweisung format=sort blank(.5) wird bewirkt, daß die Indikatoren entsprechend ihrer Zugehörigkeit zu den Faktoren sortiert werden. Die Zusatzspezifikation blank(.5) bewirkt, daß nur Faktorladungen mit einem Absolutbetrag größer 0.5 ausgegeben werden.
- 16 Die Zahl der Faktoren wird gleich 3 gesetzt.
- 17 Als Verfahren zur Bestimmung der Eigenwerte wird die Hauptkomponentenmethode vereinbart, d.h. es wird mit einer Kommunalitätenschätzung von 1.0 gearbeitet.
- 18 Die Varimax-Rotation wird als Rotationsverfahren definiert.
- 19 entspricht Zeile 14
- 20 entspricht Zeile 15
- 21 entspricht Zeile 16

- 22 entspricht Zeile 17  
23 Anstelle der Varimax-Rotation wird die schiefwinkelige Oblimin-Rotation definiert.

Beide Rotationsverfahren liefern ähnliche Ergebnisse. Dies ist darauf zurückzuführen, daß bei der schiefwinkligen OBLIMIN-Rotation nur sehr geringe Korrelationen zwischen den Faktoren berechnet werden. Eine Nullkorrelation bedeutet, daß die Faktoren rechtwinklig aufeinander stehen. Der einzige Unterschied zwischen beiden Lösungen besteht darin, daß der Indikator »Anteil der in der Land- und Forstwirtschaft Erwerbstätigen« (Kurzbezeichnung: LWANTE) bei der VARIMAX-Rotation auf den ersten beiden Faktoren lädt. Inhaltlich können die Ergebnisse wie folgt interpretiert werden:

- Der erste Faktor könnte als »Mobilitätsdimension« mit den Polen »hoher Landwirtschaftsanteil« (= geringe Mobilität) und »hohe Abwanderungs- und Zuwanderungsrate« (= hohe Mobilität) bezeichnet werden.
- Für den zweiten Faktor könnte die Bezeichnung »Industrialisierung im Textilgewerbe« mit den Polen »starke Industrialisierung im Textilgewerbe« und »hoher Anteil von in der Landwirtschaft selbständig und/oder unselbständig Erwerbstätigen« (= geringe Industrialisierung im Textilgewerbe) gewählt werden. Bei der OBLIMIN-Rotation wird der eine Pol des Faktors nur von dem »Anteil der in der Landwirtschaft selbständig Erwerbstätigen« gebildet.
- Der dritte Faktor schließlich könnte als »Industrialisierung im Bergbau und Hüttenwesen« interpretiert werden. Einen Pol des Faktors (= starke Industrialisierung) bilden ein »hoher Anteil von Erwerbstätigen im Hüttenwesen und Bergbau« und ein »hoher Anteil von Tagelöhnern unter den in der Land- und Forstwirtschaft Erwerbstätigen«, den anderen Pol (= geringe Industrialisierung im Hüttenwesen und Bergbau) bildet ein »hoher Anteil von selbständig Erwerbstätigen in Industrie und Gewerbe«.
- Die (beinahe) Nullkorrelationen zwischen den Faktoren bedeuten inhaltlich, daß die Ausprägungen in den Faktoren unabhängig sind: Es gibt Untersuchungseinheiten mit einer geringen Mobilität, aber gleichzeitig mit einer starken Industrialisierung im Textilgewerbe und/oder im Bergbau und Hüttenwesen und umgekehrt. Auf der anderen Seite haben Untersuchungseinheiten eine hohe Mobilität und einen geringen Industrialisierungsgrad im Textilgewerbe und/oder im Bergbau und Hüttenwesen. Es kann sich hier um Auswanderungsgebiete handeln.

### 5.8.23 Berechnung von Faktorwerten

Empirische Anhaltspunkte über die Verteilung der Untersuchungseinheiten auf den Faktoren können wie folgt gewonnen werden: Die Skalenwerte (Faktorwerte) der Untersuchungseinheiten auf den Faktoren werden berechnet. Sie werden anschließend graphisch dargestellt.

SPSS-PC+ enthält wiederum eine Reihe von Schätzverfahren (NORUSIS 1985: 148), von denen hier nur die sogenannte Regressionsschätzung behandelt werden soll. Bei der Regressionsschätzung werden Gewichte für die Indikatoren gesucht, sodaß die Fehlerquadratsumme

$$\sum(u(k,g))^2$$

in

$$y(k,g) = \beta(k,1) \cdot x(1,g) + \dots + \beta(k,m) \cdot x(m,g) + u(k,g)$$

mit

$y(k,g)$  = Faktorwert der Untersuchungseinheit g auf dem Faktor k

$\beta(k,i)$  = Regressionsgewicht des Indikators i zur Berechnung der Faktorwerte auf dem Faktor k

$x(i,g)$  = Ausprägung der Untersuchungseinheit g im Indikator g

$u(k,g)$  = »Prognosefehler« der Untersuchungseinheit g im Faktor k

ein Minimum wird. In SPSS-PC+ kann die Regressionsschätzung durch folgendes Programm durchgeführt werden:

```

1      data list file = 'hml880z.asc' free/nummer status(A)
2          bezirk(A) land(A) lqm25 lqm45 bevzahl sexprop
3          zuwand abwand lwanteil bergante textante taglohn1
4          selbstlw selbstig lq25z,lq45z,region.
5      missing value zuwand to abwand(-1).
6      compute zuwand = 1-(zuwand/1000).
7      compute abwand = 1-(abwand/1000).
8      compute lwanteil = lwanteil/10000.
9      compute bergante = bergante/10000.
10     compute textante = textante/10000.
11     compute taglohn1 = taglohn1/1000.
12     compute selbstlw = selbstlw/1000.
13     compute selbstig = selbstig/1000.
14     factor var = zuwand to selbstig
15         /criteria = factors(3)
16         /extraction = pa 1
17         /rotation = varimax
18         /save = reg (all dim).
```

Die Zeilen bedeuten:

- 1 - 13 siehe Programm des Abschnitts 5.8.1
- 14 Aufruf der Prozedur FACTOR. Mit dem Befehl var = zuwand to selbstig werden die zu analysierenden Variablen definiert.
- 15 Die Zahl der Faktoren wird gleich 3 gesetzt.
- 16 Als Verfahren zur Bestimmung der Eigenwerte wird die Hauptkomponentenmethode vereinbart.
- 17 Die Varimax-Rotation wird als Rotationsverfahren definiert.
- 18 Durch diese Anweisung wird dem Programm mitgeteilt, daß die Faktorwerte mittels der Regressionsmethode (reg) berechnet und zwischengespeichert werden sollen. Die Zusatzspezifikation all bewirkt, daß für alle drei Faktoren die Faktorwerte berechnet werden. In der zweiten Zusatzspezifikation (= dim) wird der Variablenname für die Faktorwerte definiert. SPSS-PC erzeugt dann folgende Variablen:  
dim1 = Faktorwerte des ersten Faktors  
dim2 = Faktorwerte des zweiten Faktors  
dim3 = Faktorwerte des dritten Faktors.

Nach Ablauf des Programms stehen die Faktorwerte der Untersuchungseinheiten als Variablen in der SPSS-PC+ Arbeitsdatei zur Verfügung. Sie können wie alle anderen Variablen behandelt werden. Eine graphische Darstellung der Verteilung der Untersuchungseinheiten auf den Faktoren erhält man dadurch, daß das obige Programm durch folgende Zeilen ergänzt wird:

- 19 plot plot=dim1 with dim2/plot=dim1 with dim3
- 20 /plot - dim2 with dim3.

In der Abbildung 5-14 ist die Verteilung der Untersuchungseinheiten auf den beiden ersten Faktoren wiedergegeben.

Der Abbildung ist zu entnehmen, daß sich der Großteil der Untersuchungseinheiten um den Mittelwert (= 0) bzw. leicht links vom Mittelwert verteilt. Ferner gibt es drei »Ausreißergruppen«:

- Untersuchungseinheiten mit einer geringen Industrialisierung im Textilgewerbe, aber mit hoher Mobilität (Ausreißergruppe A1).
- Untersuchungseinheiten mit einer mittleren Mobilität, aber einer hohen Industrialisierung im Textilgewerbe (Ausreißergruppe A2).
- Eine Untersuchungseinheit mit einer hohen Industrialisierung im Textilgewerbe, aber einer geringen Mobilität (Ausreißergruppe A3).

#### 5.8.2.4 Zusammenfassung der Ergebnisse der explorativen Faktorenanalyse

Zusammenfassend ergibt die explorative Faktorenanalyse folgendes Bild:

- Die Streuung in den Indikatoren kann durch eine einfaktorielles oder dreifaktorielle Lösung erklärt werden. Für welche der beiden Lösungen man sich entscheidet, hängt letzten Endes von den inhaltlichen Forschungszielen ab, da beide Lösungen inhaltlich interpretierbar sind.
- Die dreifaktorielle Lösung erbringt drei relativ unabhängige Faktoren, die als Mobilitätsdimension sowie als Dimension der Industrialisierung im Textilgewerbe und im Hüttenwesen und Bergbau interpretiert werden können.
- Die Mehrdimensionalität selbst läßt drei Interpretationen zu:
  - Es liegt echte Mehrdimensionalität vor. Das heißt, die dreifaktorielle Struktur gilt für jede Untersuchungseinheit.
  - Die Mehrdimensionalität ist das Ergebnis von Ausreißern. Diese Ausreißer sind eventuell für weitere Analysen zu eliminieren.
  - Die Mehrdimensionalität ist das Ergebnis einer heterogenen Population. Die Ausreißer werden hier als eigenständige Populationen interpretiert.
- Die Entscheidung, ob eine Ausreißergruppe als eine eigenständig heterogene Population interpretiert wird, kann formal von der Größe der Ausreißergruppe abhängig gemacht werden. Besitzt eine Ausreißergruppe einen kleinen Anteil an der Gesamtpopulation (z.B. 3%), wird sie als Ausreißer interpretiert, bei einem hohen Anteilswert (z.B. 30%) dagegen als eigenständige Gruppe.
- Zu beachten ist, daß für jede der eben angeführten Interpretationen weitere empirische Evidenz gesammelt werden kann, daß aber letztlich die Entscheidung für eine bestimmte Interpretation durch den Forscher getroffen werden muß. So z.B. kann die »echte« Mehrdimensionalität dadurch »geprüft« werden, daß für einzelne Gruppen getrennt Faktorenanalysen gerechnet werden. Gruppierungsmerkmale können sein:
  - zufällige Trennung
  - inhaltliches Gruppierungsmerkmal, z.B. die Variable LAND mit den Ausprägungen »BÖHMEN«, »heutiges Österreich«, ...Liegt »echte« Mehrdimensionalität vor, muß in jeder Gruppe dieselbe Faktorstruktur reproduziert werden.
- Eine weitere Prüfmöglichkeit auf »echte« Mehrdimensionalität besteht darin, für die Untersuchungseinheiten eine Clusteranalyse zu rechnen. Mögliche Ergebnisse der Clusteranalyse können sein:
  - Es wird nur ein einziges großes Cluster gefunden. In diesem Fall kann die Mehrdimensionalität nicht auf Heterogenität der Population zurückgeführt werden.

- Es werden mehrere Cluster gefunden. Bis auf ein Cluster sind aber alle anderen Cluster nur schwach besetzt. In diesem Fall wird man die schwach besetzten Cluster als Ausreißer interpretieren.
- Es werden zumindest zwei relativ stark besetzte Cluster berechnet. Dies ist möglicherweise ein Hinweis auf Heterogenität. Aus der Verteilung der Cluster auf den Indikatoren ergeben sich Hinweise auf die Art der Heterogenität (siehe Abbildung 5-15).

## 5.9 Faktorenanalyse für nicht intervall- und ratioskalierte Indikatoren

Die Faktorenanalyse setzt strenggenommen intervall- und ratioskalierte Indikatoren voraus. In der Forschungspraxis wird sie i.d.R. aber auch für dichotome oder polytome ordinalskalierte Indikatoren verwendet, in der Soziologie beispielsweise zur Analyse von Fragebatterien oder in der Psychologie zur Analyse eines Wissenstests. Die Gründe dafür sind:

- Die Standardstatistikprogramme (BMDP, SAS, SPSS-X) bieten neben der Faktorenanalyse kaum andere Skalierungsverfahren an.
- Auch Einzelprogramme zur Skalierung von polytomen ordinalskalierten Variablen, wie z.B. polytome Raschmodelle, sind erst jüngst erhältlich.
- Bei der Faktorenanalyse werden im allgemeinen inhaltlich interpretierbare Ergebnisse berechnet.

Die Tatsache, daß inhaltlich interpretierbare Ergebnisse erzielt werden, schließt aber nicht aus, daß Artefakte interpretiert werden. Deshalb wurde durch Simulationsexperimente versucht, die Forschungspraxis zu rechtfertigen. Der Vorteil von Simulationsexperimenten gegenüber empirischen Analysen besteht darin, daß die »wahre« Struktur der Daten bekannt ist. Die Simulationsstudien untersuchten zum einen die Verzerrung der Korrelation und zum anderen die Effekte auf die Faktorenanalyse (siehe zusammenfassend BACHER 1986: 63-85):

### Verzerrung der »wahren« Modellkorrelation bei Verwendung von ordinalskalierten Indikatoren:

- Je höher die »wahren« Modellkorrelationen sind, desto stärker ist die Unterschätzung der »wahren« Modellkorrelationen.
- Die Unterschätzung nimmt mit der Zahl der Kategorien ab.
- Zuordnungsfehler (der latente Skalenwert wird einer falschen Antwortkategorie zugeordnet) führen zu einer weiteren Unterschätzung der »wahren« Modellkorrelationen.
- Unterschiedliche Schwierigkeitsgrade führen ebenfalls zu einer Unterschätzung der »wahren« Modellkorrelationen.

### Effekte auf die Faktorenanalyse:

- Trotz dieser Effekte erweist sich die Faktorenanalyse als relativ robustes Verfahren: Die Faktorstruktur wird selbst in dem ungünstigsten Fall dichotomer Variablen i.d.R. reproduziert, wenn (a) die Items nicht allzu unterschiedliche Schwierigkeitsgrade besitzen, (b) die Items nicht perfekt die latente Meßdimension messen und (c) Mehrdimensionalität vorliegt.

Neben dem durch Simulationsexperimente versuchten Nachweis, daß die Faktorenanalyse für dichotome und polytome ordinalskalierte Variablen brauchbar ist, wurden spezielle Verfahren entwickelt, um die Unterschätzung der »wahren« Korrelation zu beseitigen. LISREL bietet dazu die Möglichkeit der Berechnung von polychorische Korrelationskoeffizienten an (JÖRESKOG/SÖRBOM 1984: IV.1-IV.19; zur Faktorenanalyse dichotomer Variablen siehe z.B. BARTHOLOMEW 1987; MUTHEN 1989; SCHOENEBERG/ARMINGER 1989; TUCKER 1983).

## 5.10 Exkurs: Das Programmsystem LISREL

In diesem Abschnitt soll kurz die Eingabe in das Programmsystem LISREL (JÖRESKOG/SÖRBOM 1984) beschrieben werden. Eine ausführliche Einführung geben beispielsweise PFEIFER/SCHMIDT (1987). Dazu ist es zunächst vorteilhaft, sich mit der Schreibweise in LISREL vertraut zu machen (siehe Abbildung 5-16). LISREL unterscheidet auf der latenten und manifesten Ebene zwischen unabhängigen und abhängigen Variablen. Die x-Variablen sind dabei die Indikatoren der KSI-Variablen (= unabhängigen latenten Variablen), die y-Variablen die Indikatoren der ETA-Variablen (= abhängige Variablen). Die x-Variablen sind mit den KSI-Variablen durch die Faktorladungsmatrix LAMBDA-X verknüpft, die y-Variablen mit den ETA-Variablen durch die Faktorladungsmatrix LAMBDA-Y. Zusätzlich können die Messungen in den Indikatoren von Meßfehlern überlagert sein. Die Meßfehler der x-Variablen werden in LISREL als DELTA-Variablen bezeichnet, die der y-Variablen als EPSILON-Variablen. Neben der Verknüpfung der latenten und manifesten Ebene durch die Faktorladungsmatrizen bestehen noch kausale und korrelative Beziehungen auf der latenten Ebene. Die Korrelationen bzw. Kovarianzen zwischen den unabhängigen latenten Variablen KSI werden in der Matrix PHI spezifiziert. Der Einfluß der unabhängigen Variablen KSI auf die abhängigen ETA-Variablen wird durch die GAMMA-Matrix spezifiziert, die kausale Beziehung zwischen den abhängigen Variablen durch die BETA-Matrix. Schließlich enthält das Modell noch nicht erfaßte Einflußfaktoren, die ZETA-Variablen, die auf die abhängigen ETA-Variablen



wirken. Der Einfluß der Meßfehler DELTA und EPSILON auf die Messungen in den Indikatoren sowie der Einfluß der nicht erfaßten Einflußfaktoren ZETA auf die abhängigen Variablen wird in LISREL - wie beim faktorenanalytischen Modell - durch die Varianzen und Kovarianzen dieser Variablen erfaßt. Die zu schätzenden Modellparameter sind:

LAMDA-X (= LX)	Beziehung zwischen KSI- und x-Variablen (Faktorladungsmatrix der x-Variablen)
THETA-D (= TD)	Kovarianzmatrix der Meßfehler der x-Variablen
PHI (= PH)	Kovarianzmatrix der KSI-Variablen
GAMMA (= GA)	Einfluß der KSI-Variablen auf die ETA-Variablen
BETA (= BE)	Einfluß von ETA-Variablen auf andere ETA-Variablen
PSI (= PS)	Kovarianzmatrix der nicht erfaßten Einflußfaktoren ZETA
LAMDA-Y (= LY)	Beziehung zwischen den ETA- und y-Variablen.
THETA-E (= TE)	Kovarianzmatrix der Meßfehler der y-Variablen.

Diese Parameter werden in LISREL in einer Modellkarte spezifiziert. In dieser muß auch festgelegt werden, wieviele x-, KSI-, y- und ETA-Variablen das Modell enthält, damit die Matrizen aufgebaut werden können. Dazu werden folgende Kurzbezeichnungen verwendet:

NX	=	Zahl der x-Variablen
NK	=	Zahl der KSI-Variablen
NY	=	Zahl der y-Variablen
NE	=	Zahl der ETA-Variablen

Zur Darstellung der Logik der Parametereingabe in LISREL soll das Modell in der Abbildung 5-17 ausführlich in LISREL angeschrieben werden.

Das Modell besteht aus

NX	=	1 x-Variablen
NK	=	1 KSI-Variablen
NY	=	2 y-Variablen
NE	=	2 ETA-Variablen.

Da nur eine x- und KSI-Variable vorliegt, besteht die Faktorladungsmatrix LAMBDA-x (= LX) nur aus einem Element

LX(1,1)	=	Ladung der ersten x-Variablen auf der ersten KSI-Variablen.
---------	---	---

Die THEATA-D- und PHI-Matrix enthalten ebenfalls nur ein Element:

$TD(1,1)$  = Varianz des Meßfehlers der ersten x-Variablen

$PH(1,1)$  = Varianz der ersten KSI-Variablen.

Die Faktorladungsmatrix LAMBDA-Y (= LY) umfaßt dagegen vier Elemente:

$LY(1,1)$  = Faktorladung der ersten y-Variablen auf der ersten ETA-Variablen.

$LY(2,1)$  = Faktorladung der zweiten y-Variablen auf der ersten ETA-Variablen.

$LY(1,2)$  = Faktorladung der ersten y-Variablen auf der zweiten ETA-Variablen.

$LY(2,2)$  = Faktorladung der zweiten y-Variablen auf der zweiten ETA-Variablen.

Die THETA-E-, BETA- und PSI-Matrix besteht ebenfalls aus vier Elementen:

$TE(1,1)$  = Varianz der Meßfehler der ersten y-Variablen.

$TE(2,2)$  = Varianz der Meßfehler der zweiten y-Variablen.

$TE(1,2)$  = Kovarianz der Meßfehler der ersten y- und zweiten y-Variablen.

$TE(2,1)$  = Kovarianz der Meßfehler der ersten y- und zweiten y-Variablen. THETA-E ist eine symmetrische Matrix, d.h.  $TE(1,2) = TE(2,1)$

$BE(1,1)$  = Einfluß der ersten ETA-Variablen auf sich selbst. Da hier nicht sinnvoll von einer kausalen Wirkung gesprochen werden kann, wird man  $BE(1,1)$  i.d.R. gleich 0 setzen (Ausnahme: Feedbackmodelle; PFEIFER/SCHMIDT 1987: 82-87).

$BE(2,2)$  = Einfluß der zweiten ETA-Variablen auf sich selbst. Da hier nicht sinnvoll von einer kausalen Wirkung gesprochen werden kann, wird man  $BE(2,2)$  i.d.R. gleich 0 setzen (Ausnahme: Feedbackmodelle; PFEIFER/SCHMIDT 1987: 82-87).

$BE(1,2)$  = Einfluß der ersten ETA-Variablen auf die zweite ETA-Variablen.

$BE(2,1)$  = Einfluß der zweiten ETA-Variablen auf die erste ETA-Variablen.

$PS(1,1)$  = Varianz der Meßfehler der ersten y-Variablen.

$PS(2,2)$  = Varianz der Meßfehler der zweiten y-Variablen.

$PS(1,2)$  = Kovarianz der Meßfehler der ersten y- und zweiten y-Variablen.

$PS(2,1)$  = Kovarianz der Meßfehler der ersten y- und zweiten y-Variablen. Die PSI-Matrix ist ebenfalls eine symmetrische Matrix, d.h.  $PS(1,2) = PS(2,1)$

Die GAMMA-Matrix schließlich besteht aus zwei Elementen:

GA(1,1) - Einfluß der ersten KSI-Variablen auf die erste ETA-Variable

GA(1,2) = Einfluß der ersten KSI-Variablen auf die zweite ETA-Variable.

In LISREL könnten diese Parameter in der Modellkarte wie folgt definiert werden:

```
1 MO NXX=1 NIK=1 NE=2 NY=2 C
2 LX = FU,FR TD = DI,FR PH = DI,FR C
3 LY = FU,FI TE = DI,FI PS = SY,FI C
4 GA = FU,FR BE = FU,FR
```

Die Zeilen bedeuten:

- 1 Durch die Anweisung MO wird dem Programm der Beginn der Modellkarte mitgeteilt. Die Zahl der x-, KSI-, ETA- und y-Variablen wird definiert. Die Anweisung C bedeutet, daß die Modellkarte fortgesetzt wird.
- 2 Durch die Anweisung LX = FU,FR wird dem Programm mitgeteilt, daß die LAMBDA-x-Matrix eine »volle« (rechteckige) Matrix (= FU) ist und alle Werte der Matrix geschätzt werden sollen (= FR). Durch die Anweisung TD = DI,FR wird die Matrix TD als Diagonalmatrix (= DI) definiert. Das bedeutet, daß alle Elemente außerhalb der Diagonalen gleich Null sind. Die zusätzliche Spezifikation FR bedeutet, daß die Diagonalelemente zu schätzen sind. Die PHI-Matrix (= PH) wird ebenfalls als Diagonalmatrix mit zu schätzenden Diagonalelementen definiert. Die Anweisung C bedeutet, daß die Modellkarte in der nächsten Zeile fortgesetzt wird.
- 3 Durch die Anweisung LY = FU,FI wird die LAMBDA-y-Matrix (= LY) als »volle« (rechteckige) Matrix definiert. Die Spezifikation FI bedeutet, daß alle Elemente gefixt sind (Voreinstellung: Wert 0). Durch die Anweisung TD = DI,FI wird die THETA-D-Matrix als Diagonalmatrix mit gefixten Elementen definiert. Die Anweisung PS = SY,FI bewirkt, daß die PSI-Matrix als symmetrische Matrix mit gefixten Elementen definiert wird. Durch die Anweisung C wird dem Programm mitgeteilt, daß die Modellkarte fortgesetzt wird.
- 4 Durch diese beiden Anweisungen werden die GAMMA- und BETA-Matrix als »volle« rechteckige Matrix (= FU) mit zu schätzenden Parametern (= FR) definiert.

Vergleicht man die bisher durchgeführten Spezifikationen mit dem Modell der Abbildung 5-17, so zeigt sich, daß das Modell noch nicht ganz in ein LISREL-Modell übersetzt ist. So z.B. übt die KSI-Variable keinen Einfluß

auf die zweite ETA-Variable aus. Die Definition  $GA = FU$ ,  $FR$  würde aber dazu führen, daß auch ein direkter Einfluß der KSI-Variablen auf die zweite ETA-Variable geschätzt wird. Es sind deshalb weitere Spezifikationen bezüglich einzelner Elemente der Matrizen erforderlich. Wie in PAN-MARK sind in LISREL zwei Parameterrestriktionen möglich:

- Einzelne Parameter können gleich einem bestimmten Wert gesetzt werden.
- Mehrere Elemente können gleich gesetzt werden.

Die Restriktionen des Modells der Abbildung 5-18 können in LISREL wie folgt übersetzt werden:

1	FI	LX(1,1)
2	VA	1.0 LX(1,1)
3	FI	TD(1,1)
4	VA	0.0 TD(1,1)
5	FI	GA(1,2)
6	VA	0.0 GA(1,2)
7	FR	TE(1,1)TE(2,2)
8	FR	PS(1,1)PS(2,2)
9	FI	BE(2,1) BE(2,2) BE(1,1)
10	VA	0.0 BE(2,1) BE(2,2) BE(1,1)
11	FR	LY(1,1)LY(2,2)
12	EQ	LY(1,1)LY(2,2)

Die Zeilen bedeuten:

- 1 Die Faktorladung LX(1,1) wird gefixt. Dem Programm wird mitgeteilt, daß dieser Parameter nicht zu schätzen ist.
- 2 Dem Programm wird mitgeteilt, daß der Wert (VALUE) von LX(1,1) gleich 1.0 ist.
- 3 Die Fehlervarianz der ersten x-Variablen TD(1,1) wird gefixt.
- 4 Dem Programm wird mitgeteilt, daß TD(1,1) gleich Null ist. Diese Zeile wäre nicht erforderlich, da LISREL für gefixte Parameter automatisch einen Wert von 0.0 annimmt.
- 5 Der Einfluß der KSI-Variablen auf die zweite ETA-Variable wird gefixt.
- 6 Der Wert von GA(1,2) wird gleich Null gesetzt. Diese Spezifikation ist wiederum nicht erforderlich, da LISREL automatisch einen Wert von 0 annehmen würde.
- 7 Dem Programm wird mitgeteilt, daß die Diagonalelemente der Kovarianzmatrix der THETA-E-Matrix frei variieren und zu schätzen sind.
- 8 Die Diagonalelemente der PSI-Matrix werden zur Schätzung freigegeben.

- 9 Die Parameter BE(2,1), BE(1,1) und BE(2,2) werden gefixt, da nur ein direkter Einfluß von ETA(1) auf ETA(2) (= Parameter BE(1,2)» angenommen wird.
- 10 Die Parameter BE(2,1), BE(1,1) und BE(2,2) werden gleich Null gesetzt. Diese Zeile ist wiederum nicht erforderlich, da LISREL bei fixten Parametern automatisch einen Wert von 0 annimmt.
- 11 Die Parameter LY(1,1) und LY(2,2) werden freigegeben, da die erste y-Variable die erste ETA-Variable mißt und die zweite y-Variable die zweite ETA-Variable. Diese Freigabe der Parameter zur Schätzung ist erforderlich, da in der Modellkarte die gesamte LY-Matrix gefixt wurde.
- 12 Dem Programm wird mitgeteilt, daß der geschätzte Wert von LY(1,1) gleich LY(2,2) sein soll.

Durch diese zusätzlichen Anweisungen ist das Modell der Abbildung 5-17 vollständig in ein LISREL-Modell übersetzt. Um eine Schätzung durchführen zu können, müssen dem Programm u.a. noch folgende Informationen mitgeteilt werden:

- Wie groß ist die Zahl der Untersuchungseinheiten?
- Auf welchem File steht die Kovarianz- oder Korrelationsmatrix der Indikatoren?
- Handelt es sich dabei um eine Kovarianz- oder Korrelationsmatrix?
- Welches Schätzverfahren soll verwendet werden?
- Welche Ergebnisgrößen sollen ausgegeben werden?

Diese Informationen werden in LISREL ebenfalls in Form von »Karten« eingegeben. Diese sind:

- Daten karte: Diese beginnt mit DA und enthält u.a. folgende Parameter:  
NO = Zahl der Beobachtungen  
MA = der Matrixtyp, der untersucht werden soll, z.B. MA=KM  
(= Korrelationsmatrix).  
NI — Zahl der Variablen

Die DA-Karte könnte wie folgt aussehen:

1 DA NI-3 NO - 100 MA = KM

Dem Programm wird mitgeteilt, daß eine Korrelationsmatrix für drei Variablen eingelesen wird. Die Zahl der Beobachtungen, die der Berechnung der Korrelationsmatrix zugrundeliegen, ist 100.

- Definition von Eingabefiles. Hier wird der Typ des Eingabefiles, dessen Name sowie das Format der Daten festgelegt. Eine derartige Definition könnte wie folgt aussehen:

- 1 CM FU FI - TEST.COV FO
- 2 (3F10.8)

Dem Programm wird mitgeteilt, daß eine Kovarianzmatrix aus dem File TEST.COV eingelesen werden soll. Die Matrix ist im »vollen« Format gespeichert, d.h. das untere und obere Dreieck einschließlich der Diagonalen liegt vor. Eine Zahl ist im Format 10.8 gespeichert. Sie besetzt aus 10 Spalten mit 8 Stellen hinter dem Dezimalpunkt.

Anmerkung: Die erforderlichen Inputfiles können zum Teil durch SPSS-PC+ Programme erzeugt werden. I.d.R. ist dabei eine Nachbearbeitung mit einem Editor erforderlich.

Inputfiles: Erzeugungsmöglichkeit mit SPSS-PC +:

- Korrelationsmatrix - durch Option in den Prozeduren CORRELATION, FACTOR, REGRESSION
- Kovarianzmatrix - durch Option in der Prozedur REGRESSION
- Vektor der Standardabweichungen - durch Aggregation über alle Datensätze
- Vektor der Mittelwerte - durch Aggregation über alle Datensätze
- Die Outputkarte. Hier wird das Schätzverfahren definiert sowie die gewünschten Ergebnisse. Eine Outputkarte könnte wie folgt aussehen:

- 1 OU ML ALL

Dem Programm wird mitgeteilt, daß eine Maximum-Likelihood-Schätzung (= ML) durchgeführt und alle Ergebnisgröße (= ALL) ausgegeben werden sollen.

Zu beachten ist, daß die Karten in einer bestimmten Reihenfolge stehen müssen:

- DA-Karte
- Definition von Eingabefiles
- MO-Modellkarte
- Karten mit Parameterrestriktionen
- OU-Karte

Zusätzlich können den manifesten und latenten Variablen noch Namen gegeben werden sowie dem gesamten Programm ein Titel (siehe z.B. PFEIFER/SCHMIDT 1987).

Abbildung 5-1: Logik des Meßmodells "Industrialisierung"

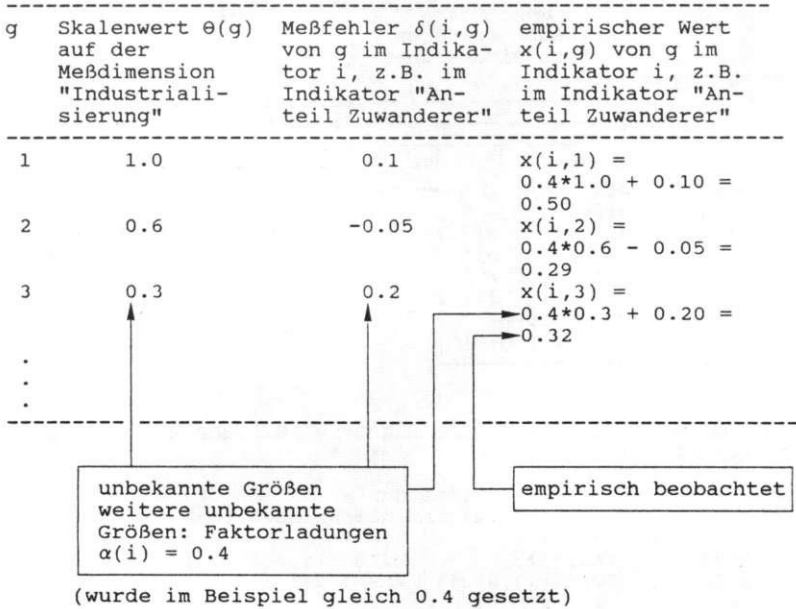
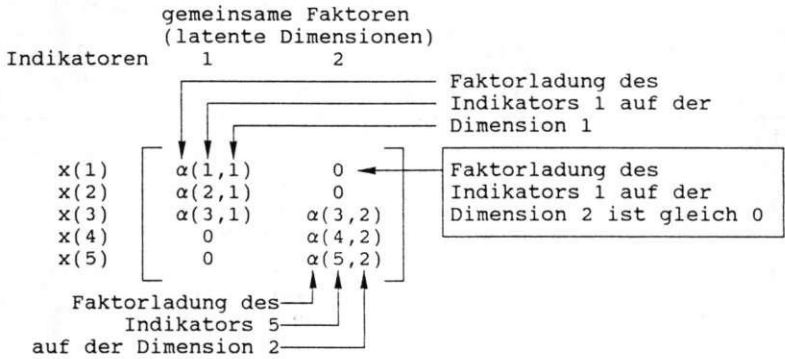
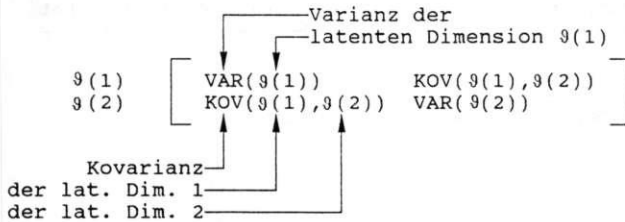


Abbildung 5-2: Darstellung eines faktorenanalytischen Modells mit zwei gemeinsamen Faktoren in Matrixschreibweise

Faktorenladungsmatrix:



Kovarianzmatrix der latenten Dimensionen (gemeinsamen Faktoren):



Kovarianzmatrix der Meßfehler  $\delta(i)$ :

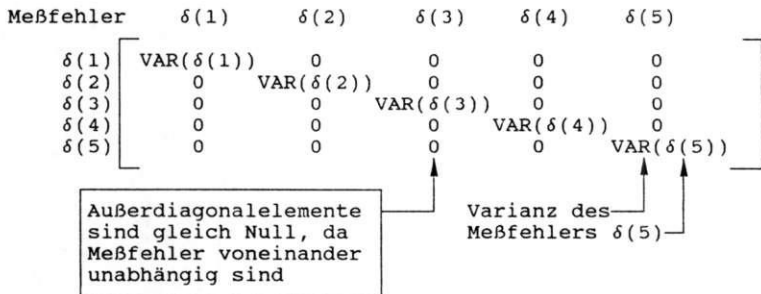




Abbildung 5-2a: Graphische Darstellung der Abbildung 5-2

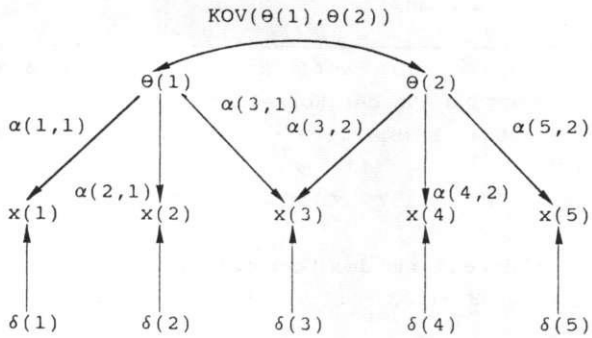
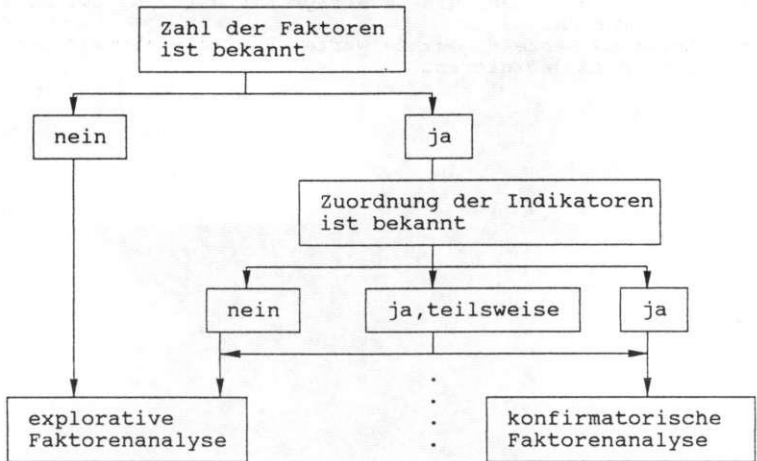


Abbildung 5-3: Entscheidung für konfirmatorische oder explorative Faktorenanalyse



.....  
 . Entscheidung ist u.a. abhängig .  
 . davon, für wieviele Indikatoren .  
 . genaue Vorstellungen bestehen. .  
 .....

Abbildung 5-4: Ergebnisse der LISREL-Schätzung für die Indikatoren der Industrialisierung

	x(1)	x(2)	x(3)	x(4)	x(5)	x(6)	x(7)	x(8)
ML-Schätzung aus der Korrelationsmatrix								
LAMBDA X	<u>1.000</u>	.624	-1.351	.199	.656	.274	-.732	-.049
VAR( $\theta$ )	.638							
VAR( $\delta$ )	.385	.763	-.157	.975	.721	.952	.659	.996
GLS-Schätzung aus der Korrelationsmatrix								
LAMBDA X	<u>1.000</u>	.708	-1.325	.229	.667	.374	-.835	-.147
VAR( $\theta$ )	.580							
VAR( $\delta$ )	.115	.326	-.036	.647	.302	.641	.357	.590
ULS-Schätzung aus der Korrelationsmatrix								
LAMBDA X	<u>1.000</u>	.955	-1.215	.258	.381	.626	-.921	-.506
VAR( $\theta$ )	.582							
VAR( $\delta$ )	.418	.469	.141	.961	.915	.772	.507	.851

LAMBDA X = in LISREL Matrix der Faktorladungen der x-Variablen, die im Skript mit  $\alpha$  (i) symbolisiert wurden.

unterstrichene Werte = gefixte Werte, um Skaleneinheit der Meßdimension zu definieren.

Abbildung 5-5: Modellanpassungstests und -indizes für die Indikatoren der Industrialisierung

-----  
ML-Schätzung aus der Korrelationsmatrix

CHI-SQUARE WITH 20 DEGREES OF FREEDOM = 570.21 (P = .000)  
GOODNESS OF FIT INDEX = .712  
ADJUSTED GOODNESS OF FIT INDEX = .481

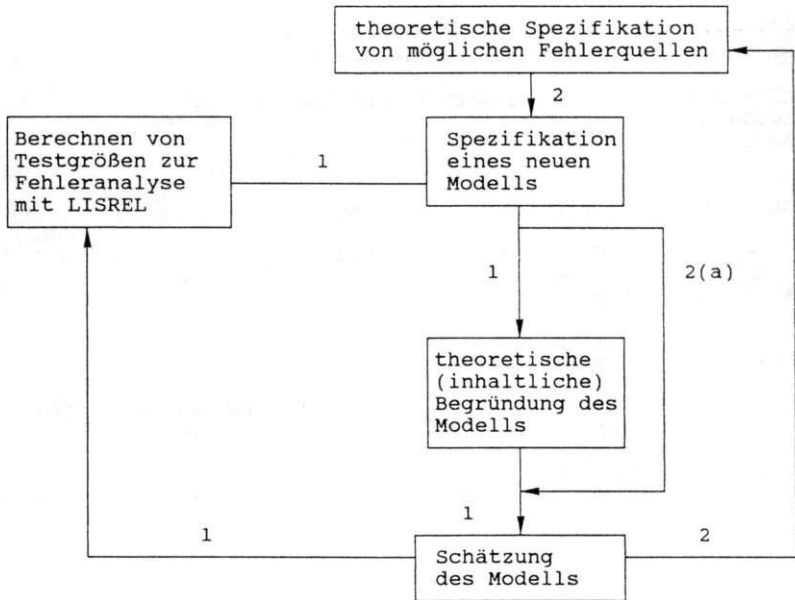
GL-Schätzung aus der Korrelationsmatrix

CHI-SQUARE WITH 20 DEGREES OF FREEDOM = 260.84 (P = .000)  
GOODNESS OF FIT INDEX = .817  
ADJUSTED GOODNESS OF FIT INDEX = .670

UL-Schätzung aus der Korrelationsmatrix

CHI-SQUARE WITH 20 DEGREES OF FREEDOM = 466.94 (P = .000)  
GOODNESS OF FIT INDEX = .946  
ADJUSTED GOODNESS OF FIT INDEX = .903  
ROOT MEAN SQUARE RESIDUAL = .110  
-----

Abbildung 5-6: Strategien der Fehleranalyse



- 1 = "empirische" Zugangsweise
- 2 = "theoretische" Zugangsweise
- a = bei der "theoretischen" Zugangsweise nicht erforderlich, da das Modell durch die Spezifikation von Fehlerquellen bereits inhaltlich begründet ist.

Abbildung 5-7: Auswahl von Verfahren zur Überprüfung von Fehlerquellen

Fehlerquelle	Verfahren
methodisch	Inspektion der eindimensionalen Verteilungen
inhaltlich	
- Faktorenanalyse kein geeignetes Modell	Verwendung eines anderen Skalierungsverfahrens, z.B. der nichtlinearen Faktorenanalyse von McDONALD (1983)
- Heterogenität bzgl. Verteilung (b)	Inspektion der eindimensionalen Verteilungen
- Heterogenität bzgl. Struktur (b)	-Gruppierungsmerkmal (a) bekannt und präzise Vorstellung über Modell ---> Anwendung konfirmatorischer Faktorenanalyse. -Gruppierungsmerkmal (a) bekannt, aber keine präzisen Vorstellungen über Modell ---->explorative Faktorenanalyse. -Gruppierungsmerkmal (a) nicht bekannt ----> Clusteranalyse
- Mehrdimensionalität	- präzise Vorstellungen bzgl. Modell ----> konfirmatorische Faktorenanalyse - keine präzisen Vorstellungen bzgl. Modell ----> explorative Faktorenanalyse

(a) = Merkmal, das Untersuchungspopulation in heterogene Gruppen unterteilt.

(b) Zu diesen beiden Begriffen siehe Abschnitt 5.8.1

Abbildung 5-8: Verteilung des Indikators "Anteil der Zuwanderer"

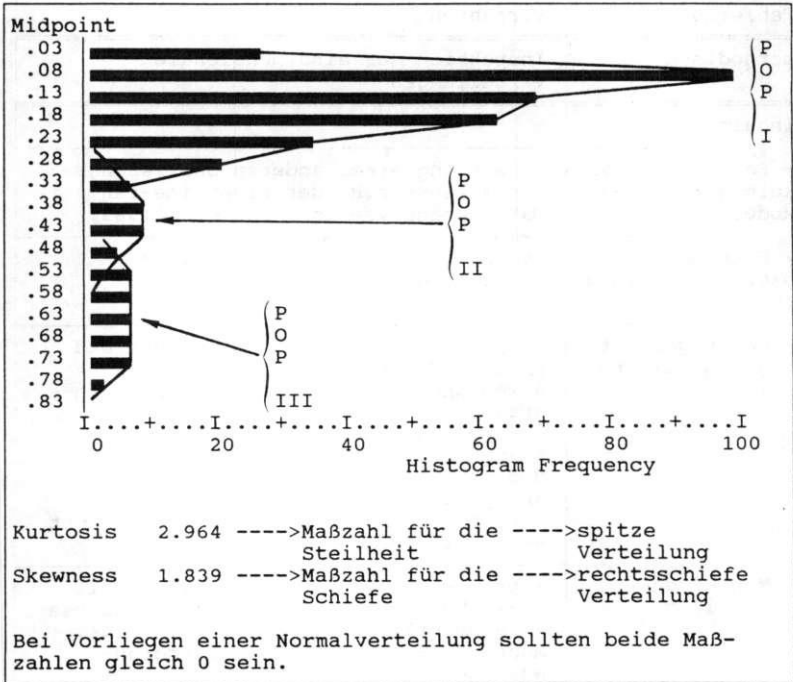


Abbildung 5-9: Verteilung der transformierten Werte des Indikators "Anteil der Zuwanderer"

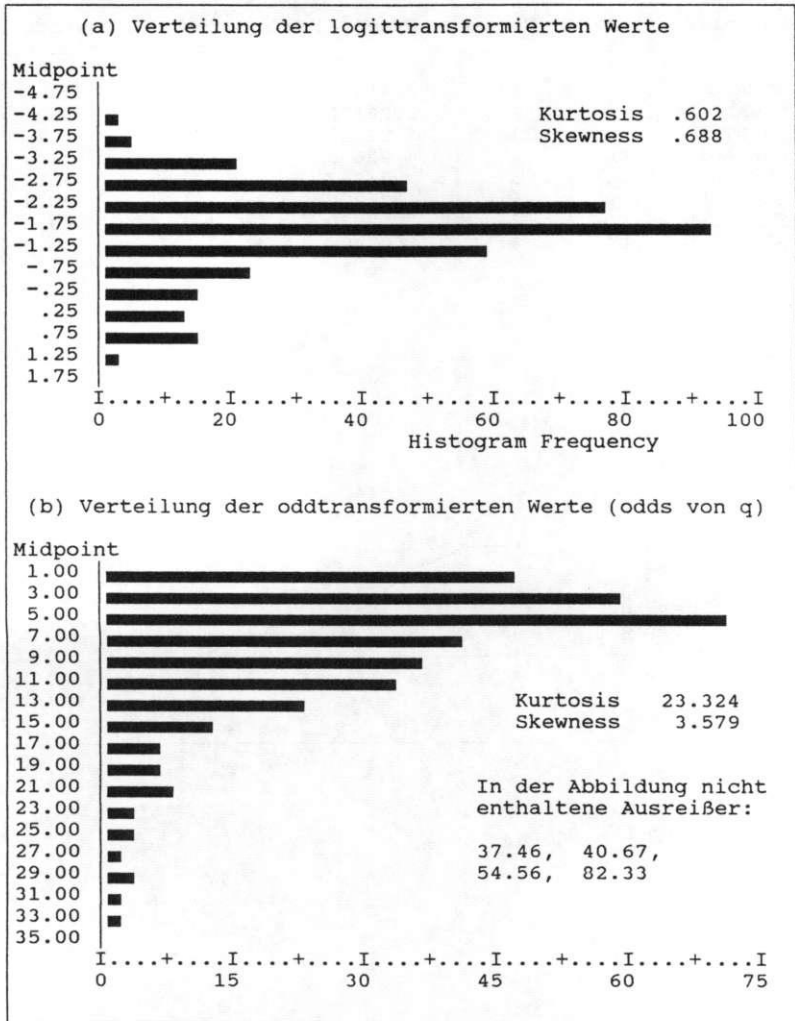


Abbildung 5-10: Ergebnisse der explorativen Faktorenanalyse

Variable	Commun* ality *	Factor	Eigenvalue	Pct of Factor	CumPct of Fact.
ZUWAND	1.00 *	1	3.41702	42.7	42.7
ABWAND	1.00 *	2	1.28122	16.0	58.7
LWANTEIL	1.00 *	3	1.12915	14.1	72.8
BERGANTE	1.00 *	4	.92693	11.6	84.4
TEXTANTE	1.00 *	5	.56057	7.0	91.4
TAGLOHNL	1.00 *	6	.35075	4.4	95.8
SELBSTLW	1.00 *	7	.27952	3.5	99.3
SELBSTIG	1.00 *	8	.05484	.7	100.0

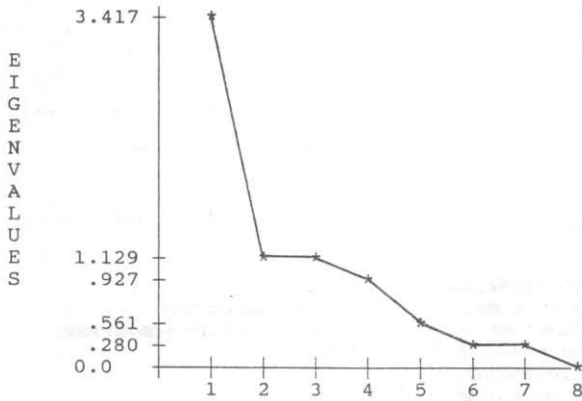




Abbildung 5-11: Ergebnisse der explorativen Faktorenanalyse (Faktorladungen der Indikatoren auf die ersten drei Faktoren bei der unrotierten Lösungen)

- - - - FACTOR ANALYSIS - - - -

-----  
Factor Matrix:

	FACTOR 1	FACTOR 2	FACTOR 3
ZUWAND	.79451	-.12518	.41618
ABWAND	.78258	-.23486	.37467
LWANTEIL	-.89217	-.26454	-.17655
BERGANTE	.25877	-.18981	-.71207
TEXTANTE	.35855	.81759	-.20114
TAGLOHNL	.58456	-.21989	-.25024
SELBSTLW	-.77516	-.26689	.31757
SELBSTIG	-.48916	.56245	.27096

|  
nur bei einfaktorieller  
Lösung ist erster Faktor  
inhaltlich interpretierbar.  
Faktoren 2 und 3 werden dann  
vernachlässigt.

-----  
Kennzahlen der explorativen Faktorenanalyse:

Zeilensummen der quadrierten  
Faktorladungen=  
Kommunalitäten  
der Indikatoren  
bei drei Faktoren

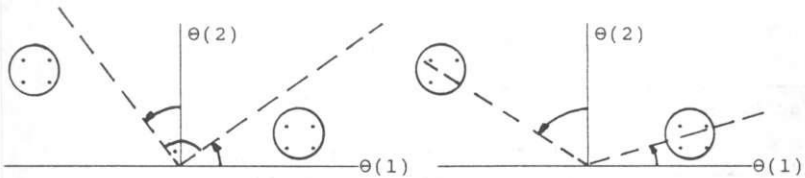
ZUWAND	$(.79451)^2$	$+(-.12518)^2$	$+ (.41618)^2$	$= .8201$
ABWAND	$+ (.78258)^2$	$+(-.23486)^2$	$+ (.37467)^2$	
LWANTEIL	$+(-.89217)^2$	$(-.26454)^2$	$+(-.17655)^2$	
BERGANTE	$+ (.25877)^2$	$(-.18981)^2$	$+(-.71207)^2$	
TEXTANTE	$+ (.35855)^2$	$(.81759)^2$	$+(-.20114)^2$	
TAGLOHNL	$+ (.58456)^2$	$(-.21989)^2$	$+(-.25024)^2$	
SELBSTLW	$+(-.77516)^2$	$(-.26689)^2$	$+ (.31757)^2$	
SELBSTIG	$+(-.48916)^2$	$(.56245)^2$	$+ (.27096)^2$	
	3.41702	1.28122	1.12915	
		Spaltensummen der quadrierten Faktorladungen = Eigenwerte der Faktoren		

Gesamtstreuung  
der Indikatoren  
bei Verwendung  
der Korrelations-  
matrix = Zahl der  
Indikatoren

====>

erklärte Streuung des  
ersten Faktors:  
 $100*(3.41702/8) = 42.7\%$

Abbildung 5-12: Schiefwinkelige und rechtwinkelige Rotation bei der explorativen Faktorenanalyse



rechtwinkelige Rotation: Die Faktoren werden so gedreht, daß sie durch die von den Indikatoren gebildeten Punktwolken gehen, aber winkelig sind.

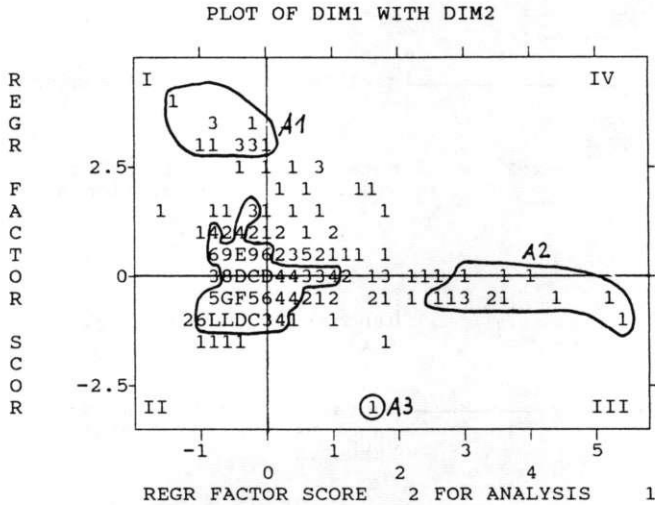
schiefwinkelige Rotation: Die Faktoren werden so gedreht, daß sie durch die von den Indikatoren gebildeten Punktwolken gehen. Die Faktoren müssen nicht rechtwinkelig sein.

Abbildung 5-13: Ergebnisse der VARIMAX und OBLIMIN-Rotation

	VARIMAX Faktor			OBLAMIN Faktor		
	1	2	3	1	2	3
-----						
Faktorladungen(a)						
ZUWAND	.90			.92		
ABWAND	.88			.90		
LWANTE	-.77	-.54		-.75		
TEXTANTE		.91		.92		
SELBSTLW		-.64			-.57	
BERGANTE			.75		-.80	
SELBSTIG			-.68			.65
TAGLOHNL			.55			-.51
-----						
Korrelation der Faktoren						
Faktor 1	1.00			1.00		
2	.00	1.00		-.16	1.00	
3	.00	.00	1.00	.29	-.10	1.00
-----						

(a) Bei der OBLIM-Rotation stehen in SPSS-PC+ die Faktorladungen in der PATTERN MATRIX.

Abbildung 5-14: Verteilung der Untersuchungseinheiten auf die Faktoren



DIM1 = Mobilität (vertikale Achse),  
 DIM2 = Industrialisierung im Textilgewerbe  
 (horizontale Achse)

Quadrate :

- I = hohe Mobilität, geringe Industrialisierung im Textilgewerbe
- II = geringe Mobilität, geringe Industrialisierung im Textilgewerbe
- III = geringe Mobilität, hohe Industrialisierung im Textilgewerbe
- IV = hohe Mobilität, hohe Industrialisierung im Textilgewerbe

Abbildung 5-15: Heterogenität bezgl. Struktur und Verteilung (fiktive Ergebnisse)

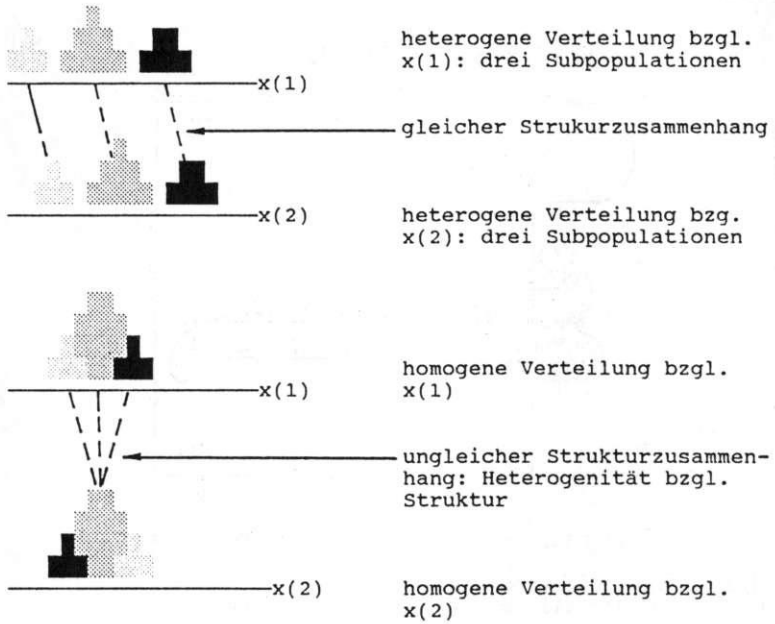


Abbildung 5-16: LISREL-Konzept und Schreibweise

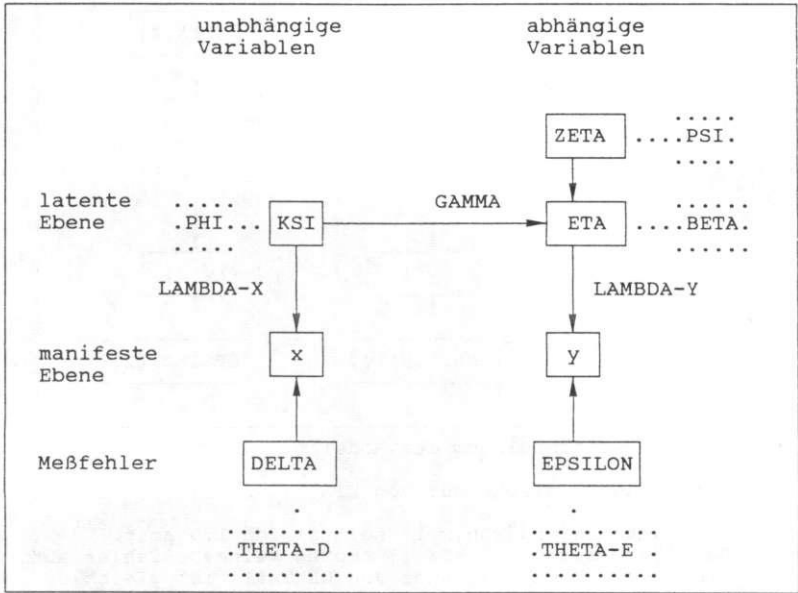
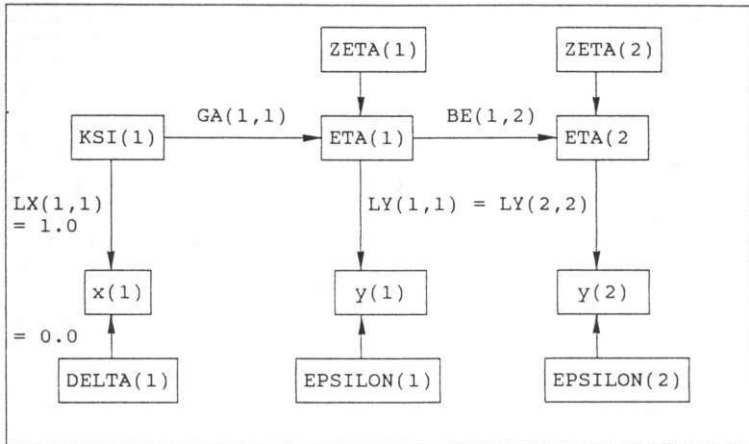


Abbildung 5-17: Ein einfaches LISREL-Modell



Verbale Kurzbeschreibung des Modells:

- $ETA(2)$  hängt direkt nur von  $ETA(1)$  ab,  $ETA(1)$  von  $KSI(1)$
- Die Faktorladung von  $x(1)$  soll gleich 1.0 sein.
- Bei der Messung von  $KSI(1)$  treten keine Meßfehler auf, d.h. die Meßfehlervarianz von  $DELTA(1)$  ist gleich 0.
- Die Meßfehler der  $ETA$ -Variablen sind unkorreliert.
- Die Variable  $y(1)$  mißt die latente Variable  $ETA(1)$ , Die Variable  $y(2)$  die latente Variable  $ETA(2)$ .
- Die Faktorladungen  $LY(1,1)$  und  $LY(2,2)$  sollen gleich sein.
- Die nicht erfaßten Einflußfaktoren sind unkorreliert.

## 6. Anmerkungen

### Anmerkungen zu Kapitel 2:

- 1) Bei der Fehlerzählweise von **GUTTMAN** selbst (BORG/SCHAUFENBIEL 1989: 126) wird die Reproduzierbarkeit noch zusätzlich dadurch überschätzt, daß nur die Zahl der Vertauschungen gezählt wird, die erforderlich sind, um eine Guttmanskala zu erhalten. Dies führt i.d.R. dazu, daß die berechnete Fehlersumme nur halb so groß ist wie bei der Zählweise von GOODENOUGH. Die von GUTTMAN empfohlene Zählweise wird beispielsweise in dem Einführungsbuch in die empirische Sozialforschung von FRIEDRICHS (1985: 179-184) verwendet.
- 2) Die Verwendung des Reproduktionskoeffizienten würde zu verzerrten Ergebnissen führen, da die Antwortvektoren, die lauter Nullen oder Einsen enthalten, nicht mitanalysiert werden. Theoretisch könnten Personen mit diesen Antwortmustern aber auch dieser Personengruppe angehören.
- 3) Die Verweise auf die Statistikprogramme BMDP, SAS und SPSS-X beziehen sich auf folgende Versionen: DIXON, W.J. (Ed.), 1983: BMDP Statistical Software. 1983 Printing with Additions. Berkley. SAS Institute Inc. (Ed.), 1988: SAS/STAT User's Guide. Release 6.03 Edition. Cary. SPSS Inc. (Ed.), 1986: SPSS-X User's Guide. 2nd Edition. Chicago.
- 4) Das Programm kann beim Autor angefordert werden.
- 5) Version SPSS-PC+ 2.0 (1986). Alle im Skript beschriebene Programme können auf Diskette beim Autor angefordert werden.

### Anmerkung zu Kapitel 4:

- 1) In diesem Kapitel werden die zur Darstellung der Guttmanskala verwendeten Daten analysiert. Zur Beschreibung der Daten siehe Abschnitt 2.1 und 2.2.

Anmerkungen zu Kapitel 5:

- 1) Werden mehrere Gruppen (z.B. Frauen und Männer) faktorenanalytisch untersucht, können mit LISREL auch sogenannte latente Gruppenmittelwerte berechnet werden. Das heißt, die Annahme  $E(x(i)) = 0$  ist dann nicht mehr erforderlich.
- 2) Bei der Maximum-Likelihood-Schätzung ist die Annahme einer multidimensionalen Normalverteilung für die Ableitung der Schätzfunktion erforderlich.
- 3) Die Modellgleichungen können leicht abgeleitet werden:  
Zunächst gilt wegen  $E(x(i)) = 0$ :

$$\text{VAR}(x(i)) = E(x(i)^2)$$

und

$$\text{KOV}(x(i), x(j)) = E(x(i) \cdot x(j))$$

Einsetzen von  $x(i) = \alpha(i) \cdot \Theta + \delta(i)$  und  $x(j) = \alpha(j) \cdot \Theta + \delta(j)$  ergibt:

$$\begin{aligned} E((\alpha(i) \cdot \Theta + \delta(i)) \cdot (\alpha(i) \cdot \Theta + \delta(i))) &= \\ E(\alpha(i) \cdot \alpha(i) \cdot \Theta \cdot \Theta + 2 \cdot \alpha(i) \cdot \Theta \cdot \delta(i) + \delta(i) \cdot \delta(i)) &= \\ \alpha(i)^2 \cdot E(\Theta^2) + 2 \cdot \alpha(i) \cdot E(\Theta \cdot \delta(i)) + E(\delta(i)^2) \end{aligned}$$

da  $\alpha(i)$  eine Konstante ist, wurde sie vor den Erwartungswertoperator geschrieben. Da  $E(\Theta) = 0$  und  $E(\delta(i)) = 0$  gilt, ist

$$\begin{aligned} E(\Theta^2) &= \text{VAR}(\Theta) \\ E(\Theta \cdot \delta(i)) &= \text{KOV}(\Theta, \delta(i)) \\ E(\delta(i)^2) &= \text{VAR}(\delta(i)) \end{aligned}$$

Wegen der statistischen Unabhängigkeit von  $\Theta$  und  $\delta(i)$ , die im Modell angenommen wird, ist  $\text{KOV}(\Theta, \delta(i)) = 0$ . Damit ergibt sich die Modellgleichung:

$$\text{VAR}(x(i)) = \alpha(i)^2 \cdot \text{VAR}(\Theta) + \text{VAR}(\delta(i))$$

Für die Kovarianz  $\text{KOV}(x(i), x(j)) = E(x(i) \cdot x(j))$  ergibt sich analog:

$$\begin{aligned} E((\alpha(i) \cdot \Theta + \delta(i)) \cdot (\alpha(j) \cdot \Theta + \delta(j))) &= \\ E(\alpha(i) \cdot \alpha(j) \cdot \Theta \cdot \Theta + \alpha(i) \cdot \Theta \cdot \delta(j) + \\ + \alpha(j) \cdot \Theta \cdot \delta(i) + \delta(i) \cdot \delta(j)) &= \\ \alpha(i) \cdot \alpha(j) \cdot E(\Theta^2) + \alpha(i) \cdot E(\Theta \cdot \delta(j)) + \\ \alpha(j) \cdot E(\Theta \cdot \delta(i)) + E(\delta(i) \cdot \delta(j)) \end{aligned}$$

da  $\alpha(i)$  eine Konstante ist, wurde sie vor den Erwartungswertoperator geschrieben. Da  $E(\Theta) = 0$  und  $E(\delta(i)) = 0$  gilt, ist



$$\begin{aligned} E(\Theta^2) &= \text{VAR}(\Theta) \\ E(\Theta \cdot \delta(j)) &= \text{KOV}(\Theta, \delta(j)) \\ E(\Theta \cdot \delta(i)) &= \text{KOV}(\Theta, \delta(i)) \\ E(\delta(i), \delta(j)) &= \text{KOV}(\delta(i), \delta(j)) \end{aligned}$$

Wegen der statistischen Unabhängigkeit von  $\Theta$ ,  $\delta(i)$  und  $\delta(j)$  ist  $\text{KOV}(\Theta, \delta(i)) = \text{KOV}(\Theta, \delta(j)) = \text{KOV}(\delta(i), \delta(j)) = 0$ .

$$\text{KOV}(\Theta, \delta(i)) = \text{KOV}(\Theta, \delta(j)) = \text{KOV}(\delta(i), \delta(j)) = 0.$$

$$\text{KOV}(x(i), x(j)) = \alpha(i) \cdot \alpha(j) \cdot \text{VAR}(\Theta)$$

$$\text{KOV}(x(i), x(j)) = \alpha(i) \cdot \alpha(j) \cdot \text{VAR}(\Theta) \quad \text{g:}$$

- 4) Diese Tatsache läßt sich leicht verdeutlichen. Betrachten wir dazu ein faktorenanalytisches Modell mit drei Indikatoren. Die Modellgleichungen sind:

$$\text{VAR}(x(1)) = \alpha(1)^2 \cdot \text{VAR}(\Theta) + \text{VAR}(\delta(1))$$

$$\text{VAR}(x(2)) = \alpha(2)^2 \cdot \text{VAR}(\Theta) + \text{VAR}(\delta(2))$$

$$\text{VAR}(x(3)) = \alpha(3)^2 \cdot \text{VAR}(\Theta) + \text{VAR}(\delta(3))$$

$$\text{KOV}(x(1), x(2)) = \alpha(1) \cdot \alpha(2) \cdot \text{VAR}(\Theta)$$

$$\text{KOV}(x(1), x(3)) = \alpha(1) \cdot \alpha(3) \cdot \text{VAR}(\Theta)$$

$$\text{KOV}(x(2), x(3)) = \alpha(2) \cdot \alpha(3) \cdot \text{VAR}(\Theta)$$

Wendet man das bei der Analyse latenter Klassen beschriebene Vorgehen der Identifikationsprüfung an, wird man zunächst versuchen, aus den Modellgleichungen für die Kovarianzen eine Darstellung für die Faktorladungen  $\alpha(i)$  und die Varianz  $\text{VAR}(\Theta)$  zu gewinnen, da diese die einfachere Struktur besitzen. Es ergeben sich folgende Darstellungen:

$$\alpha(1)^2 \cdot \text{VAR}(\Theta) = \frac{\text{KOV}(x(1), x(2)) \cdot \text{KOV}(x(1), x(3))}{\text{KOV}(x(2), x(3))}$$

$$\alpha(2)^2 \cdot \text{VAR}(\Theta) = \frac{\text{KOV}(x(1), x(2)) \cdot \text{KOV}(x(2), x(3))}{\text{KOV}(x(1), x(3))}$$

$$\alpha(3)^2 \cdot \text{VAR}(\Theta) = \frac{\text{KOV}(x(1), x(3)) \cdot \text{KOV}(x(2), x(3))}{\text{KOV}(x(1), x(2))}$$

Es können somit nur Darstellungen für die quadrierten Faktorladungen multipliziert mit der Varianz von  $\Theta$  gewonnen werden, nicht aber für die (quadrierten) Faktorladungen und für  $\text{VAR}(\Theta)$  getrennt. Das bedeutet, daß die Faktorladungen und die Varianz von  $\Theta$  nicht identifiziert sind, sondern nur die Produkte  $\alpha(1)^2 \cdot \text{VAR}(\Theta)$ ,  $\alpha(2)^2 \cdot \text{VAR}(\Theta)$ ,  $\alpha(3)^2 \cdot \text{VAR}(\Theta)$ . Wird nun beispielsweise  $\text{VAR}(\Theta) = 1.0$  gesetzt, sind die Faktorladungen bis auf das Vorzeichen identifiziert.

- 5) Diese Tatsache läßt sich leicht beweisen. Verwenden wir dazu die in der Anmerkung 4 angegebenen Modellgleichungen.

$$\alpha(1)^2 \cdot \text{VAR}(\Theta) = \frac{\text{KOV}(x(1),x(2)) \cdot \text{KOV}(x(1),x(3))}{\text{KOV}(x(2),x(3))}$$

$$\alpha(2)^2 \cdot \text{VAR}(\Theta) = \frac{\text{KOV}(x(1),x(2)) \cdot \text{KOV}(x(2),x(3))}{\text{KOV}(x(1),x(3))}$$

$$\alpha(3)^2 \cdot \text{VAR}(\Theta) = \frac{\text{KOV}(x(1),x(3)) \cdot \text{KOV}(x(2),x(3))}{\text{KOV}(x(1),x(2))}$$

Da vorausgesetzt wird, daß entweder eine Faktorladung  $\alpha(i)$  oder die Varianz  $\text{VAR}(\Theta)$  von  $\Theta$  gleich 1.0 gesetzt wird, ergeben sich folgende Darstellungen der Parameter:

- a)  $\text{VAR}(\Theta)$  wird gleich 1.0 gesetzt:

$$\alpha(1)^2 = \frac{\text{KOV}(x(1),x(2)) \cdot \text{KOV}(x(1),x(3))}{\text{KOV}(x(2),x(3))}$$

$$\alpha(2)^2 = \frac{\text{KOV}(x(1),x(2)) \cdot \text{KOV}(x(1),x(3))}{\text{KOV}(x(2),x(3))}$$

$$\alpha(3)^2 = \frac{\text{KOV}(x(1),x(2)) \cdot \text{KOV}(x(2),x(3))}{\text{KOV}(x(1),x(3))}$$

Da die Faktorladungen bis auf das Vorzeichen bekannt sind, können sie zur Identifikation der Fehlervarianzen verwendet werden:

$$\text{VAR}(\delta(1)) = \text{VAR}(x(1)) - \alpha(1)^2$$

$$\text{VAR}(\delta(2)) = \text{VAR}(x(2)) - \alpha(2)^2$$

$$\text{VAR}(\delta(3)) = \text{VAR}(x(3)) - \alpha(3)^2$$

Damit ist das Modell mit drei Indikatoren, die einen gemeinsamen Faktor messen, identifiziert.

- b) Eine Faktorladung  $\alpha(i)$  wird gleich 1.0 gesetzt. Im folgenden soll angenommen werden, daß  $\alpha(1)$  gleich 1.0 gesetzt wird. Damit ist

$$\text{VAR}(\Theta) = \frac{\text{KOV}(x(1),x(2)) \cdot \text{KOV}(x(1),x(3))}{\text{KOV}(x(2),x(3))}$$

Da die Varianz  $\text{VAR}(\Theta)$  identifiziert ist, kann sie zur Identifikation der anderen Parameter verwendet werden:

$$\alpha(2)^2 = \frac{\text{KOV}(x(1),x(2)) \cdot \text{KOV}(x(2),x(3))}{\text{KOV}(x(1),x(3)) \cdot \text{VAR}(\Theta)}$$

$$\alpha(3)^2 = \frac{\text{KOV}(x(1),x(3)) \cdot \text{KOV}(x(2),x(3))}{\text{KOV}(x(1),x(2)) \cdot \text{VAR}(\Theta)}$$

$$\text{VAR}(\delta(1)) = \text{VAR}(x(1)) - \text{VAR}(\Theta)$$

$$\text{VAR}(\delta(2)) = \text{VAR}(x(2)) - \alpha(2)^2 \cdot \text{VAR}(\Theta)$$

$$\text{VAR}(\delta(3)) = \text{VAR}(x(3)) - \alpha(3)^2 \cdot \text{VAR}(\Theta)$$

- 6) Wird die Stichprobe um den Faktor  $a$  vergrößert, vergrößert sich der Chiquadrat-Wert um den Faktor  $\sqrt{a}$ . Wird also beispielsweise die

Stichprobe vervierfacht, verdoppelt sich der Chi-Quadrat-Wert. In Abhängigkeit von der HO-Hypothese hat dies folgende Konsequenzen:

HO-Hypothese: Modell trifft zu:

Große Stichproben  $\rightarrow$  HO (Modell) wird tendenziell immer verworfen.

Kleine Stichproben  $\rightarrow$  HO (Modell) wird tendenziell nie verworfen

HO-Hypothese: Modell trifft nicht zu:

Große Stichproben  $\rightarrow$  HO wird tendenziell immer verworfen, das Modell also akzeptiert.

Kleine Stichproben  $\rightarrow$  HO wird tendenziell nie verworfen, das Modell also nie akzeptiert.

Diese Effekte gelten für alle in den einzelnen Kapiteln besprochenen Chi-Quadrat-verteilten Testgrößen.

- 7) Strenggenommen ist die Hauptkomponentenanalyse keine Faktorenanalyse (siehe BACHER 1986: 15-17). Für explorative Skalierungsaufgaben kann diese Unterscheidung vernachlässigt werden.

## 7. Literaturverzeichnis

Das nachfolgende Literaturverzeichnis wurde in fünf Gruppen unterteilt. Zu beachten ist, daß das Literaturverzeichnis nur die im Text verwendete Literatur enthält und folglich die bei den einzelnen Gruppen und Untergruppen angegebene Literatur keine vollständige Bibliographie darstellt.

### 7.1 Einführende Lehrbücher und Überblicksarbeiten zu den Methoden der empirischen Sozialforschung

- BENNINGHAUS, H., 1990: Einführung in die sozialwissenschaftliche Datenanalyse. München-Wien.
- Borz, G./C. FLECK/A. MÜLLER/M. THALLER, 1988: Qualität und Quantität. Zur Praxis der Methoden der historischen Sozialforschung. Frankfurt a.M.
- DENZ, H., 1989: Einführung in die empirische Sozialforschung. Ein Lern- und Arbeitsbuch mit Diskette. Wien-New York.
- FRIEDRICH, J., 1985: Methoden empirischer Sozialforschung. 13. Auflage. Opladen.

### 7.2 Einführende Lehrbücher und Überblicksarbeiten zu Skalierungsverfahren

- BAGOZZI, R.P./C. FORNELL, 1982: Theoretical Concepts, Measurement, and Meaning. In: C. FORNELL (Ed.): A Second Generation of Multivariate Analysis. Volume 2. New York, pp 24-38.
- BORG, I./Th. STAUFENBIEL, 1989: Theorien und Methoden der Skalierung. Bern.
- COOMBS, Cl. H., 1964: A Theory of Data. New York-London-Sydney.
- MCIVER, J./E.G. CARMINES, 1981: Unidimensional Scaling. Beverly Hills.
- SLXTL, F., 1982: Meßmethoden der Psychologie. 2. Auflage. Weinheim.

## 7.3 Weiterführende Arbeiten zu einzelnen Skalierungsverfahren

### 7.3.1 Analyse latenter Klassen/Strukturen

- CLOGG, C.C., 1988: Latent Class Models for Measuring. In: R. LANGEHEINE/J. ROST (Eds.): Latent Trait and Latent Class Models. New York, pp 173-206.
- DENZ, H., 1982: Analyse latenter Strukturen. München.
- DENZ, H., 1989a: Analyse latenter Strukturen. In: K. HOLM: ALMO Statistik-System. Linz, pp 286-299.
- LANGEHEINE, R., 1988: New Developments in Latent Class Theory. In: R. LANGEHEINE/J. ROST (Eds.): Latent Trait and Latent Class Models. New York-London, pp 77-108.
- LAZARSFELD, P.F./N.W. HENRY, 1968: Latent Structure Analysis. New York.
- DEMPSTER, A.P./N.M. LAIRD/D.B. RUBIN, 1977: Maximum Likelihood Estimation for Incomplete Data via the EM-Algorithm. Journal of Royal Statistical Society, Vol. 39, pp 1-22.
- POL, F., van de/P. de LEEUW, 1986: A Latent Markov Model to Correct Measurement Error in Categorical Data. Sociological Methods and Research, Vol. 15, pp 118-141.

### 7.3.2 Faktorenanalyse

- ARMINGER, G., 1979: Faktorenanalyse. Stuttgart.
- BACHER, J., 1986: Faktorenanalyse und Modelle des Antwortverhaltens. Wien.
- BARTHOLOMEW, R.J., 1987: Latent Variable Models and Factor Analysis. London-New York.
- BROWNE, M.W., 1982: Covariance Structures. In: D.W. HAWKINS (Ed.): Topics in Multivariate Analysis. Cambridge, pp 72-141.
- BROWNE, M.W., 1984: Asymptotic Distribution Free Methods in Analysis of Covariance Structures. British Journal of Mathematical Psychology, Vol. 37, pp 62-83.
- HOLM, K., 1976: Die Faktorenanalyse - ihre Anwendung auf Fragebatterien. In: K. HOLM (Hg.): Die Befragung 3. München pp 11-268.
- JÖRESKOG, K., 1983: Factor Analysis as an Error-in-Variables Model. In: H. WAINER/S. MESSICK (Eds.): Principals of Modern Psychological Measurement. Hillsdale, pp 185-196.
- LONG, J.S., 1983: Confirmatory Factor Analysis. A Preface to LISREL. Beverly Hills.

- MUTHFEN, B. O., 1989: Dichotomous Factor Analysis of Symptom Data. Sociological Methods and Research, Vol. 18, pp 19-67.
- PFEIFER, A./P. SCHMIDT, 1987: LISREL. Die Analyse komplexer Strukturgleichungsmodelle. Stuttgart.
- SCHOENEBERG, R./G. ARMINGER, 1989: Latent Variable Models of Dichotomous Data. The State of the Method. Sociological Methods and Research, Vol. 18, pp 19-67.
- IXJCKER, L.R., 1983: Searching for Structure in Binary Data. In: H. WAINER/S. MESSICK (Eds.): Principals of Modern Psychological Measurement. Hillsdale, pp 215-236.

### **7.33 Guttmanskalierung**

- CLIFF, N., 1983: Evaluating Guttman Scale: Some Old and New Techniques. In: H. WAINER/S. MESSICK (Ed.): Principals of modern Psychological Measurement. Hillsdale, pp 283-302.
- GOODENOUGH, W.H., 1944: A Technique for Scale Analysis. Educational and Psychological Measurement, Vol. 4, pp 179-190.
- GUTTMAN, L., 1950: The Basis for Scalogram Analysis. In: S.A. STOUFFER et al. (Ed.): Measurement and Prediction. Princeton, pp 60-90.
- LOEVINGER, J.A., 1948: The Technique of Homogeneous Test Compared with Some Aspects of »Scale Analysis«. Psychological Bulletin, Vol. 45, pp 507-530.

### **7.3.4 Mehrdimensionale Skalierung**

- BORG, I., 1981: Anwendungsorientierte Multidimensionale Skalierung. Berlin-Heidelberg-New York.
- COXON, A.P.M., 1982: The User's Guide to Multidimensional Scaling. London.

### **7.3.5 Unfoldingverfahren**

siehe COOMBS (1964); MCIVER/CARMINES (1981); SIXTL (1982).

### **7.3.6 Verfahren, die als Weiterentwicklung der Guttmanskala gedacht werden können**

- ANDRICH, D., 1985: An Elaboration of Guttman Scaling with Rasch Models. In: N.B. **TJMA** (Ed.): Sociological Methodology 1985. San Francisco, pp 33-80.
- ANDERSEN, E.B., 1983: A General Latent Structure Model for Contingency Table Data. In: H. WAINER/S. MESSICK (Eds.): Principals in Modern Psychological Measurement. Hillsdale, pp 117-138.

- DAYTON, C.M./G.B. MacREADY, 1976: A Probabilistic Model for Validation of Behavioral Hierarchies. *Psychometrika*, Vol 41, pp 189-204.
- GOODMAN, L.A., 1975: A New Model for Scaling Response Patterns: An Application of the Quasi-Independence Concept. *Journal of the American Statistical Association*, Vol. 30, pp 755-768.
- KINGMA, J./T. TAERUM, 1989: SPSS-X Procedure and Standalone Programs for the Mokken Scale Analysis: A Nonparametric Item Response Theory Model. *Educational and Psychological Measurement*, Vol. 49, pp 101-136.
- MOKKEN, R.J., 1971: A Theory and Procedure of Scale Analysis. Hawthorne-New York.
- POTUSCHAK, H., 1989: Rasch-Skalierung und probabilistisches Unfolding. In: K. HOLM: ALMO Statistik-System. Linz, pp 277-285.
- PROCTOR, C. H., 1970: A Probabilistic Formulation and Statistical Analysis of Guttman Scaling. *Psychometrika*, Vol. 35, pp 73-78.
- RASCH, G., 1966: An Individualistic Approach to Item Analysis. In: P.F. LAZARFELD/N.W. HENRY (Eds.): *Readings in Mathematical Social Science*. Chicago, pp 89-107.
- ROST, J., 1988: *Quantitative und qualitative probabilistische Testtheorie*. Bern.
- SAMEJIMA, F., 1974: Normal Ogive on the Continuous Response Level in Multidimensional Space. *Psychometrika*, Vol. 29, pp 461-467
- WOLLENBERG, A., van den, 1988: Testing a Latent Trait Model. In: R. LANGEHEINE/J. ROST (Eds.): *Latent Trait and Latent Class Models*. New York-London, pp 31-50.

## 7.4 Statistikprogramme

- DIXON, W.J. (Ed.), 1983: *BMDP Statistical Software*. 1983 Printing with Additions. Berkley.
- HOLM, K., 1989: *ALMO-Statistiksystem*. Linz.
- JÖRESKOG, K.G./D. SÖRBOM, 1984: *LISREL VI. Analysis of Linear Structural Relationships by the Method of Maximum Likelihood*. Uppsala.
- NORUSIS, M.J., 1985: *Advanced Statistics Guide SPSS-X*. Chicago.
- POL, F., van de/R. LANGEHEINE/W. de JONG, 1989: *Panmark User Manual*. Voorburg.
- SAS Insitute Inc. (Ed.), 1988: *SAS/STAT User's Guide*. Release 6.03 Edition. Cary.
- SPSS Inc. (Ed.), 1986: *SPSS-X User's Guide*. 2nd Edition. Chicago.

## 7.5 Weitere verwendete Literatur

BACHER, J., 1990: Nominale Meßfehler - Einige Konsequenzen und Schätzmöglichkeiten. *Wisdom*, Vol 4, pp 1-34.

HALLER, M./K. HOLM (Hg.), 1987: *Werthaltungen und Lebensformen in Österreich. Ergebnisse des sozialen Survey 1986*. München-Wien.