

SOZIALWISSENSCHAFTLICHES INSTITUT DER BUNDESWEHR

Berichte

Heft 64

Michael Bühner, Heinz-Ulrich Kohr, Georg Leyh

ITEMPERFECT

Benutzerorientierte PC-Programme zur flexiblen
Analyse von Fragebögen, Skalen und Tests in der
empirischen Sozialforschung

München 1994

VORWORT DES HERAUSGEBERS

BERICHT 64 des Sozialwissenschaftlichen Instituts der Bundeswehr ist ein Projekt der Grundlagenarbeit. Für die empirische Forschungstätigkeit müssen Befragungsinstrumente entwickelt und nach wissenschaftlichen Kriterien und Verfahren auf ihre Brauchbarkeit hin überprüft werden. Die statistische Datenanalyse hat in der empirischen Sozialforschung herausragendes Gewicht. Die einschlägigen Analyseverfahren werden zusammenfassend als "Itemanalyse" bezeichnet.

Mit den Berichtsbänden 6 und 40 aus den Jahren 1976 und 1986 hatte das Institut der wissenschaftlichen Öffentlichkeit die beiden Computerprogramme ITAMIS und ASIA vorgestellt. Beide Programme, für Großrechenanlagen geschrieben, sind von empirisch arbeitenden Soziologen, Psychologen, Pädagogen und Politikwissenschaftlern in Universitäten und sonstigen Forschungseinrichtungen genutzt worden. BERICHT 6 mußte wegen der großen Nachfrage in 1978 nachgedruckt werden. Im Zuge der Umstellung auf Personalcomputer entstand indessen in zunehmendem Maße der Wunsch nach einer PC-gerechten Adaptation der Programme.

BERICHT 64 beschreibt die nunmehr modifizierte Programmversionen zur PC-gestützten Analyse von Fragebögen, Skalen und Tests.

Die Verantwortung für den Inhalt liegt bei den Autoren.

Copyright by
SOWI 1994
Alle Rechte vorbehalten
ISSN 0342-2569

Sozialwissenschaftliches
Institut der Bundeswehr
Winzererstraße 52
80797 MÜNCHEN
Tel.: 089 / 120 03-1

Interessenten an ITEMPERFECT stehen die Programme gegen Erstattung der Material- und Portokosten zur Verfügung. Das Sozialwissenschaftliche Institut der Bundeswehr betrachtet dieses quasi kostenfreie Angebot als eine kollegiale Serviceleistung für die "scientific community" der empirischen Sozialforschung.

Als Herausgeber der wissenschaftlichen Schriften des Instituts danke ich den drei Programmautoren für ihre Arbeit und erhoffe mir für ITEMPERFECT eine ebenso intensive Nutzung und Nachfrage wie seinerzeit bei den Vorgängern ITAMIS und ASIA.

München, im Oktober 1994

Bernhard Fleckenstein
Direktor und Professor

Inhaltsverzeichnis:

	Seite
I. <u>Vorbemerkungen</u>	1
II. <u>ASIA</u>	4
1 Algorithmus und Formeln	4
1.1 Zweck des Programms	4
1.2 Der ASIA-Algorithmus	5
1.2.1 Vorgehensweise des Programms	5
1.2.2 Invertierung	8
1.2.3 Steuerung des Algorithmus	9
1.3 Zentrale Formeln	12
1.4 Missing-Data-Verarbeitung	16
1.5 ASIA-Programmablauf	17
2 ASIA-Steueranweisungen (Überblick)	21
2.1 SPSS-kompatible Steueranweisungen	22
2.2 SPSS-ähnliche Steueranweisungen	22
2.3 Reihenfolge der Steueranweisungen	22
2.4 Das Format der Steueranweisungen	24
2.5 Hinweise zur Notation	28
3 Die einzelnen Steueranweisungen	30
3.1 Beschreibung der Eingabedaten	30
3.1.1 VARIABLE LIST	30
3.1.2 INPUT FORMAT	31
3.1.2.1 Das Schlüsselwort FREEFIELD	32
3.1.2.2 Das Schlüsselwort FIXED	33
3.1.2.3 Unterschiede zu SPSS	36
3.1.3 N OF CASES	37

3.2	Erweiterte Beschreibung der Eingabedaten	37
3.2.1	FILE NAME	38
3.2.2	VAR LABELS	39
3.2.3	VALUE LABELS	40
3.2.4	MISSING VALUES	40
3.2.5	RANGE	43
3.2.6	PRINT FORMATS	44
3.3	Zur Form der Ein- und Ausgabe	44
3.3.1	EDIT	45
3.3.2	RUN NAME	46
3.3.3	TASK NAME	46
3.3.4	COMMENT	47
3.3.5	PRINT BACK	47
3.3.6	NUMBERED	48
3.3.7	PAGESIZE	49
3.4	Datenselektion	49
3.4.1	Permanente/temporäre Selektion	49
3.4.2	ONLY IF und NOT IF	50
3.4.3	Allgemeines Format der Selektionskriterien	53
3.5	ASIA-Systemdateien	55
3.5.1	SAVE FILE	56
3.5.2	SAVE MATRIX	57
3.5.3	OLD FILE	58
3.5.4	OLD MATRIX	58
3.6	DEFINE MATRIX	59
3.6.1	Das Schlüsselwort VARIABLES	59
3.6.2	Status und Position von DEFINE MATRIX ..	60
3.7	Die Prozeduren	61
3.7.1	CLUSTER	61
3.7.2	WRITE SCORES	68
3.8	Kommandos zur Ablaufkontrolle	71
3.8.1	READ INPUT DATA	72

3.8.2	END INPUT DATA	73
3.8.3	FINISH	73
4	Überblick über die Kommando-Syntax	74
4.1	Beschreibung der Eingabedaten	74
4.2	Erweiterte Beschreibung der Daten	74
4.3	Zur Form der Ein- und Ausgabe	74
4.4	Daten-Selektion	75
4.5	System-Dateien und Output	75
5	Literatur	77
6	Stichwortverzeichnis	78
III.	<u>ITAMIS</u>	81
1	Vorbemerkung	81
2	Leistungen des Programmsystems	83
2.1	Allgemeine Hinweise	83
2.2	Die wesentlichen Funktionen des Programmsystems	84
2.2.1	Datenmanagement	84
2.2.2	Itemanalyse	87
2.2.3	Hilfsfunktionen	91
2.2.4	Sonstiges	94
2.2.5	Anwendung des Systems auf "pathologische Daten"	95
2.3	Fehlerdiagnostik	96
3	Eingabebeschreibung	99
3.1	Anordnung und Reihenfolge der Steuerkarten für ITAMIS: Schematische Übersicht	99

3.2	Beschreibung der Steuerkarten	101
3.2.1	Textkarte	101
3.2.2	Parameterkarte	101
3.2.3	Einleseformatkarte	108
3.2.3.1	Eingabefälle	109
3.2.3.2	Allgemeine Gesichtspunkte, die bei der Anfertigung der Einleseformat- karte(n) zu beachten sind	112
3.2.3.3	Allgemeine Gesichtspunkte, die bei der Anfertigung von Schablonen zu beachten sind	113
3.2.4	Zusatzlabelkarte	113
3.2.5	Druckformatkarte	114
3.2.6	Hinweise zur Anfertigung von Schablonen	115
3.2.7	Subskalenschablone	116
3.2.8	Invertierungsschablone	117
3.2.9	Richtig-Falsch-Schablone	118
3.2.10	Dichotomisierungsschablone	119
3.2.11	Minimax-Schablone	120
3.2.12	Missing-Data-Schablone	121
3.2.13	Korrelationsschablone	123
3.2.14	Subskalenkennzeichnungskarten	125
3.2.15	Itemkennzeichnungskarten	125
3.2.16	Finish-Karte	126
4	Eingabebeispiel	127
5	Literatur	133
6	Stichwortverzeichnis	134

IV.	<u>Allgemeine Hinweise zu den Programmen</u>	136
1	Voraussetzungen an die Hardware	136
2	Start der Programme	137
2.1	ASIA	137
2.2	Beispiel für ein ASIA-Steuerfile	138
2.3	ITAMIS	139
2.4	Beispiel für ein ITAMIS-Steuerfile	140
3	Zum Umgang mit Ausgabelisten	141
4	Bezug der Programme	142

I. Vorbemerkungen

Die Analyse von Fragebogen und Tests wird gemeinhin unter dem Oberbegriff "Itemanalyse" zusammengefaßt. Im Detail zählen dazu vor allem die Berechnungen von Trennschärfe- und Schwierigkeitsindizes sowie die Berechnung von Reliabilitätskoeffizienten und Koeffizienten der Internen Konsistenz (vgl. dazu z.B. Diehl & Kohr 1977, Kap. 1 C).

Im SPSS-Programmpaket, dem wohl am meisten verbreiteten und genutzten System zur sozialwissenschaftlich-statistischen Datenanalyse, wird dem Benutzer hierzu das Modul RELIABILITY zur Verfügung gestellt, mit dem sich die "Standardanalysen" durchführen lassen. Wozu also noch ein weiteres Angebot?

Die Überlegungen, die uns zur PC-gerechten Überarbeitung zweier vorliegender, vom Sozialwissenschaftlichen Institut der Bundeswehr (SOWI) publizierter Großrechner-Programme (ASIA und ITAMIS) veranlaßt haben, sind folgende:

- 1) SPSS - RELIABILITY bietet zwar ebenfalls die oben genannten elementaren Analysemöglichkeiten. Das in der Regel aber vor der Durchführung von Itemanalysen erforderliche Datenmanagement (wie z.B. Invertierung von Items, Rekodierung von "richtigen Lösungen" in Leistungs- bzw. Kenntnistests, flexible Behandlung von "Missing Values", Bildung von Skalensummenwerten/Indexwerten) ist zwar möglich, aber es muß gewissermaßen "außerhalb" der eigentlichen Analysen im Datenmanagement - Modul vorgenommen werden.

Für den Benutzer ist der Nachteil vor allem darin zu sehen, daß Dokumentation und eigentlicher Analyselauf voneinander getrennt erfolgen müssen. Demgegenüber kann man mit den von uns vorgestellten Modulen sämtliche (gängigen) Arbeiten durchführen, die im Kontext der Itemanalyse gewöhnlich anfallen, wobei das meist recht aufwendige Datenmanagement eingeschlossen ist.

2) Itemanalysen und Index- bzw. Skalenbildungen sind gewöhnlich ein Prozeß, der dem Muster "trial-and-error" folgt. Bei der Konstruktion von Tests, Skalen und Indizes sucht man nach einer möglichst optimalen Zuordnung der Items. Diese Arbeit ist aufwendig und zeitraubend. Wir haben deshalb auf der Grundlage der FORTAN-Programme DATANA und ITAMIS ein Programm zur "Automatischen Skalen- und Item-Analyse" entwickelt, das den Namen ASIA als Akronym erhielt. Die besonderen Eigenschaften dieses Programms, das auf den bei Kohr 1978 publizierten Algorithmen basiert und eine SPSS-analoge Eingabe ermöglicht (siehe Bühner 1985), bestehen darin, daß ASIA sowohl die Durchführung "klassischer" Itemanalysen erlaubt als auch "automatisierte" Suche nach Clustern von Items/Variablen, die optimale Trennschärfekoeffizienten aufweisen. Dabei können Items/Variablen - wie üblich - jeweils einer Skala/einem Cluster zugeordnet werden, ASIA kann aber auch die Mehrfachzuordnung von Items/Variablen zulassen. Auf diesem Wege läßt sich z.B. feststellen, zu welchem Cluster ein Item optimal paßt. Ferner kann der Benutzer Clusterkerne vorgeben und den Algorithmus zur Suche nach "bestpassenden" Items zu den Clusterkernen instruieren. Wenn die Datenstruktur so beschaffen ist, daß darin Cluster enthalten sind, dann führt der Algorithmus zu Lösungen, die denjenigen ziemlich ähnlich sind, die man mittels Faktorenanalyse (Hauptkomponentenanalyse, VARIMAX-Rotation) erhalten würde. Im günstigsten Fall nähern sich die ASIA-Cluster dem Ideal der Eindimensionalität. Dies drückt sich in der Höhe der Trennschärfekoeffizienten der Itemanalysen aus, die ASIA für die neugebildeten Cluster durchführt, ferner natürlich auch in den Koeffizienten zur Abbildung der internen Konsistenz.

Das Programm ITAMIS haben wir ebenfalls für PC adaptiert, weil aus dem Benutzerkreis der Großrechnerversion eine Reihe von Anfragen nach einer PC-Version kamen. Wir haben längere Diskussionen darüber geführt, ob die PC-Anpassung überhaupt sinnvoll ist, denn es handelt sich bei ITAMIS um ein Programm, das noch aus der Zeit der Lochkarten- und Stapelverarbeitung stammt und hinsichtlich der Ein-

gabesyntax entsprechend schwerfällig und rigide ist.

Einerseits wegen der Nachfragen, andererseits wegen der in ITAMIS besonders aufwendig und flexibel realisierten Verarbeitung von Missing Values haben wir uns schließlich zur "Wiederbelebung des Sauriers" entschlossen - wieweit bei diesem Entschluß auch Nostalgie eine entscheidende Rolle gespielt hat, überlassen wir dem hoffentlich milden Urteil der Benutzer. Wegen dieser Überlegungen haben wir die PC-Anpassung von ITAMIS auf das unbedingt erforderliche reduziert und bitten die Benutzer dafür um Nachsicht.

Unseres Erachtens repräsentiert ASIA, das gewiß auch nicht mehr den modernsten Ansprüchen an Programmierertechnik und Benutzerfreundlichkeit genügen kann, für die in den Sozialwissenschaften wie Psychologie, Pädagogik und Soziologie anfallenden Analyseprobleme eine befriedigende Lösung.

Freilich ist ITEM-PERFECT eine nicht gerade bescheidene Bezeichnung; als Rechtfertigung dafür können wir anführen, daß uns die Bezeichnung wegen der Eingängigkeit und des Assoziationsumfeldes gefiel - und daß es unter den uns bekannten Programmen, die sich auf die Logik der "Klassischen Testtheorie" beziehen, keine Programme mit gleicher Leistungsbreite gibt.

Für kritische Anmerkungen, Hinweise auf "bugs" und für Verbesserungsvorschläge sind wir dankbar. Die Programme stehen allen Interessenten kostenfrei zur Verfügung; wie sie bestellt werden können, ist unter IV.4 beschrieben. Garantien für die fehlerfreie Funktion der Programme können trotz aufwendiger Tests weder von den Autoren noch vom SOWI übernommen werden.

II ASIA

(Automatische Skalenbildung für itemanalytische Anwendungen)

1 Algorithmus und Formeln

1.1 Zweck des Programms

Die klassische Testtheorie geht davon aus, daß Items linear kombiniert werden können; die Summe dieser Einzelvariablen wird dann als "Meßwert" eines theoretischen Konstrukts verwendet, der das gewünschte Merkmal genauer und ökonomischer "mißt" als die jeweils einzelnen Elemente. Ein grundlegendes Problem dabei ist, Items/Variablen zu finden, die sinnvoll kombiniert werden können.

Herkömmlicherweise versucht man das Problem der Variablenkombination zu lösen, indem man die vorhandenen Informationen (Items, Variablen) auf dem Hintergrund der theoretischen Erwartungen zu sogenannten Indizes oder "Skalen" ordnet, diese Zuordnung dann empirisch durch itemanalytische Berechnungen (Trennschärfe, interne Konsistenz, Reliabilität) überprüft und möglicherweise modifiziert. Das Ziel dabei ist es, möglichst homogene und konsistente Item-Kombinationen zu erhalten; die einzelnen Elemente der Kombinationen sollten untereinander möglichst hoch korrelieren (möglichst viel "miteinander zu tun haben"). Die Praxis zeigt, daß man sich diesem Ziel meist erst nach mehreren Versuchen nähert.

ASIA kann nun nicht etwa die theoretischen Überlegungen ersetzen oder gar automatisch nach

inhaltlichen Gesichtspunkten befriedigende Skalen bilden: Aber es kann das beschriebene "Versuch und Irrtum"-Verfahren der Skalenkonstruktion abkürzen und bequemer und effizienter gestalten.

1.2 Der ASIA-Algorithmus

In diesem Kapitel soll der von ASIA verwendete Algorithmus kurz erläutert werden. Im Rahmen dieses Benutzer-Handbuchs ist es leider nicht möglich, das Konzept der klassischen Testtheorie darzulegen oder die Zusammenhänge des ASIA-Algorithmus (vgl. Kohr 1978; Bühner und Kohr 1985) mit anderen Clusterverfahren zu untersuchen; dazu sei auf die entsprechende Literatur verwiesen (etwa Magnusson 1975).

Im ersten Unterabschnitt folgt ein grober Überblick über das Vorgehen des Programms bei der Skalenkonstruktion, der im Anschluß daran durch die entsprechenden Befehle zur Steuerung des Programms ergänzt wird. Im dritten Unterabschnitt dann werden einige der zentralen Formeln und Rechenschritte bei der Clusterung und Analyse der Skalen dargestellt und erläutert.

1.2.1 Vorgehensweise des Programms

ASIA versucht, eine Menge von m Items oder Variablen so zu ordnen, daß homogene und konsistente Skalen entstehen, deren Elemente untereinander möglichst hoch korrelieren (Produkt-Moment Korrelationskoeffizienten nach BRAVAIS/PEARSON). Das Programm geht dazu schrittweise vor:

S1: Im ersten Schritt sucht ASIA aus der Menge der m Variablen das Variablenpaar X und Y mit dem größten Zusammenhang (d.h. mit der größten

Korrelation). Dieses Variablenpaar bildet den Kern der Skala.

S2: Im zweiten Schritt nun sucht ASIA unter den restlichen $m-2$ Variablen diejenige aus, die zum bestehenden Kern "am besten paßt". Dazu wird für jede dieser Variablen die Korrelation zur Summe der beiden Kernvariablen errechnet und die Variable in die Skala aufgenommen, die zum bestehenden Kern maximal korreliert. Die Skala besteht jetzt also aus den 3 Variablen X, Y und Z.

S3: Im nächsten Schritt wird dann nach derselben Methode aus den verbleibenden $m-3$ Variablen diejenige ausgesucht, die am höchsten zur Linearkombination (d.h. zur Summe) der bisher zugeordneten Variablen korreliert und als vierte Variable in die Skala mit aufgenommen.

S_n: Nach dieser Strategie kann die Skala bei jedem Schritt um die jeweils "bestpassende" Variable erweitert werden.

Eine "natürliche" Grenze dieses Algorithmus ist dann erreicht, wenn alle m Variablen nacheinander in die Skala aufgenommen wurden. Da dies im Normalfall aber den Intentionen des Anwenders nicht gerecht wird, weil alle Variablen allen Skalen zugeordnet würden bzw. im Fall der Einfachzuordnung nur eine Skala entstünde, in die alle Variablen eingehen, gilt es, sinnvolle Abbruchkriterien zu definieren. Es müssen also Bedingungen formuliert werden, bei deren Erfüllung der Algorithmus abgebrochen wird und die Skalenbildung als beendet gilt. Zum Beispiel könnte man fordern, daß jede Skala nur aus maximal 10 Elementen bestehen soll oder daß die Trennschärfe (also die Korrelation eines Elementes zur Skalensumme) einen bestimmten Wert nicht unterschreiten darf.

Wenn mit dieser Strategie die erste Skala gebildet und abgeschlossen ist, kann mit demselben Verfahren eine zweite Skala extrahiert werden usw. Der in ASIA realisierte Algorithmus erlaubt zwei Varianten: Entweder können die bisher zugeordneten Variablen aus der Analyse ausgeschlossen (Einfachzuordnung) oder wieder in die Menge der m Variablen aufgenommen werden (Mehrfachzuordnung). Im zweiten Fall gilt dann lediglich die Einschränkung, daß Variablen, die schon einmal Kernvariablen waren, also im Schritt S1 ausgewählt wurden, nicht noch einmal Kernvariablen werden dürfen. Der Algorithmus erlaubt es aber durchaus, daß eine Variable in zwei (oder mehrere) unterschiedliche Skalen aufgenommen wird (eventuell sogar mit unterschiedlichem Vorzeichen).

Außerdem ist leicht einsichtig, daß der Algorithmus auch dann funktioniert, wenn man den ersten Schritt (oder die ersten Schritte) von vorneherein selbst festlegt, wenn man dem Programm also einen Skalenrumpf vorgibt, in dem eine oder mehrere theoretisch wichtige Variablen enthalten sind, und den Algorithmus dazu nutzt, diesen Rumpf um passende Items zu erweitern.

Die Größe des Skalenrumpfes ist dabei nicht beschränkt. Deshalb kann ASIA auch zur Itemanalyse benutzt werden. Durch die Wahl geeigneter Abbruchkriterien (oder durch Angabe des Schlüsselworts METHOD=ANALYZE) kann die Skalenerweiterung ganz übersprungen werden. ASIA gibt dann nur die itemanalytischen Kennwerte der definierten Skalen aus.

1.2.2 Invertierung

Wenn ASIA nach dem oben beschriebenen Verfahren Skalen bildet bzw. erweitert, ignoriert es zunächst das Vorzeichen des Zusammenhangs (also der Korrelationen), weil ein starker negativer Zusammenhang eben auch darauf hinweist, daß das entsprechende Item zur Skala "paßt", weil es "etwas mit der bereits gebildeten Skala zu tun hat". Bei Itembatterien liegt eine solche negative Korrelation meist daran, daß nicht alle Fragen "in dieselbe Richtung" formuliert wurden. Solche Fragen müßten nun einfach invertiert, also die Kodierung "umgedreht" werden. Wenn allerdings ein Item (bei Mehrfachzuordnung) in mehrere Skalen aufgenommen wird, kann es durchaus passieren, daß es in unterschiedlichen Skalen mit unterschiedlichem Vorzeichen vorkommt. Eine A-priori-Invertierung würde dann nicht zum angestrebten Ziel führen.

ASIA invertiert deshalb automatisch jede Variable, wenn sie eine negative Trennschärfe zum Skalenrumpf aufweist (und markiert dieses Item bei der Ausgabe durch ein Sternchen in der Spalte INV); diese Invertierung gilt allerdings nur für die aktuelle Skala. Bei der Automatischen Skalenbildung und -erweiterung gibt es keine Möglichkeit, für einzelne Variablen das Vorzeichen festzulegen. Lediglich bei der Bildung von Skalenrümpfen kann angegeben werden, ob das jeweilige Element positiv oder negativ in die Skala aufgenommen werden soll (s. die Schlüsselwörter SCALE und INVERT).

Der Rechengang bei der Invertierung von Items verlangt nicht, daß die Rohwerte jedes Falles invertiert sein müssen. Da ASIA mit der Varianz-/ Kovarianz-Matrix arbeitet, genügt es, lediglich die entsprechende Spalte und Reihe dieser Matrix mit -1 zu multiplizieren. Um allerdings den Skalenmittelwert errechnen und später den Skalensummen-

wert pro Fall ausgeben zu können, braucht das Programm Informationen über das theoretische Minimum und Maximum der Variablen. Für alle Variablen, die eventuell invertiert werden müssen, sollte ASIA diese Angabe über die Anweisung RANGE mitgeteilt werden.

1.2.3 Steuerung des Algorithmus

Die Basis des Algorithmus ist ein Variablenpool, in dem alle Items vorhanden sind, die zur Clusterung (Skalenbildung und -erweiterung) zur Verfügung stehen. Diese Variablenmenge wird dem Programm mitgeteilt durch das Schlüsselwort **VARIABLES** ="varlist".

Falls Skalenrümpfe definiert werden sollen, kann das geschehen durch die Schlüsselwörter **SCALE** und **INVERT**. Zuerst gibt man der Skala einen Namen (und eventuell eine erläuternde Etiketete) und spezifiziert, welche Variablen aus dem Pool in diesen Rumpf aufgenommen werden sollen.

Beispiel:

```
SCALE(POL,POLITISCHES INTERESSE)=POL1,POL2,POL5/  
SCALE(SOZ,SOZIALE AKTIVITÄTEN)=SOZ1 TO SOZ5 usw.
```

Durch das Schlüsselwort **INVERT** können jetzt gegebenenfalls einzelne dieser gesetzten Variablen als zu invertierende markiert werden.

Beispiel:

```
... INVERT(POL)=POL2/  
INVERT(SOZ)=SOZ3,SOZ4 usw.
```

Als nächstes sollte man sich überlegen, ob der Algorithmus Mehrfachzuordnung zulassen oder jedes Item maximal einer Skala zuordnen soll. Mehrfachzuordnung ist voreingestellt, Einfachzuordnung muß über das Schlüsselwort `/SINGLE` angefordert werden.

Dann gilt es, sinnvolle Abbruchkriterien zu definieren, die u.U. für jede Skala anders aussehen können. ASIA bietet vier Möglichkeiten:

- durch `RKRIT=value` kann die minimale Trennschärfe festgelegt werden, mit der ein Item/eine Variable zum bestehenden Kern korrelieren muß, um noch in die Skala aufgenommen zu werden. Wenn also die nächste Variable, die aufgenommen werden würde, eine Trennschärfe kleiner "value" aufweist, gilt die Skalenerweiterung als beendet.
- durch `DKRIT=value` kann dieses absolute Kriterium vermieden und durch ein relatives ersetzt werden. Unter der Annahme, daß einige wenige Variablen des Pools sehr gut "zusammenpassen" und der Rest wenig oder nichts mit diesem Kern zu tun hat, müßte, nachdem die letzte gut passende Variable aufgenommen wurde, ein deutlicher Abfall der Trennschärfe zu beobachten sein. Mit `DKRIT` kann festgelegt werden, wie groß die Differenz zur vorherigen Trennschärfe maximal sein darf, damit eine Variable noch aufgenommen wird.
- ein anderes Kriterium kann durch `ALFAD=value` festgesetzt werden; `ALFAD` legt fest, um wieviel die interne Konsistenz der Skala (gemessen durch Cronbachs alpha) durch die Hinzunahme eines Items mindestens ansteigen muß. Voreingestellt ist `ALFAD=0`; das bedeutet, daß die interne Konsistenz zumindest nicht abfallen darf. Durch `ALFAD=-.1` könnte man auch fordern, daß die interne Konsistenz um nicht mehr als 0.1 abfallen darf. Da

Cronbachs alpha bei zwei Skalen mit unterschiedlicher Itemzahl nicht vergleichbar ist, wird dieser Meßwert auf die Skalengröße $n=10$ standardisiert (vgl. 1.3. Zentrale Formeln).

- schließlich kann durch `MXSIZE=n` festgelegt werden, wieviele Elemente/Items die Skala maximal beinhalten darf.

Alle Kriterien werden durch `ODER` verbunden, d.h., die Skalenbildung gilt als beendet, wenn wenigstens eines der Abbruchkriterien erfüllt ist; ASIA gibt den Namen des verursachenden Abbruchkriteriums aus.

Beispiel:

```
RKRIT(POL)=.50/MXSIZE(ALL)=15/
MXSIZE(POL,SOZ)=20/
ALFAD=.01/DKRIT(S3,S4)=.20 usw.
```

Mit dem Schlüsselwort `METHOD=keyword` kann der Algorithmus zur Skalenerweiterung bestimmt werden. Voreingestellt ist `METHOD=CORRMAX`: bei jedem Schritt wird die Variable ausgewählt, die am höchsten zum bestehenden Kern korreliert (s.o.). Unter bestimmten Bedingungen ("Suppressor-Effekte") kann es vorkommen, daß die Trennschärfe bei Schritt $n+1$ größer wird, als sie es beim Schritt n war. Wenn die Trennschärfe "n" unter dem Abbruchkriterium `RKRIT` lag, werden u.U. "sinnvolle" Variablen nicht mehr in die Skala aufgenommen. (Man kann das an der Summary-Tabelle sehen, wenn eine Fremdtrennschärfe größer ist als das Abbruchkriterium). Teilweise läßt sich dies verhindern, wenn man bei den Schritten 2 bis n nicht nach dem Item mit der größten Trennschärfe sucht, sondern nach dem Item, das die interne Konsistenz maximiert. ASIA stellt diese Möglichkeit, bei der die interne Konsistenz durch Cronbachs alpha gemessen wird, durch das Schlüsselwort `METHOD=ALFAMAX` zur Verfügung.

Mit METHOD=ANALYZE wird ASIA veranlaßt, die Skalen-erweiterung zu überspringen und für die gesetzten Skalenrumpfe lediglich die itemanalytischen Kennwerte auszugeben. Dies ist natürlich nur für Skalen sinnvoll, die mit dem Schlüsselwort SCALE (s.o.) gebildet wurden.

Wenn keine Skalenrumpfe festgelegt wurden, sondern ASIA die Skalen selbst bilden soll, ist es - vor allem bei Mehrfachzuordnung - sinnvoll, die maximale Clusteranzahl festzulegen; dies geschieht durch das Schlüsselwort MAXCL=n.

Mit den Schlüsselwörtern LIST, MATRIX und SUMMARY kann ASIA veranlaßt werden, zusätzlichen Output zu produzieren. LIST erzeugt eine Liste aller Poolvariablen mit Mittelwert, Standardabweichung und Anzahl gültiger Fälle; MATRIX gibt die zugrundeliegende Kovarianz-/Korrelationsmatrix aus (Korrelationen oberhalb, Kovarianzen unterhalb der Diagonalen); SUMMARY produziert erstens eine Tabelle mit den (korrigierten) Trennschärfen aller Variablen zu allen Skalen und zweitens eine Matrix mit den Skaleninterkorrelationen.

1.3 Zentrale Formeln

Die Basis für alle ASIA-Berechnungen ist die Varianz-/Kovarianz-Matrix der Poolvariablen. Die Kovarianz der Variablen X und Y wird im folgenden als Cxy bezeichnet, ihre (Pearson Produkt-Moment) Korrelation als Rxy.

Im Schritt S1 benötigt ASIA das Variablenpaar mit der maximalen Korrelation Rxy. Da die Korrelation lediglich die auf die beiden Standardabweichungen standardisierte Kovarianz der beiden Variablen ist, ergibt sich

$$(1) R_{xy} = \frac{C_{xy}}{\sqrt{C_{xx} * C_{yy}}}$$

In den weiteren Schritten wird jeweils die Korrelation einer Variablen mit einer Linearkombination von mehreren Variablen benötigt. Analog zu Formel (1) ergibt sich für die Korrelation einer Variablen v mit einer Skala s

$$(2) R_{vs} = \frac{C_{vs}}{\sqrt{V_v * V_s}}$$

dabei ist V_v die Varianz der Variablen v, V_s die Varianz der Skala s und C_{vs} die Kovarianz von Skala und Variable. Wie sich zeigen läßt (vgl. Kohr 1976), ist das gleichbedeutend mit

$$(2a) R_{vs} = \frac{\left(\begin{array}{l} \text{Summe der Kovarianzen der Variablen v} \\ \text{mit allen Skalenvariablen} \end{array} \right)}{\sqrt{\text{Varianz - v} * \left(\begin{array}{l} \text{Summe der Kovarianzen} \\ \text{der Skalenvariablen} \end{array} \right)}}$$

"Rvs" wird auch als (korrigierte) Trennschärfe bezeichnet und von ASIA unter dieser Rubrik ausgegeben. Falls diese Trennschärfe negativ ist, wird die Variable invertiert in die Skala aufgenommen; ASIA invertiert eine Variable, indem es Zeile und Spalte "v" der Kovarianzmatrix invertiert.

Als weiteres Kriterium, welche Variable am besten zum bestehenden Kern paßt, errechnet ASIA in den Schritten 2 bis n die interne Konsistenz der (erweiterten) Skala durch Cronbachs alpha (bzw. Kuder-Richardson 20):

$$(3) \alpha = \frac{n}{n-1} * \frac{Vs - \text{Summe } Vi}{Vs}$$

wobei Vs für die Varianz der Skala und Vi für die Varianz der Einzelitems steht; n ist die Anzahl der Items in der Skala.

Diesen Wert gibt ASIA aus unter der Rubrik "(unkorrigiertes) Cronbach Alpha". Gleichzeitig errechnet es einen "korrigierten" Wert. Diese Korrektur ist nötig, weil mehrere Alphakoeffizienten nur dann vergleichbar sind, wenn die Itemanzahl n annähernd gleich ist. Dies ist aber beim schrittweisen Vorgehen eben nie der Fall. Mit der bekannten "Spearman-Brown-Formel" (4) wird deshalb jedes alpha auf eine konstante Testlänge von 10 Items standardisiert: $K=n/10$ (vgl. Schmolck 1984)

$$(4) \alpha_{\text{korr}} = \frac{k * \alpha}{1 + (k-1) * \alpha}$$

Als weitere Kennwerte gibt ASIA bei jedem Schritt den Mittelwert und die Standardabweichung der Skala aus; die Standardabweichung ist die Quadratwurzel aus der Skalenvarianz (s. Formel (2a)), der Mittelwert errechnet sich einfach als die Summe der n Einzelmittelwerte (eventuell invertiert).

Bei der Item-Analyse (nach Beendigung der Skalenbildung) gibt ASIA dieselben Werte für die endgültige Skala aus; Cronbachs alpha wird allerdings hier nicht mehr fest auf die Testlänge 10, sondern auf die Länge der ersten Skala standardisiert.

Als zusätzliches Reliabilitätsmaß errechnet ASIA dann die Odd-even-split-half-Korrelation der Test-

hälften. Wieder in Analogie zu Formel (2) ergibt sich

$$(5) Roe = \frac{(\text{Summe der Kovarianzen odd - even})}{\sqrt{\left(\frac{\text{Summe der Cov odd}}{\text{Cov odd}}\right) * \left(\frac{\text{Summe der Cov even}}{\text{Cov even}}\right)}}$$

Als Korrekturschätzungen errechnet ASIA daraus noch die Koeffizienten nach Spearman-Brown (SB) und Flanagan (F):

$$(6) SB = \frac{2 * Roe}{1 - Roe}$$

$$(7) F = \frac{4 * Coe}{2 * Coe + Vo + Ve}$$

dabei bedeuten:

Roe: Odd-Even-Korrelation (s. Formel 5)

Coe: Kovarianzen (Odd-even)

Vo : Varianz der Testhälfte (odd)

Ve : Varianz der Testhälfte (even)

Zur Berechnung der durchschnittlichen Trennschärfen werden die Absolutbeträge aller Korrelationskoeffizienten mittels Fishers z -Transformation umgewandelt, aufsummiert, gemittelt und wieder rücktransformiert. In den ersten Wert ("korr.") gehen alle Skalenitems ein, in den zweiten ("ges.") alle Variablen des Pools.

Wenn man mit dem Schlüsselwort SUMMARY eine Summary-Tabelle anfordert, gibt ASIA erstens noch einmal ein Überblicksschema aus, in dem für jede Va-

riable des Pools die (korrigierte) Trennschärfe zu allen gebildeten Skalen und für jede Variable die durchschnittliche Trennschärfe zu allen Skalen aufgezeigt werden. Zweitens errechnet ASIA dann eine Skalen-Interkorrelationsmatrix, in der die Korrelationen sowohl unkorrigiert (erste Zeile), als auch korrigiert (zweite Zeile) angegeben sind.

Die Korrektur ist nur bei Mehrfachzuordnung nötig; dann nämlich müssen Variablen, die in beiden Skalen vorkommen, aus der größeren Skala entfernt werden, um den Varianzanteil der gemeinsamen Variablen auszusondern. In der dritten Zeile steht dann der Prozentsatz gemeinsamer Variablen in bezug auf die kleinere Skala. Steht hier also der Wert 0.00, gab es keine gemeinsamen Variablen, steht dagegen in dieser Zeile der Wert 1.00, waren alle Variablen der kleineren Skala auch in der größeren enthalten.

Die Korrelation zweier Skalen S und T errechnet sich analog zu Formel (5):

$$(8) R_{st} = \frac{C_{st}}{\sqrt{V_s * V_t}}$$

1.4 Missing-Data-Verarbeitung

Wie oben ausgeführt, ist die Basis aller Berechnungen eine zweidimensionale Varianz-/Kovarianz-Matrix. Da ASIA standardmäßig fehlende Werte nur pro jeweiliger Variablenkombination aus den Berechnungen ausschließt, können aufgrund "unsauberer" Matrizen gravierende Fehler entstehen.

Der Algorithmus (bzw. die Berechnungen) funktioniert optimal nur dann, wenn keine Missing-Data im Datensatz vorkommen, d.h. wenn alle Zellen der Matrix aufgrund derselben Fallzahl errechnet werden.

Nur dann nämlich basieren zum Beispiel die Diagonalelemente der Matrix (die Varianzen der Einzelitems) auf denselben Fällen wie irgendeine Kovarianz-Zelle; und im strengen Sinne sind nur dann die Formeln (1) und (2) anwendbar.

Der Fehler, der begangen wird, wenn dennoch mit Matrizen gerechnet wird, die pro Zelle auf unterschiedlichen Fallzahlen beruhen, hängt davon ab, inwiefern die fehlenden Werte die Standardabweichungen und die Kovarianzen verändern. Wenn die Missing-Data nur "zufällig" fehlen, ist der Fehler sicher sehr klein und für den Zweck des Programmes vernachlässigbar.

1.5 ASIA - Programmablauf

1. Durch die Datendefinitionskommandos (VARIABLE LIST und INPUT FORMAT) wird ASIA mitgeteilt, welche Variablen in welchem Format gelesen werden sollen.
2. Mit den datenbeschreibenden Kommandos (MISSING VALUES, RANGE, VAR LABELS) können dann die Missing-Data-Codes, Minimum und Maximum und erläuternde Label der einzelnen Variablen definiert werden.
3. Die Schritte 1 und 2 können entfallen, wenn bereits ein ASIA-System-File vorliegt, auf dem all diese Informationen zusammen mit den Daten abgespeichert sind (vgl. SAVE FILE und OLD FILE).
4. Falls nur ein Teil des Input-Datensatzes verarbeitet werden soll, kann das in manchen Fällen durch das Kommando N OF CASES festgelegt werden. Wenn die gewünschten Fälle allerdings nicht in der richtigen Reihenfolge am Anfang

des Files liegen, müssen (mit dem Kommando ONLY IF bzw. NOT IF) Selektionskriterien festgelegt werden.

5. Unter Umständen kann es notwendig sein, mit der Anweisung DEFINE MATRIX festzulegen, welche Variablen in die Varianz-/Kovarianz-Matrix aufgenommen werden sollen. DEFINE MATRIX ist allerdings nur dann erforderlich, wenn in einem Lauf mehrere Analysen mit unterschiedlichen Matrizen durchgeführt werden; andernfalls können alle diese Angaben auch bei der CLUSTER-Anweisung angegeben werden. (Tip: Es erhöht die Übersichtlichkeit, wenn man die DEFINE MATRIX-Anweisung dennoch immer verwendet)
6. Erst mit der ersten CLUSTER-Anweisung beginnt ASIA damit, tatsächlich "etwas zu tun":
 - a) ASIA liest nun für jeden Fall die Rohdaten im entsprechenden Format und überprüft dabei, ob eventuell spezifizierte permanente Selektionskriterien erfüllt sind. Sind sie nicht erfüllt, überspringt es den aktuellen Fall und liest den nächsten ein.
 - b) ASIA überprüft dann für jede Variable, für die (mit dem Kommando RANGE) ein Wertebereich angegeben wurde, ob der aktuelle Wert innerhalb dieser Grenzen liegt und ersetzt eventuelle "Ausreißer" durch einen als fehlend deklarierten Code.
 - c) Jetzt wird jeder Fall für spätere Einlesevorgänge binär auf einen Zwischenspeicher-File geschrieben (Achtung: auf diesem File stehen nur die Fälle, die permanente Selektionskriterien erfüllen!).

- d) Anschließend prüft ASIA, ob der eingelesene Fall von der Verarbeitung ausgeschlossen werden muß, weil entweder ein temporäres Selektionskriterium nicht erfüllt ist oder der Fall aufgrund fehlender Werte nicht verarbeitet wird.
 - e) ASIA errechnet nun nach und nach die Kovarianz- bzw. Korrelationsmatrix. Die Größe dieser Matrix richtet sich nach den Angaben in der DEFINE MATRIX-Anweisung bzw. - falls keine solche vorliegt - nach der VARIABLES-Angabe der CLUSTER-Anweisung. Jede Zelle dieser Matrix basiert auf der Anzahl jeweils gültiger Wertepaare.
7. Aus der bei der ersten CLUSTER-Anweisung errechneten Matrix können nun mehrere CLUSTER-Kommandos Skalen extrahieren, ohne daß die Rohdaten jeweils neu eingelesen werden müssen. Ein neuer Matrix-Aufbau, bei dem die Rohdaten dann vom binären Zwischenspeicher gelesen werden (die Schritte 6a bis 6c entfallen), wird nur nötig durch
 - eine DEFINE MATRIX-Anweisung
 - eine MISSING VALUES-Anweisung
 - ein Selektionskommando (*ONLY IF, *NOT IF)
 - eine N OF CASES-Anweisung
 8. Nach jeder CLUSTER-Anweisung kann ASIA veranlaßt werden, für jeden Fall, der beim Matrix-Aufbau beteiligt war, Variablen- und Skalensummenwerte auszuschreiben (WRITE SCORES).
 9. Am Ende des Programms kann ein Systemfile angelegt werden (SAVE FILE), in dem außer den Rohdaten auch alle Informationen zur Datendefinition und zur Datenbeschreibung gespeichert werden.

10. Da das "Herz" von ASIA, der Cluster-Algorithmus, keine Rohdaten, sondern nur die Kovarianzmatrix benötigt, kann ASIA auch mit vorgefertigten Matrizen rechnen. Diese Matrizen müssen von ASIA in einem früheren Lauf errechnet worden sein (SAVE MATRIX; OLD MATRIX).
11. Am Ende des Steuerfiles sollte das Kommando FINISH stehen, das ASIA anzeigt, daß keine weiteren Kommandos folgen.

2 ASIA Steueranweisungen (Überblick)

ASIA verwendet aus Gründen der Kompatibilität die gleiche Befehls-Syntax wie SPSS (Version 9), eines der meistbenutzten Statistiksysteme in den Sozialwissenschaften. Die meisten Anweisungen, v.a. zur Datendefinition und -beschreibung, sind auch nominal bei beiden Programmen identisch.

Im Unterschied zur Version 9 von SPSS benutzt ASIA schon die "Freefield"-Konvention. Das bedeutet, daß Parameter-Angaben nicht unbedingt erst ab Spalte 16 beginnen dürfen (was mitunter bei der Arbeit recht lästig sein konnte), sondern vom Kommando-Schlüsselwort lediglich durch mindestens ein Leerzeichen (BLANK) getrennt sein müssen.

Für den Benutzer bedeutet das: ASIA ist "aufwärtskompatibel" zu SPSS; die SPSS-Syntax wird auch von ASIA verstanden, allerdings nicht notwendigerweise auch umgekehrt. (Vorsicht ist bei Fortsetzungszeilen von Label-Anweisungen geboten: ASIA "übersieht" eventuell erwünschte Leerzeichen oder Kommata, wenn sich das Label auf zwei Eingabezeilen verteilt!)

Die wichtigsten Steueranweisungen zur Datenbeschreibung wurden von SPSS übernommen, neue Kommandos wurden soweit möglich an die SPSS-Terminologie angepaßt. In der folgenden Übersicht sind alle ASIA-Kommandos mit ihren Schlüsselwörtern verzeichnet. In vielen Fällen wird diese Information für geübte SPSS-Anwender bereits ausreichen; im nächsten Abschnitt allerdings finden Sie zu jedem Kommando detaillierte Erläuterungen und in Abschnitt 4 einen Überblick über die Syntax der einzelnen Kommandos.

2.1 SPSS kompatible Steueranweisungen

VARIABLE LIST	INPUT FORMAT
N OF CASES	MISSING VALUES
VAR LABELS	VALUE LABELS (*)
PAGESIZE (*)	PRINT FORMATS (*)
EDIT	NUMBERED
PRINT BACK	RUN NAME
TASK NAME	COMMENT
READ INPUT DATA	END INPUT DATA
FINISH	

(*) Die markierten Kommandos sind implementiert wie COMMENT; alle Angaben werden ignoriert.

2.2 SPSS-ähnliche Steueranweisungen

DEFINE MATRIX	CLUSTER
OLD FILE	OLD MATRIX
SAVE FILE	SAVE MATRIX
RANGE	ONLY IF (*)
NOT IF (*)	AND IF (*)
AND NOT (*)	BUT NOT (*)
WRITE SCORES	

(*) Die markierten Kommandos gibt es auch in einer temporären Version (*ONLY IF, usw.).

2.3 Reihenfolge der Steueranweisungen

1. Die Kommandos EDIT, NUMBERED, PRINT BACK, FILE NAME, RUN NAME, TASK NAME und COMMENT (bzw. alle als Kommentar implementierten Kommandos) können an jeder Stelle und wiederholt im Steuerfile vorkommen, sie dürfen allerdings nicht zwischen der ersten CLUSTER- und der READ INPUT DATA-Anweisung oder zwischen den Anweisungen READ INPUT DATA und END INPUT DATA stehen.

2. Auf jeden Fall notwendig sind entweder OLD FILE bzw. OLD MATRIX oder VARIABLE LIST und INPUT FORMAT; OLD FILE/OLD MATRIX auf der einen Seite und VARIABLE LIST und INPUT FORMAT auf der anderen Seite schließen sich gegenseitig aus.
3. N OF CASES ist nur bedingt notwendig; wenn die Angabe fehlt, gilt N OF CASES UNKNOWN.
4. READ INPUT DATA muß direkt hinter der ersten CLUSTER-Anweisung stehen. Datenrecords müssen direkt auf die READ-INPUT-DATA-Anweisung folgen.
5. Wenn die Anzahl der Fälle nicht bekannt ist oder nicht angegeben wurde, muß nach dem letzten Datenrecord END INPUT DATA stehen. END INPUT DATA ist die einzige ASIA-Anweisung, die in Großbuchstaben geschrieben sein muß, alle anderen Anweisungen erkennt ASIA auch, wenn sie klein geschrieben sind.
6. Die datenbeschreibenden Kommandos (MISSING VALUES, RANGE und VAR LABELS) können jederzeit (neu) eingegeben werden (allerdings muß ASIA natürlich die Variablennamen kennen, was erst nach VARIABLE LIST bzw. OLD FILE/OLD MATRIX der Fall ist).
7. Permanente Selektionskommandos (ONLY IF, AND IF usw.) können nur vor der ersten CLUSTER-Anweisung stehen; die temporären Formen (*ONLY IF, *AND IF usw.) sind auch später möglich; allerdings ist die Zahl der Selektionskriterien pro Matrix-Aufbau auf 5 beschränkt.
8. Die Anweisung CLUSTER ist beliebig oft erlaubt, sobald die Eingabedaten definiert sind (also

nach OLD FILE/OLD MATRIX bzw. VARIABLE LIST und INPUT FORMAT).

9. WRITE SCORES ist erst nach der ersten CLUSTER-Anweisung sinnvoll und erlaubt.
10. SAVE FILE und SAVE MATRIX müssen unmittelbar vor FINISH stehen; sie schließen sich gegenseitig aus.
11. Anweisungen, die hinter dem Kommando FINISH stehen, werden nicht mehr verarbeitet.

2.4 Das Format der Steueranweisungen

ASIA-Anweisungen bestehen aus zwei verschiedenen Feldern: dem Kontrollfeld mit dem "Namen" des Kommandos und dem Parameterfeld mit den nötigen Angaben.

Das Kontrollfeld muß in Spalte 1 mit dem Namen des Kommandos beginnen. Alle Eingabezeilen, in denen Spalte 1 nicht besetzt ist, werden als Fortsetzungszeile des vorhergehenden Kommandos interpretiert. Der Kommandoname muß vollständig und korrekt (ab Spalte 1) geschrieben sein. Groß- und Kleinschreibung spielt dabei allerdings - mit Ausnahme der Anweisung END INPUT DATA - keine Rolle. Das Kontrollfeld muß durch ein oder mehrere Trennzeichen (Kommata oder Leerzeichen) vom Parameterfeld getrennt werden.

Das Parameterfeld geht bis Spalte 80: durch die NUMBERED-Anweisung können die Spalten 73-80 allerdings für Kennzeichnungen reserviert werden. Innerhalb des Parameterfeldes ist die Position der Parameter bzw. der Parameterelemente völlig beliebig, sofern bei Fortsetzungszeilen Spalte 1 frei bleibt.

Parameter bestehen aus 5 Elementen: Namen, Schlüsselwörter, Label und Trennzeichen:

- Namen dienen dazu, Variablen, Skalen oder Files zu bezeichnen. Sie müssen aus 1 bis maximal 8 alphanumerischen Zeichen bestehen, wobei das erste Zeichen immer ein Buchstabe sein muß.
- Werte (oder Konstanten) können entweder numerisch oder alphanumerisch sein. Numerische Werte können ganzzahlig sein, mit einem Dezimalpunkt versehen oder in wissenschaftlicher Schreibweise angegeben werden; als Vorzeichen sind "+" und "-" erlaubt (Beispiele: 1.5, 2.5, -2, .14E-2, +.5). Alphanumerische Werte müssen in Hochkommata eingeschlossen werden; wenn sie länger als 4 Zeichen sind, wird von rechts abgeschnitten (Beispiele: 'ABC', 'X', '01').
- Schlüsselwörter sind ASIA-eigene "Namen" mit spezieller Bedeutung; sie müssen vollständig und korrekt angegeben werden. Ihre Bedeutung wird im Detail bei den einzelnen Kommandos beschrieben.
- Label (Etiketten) dienen dazu, Files, Skalen und Variablen ausführlichere Texte zuzuordnen; sie können bis zu 40 Zeichen lang sein und (mit Ausnahme der besonderen Trennzeichen) jedes druckbare Zeichen enthalten.
- Trennzeichen sind erstens die allgemeinen Trennzeichen Blank und Komma; sie sind untereinander austauschbar und können beliebig oft zwischen Parameter-Elemente eingestreut werden, um die Lesbarkeit zu erhöhen. Sie sind notwendig, um die Parameter-Elemente untereinander zu trennen, sofern sie nicht ohnehin durch eines der besonderen Trennzeichen (s.u.) unterschieden sind. Zweitens gibt es die besonderen Trennzeichen "(", ")", "/", und "=". Diese Trennzeichen sind nur nötig, wenn

sie von der Parameter-Syntax ausdrücklich verlangt werden.

Mehrere ASIA-Kommandos verlangen oder erlauben einen Parameter "varlist" (bzw. "slist"). Eine solche Variablenliste besteht im Prinzip einfach aus einer Folge von Variablen- (oder Skalen-) Namen, die durch mindestens ein Trennzeichen (Blank oder Komma) getrennt sind. So könnte also "varlist" durch folgende Angabe ersetzt werden:

```
GESCHL,EINKOMM,FRAG1,FRAG2,FRAG3
```

Wenn allerdings zum Beispiel bei der CLUSTER-Anweisung 200 Variablen in den Itempool aufgenommen werden sollen, wäre es ziemlich viel Schreiarbeit, alle 200 Variablennamen einzeln einzugeben. In ASIA gibt es deshalb die Möglichkeit, sogenannte inklusive Listen zu bilden. Die Angabe

```
GESCHL TO BILDUNG
```

wird von ASIA folgendermaßen interpretiert: nimm alle Variablen in die Liste auf, die zwischen "GESCHL" und "BILDUNG" (jeweils einschließlich) liegen. Da es keine "natürliche" Ordnung von Variablen gibt, muß man aber genau darauf achten, auf welche Reihenfolge sich diese "TO-Konvention" bezieht:

- Die Reihenfolge aller Variablen wird normalerweise durch die VARIABLE-LIST-Anweisung festgelegt; auf diese Anordnung (die sich auch durch SAVE FILE bzw. OLD FILE nicht ändert) beziehen sich zum Beispiel die Variablenlisten der Kommandos MISSING VALUES oder RANGE.
- Die Reihenfolge der Variablen allerdings, die in die Berechnung einer Matrix eingehen, wird durch das Kommando DEFINE MATRIX festgelegt; auf diese

Anordnung bezieht sich zum Beispiel die Variablenliste des Kommandos CLUSTER. Es ist keine schlechte Idee, bei der Definition einer Matrix die ursprüngliche Reihenfolge der Variablen beizubehalten, wenn nicht sehr gute Gründe dagegen sprechen!

- Die Reihenfolge der Skalen wird bei jeder CLUSTER-Anweisung neu festgelegt, und zwar entsprechend der Reihenfolge der Definition durch das Schlüsselwort SCALE.

ACHTUNG: inklusive Listen haben nichts mit einem eventuellen Index der Variablennamen zu tun! Bei der

```
VARIABLE LIST VAR1,VAR2,SEX,VAR3,VAR5,VAR4
```

bedeutet die "varlist"

```
VAR1 TO VAR5
```

nicht, daß die Variablen VAR1, VAR2, VAR3, VAR4 und VAR5 in die Liste aufgenommen werden, sondern alle Variablen, die in der obigen Reihenfolge zwischen VAR1 und VAR5 liegen. Die Variable "VAR4" fällt also weg, dagegen wird "SEX" mit in die Liste aufgenommen!

"Normale" und inklusive Listen können beliebig gemischt werden; folgendes sind zum Beispiel gültige "varlist"-Angaben:

```
GESCHL
FRAG1 TO FRAG10
ALTER TO BERUF, KINDER
V1,V5,V8,X80 TO X105
```

Eine weitere Vereinfachung der Schreibweise von Variablen-Listen bietet das Schlüsselwort ALL, das

immer dann für eine "varlist" eingesetzt werden kann, wenn alle definierten Elemente angesprochen werden sollen.

Beispiele:

```
MISSING VALUES  ALL(-9)
DEFINE MATRIX    VARS=ALL
CLUSTER         .../ RKRIT(ALL)=.50
```

2.5 Hinweise zur Notation

In der folgenden Darstellung der einzelnen Kommandos und ihrer Parameter gelten folgende Vereinbarungen:

1. Angaben in spitzen Klammern (<>) sind optional; wenn sie weggelassen werden, setzt ASIA einen voreingestellten Wert ein, der im Text erläutert wird.
2. Das Zeichen & bedeutet, daß zwischen den einzelnen Elementen innerhalb zweier Ampersands gewählt werden kann. Die einzelnen Elemente werden durch ! getrennt, die unterstrichenen Elemente sind voreingestellt.

Beispiel:

```
NUMBERED <& YES ! NO &>
```

3. Angaben in GROSSBUCHSTABEN sind Schlüsselwörter und müssen exakt übernommen werden; Angaben in Kleinbuchstaben sind Beschreibungen des Parameters und müssen ersetzt werden, eventuell werden sie im Text ausführlicher erläutert. Die Angabe "varlist" muß zum Beispiel durch eine Variablenliste, "value" durch einen Wert ersetzt werden.

4. Die besonderen Trennzeichen "()/=" müssen immer übernommen werden.
5. Eine Ellipse (Auslassungszeichen) "..." innerhalb spitzer Klammern deutet an, daß der Klammereinhalte beliebig oft wiederholt werden kann.

Wenn ASIA einen Fehler findet, gibt es unter der entsprechenden Spalte (die allerdings nicht genau die fehlerhafte sein muß!) ein Fragezeichen ("?") aus. Normalerweise reicht das aus, um den Fehler zu finden. Da die Fehler meist nicht vorsätzlich eingebaute "Fallen" für den Computer sind, sondern oft nur Tippfehler, die ohnehin ins Auge springen, wurde nicht allzuviel Aufwand für die Fehlerrückverfolgung betrieben; in den meisten Fällen müßte die Position des Fragezeichens genügend Hinweise für den Grund des Fehlers bieten.

3 Die einzelnen Steueranweisungen

3.1 Beschreibung der Eingabedaten

Zur einfachen Beschreibung der Eingabedaten sind folgende Anweisungen notwendig:

Mit VARIABLE LIST wird für jede einzulesende Variable ein eindeutiger Name definiert, unter dem sie dann später angesprochen werden kann. Gleichzeitig legt diese Angabe die Anzahl und Reihenfolge der Variablen im Datensatz fest. Die INPUT FORMAT-Anweisung spezifiziert dann den Dateityp (festes oder freies Format), die Zahl der Records pro Fall, in welchen Spalten welche Variablen stehen, ob numerisch oder alphanumerisch, ganzzahlig oder mit Dezimalbruch.

Die N OF CASES-Anweisung legt die Anzahl der Fälle fest, die eingelesen werden sollen.

3.1.1 VARIABLE LIST

Die VARIABLE LIST-Anweisung gibt jeder einzulesenden Variablen einen Namen. Sie legt damit implizit auch die Zahl und die Reihenfolge der Variablen fest.

VARIABLE LIST variablenliste

Die Variablenliste besteht aus einer Folge von Namen und/oder inklusiven Listen mit dem Schlüsselwort TO. Namen können aus 1 bis maximal 8 Buchstaben oder Ziffern bestehen; das erste Zeichen eines

Namens muß ein Buchstabe sein. Eine inklusive Liste hat die allgemeine Form

XXXnn TO XXXmm

ASIA kreiert dabei automatisch mm-nn+1 Variablen mit dem Präfix XXX und dem Index i=nn bis mm. Der Index wird mit so vielen führenden Nullen kodiert, wie bei nn angegeben wurden; mm und nn müssen ganze, positive Zahlen, mm muß größer als nn sein.

Beispiel:

VAR1 TO VAR10 = VAR1,VAR2, ..., VAR9,VAR10
VAR01 TO VAR10 = VAR01,VAR02, ..., VAR09,VAR10

Kein Name darf zweimal vorkommen; die reservierten "Namen" ALL und TO dürfen nicht als Variablennamen verwendet werden. Weder implizit noch explizit dürfen mehr als 250 Variablen definiert werden.

Gültige Beispiele für die VARIABLE LIST-Anweisung:

VARIABLE LIST AGE,SEX,INCOME,COUNTY,FAMSTAT
VARIABLE LIST V1 TO V257
VARIABLE LIST A1 TO A10 VPB CRD A11 BERUF
 Z01 TO Z21 Q1 TO Q111
 AGE PARTY NONSENSE

3.1.2 INPUT FORMAT

Die INPUT FORMAT-Anweisung gibt an, ob die Daten in festem oder freiem Format kodiert sind. Beim festen Format müssen darüber hinaus in einer Formatliste die Zahl der Records pro Fall und die Zuordnung der einzelnen Werte in den Spalten zu den Variablen der VARIABLE LIST festgelegt werden.

Die INPUT FORMAT-Anweisung kennt also zwei verschiedene Modi, die durch ein Schlüsselwort ausgewählt werden können: FREEFIELD oder FIXED.

3.1.2.1 Das Schlüsselwort FREEFIELD

```
INPUT FORMAT  FREEFIELD
```

Die einfachste Form der Datenbeschreibung - bei großen Datenmengen leider aber auch die fehleranfälligste - ist die unformatierte Form. Im Anweisungsfeld der INPUT FORMAT-Anweisung muß dann lediglich das Schlüsselwort FREEFIELD stehen. Die Daten können einfach hintereinander aufgeführt werden, der Benutzer muß nur darauf achten, daß erstens die Reihenfolge der Werte der Reihenfolge der Variablen entspricht und zweitens jeder Wert vom anderen durch ein oder mehrere Leerzeichen oder Kommata getrennt ist.

Bei der FREEFIELD-Konvention ist zu beachten:

- Die Daten müssen fallweise in der (durch VARIABLE LIST) vereinbarten Reihenfolge stehen.
- Alphanumerische Variablen müssen in Hochkommata (""") eingeschlossen werden und dürfen höchstens 4 Zeichen lang sein.
- Kein Wert darf in zwei verschiedenen Zeilen stehen
- Nullen in numerischen Variablen müssen explizit geschrieben werden! Das Schlüsselwort BLANK ist

im Zusammenhang mit unformatierter Eingabe natürlich nicht sinnvoll!

- Bei der unformatierten Eingabe werden immer nur die Spalten 1 bis 80 interpretiert; wenn die NUMBERED-Anweisung in Kraft ist, werden zudem die Spalten 73 bis 78 ignoriert!

3.1.2.2 Das Schlüsselwort FIXED

```
INPUT FORMAT  <FIXED> (formatliste)
```

Wenn das Schlüsselwort FIXED angegeben ist (bzw. das Schlüsselwort ganz wegfällt), dann beschreibt die "formatliste" den Aufbau eines exemplarischen Falles der Datei. Für jede einzulesende Variable muß ein Formatelement angegeben werden, das den Typ und die Spaltenposition der Variablen festlegt. ASIA interpretiert die "formatliste" zusammen mit den Angaben der VARIABLE LIST von links nach rechts; das heißt, das erste Formatelement beschreibt die erste Variable der VARIABLE LIST, das zweite Formatelement die zweite Variable usw.

In der "formatliste" dürfen die FORTRAN-äquivalenten Formatelemente F, A, X, T und / vorkommen:

- Das F-Format (rFw.d) beschreibt eine numerische Variable; "w" ist die Feldweite oder die Anzahl reservierter Spalten für diesen Wert; "d" ist die Anzahl der Ziffern rechts vom Dezimalpunkt, für ganze Zahlen ist d=0. Falls ein Dezimalpunkt tatsächlich abgelocht ist, hat dessen Position Vorrang über die Formatangabe. "r" ist ein Wiederholungsfaktor, der für r=1 entfallen kann.

Hinweis: Da in den Sozialwissenschaften hauptsächlich mit ganzen Zahlen gearbeitet wird, erlaubt es ASIA im Unterschied zu SPSS, die Angabe Fw.0 abzukürzen (in der Form Fw. oder Fw).

- Das A-Format (rAw) beschreibt alphanumerische Variablen. Alphanumerische Werte können Folgen von maximal 4 beliebigen Zeichen sein. "r" ist der Wiederholungsfaktor, "w" die Anzahl der einzulesenden Zeichen (intern wird rechts mit Leerzeichen auf 4 Stellen aufgefüllt).
- Mit dem X-Format (wX) können "w" einzulesende Spalten (leere oder nicht gewünschte) übersprungen werden.
- Das T-Format (Tn) erlaubt es, die nächste einzulesende Spalte direkt anzugeben. Zusätzlich zu den Möglichkeiten des X-Formats können damit auch bereits eingelesene Spalten erneut angesprochen werden.
- Der Schrägstrich (/) befiehlt dem Programm, die Bearbeitung des aktuellen Records zu beenden und den nächsten einzulesen. Wenn ein Fall aus n Records besteht, müssen in der Formatliste genau n-1 Schrägstriche vorkommen.

Die Formatliste kann oft wesentlich vereinfacht werden, wenn man die Möglichkeiten der Zusammenfassung nutzt:

- Wenn beispielsweise 20 einstellige Variablen hintereinanderstehen, braucht man nicht 20 mal F1.0 zu schreiben, sondern kann durch 20F1.0 den Wiederholungsfaktor "r" ausnutzen.
- Kompliziertere Strukturen können durch Klammerung schematisiert werden: 20(F1.0,1X,A2) zum Beispiel könnte angegeben werden, wenn für 20 Fragen eines

Fragebogens hintereinander ein einstelliger numerischer Code eingelesen, eine uninteressante (leere) Spalte übersprungen und wieder ein zweistelliger (alphanumerischer) Code eingelesen werden soll.

Beispiel:

```
INPUT FORMAT (/10X,5F1.0,2X,A2/10X,5F1.0,T70,F5.4)
```

Dieses Format wird folgendermaßen interpretiert:

1. Überspringe den ersten Record (wegen des "/")
2. Überspringe 10 Spalten (10X; nächste Spalte=11)
3. Lies 5 einstellige ganze Zahlen aus den Spalten 11 bis 15 (5F1.0)
4. Überspringe die Spalten 16 und 17 (wegen 2X)
5. Lies eine zweistellige alphanumerische Variable aus den Spalten 18 und 19 (A2)
6. Springe zum nächsten (dritten) Record (wegen "/")
7. Überspringe die ersten 10 Spalten (10X)
8. Lies 5 einspaltige ganze Zahlen (Spalten 11-15)
9. Springe zur Spalte 70 (wegen T70)
10. Lies aus den 5 Spalten 70-74 eine Zahl mit insgesamt 5 Stellen und interpretiere die rechten 4 Stellen als Nachkommateil (es sei denn, in diesen 5 Spalten ist ein Dezimalpunkt abgeleuchtet) (F5.4)

Die Schritte 1 bis 10 werden für jeden Fall der Eingabedatei wiederholt; jeder Fall muß aus drei Records bestehen.

Zur Kontrolle gibt ASIA eine Korrespondenzliste aus, an der abzulesen ist, welche Variablen in welchem Format aus welchen Spalten gelesen werden:

VARIABLE LIST V1 TO V10, KENNUNG, V11 TO V20, RATIO
 INPUT FORMAT FIXED(/10X, 5F1, 2X, A2/10X, 5F1,
 T70, F5.4)

PRO FALL WERDEN 3 "KARTE(N)" MIT
 MAXIMAL 74 "SPALTEN" EINGELESEN

VARIABLE RECORD COLUMNS FORMAT

V1	2	11-11	F1.0
V2	2	12-12	F1.0
V3	2	13-13	F1.0
.....			
KENNUNG	2	18-19	A2
V11	3	11-11	F1.0
.....			
RATIO	3	70-74	F5.4

3.1.2.3 Unterschiede zu SPSS

- Die Angabe Fw.d kann bei ASIA für d=0 abgekürzt werden zu Fw. oder Fw
- Die Schreibweise "n/" ist in ASIA erlaubt und äquivalent zu "n(/)"
- Die maximale Verschachtelungstiefe von Klammern ist 5
- Die Anzahl der Formatelemente und die Anzahl der Variablen brauchen nicht unbedingt übereinzustimmen (ASIA warnt dann lediglich mit der Meldung ?NOMATCH)
- Geklammerte Ausdrücke ohne Wiederholungsfaktor werden solange wiederholt, bis alle Variablen abgearbeitet sind. Am Beginn der Klammer Oter Ordnung wird ein neuer Record angefordert

- Die Trennzeichen Blank oder Komma sind auch innerhalb von Formatanweisungen völlig austauschbar und dürfen beliebig oft zwischen Formatelemente eingestreut werden

3.1.3 N OF CASES

N OF CASES	<& n <u>UNKNOWN</u> &>
------------	--------------------------

Die N OF CASES-Anweisung gibt die Anzahl der Fälle an, die eingelesen bzw. verarbeitet werden sollen. Das Schlüsselwort UNKNOWN und die Angabe N OF CASES 0 sind gleichbedeutend. Sie sind nur erlaubt vor der ersten CLUSTER-Anweisung und auch dann nur, wenn nicht mit OLD FILE ein ASIA-System-File eingelesen wurde. In allen anderen Fällen muß "n" (eine positive, ganze Zahl größer 0) angegeben werden, n muß dabei kleiner oder gleich der Anzahl der beim ersten Mal eingelesenen Fälle (bzw. der Anzahl der Fälle im System-File) sein.

Wird das Schlüsselwort UNKNOWN verwendet, liest ASIA alle Fälle bis zum Ende der Eingabedaten, das durch die END INPUT DATA-Anweisung definiert sein muß.

Das Schlüsselwort UNKNOWN ist voreingestellt, wenn keine N OF CASES-Angabe vorliegt.

3.2 Erweiterte Beschreibung der Eingabedaten

Mit den folgenden Datendefinitionsanweisungen können ASIA weitere Informationen über die Variablen

mitgeteilt werden. FILE NAME und VAR LABELS dienen der Dokumentation und erhöhen die Lesbarkeit der ASIA-Ausgabe; sie sollten deshalb ausführlich genutzt werden.

Mit den MISSING VALUES-Angaben können für jede Variable Werte oder Wertebereiche festgelegt werden, die nicht in die Analyse aufgenommen werden sollen.

Durch die RANGE-Anweisung schließlich wird für jede Variable der theoretische Wertebereich festgelegt. Diese Angaben sind nötig, wenn eine Variable bei der Skalenbildung invertiert werden muß (s. Invertierung).

SPSS kennt in diesem Zusammenhang noch die Anweisungen PRINT FORMATS und VALUE LABELS; da ASIA diese Informationen nicht benötigt, diese Anweisungen aber oft in SPSS-Steuerfiles vorkommen, sind diese Anweisungen in ASIA als COMMENT implementiert.

3.2.1 FILE NAME

```
FILE NAME < name <label> >
```

Die FILE NAME-Anweisung gibt einer Datei einen Namen von maximal 8 Zeichen; "name" darf nur aus alphanumerischen Zeichen bestehen, das erste Zeichen muß ein Buchstabe sein. Die Dateietikette "label" darf alle Zeichen enthalten und maximal 72 Zeichen lang sein.

Name und Etikette des Datensatzes werden am Kopf jeder Seite von ASIA ausgegeben und bei einem SAVE

FILE-Kommando zusammen mit den Daten und den übrigen Datendefinitionen im Systemfile abgespeichert.

Wenn keine FILE NAME-Anweisung vorliegt, setzt ASIA standardmäßig den Namen NONAME und eine leere Etikette ein; ein FILE NAME ohne Parameter allerdings löscht die vorherige Definition nicht.

ACHTUNG: Keine Fortsetzungszeile(n) erlaubt!

3.2.2 VAR LABELS

```
VAR LABELS varnam,varlabel </varnam,varlabel...>
```

Dieses Kommando dient dazu, den Variablen außer ihrem Namen "varnam" noch eine maximal 40 Zeichen lange Etikette "varlabel" zuzuordnen, die bei der Analyse ausgegeben wird und die Lesbarkeit erhöht, bzw. vor Fehlinterpretationen schützt.

"varnam" ist der Name der Variablen, wie er mit der VARIABLE LIST-Anweisung definiert wurde; "varlabel" ist ein beliebiger Text, der allerdings keinen Schrägstrich ("/") enthalten darf.

"varnam" und "varlabel" müssen durch mindestens ein Trennzeichen (Blank oder Komma) voneinander getrennt werden. "varlabel" darf beliebig lang sein und sich über mehrere Fortsetzungszeilen erstrecken; verarbeitet werden allerdings immer nur die ersten 40 Zeichen. Mehrere "varnam,varlabel"-Angaben müssen voneinander durch "/" getrennt werden.

3.2.3 VALUE LABELS

```
VALUE LABELS  beliebiger Text
```

Nur aus Gründen der SPSS-Kompatibilität implementiert; das Parameterfeld wird vollständig ignoriert.

Vergleiche auch PRINT FORMATS und PAGESIZE

3.2.4 MISSING VALUES

```
MISSING VALUES varlist(codes)
                </varlist(codes)...>
```

MISSING VALUES sollten immer definiert werden. Grundsätzlich läuft ASIA auch ohne eine MISSING VALUES-Angabe, aber es besteht dann die Gefahr, daß die internen Berechnungen auf Grund von Fließkommafehlern abgebrochen werden.

Die Angaben bestehen aus einer Variablenliste "varlist" und einer Codeliste "codes", die in runde Klammern eingeschlossen werden muß. Die "varlist" kann auch durch das Schlüsselwort ALL ersetzt werden, wenn alle Variablen des Datensatzes angesprochen werden sollen. Weitere Definitionen können in derselben Zeile folgen, wenn sie durch einen Schrägstrich ("/") abgetrennt werden.

Die in Klammer eingeschlossene Codeliste kann folgende Formen annehmen:

- a) (x <,y> <,z>)
- b) (x THRU y <,z>)
- c) (x, y THRU z)
- d) ()

Man kann also pro Variable entweder (a) einen bis maximal drei einzelne Werte angeben oder (b und c) einen Wertebereich und zusätzlich einen einzelnen Wert. Dabei kann anstelle eines Wertes auch das Schlüsselwort BLANK eingesetzt werden, wenn zwar leere Spalten, nicht aber die Zahl 0 als fehlend zu deklarieren sind.

Bei der Angabe eines Intervalls können als Intervallgrenzen auch die Schlüsselwörter LOWEST (oder LO) bzw. HIGHEST (oder HI) verwendet werden, die von ASIA durch die kleinste bzw. größte maschinenintern darstellbare Zahl ersetzt werden.

Wenn die Definition von Missing-Data-Codes wieder aufgehoben werden soll, muß die Codeliste zwischen den Klammern leer bleiben (Form d); diese Möglichkeit ist vor allem nützlich, um Missing-Data-Definitionen in System-Files wieder aufzuheben.

ACHTUNG: ASIA kann keine Missing-Data-Definitionen für alphanumerische Variablen abspeichern! Allerdings führt eine syntaktisch richtige Angabe nicht zum Programmabbruch, sondern nur zu einer Warnung; SPSS-Steueranweisungen können also ohne Änderungen übernommen werden.

Beispiele:

```
MISSING VALUES ALTER(-1)
MISSING VALUES V1 TO V11,BETA,NEU(-9)/
```

```

EINKOMM,ERZIEH(0)/
SKALA1 TO SKALA10(0,8 THRU HIGHEST)
MISSING VALUES SKAL55 TO SKAL60(LOWEST THRU 5)/
SKAL61()
MISSING VALUES ALL (BLANK)/ALPHA ('KA')

```

ASIA verarbeitet Missing Values bzw. Out Of Range Values auf folgende Weise:

Bei der Berechnung von Item-Mittelwerten, Standardabweichungen etc. werden zur Berechnung die von Missing bzw. Out Of Range verschiedenen Werte herangezogen ($\text{valid } n \leq n$). Die Kovarianzen und Korrelationen werden stets anhand der vorhandenen Meßwertpaare berechnet.

Wenn viele Missing Values auftreten und wenn $\text{valid } n$ von Itempaar zu Itempaar stark variiert, dann können schwer vorhersehbare Fehler auftreten, da der ASIA-Algorithmus von einer Matrix ausgeht, die auf vollständigen Daten basiert. In solchen Fällen sollte man über die vorherige Bereinigung der Datenmatrix nachdenken - z.B. könnte man die fehlenden Daten ersetzen durch

- den je Item bestimmten interindividuellen Modalwert, also die häufigste in der Stichprobe auftretende Reaktion,
- den pro Person bestimmten intraindividuellen Modalwert, also die häufigste Reaktion dieser Person innerhalb der Items einer Skala; dies wird freilich nur bei Itemanalyse möglich sein.

Siehe dazu auch Abschnitt 1.4.

3.2.5 RANGE

```
RANGE varlist(min,max) </ varlist(min,max)...>
```

Auch die RANGE-Anweisung sollte in einem ASIA-Steuerfile nicht fehlen. RANGE definiert für jede Variable in "varlist" die Werte "min" und "max" als theoretisches Minimum bzw. Maximum der Variablen. Diese Angaben benötigt ASIA, um für invertierte Items (\rightarrow Invertierung) die Skalensumme bzw. den Skalenmittelwert korrekt berechnen zu können. Allerdings ist dies für den "eigentlichen" Algorithmus nicht von Bedeutung. Falls deshalb eine Variable invertiert in eine Skala aufgenommen werden muß, für die keine RANGE-Angabe gemacht wurde, ist lediglich der Skalenmittelwert nicht mehr interpretierbar; ASIA macht dies bei der Ausgabe kenntlich, indem es hinter den entsprechenden Wert ein Ausrufezeichen setzt. Bei der Ausgabe von Skalensummenwerten allerdings (s. WRITE SCORES), wird bei allen Personen für diesen Wert der Missing-Data-Code ausgegeben.

Liegt eine RANGE-Angabe vor, überprüft ASIA beim Einlesen der Daten, ob der eingelesene Wert innerhalb des gültigen Bereichs liegt. Falls dies nicht der Fall ist, gibt es eine Meldung aus und ersetzt diesen Wert durch den ersten Wert der MISSING VALUES-Angabe für diese Variable. Wurde für diese Variable kein Missing-Data-Code definiert, gilt automatisch dieser "Ausreißerwert" als Missing-Data Code.

ACHTUNG: "max" muß natürlich größer oder gleich "min" sein; für alphanumerische Variablen sind RANGE-Angaben nicht sinnvoll, man sollte deshalb

vorsichtig sein, wenn man das Schlüsselwort ALL benutzt!

Beispiel:

```
RANGE SKAL1 TO SKAL9(1,9)/A,B,C (5,11)
RANGE ALL(0,9)
```

3.2.6 PRINT FORMATS

```
PRINT FORMATS  beliebiger Text
```

Nur aus Gründen der SPSS-Kompatibilität implementiert; das Parameterfeld wird vollständig ignoriert.

Vergleiche auch VALUE LABELS und PAGESIZE

3.3 Zur Form der Ein- und Ausgabe

Neben den bisher beschriebenen Datendefinitionsanweisungen gibt es noch eine Reihe von Kommandos, die einige Modalitäten der Ein- und Ausgabe von ASIA festlegen:

- das EDIT-Kommando dient der Fehlersuche
- RUN NAME, TASK NAME und COMMENT sind nützlich, um die Ausgabe besser dokumentieren zu können
- NUMBERED und PRINT BACK beeinflussen die Form der Eingabe bzw. Ausgabe von Kommandozeilen

3.3.1 EDIT

```
EDIT  <beliebiger Text>
```

Die EDIT-Anweisung ermöglicht es, ASIA-Steueranweisungen auf formale Richtigkeit zu prüfen. Dies ist besonders nützlich bei langen Batchjobs und/oder bei Rechnern mit geringer Rechenleistung.

Im Anschluß an die EDIT-Anweisung werden keine Daten mehr eingelesen und keine Berechnungen mehr angestellt. Das bedeutet für den Aufbau des Steuerfiles:

- Zwar muß die READ INPUT DATA-Anweisung vorhanden sein, es dürfen ihr aber weder Eingabedaten noch die END INPUT DATA-Anweisung folgen.
- Bei der Eingabe von ASIA-Systemfiles muß der File vorhanden sein, weil ASIA die Variablen und die Datendefinitionen lesen muß.

Im Unterschied zu SPSS kann die EDIT-Anweisung an jeder beliebigen Stelle im Steuerfile stehen (allerdings gilt sie natürlich auch erst ab dieser Stelle). Eventuelle Parameter (auch in Fortsetzungszeilen) werden als Kommentar behandelt und ignoriert.

3.3.2 RUN NAME

```
RUN NAME <beliebiger Text>
```

Der Text ab dem ersten Zeichen, das von Komma und Blank verschieden ist, bis zum Ende der Zeile (Fortsetzungszeilen sind nicht erlaubt!) wird am Kopf jeder Seite ausgegeben.

Unterschiede zu SPSS: kann an jeder beliebigen Stelle und beliebig oft benutzt werden; speichert max. 72 Zeichen.

3.3.3 TASK NAME

```
TASK NAME <beliebiger Text>
```

TASK NAME entspricht im wesentlichen der Funktion der RUN NAME-Anweisung: Der maximal 72 Zeichen lange Text wird am Kopf jeder Seite (in der zweiten Zeile) ausgegeben; der Text kann jederzeit durch eine neue TASK NAME-Anweisung überschrieben werden. Normalerweise wird man eine TASK NAME-Anweisung vor jedes CLUSTER-Kommando setzen, um die Ausgabe später besser identifizieren zu können. Etwaige Selektionskriterien sollten in TASK NAME nochmals aufgeführt werden, damit diese am Kopf jeder Tabelle und nicht nur in den Kommandozeilen abgelesen werden können.

ACHTUNG: keine Fortsetzungszeilen erlaubt!

3.3.4 COMMENT

```
COMMENT <beliebiger Text>
```

COMMENT-Zeilen können (fast) beliebig in den ASIA-Steuerfile eingefügt werden; sie beeinflussen die Funktion von ASIA nicht, sondern werden nur zusammen mit den anderen Kommandozeilen ausgegeben. COMMENT-Zeilen können fortgesetzt werden; sie dürfen nicht stehen

- zwischen einer Steueranweisung und ihren Fortsetzungszeilen,
- zwischen der ersten Prozedur (CLUSTER) und der READ INPUT DATA-Anweisung,
- zwischen der READ INPUT DATA- und der END INPUT DATA-Anweisung!

3.3.5 PRINT BACK

```
PRINT BACK <& NO ! CONTROL ! FORMAT &>
```

Die PRINT BACK-Anweisung steuert die Ausgabe der eingelesenen Kommandozeilen und der Formatreferenz-tabelle (s. INPUT FORMAT FIXED), die die Angabe der VARIABLE LIST-Anweisung und der INPUT FORMAT-Anweisung gegenüberstellt.

Ursprünglich - und durch eine PRINT BACK-Anweisung ohne Parameter - werden beide Informationen ausgegeben. Durch das Schlüsselwort NO kann beides unterdrückt werden. Das Schlüsselwort CONTROL bewirkt, daß nur die Kommandos aufgelistet werden, beim Schlüsselwort FORMAT erscheint in der Ausgabe nur die Korrespondenztabelle.

3.3.6 NUMBERED

```
NUMBERED  <& YES ! NO &>
```

Die NUMBERED-Anweisung beschränkt das Parameterfeld der ASIA-Steueranweisungen auf die Spalten 1 bis 72; die Spalten 73 bis 80 stehen dann zur Numerierung bzw. Kennzeichnung der Anweisungen zur Verfügung. Zu Beginn des Programms ist NUMBERED NO vor eingestellt, d.h., es werden alle 80 Spalten interpretiert.

ACHTUNG: Wenn bei INPUT FORMAT das Schlüsselwort FREEFIELD benutzt wird, werden bei NUMBERED YES auch in den Datenzeilen die Spalten 73 bis 80 als Datenbereich ignoriert!

3.3.7 PAGESIZE

```
PAGESIZE  <beliebiger Text>
```

Nur aus Gründen der SPSS-Kompatibilität implementiert; das Parameterfeld wird vollständig ignoriert.

Vergleiche auch VALUE LABELS und PRINT FORMATS.

3.4 Datenselektion

ASIA kennt im Unterschied zu SPSS keine Kommandos zur Datenmodifikation und -selektion (COMPUTE, IF, COUNT, RECODE, SELECT IF usw.). Dennoch kommt es häufig vor, daß man nicht den gesamten Datensatz analysieren möchte, sondern nur eine bestimmte Subpopulation. Ein Lösungsweg dafür wäre es, die gewünschten Fälle mit einem anderen Programm (z.B. SPSS) zu extrahieren und auf einen eigenen File zu schreiben, der dann von ASIA eingelesen werden kann. Da dies allerdings recht umständlich ist und das Problem häufig auftreten wird, bietet ASIA mit den Kommandos ONLY IF und NOT IF eine von der SPSS-Notation abweichende, in den meisten Fällen aber wohl ausreichende Alternative.

3.4.1 Permanente/temporäre Selektion

In Anlehnung an SPSS unterscheidet ASIA zwischen permanenten und temporären Formen der Kommandos ONLY IF und NOT IF; die temporäre Version wird durch ein vorangestelltes Sternchen (*) kenntlich gemacht: *ONLY IF bzw. *NOT IF.

Während die Selektionen, die durch die permanente Form der Kommandos erreicht werden, für die Dauer des gesamten Laufs gelten, gelten temporäre Selektionskriterien nur für den jeweils nächsten Matrix-Aufbau (s. ASIA-Programmablauf).

Deshalb gelten folgende Einschränkungen:

- Alle permanenten Kriterien müssen vor etwaigen temporären Kriterien angefordert werden.
- Permanente Kriterien sind nur vor der ersten CLUSTER-Anweisung erlaubt.
- Temporäre Kriterien können auch vor späteren Prozeduren formuliert werden.

Da der Programmablauf von ASIA etwas anders ist als der von SPSS, kann es leicht zu Verwechslungen kommen. Wenn man sich über die Wirkung dieser Kommandos nicht ganz im klaren ist, ist es sinnvoller, mehrere ASIA-Läufe mit nur einem Kriterium zu starten, als alle gewünschten Analysen in einen Job zu packen.

Wichtig ist allerdings, daß nur permanente Kriterien eine Auswirkung auf das Kommando SAVE FILE haben!

3.4.2 ONLY IF und NOT IF

```
<*>ONLY IF  varlist(codes)
<*>NOT IF   varlist(codes)
```

Die Auswahlkommandos ONLY IF und NOT IF bieten die Möglichkeit, Selektionskriterien in einer Form zu

definieren, die der MISSING VALUES-Definition entspricht: ein Kriterium besteht aus einer Variablenliste, der eine Codeliste zugeordnet wird. Diese Codeliste "codes" kann entweder aus bis zu drei einzelnen Werten oder aus einem Wertebereich und einem einzelnen Wert bestehen (s. das Kommando MISSING VALUES).

Beispiel:

Bei dem Kommando

```
ONLY IF SEX(1)
```

prüft ASIA bei jedem Fall, der eingelesen wird, ob die Variable SEX den Wert 1 hat, und nimmt nur solche Fälle in die Berechnungen auf, in denen dies der Fall ist. Hätte das Kommando

```
NOT IF SEX(1)
```

gelautet, würden im Gegenteil alle Fälle ausgeschlossen, bei denen die Variable SEX den Wert 1 hat.

Für die Definition von Selektionskriterien gelten folgende Einschränkungen:

- Pro Matrix-Aufbau (also vor der ersten bzw. zwischen zwei CLUSTER-Anweisungen) dürfen maximal 5 (permanente oder temporäre) Kriterien definiert werden.
- In jeder der 5 möglichen Variablenlisten "varlist" dürfen maximal 10 Variablen-Namen vorkommen (in einer inklusiven Liste A TO B kommen nur 2 Namen vor!).

Insgesamt können mit diesen Kommandos viele der normalerweise nötigen Datenselektionen realisiert

werden. Einige Beispiele (mit entsprechender SPSS-Notation) sollen zunächst einmal die Anwendungsmöglichkeiten verdeutlichen:

1. Um aus einem Datensatz nur alle Männer in die Berechnung der Matrix aufzunehmen (Variable SEX; 1=männl, 2=weibl), heißen die Kommandos

a) in SPSS: SELECT IF (SEX EQ 1)
 b) in ASIA: ONLY IF SEX(1)
 oder: NOT IF SEX(2)

Dabei ist das ONLY IF Kommando tatsächlich mit dem SPSS-Kommando äquivalent, das NOT IF Kommando u.U. nicht, da außer allen Männern auch alle Fälle analysiert werden, bei denen die Variable SEX einen ungültigen oder als fehlend deklarierten Wert besitzt.

2. Wenn aus einem Datensatz alle Jugendlichen zwischen 18 und 25 Jahren (Variable AGE) ausgesucht werden sollen, müssen folgende Kommandos eingegeben werden:

a) in SPSS: SELECT IF (AGE GE 18 AND LE 25)
 b) in ASIA: ONLY IF AGE (18 THRU 25)

3. Um nur Fälle zu bearbeiten, die auf Frage 10 mit Ja ("J") geantwortet haben, auf Frage 21 aber nicht mit Ja, sondern mit Nein oder überhaupt nicht geantwortet haben, heißen die Kommandos

a) in SPSS: SELECT IF (F10 EQ 'J' AND
 F21 NE 'J')
 b) in ASIA: ONLY IF F10('J')
 NOT IF F21('J')

4. Um fallweisen Ausschluß fehlender Werte zu simulieren (in SPSS normalerweise OPTION 2), also jeden Fall auszuschließen, der in einer der Va-

riablen einen ungültigen Wert hat, sind folgende ASIA-Kommandos nötig:

MISSING VALUES ALL(BLANK)
 NOT IF ALL(BLANK)

3.4.3 Allgemeines Format der Selektionskriterien

Am Beispiel der zwei Variablen X und Y und der drei Werte a, b, c sollen noch einmal die wesentlichen Möglichkeiten verdeutlicht werden:

1. ONLY IF X,Y(a,b,c)

bedeutet, daß nur die Fälle in die Analyse aufgenommen werden, die sowohl in der Variablen X, als auch in der Variablen Y einen der drei Werte a, b oder c annehmen. In SPSS-Notation:

SELECT IF ((X EQ a OR b OR c) AND (Y EQ a OR b OR c))

Das heißt, die Bedingung muß für alle Variablen der Variablenliste erfüllt sein; es gibt keine Möglichkeit, eine Bedingung zu formulieren, nach der ein Fall ausgeschlossen wird, wenn entweder X oder Y einen bestimmten Wert annehmen.

2. ONLY IF X,Y(a THRU b,c)

bedeutet, daß nur die Fälle aufgenommen werden, bei denen sowohl die Variable X als auch die Variable Y einen Wert besitzen, der entweder im Bereich a bis b (einschließlich) liegt oder gleich c ist. In SPSS-Notation:

SELECT IF ((X GE a AND LE b OR EQ c) AND (Y GE a AND LE b OR EQ c))

Auch hier gilt, daß die in den Klammern gebildete Bedingung für alle Variablen der Liste erfüllt sein muß.

Mehrere Kriterien werden durch das logische UND verknüpft; die ASIA-Kommandos

```
ONLY IF SEX (1)
ONLY IF AGE (18 THRU 21)
```

heißen in SPSS

```
SELECT IF (SEX EQ 1 AND (AGE GE 18 AND
LE 21))
```

Fälle werden nur dann in die Berechnung der Korrelations- bzw. Kovarianzmatrix aufgenommen, wenn alle Bedingungen erfüllt sind (ONLY IF bzw. wenn keine der Bedingungen erfüllt ist (NOT IF)).

Da mehrere Kriterien durch AND verknüpft werden, wurden aus mnemotechnischen Gründen die äquivalenten Kommandos

```
AND IF (entspricht ONLY IF)
und AND NOT bzw. BUT NOT (entspricht NOT IF)
```

eingeführt, die die Formulierung von Auswahlkriterien vereinfachen und die "Lesbarkeit" von ASIA-Kommando-Folgen erhöhen.

Beispiele:

```
ONLY IF SEX(1)
AND IF FR1 TO FR10(0 THRU 9)
```

```
*ONLY IF IDENT('AA')
*BUT NOT VERSION('X','Y')
```

```
NOT IF INCOME(15 THRU HI)
AND NOT SCHOOL(1,5,9)
```

Hinweise:

Mit diesen Selektionsmechanismen sind UND-Verknüpfungen relativ leicht zu erreichen. Schwieriger sind viele ODER-Verbindungen. Diese Methode, Auswahlkriterien zu formulieren, hat vor allem folgende Nachteile:

- Es gibt keine Möglichkeit, aus einer Nominalskala mehr als drei diskrete Elemente auszusondern (beispielsweise die Werte 1, 3, 10 oder 25).
- Es gibt keine Möglichkeit, eine Bedingung zu formulieren, die erfüllt ist, wenn entweder die Variable A einen bestimmten Wert hat oder die Variable B.
- Es ist nicht möglich, zwei Variablen zu vergleichen, also etwa eine Bedingung zu formulieren, die dann erfüllt ist, wenn A und B denselben (egal welchen) Wert besitzen (SPSS: A EQ B).

3.5 ASIA-Systemdateien

Bislang wurde davon ausgegangen, daß die Daten im Steuerfile stehen und bei jeder Analyse zusammen mit den Daten-Definitionsanweisungen und den datenbeschreibenden Kommandos neu eingelesen werden. Dies hat den Vorteil, daß der Vorgang anschaulich und vor allem für den Anfänger übersichtlich bleibt. Bei großen Datensätzen und/oder bei Rechnern mit geringer Rechenleistung kann es aber mitunter recht lange dauern, bis ASIA die notwendige Rechenarbeit erledigt hat. ASIA bietet deshalb die Möglichkeit, ähnlich wie SPSS, sogenannte Systemfiles anzulegen, in denen sowohl die Daten (in ma-

schinenlesbarer Form) als auch die Datendefinitionen abgespeichert werden.

ASIA-Dateien speichern also erstens die Daten in maschinenlesbarer (binärer) Form. Dadurch entfällt u.U. ziemlich viel Rechenzeit, da die Daten nicht jedesmal neu aufgrund der Formatangaben entschlüsselt werden müssen. Zweitens enthalten ASIA-Dateien aber auch alle Informationen, die das Programm zur Verarbeitung der Daten benötigt. Dadurch reduziert sich das Steuerfile für eine Analyse oft auf zwei Anweisungen: OLD FILE, die den ASIA-File eröffnet und CLUSTER, die die gewünschte Analyse durchführt. Alle datendefinierenden und -beschreibenden Kommandos entfallen damit.

Analog dazu können auch statt der Rohdaten die Korrelations- bzw. Kovarianzmatrizen abgespeichert werden. Dann entfällt zusätzlich noch die gesamte Rechenzeit, die für den Aufbau dieser Matrizen benötigt wird und die einen erheblichen Teil der Gesamt-Rechenzeit ausmacht.

3.5.1 SAVE FILE

```
SAVE FILE name <label>
```

SAVE FILE erstellt aus den aktuellen Datendefinitionsanweisungen (VARIABLE LIST, N OF CASES, VAR LABELS, MISSING VALUES und RANGE) und den aktuellen Rohdaten einen ASIA-Systemfile, der später mit dem Kommando OLD FILE eingelesen werden kann und es dem Rechner erspart, bei jedem Lauf alle datenbeschreibenden Kommandos erneut einlesen zu müssen.

VORSICHT: SAVE FILE ist nur nach mindestens einer CLUSTER-Anweisung erlaubt und muß unmittelbar vor der FINISH-Anweisung stehen. Temporäre Selektionskriterien haben keinen Einfluß auf SAVE FILE; abgespeichert werden alle Fälle, die einem eventuellen permanenten Selektionskriterium genügen!

ACHTUNG: ASIA- und SPSS-Systemdateien sind nicht kompatibel!

3.5.2 SAVE MATRIX

```
SAVE MATRIX name <label>
```

SAVE MATRIX hat in etwa dieselbe Funktion wie die SAVE FILE-Anweisung; allerdings werden nicht die Rohdaten abgespeichert, sondern die Korrelations- und Kovarianzmatrix. Dies spart erheblich Rechenzeit, bedeutet aber, daß immer mit derselben Matrix (oder Teilmatrix von ihr) gearbeitet werden muß. Insbesondere sind Datenselektionen nicht mehr möglich, und es können keine Skalensummenwerte für die einzelnen Fälle mehr ausgeschrieben werden.

Eine ASIA-System-Matrix kann durch das Kommando OLD MATRIX wieder eingelesen werden.

ACHTUNG: SAVE MATRIX ist nur nach mindestens einer CLUSTER-Anweisung erlaubt und muß unmittelbar vor der FINISH-Anweisung liegen. Abgespeichert wird immer nur die aktuelle Matrix!

3.5.3 OLD FILE

OLD FILE name

OLD FILE entspricht der GET FILE-Anweisung bei SPSS. Da allerdings SPSS- und ASIA-Files nicht kompatibel sind, wurde, um Verwechslungen zu vermeiden, auch ein anderer Kommandoname gewählt.

OLD FILE liest System-Informationen (Variablenamen und -etiketten, MISSING VALUES- und RANGE-Angaben) aus einem früher mit SAVE FILE erstellten Systemfile ein.

ACHTUNG: ASIA- und SPSS-Systemdateien sind nicht kompatibel!

3.5.4 OLD MATRIX

OLD MATRIX name

OLD MATRIX liest analog zu OLD FILE Informationen und Daten aus einem ASIA-Systemfile ein; ein solcher File muß in einem früheren ASIA-Lauf durch das Kommando SAVE MATRIX erstellt worden sein. In einer "Matrix-Systemdatei" werden allerdings nicht die Rohdaten abgespeichert, sondern die Korrelations- und Kovarianzmatrizen.

3.6 DEFINE MATRIX

DEFINE MATRIX <VARIABLES=varlist / >

Mit der DEFINE-MATRIX-Anweisung kann der Benutzer den Aufbau der Korrelations- bzw. Kovarianzmatrix kontrollieren. Da diese Matrix das "Herz" des Algorithmus bildet, ist diese Anweisung sehr wichtig, wenn sie allerdings auch nicht unbedingt erforderlich ist und vollständig durch die Parameter der CLUSTER-Anweisung ersetzt werden kann.

3.6.1 Das Schlüsselwort VARIABLES

Die Variablenliste "varlist" nach dem Schlüsselwort VARIABLES (mögliche Abkürzungen: VAR, VARS oder VARIABLE) legt die Anzahl (und die Reihenfolge!) der Variablen fest, die später in die Berechnung der Kovarianzmatrix eingehen sollen. In nachfolgenden CLUSTER-Anweisungen können deshalb nur solche Variablen angefordert werden, die hier angegeben wurden.

Die "varlist" kann wie immer durch das Schlüsselwort ALL ersetzt werden, wenn alle definierten Variablen auch in die Berechnungen eingehen sollen. Dies ist gleichzeitig die Voreinstellung, wenn das Schlüsselwort VARIABLES nicht angegeben wird.

Wenn in der "varlist" alphanumerische Variablen angefordert werden, werden sie von ASIA ignoriert, da Mittelwerte, Standardabweichungen oder Kovarianzen von Buchstabenfolgen nicht sehr sinnvoll wären und aufgrund der maschineninternen Darstellung meist

zum Programmabbruch führen würden, weil entstehende Produkte zu groß oder zu klein würden.

ACHTUNG: ASIA kann diese Korrektur nur durchführen, wenn die Variable mit dem A-Format im INPUT FORMAT FIXED(...) eingelesen wurde. Im FREEFIELD-Format sollte der Benutzer darauf achten, daß nur numerische Variablen in die Berechnung der Matrix eingehen!

3.6.2 Status und Position von DEFINE MATRIX

Die DEFINE MATRIX-Anweisung ist wahlfrei; sie kann vollständig durch die äquivalenten Parameter der CLUSTER-Anweisung ersetzt werden (unbedingt notwendig ist sie nur, wenn eine andere als die erste CLUSTER-Anweisung keine Teilmatrix der ersten benutzen soll: Die Variablen, die in die neue Matrix eingehen sollen, müssen ASIA dann über DEFINE MATRIX mitgeteilt werden).

Beim ersten Einlesevorgang bildet ASIA eine Matrix (s. ASIA-Programmablauf). Welche Variablen in diese Matrix aufgenommen werden, richtet sich nach der letzten DEFINE MATRIX-Anweisung. Wenn bei der ersten CLUSTER-Anweisung noch keine DEFINE MATRIX-Anweisung gelesen wurde, nimmt ASIA nur die Variablen in die Matrix auf, die beim Schlüsselwort VARIABLES der CLUSTER-Anweisung angefordert werden. Nachfolgende CLUSTER-Anweisungen können - ohne eingeschaltetes DEFINE MATRIX - nur Teilmengen dieses Variablenpools verwenden!

DEFINE MATRIX muß immer vor der (den) CLUSTER-Anweisung(en) liegen, die die neue Matrix verarbeitet.

ACHTUNG: Jede DEFINE MATRIX-Anweisung macht einen neuen Einlese-Vorgang und einen neuen Matrix-Aufbau

nötig, auch wenn nur eine Teilmenge der alten Matrix angefordert wird!

3.7 Die Prozeduren

ASIA kennt nur eine "echte" Prozedur: CLUSTER; die "Prozedur" WRITE SCORES ist nur in Verbindung mit der Prozedur CLUSTER erlaubt.

3.7.1 CLUSTER

```

CLUSTER <VARIABLES=varlist / >
      <SCALE <(name label)> = varlist / >
      <INVERT <(slist)>=varlist / >
      <RKTRIT <(slist)>=value / >
      <DKTRIT <(slist)>=value / >
      <ALFAD <(slist)>=value / >
      <MXSIZE <(slist)>=value / >
      <METHOD <(slist)>=& CORRMAX ! ALFAMAX !
                          ANALYZE & / >
      <MAXCL = value / >
      <LIST/> <MATRIX/> <SUMMARY/> <SINGLE/>
      <DIAGRAM <(slist)> = varlist/> <TURN/>

```

Bis auf die VARIABLES-Angabe sind die Angaben der CLUSTER-Anweisung wahlfrei und haben voreingestellte Werte (s.u.); "varlist" ist eine Variablenliste, "value" ein Wert und "slist" eine Skalenliste.

- o VARIABLES (Abkürzungen: VAR, VARS, VARIABLE) muß die erste Angabe der CLUSTER-Anweisung sein. Sie legt den "Variablen-Pool" fest, also die Menge der Variablen, die zur Clusterung zur Verfügung steht.

Die Reihenfolge der Variablen in "vlist" (wichtig bei der TO-Konvention!) richtet sich nach der Variablen-Liste der letzten DEFINE MATRIX-Anweisung (sofern angegeben). Deshalb können in "vlist" auch nur Variablen vorkommen, die dort angegeben wurden!

Wenn kein DEFINE MATRIX vorliegt, legt die VARIABLES-Angabe automatisch eine Matrix fest, in die alte Variablen dieser Liste in der angegebenen Reihenfolge aufgenommen werden; die Reihenfolge der Variablen bei der TO-Konvention richtet sich dann nach den Angaben der VARIABLE LIST bzw. nach der Anordnung, in der die Variablen im System-File abgespeichert wurden.

ACHTUNG: Wenn bereits eine Matrix festgelegt ist (entweder durch DEFINE MATRIX oder durch eine frühere CLUSTER-Anweisung), können immer nur echte Teilmengen dieser Matrix verarbeitet werden! Die einzige Möglichkeit, eine neue Matrix aufzubauen, besteht dann in der DEFINE MATRIX-Anweisung.

VORSICHT bei der Verwendung der TO-Konvention!
Bei den drei Anweisungen

- (a) VARIABLE LIST FR1 TO FR10,OSKAR,FR11 TO FR20
- (b) DEFINE MATRIX VARS=FR1 TO FR10,FR11 TO FR20,
OSKAR
- (c) CLUSTER VARS=FR1 TO FR20

wird die Variable OSKAR nicht mit in den Variablen-Pool aufgenommen, da sich die Reihenfolge nach der VARS-Angabe in der DEFINE MATRIX-Anweisung richtet. Wenn die Anweisung (b) allerdings wegfällt, richtet sich die Reihenfolge der Variablen nach der VARIABLE LIST-Angabe: OSKAR würde

in diesem Fall in den Variablen-Pool mitaufgenommen!

- o Vorbesetzen von Skalenrümpfen: SCALE und INVERT. Wie im Kapitel Algorithmus beschrieben, können ganze Skalen oder auch Skalenrümpfe vordefiniert und dann evtl. von ASIA erweitert und analysiert werden. Um die einzelnen Skalen ansprechen zu können, erhält jede Skala einen Namen (1-8 alpha-numerische Zeichen, das erste Zeichen muß ein Buchstabe sein) und eine Etiketle ("label" 1-40 beliebige Zeichen). Voreingestellt sind die Namen "S1" bis "Sn" (n ist die maximale Skalenzahl, also 30) und leere Label. Wenn man Skalenrümpfe definiert, ist es sinnvoll, auch diese Voreinstellung abzuändern, z.B.:

SCALE (POLWIS Polit.Wissen) = POL1 TO POL5

"slist" gibt einfach die Variablen an, die in diese Skala (oder in diesen Skalenrumpf) eingehen sollen. Falls ein oder mehrere Variablen invertiert aufgenommen werden sollen, muß das dem Programm getrennt durch das Schlüsselwort INVERT mitgeteilt werden. Die "(slist)" gibt dabei die Namen der Skalen an, in denen die Variablen in "vlist" invertiert werden sollen. ACHTUNG: Wenn eine Variable gar nicht in einem Rumpf enthalten ist, kann sie natürlich auch nicht invertiert werden. ASIA ignoriert dann diese Variable einfach; um die Schreibweise zu erleichtern, wird also keine Fehlermeldung produziert!

Voreingestellt für "(slist)" ist (ALL); dies gilt im übrigen für alle Parameter, die eine "(slist)"-Angabe vorsehen.

- o Abbruchkriterien RKRIT, DKRIT, ALFAD, MXSIZE. Die im Kapitel Algorithmus beschriebenen Abbruchkriterien können für jede Skala einzeln definiert

werden; für "(slist)", die Liste der Skalen, für die das Abbruchkriterium festgelegt werden soll, ist allerdings "(ALL)" voreingestellt, so daß normalerweise ein Abbruchkriterium für alle Skalen gilt. Die einzelnen Kriterien werden durch OR-Verbindungen verknüpft, eine Skalenbildung gilt also als beendet, wenn mindestens eines der Abbruchkriterien erfüllt ist.

Voreinstellungen:

RKRIT = 0.; minimale Trennschärfe

DKRIT = 1.; maximaler Trennschärfe-Sprung

ALFAD = 0.; maximaler Alpha-Abfall

MXSIZE = m; maximale Zahl von Skalenelementen

- o Durch das Schlüsselwort METHOD kann für jede Skala der Algorithmus festgelegt werden, mit dem die Skala erweitert wird:

CORRMAX (Voreinstellung): nimmt bei jedem Schritt die Variable auf, die am höchsten mit der Linearkombination der schon vorhandenen Variablen korreliert.

ALFAMAX: nimmt bei jedem Schritt die Variable auf, die die interne Konsistenz (gemessen durch Cronbach alpha) am meisten verbessert.

ANALYZE: nimmt keine weiteren Variablen auf und berechnet nur die Item-Analyse für die bestehende Skala (natürlich nur sinnvoll im Zusammenhang mit SCALE =).

- o Das Schlüsselwort MAXCL legt die Anzahl der Skalen fest, die gebildet, erweitert und/oder analysiert werden. Wenn MAXCL nicht angegeben ist, werden nur alle durch SCALE definierten Rumpfe bearbeitet. Sind auch keine Rumpfe definiert, gilt MAXCL = m (m = maximale Zahl möglicher Skalen, also 30).

- o Durch das Schlüsselwort SINGLE wird Einfachzuordnung angefordert, d.h., jede Variable kann nur einer Skala zugeordnet werden; das hindert den Benutzer allerdings nicht, durch SCALE-Festlegungen eine "manuelle" Mehrfachzuordnung zu simulieren. Voreingestellt ist Mehrfachzuordnung: prinzipiell kann jede Variable jeder Skala zugeordnet werden (sofern sie die Bedingungen erfüllt, die durch die Abbruchkriterien definiert wurden).

- o Kontrolle des Ausgabe-Umfangs durch LIST, MATRIX und SUMMARY. Zusätzlich zur Standard-Ausgabe können durch diese Schlüsselwörter noch folgende Features angefordert werden:

- LIST: ASIA gibt eine Liste aller Variablen des Variablenpools aus; in dieser Liste steht die Anzahl gültiger Fälle, der Mittelwert und die Standardabweichung jeder Variablen.

- MATRIX: ASIA gibt die Korrelations-/Kovarianzmatrix aus; die Kovarianzen stehen unterhalb, die Korrelationen oberhalb der Diagonalen; auf der Diagonalen stehen die Varianzen der Variablen ("Auto-Kovarianzen").

- SUMMARY: die Summary-Ausgabe besteht aus 2 Teilen: erstens eine Tabelle, in der die (korrigierten) Trennschärfen jeder Variablen mit jeder Skala verzeichnet sind, und zweitens eine Skaleninterkorrelationsmatrix, in der sowohl korrigierte als auch unkorrigierte Koeffizienten aufgeführt sind.

- o Graphische Darstellung von Zusammenhängen (DIAGRAM und TURN). Oft sind die langen Tabellen und Zahlenkolonnen, die durch eine Itemanalyse (v.a. mit Mehrfachzuordnung) produziert werden, recht verwirrend

und unübersichtlich. ASIA bietet nun die Möglichkeit, auf graphischem Wege den Zusammenhang von Skalen und Variablen anschaulicher darzustellen.

In der Grundversion spannen jeweils 2 Skalen (S1 und S2) eine Art "Koordinatensystem" auf. In dieses Koordinatensystem können nun Variablen so eingetragen werden, daß ihre Position auf der "x-Achse" der Korrelation dieser Variablen zur Skala S1 und ihre Position auf der "y-Achse" der Korrelation zur Skala S2 entspricht. Die Lage einer Variablen im "Koordinatensystem" hängt also von ihren Trennschärfen zur Skala S1 (Ordinate) und zur Skala S2 (Abszisse) ab.

ASIA skaliert die "Achsen" so, daß hohe Trennschärfen (+-1.) nahe beim Ursprung und geringe Trennschärfen (+-0.) entfernt vom Ursprung liegen. Wenn also eine Variable X in diesem Diagramm in der linken, unteren Ecke eingetragen ist, bedeutet das, daß die Variable auf beiden Skalen sehr hoch korreliert; läge sie in der rechten, oberen Ecke, hieße das, daß sie mit keiner der beiden Skalen korreliert.

Als Interpretationshilfe kann man also angeben: je weiter "links" eine Variable liegt, desto enger ist ihr Zusammenhang mit Skala 1, je weiter "unten" sie liegt, desto enger ist ihr Zusammenhang mit Skala 2.

Bei der Anweisung DIAGRAM (slist) = vlist, spannen nun alle nicht redundanten Paare aus "slist" ein Koordinatensystem auf, in das alle Variablen aus "vlist" eingetragen werden.

Beispiel: DIAGRAM (SOZ, POL, KIKI) = ...
bilde 3 Diagramme, und zwar SOZ x POL, SOZ x KIKI, POL x KIKI.

Durch das Schlüsselwort TURN kann die Logik der Diagramme "umgedreht" werden: jetzt spannen jeweils 2 Variablen aus "vlist" ein Koordinatensystem auf, in das alle Skalen aus "slist" eingetragen werden (nach demselben Prinzip).

Beschränkungen:

- 1) Sowohl die Skalen-Liste "slist" als auch die Variablen-Liste "vlist" dürfen nicht mehr als 50 Namen enthalten (eine implizite Liste A TO B enthält 2 Namen!).
- 2) Alle Variablen in "vlist" müssen in der aktuellen Matrix definiert sein; es ist aber nicht erforderlich, daß sie auch im aktuellen Variablen-Pool enthalten sind. Man kann also zum Beispiel 2 theoretische "Eckvariablen" in die Variablenliste der DEFINE MATRIX-Anweisung aufnehmen, nicht aber in die Variablenliste der CLUSTER-Anweisung. Man kann dann dennoch über diese beiden Variablen ein Koordinatensystem aufspannen und alle Skalen eintragen.

Hinweis: Unabhängig von Beschränkung Nummer 1 sollten die beiden Listen "maßvoll" definiert werden. Wenn die "aufspannende" Liste zu groß wird, werden leicht Unmengen von Diagrammen produziert; zum Beispiel führen 10 Elemente zu $10 \times (10-1)/2 = 45$ Diagrammen. Wird dagegen die Liste mit den einzutragenden Elementen zu groß, werden die Diagramme leicht unübersichtlich.

Um das Diagramm auf einem Zeilendrucker ausgeben zu können, ist eine recht grobe Skalierung nötig. ASIA spannt eine 10×10 -Matrix auf (in die erste Kategorie fallen alle Korrelationen zwischen 0 und +- 0.09, in die zweite alle Korrelationen zwischen +- 0.10 und +- 0.19 etc.). In jede Zelle dieser Matrix passen maximal 5 Einträge; sollen

mehr als 5 Einträge in einer Zelle gemacht werden, wird der 5. durch "*****" ersetzt und zeigt so an, daß in dieser Zelle mehr Einträge wären, als dargestellt werden können.

Ideal wären natürlich detailliertere, genauer skalierte Darstellungen. Wegen der großen Abhängigkeit von der Hardware bzw. der dafür notwendigen unterschiedlichen Gerätetreiber wurde allerdings von der Implementation abgesehen.

3.7.2 WRITE SCORES

```
WRITE SCORES (format)
  elemente </...>/
  <FILE = & 6 | 7 &/>
  <RECID = & <*CARDNO (c1<-c2)>> !
    <*CASENO (c1<-c2)>> !
    <varnam (c1<-c2><,f)> &>
```

Mit WRITE SCORES können Variablen- und/oder Skalen-summenwerte pro Person ausgeschrieben werden.

Als erste Angabe der WRITE SCORES-Anweisung erwartet ASIA eine Format-Angabe, in der das Aussehen der formatierten Ausgabe gesteuert wird. In der Format-Angabe sind dieselben Elemente enthalten, wie sie bei INPUT FORMAT unter dem Schlüsselwort FIXED beschrieben worden sind. Um allerdings Texte ausgeben zu können, gibt es zusätzlich die beiden Formatelemente

- nHxxxxx und
- 'xxxxx'

wobei "n" in der ersten Form die Länge des Textes angibt (Beispiel: 4HNAJA); in der zweiten Form darf der auszugebende Text keinen Apostroph (') enthalten.

Nach der (eventuellen) Format-Angabe müssen die Elemente aufgeführt werden, die mit Hilfe des Formates ausgegeben werden sollen. Da sowohl Variablen- als auch Skalen-Werte ausgeschrieben werden können (und die Namen untereinander nicht eindeutig zu sein brauchen), muß die Angabe "elemente" ersetzt werden durch entweder

```
VARIABLES = varlist (Abk: VAR, VARS, VARIABLE)
oder SCALES = slist (Abk: SCALE)
```

ACHTUNG: Die Reihenfolge der Elemente ist nicht unwesentlich! Bei der formatierten Ausgabe produziert ASIA eine Element-Format-Korrespondenz-Tabelle (ähnlich wie bei INPUT FORMAT FIXED), die man genau kontrollieren sollte, ehe man mit der Interpretation der Daten beginnt!

Wenn bei einem oder bei mehreren Items einer Skala ein Wert fehlt, dann gibt ASIA das entsprechende Feld mit BLANK aus. Bei der Summenbildung ist eine Ersetzung von BLANK's nicht möglich, diese müßte ggf. vorher auf der Einzelitem-Ebene vorgenommen werden. Siehe dazu auch Abschnitt 3.2.4.

Um Konfusionen zu vermeiden, ist es meist recht praktisch, in den Ausgabe-Zeilen ein Identitätsfeld unterzubringen, das u.a. später erst ein etwaiges Mischen von Datensätzen möglich macht. Wenn die Ausgabe pro Fall nur aus einer Zeile besteht, ist das auch leicht über die "(format)"-Angabe zu machen, bei mehreren Output-Records allerdings wird das sehr aufwendig. ASIA stellt deshalb mit dem Schlüsselwort RECID eine Möglichkeit zur Verfügung, bestimmte Informationen automatisch auf jeden Out-

put-Record zu schreiben. Insgesamt können maximal 5 Informationen abgespeichert werden (also 5 ID-Variablen oder die Fallnummer und max. 4 Variablen, usw.)

Durch die Angabe `RECID = *CASENO(c1<-c2>)` wird ASIA aufgefordert, in jedem Record in die Spalten "c1" bis "c2" (bzw. nur in die Spalte "c1") die laufende Nummer des Falls zu schreiben. Die Zählung muß nicht unbedingt fortlaufend sein, weil alle ausgeschlossenen Fälle (durch temporäre Selektionskriterien oder durch MISSING VALUES) mitgezählt, aber natürlich nicht ausgeschrieben werden.

ACHTUNG: Durch permanente Selektionskriterien ausgeschlossene Fälle werden nicht gezählt!

Diese Möglichkeit ist nur anzuraten, wenn keine eigene Fallidentifikation existiert, bzw. diese nicht eingelesen wurde.

Durch `RECID - *CARDNO (c1<-c2>)` fügt ASIA automatisch in jeder Ausgabezeile die Recordnummer in die Spalte(n) "c1" (bis "c2") ein.

Am sichersten allerdings ist es, Identifikationsvariablen mitzuführen und auf jeden Output-Record auszugeben. Dies ist mit ASIA möglich durch die Angabe `RECID = varnam (c1<-C2><,f>)`. Dabei wird bei jedem Output-Record der Inhalt der Variablen "varnam" in die Spalte(n) "c1" (bis "c2") geschrieben. Die zusätzliche Angabe "f" regelt das Format; sie kann ersetzt werden durch den Buchstaben "A" (für alphanumerische Variablen) oder durch eine der Ziffern 0 - 9 und gibt dann die Zeichen rechts vom Dezimalpunkt an. Voreingestellt für "f" ist 0 für numerische und A für alphanumerische Variablen.

Beispiel: Das Identifikationsfeld jedes Records soll die ersten 15 Spalten einnehmen; in den Spal-

ten 1-4 soll die Fallnummer, in Spalte 5 die Recordnummer stehen. In die Spalten 7-10 soll die (alphanumerische) Variable IDENT und ab 11 bis 15 die Variable SUBF übernommen werden. Die RECID-Anweisung müßte dann lauten:

```
RECID = *CASENO(1-4),*CARDNO(5),IDENT(7-10,A),
        SUBF(11-15)
```

ACHTUNG: Im "(format)" müssen die ersten 15 Spalten jedes Records (mit dem X-Element) übersprungen werden!

Durch die Angabe `FILE = 6` werden die Daten nicht auf einen eigenen File geschrieben, sondern in den Joboutput übernommen. `FILE = 7` (Voreinstellung) leitet den Output in ein eigens dafür erzeugtes ASCII-File mit Namen ASIA07.DAT.

Beispiel: Pro Fall sollen 2 Records ausgegeben werden; auf dem ersten Record sollen die 10 Variablen V1 bis V10 stehen, auf dem zweiten die 3 Skalen S1 bis S3 und die Variablen V20 bis V26. Die ersten 15 Spalten sollen das oben beschriebene Identifikationsfeld einnehmen; um das Format einfach zu halten, werden alle Werte 3-stellig angegeben:

```
WRITE SCORES (15X,10F3.0/15X,10F3.0)
        VARS=V1 TO V10/SCALES=S1,S2,S3 /
        VARS=V20 TO V26/
        RECID=*CASENO(1-4),*CARDNO(5),
        IDENT(7-10,A),SUBF(11-15)
```

3.8 Kommandos zur Ablauf-Kontrolle

Die Reihenfolge der ASIA-Kommandos ist weitgehend festgelegt; zuerst müssen dem Programm die Datendefinitionen mitgeteilt werden, dann kann die Matrix-Form bestimmt, und schließlich können Item- oder

Variablen-Clusterungen angefordert werden. Arbeitet man mit Rohdaten, müssen diese jetzt zwischen den Anweisungen READ INPUT DATA und END INPUT DATA stehen. Danach ist es möglich, Skalensummenwerte errechnen und ausgeben zu lassen und ASIA-System-Files abzuspeichern. ASIA wird beendet, wenn entweder keine Kommandos mehr eingegeben werden oder wenn es das Kommando FINISH liest.

3.8.1 READ INPUT DATA

READ INPUT DATA

ASIA liest die Rohdaten bei der Verarbeitung der ersten CLUSTER-Anweisung. Deshalb muß READ INPUT DATA unmittelbar auf diese Anweisung folgen; unmittelbar auf die READ INPUT DATA-Anweisung muß der erste Datenrecord folgen. Auch COMMENT ist an diesen Stellen nicht erlaubt!

ACHTUNG: READ INPUT DATA darf nicht fortgesetzt werden; programmintern wird allerdings nur das Schlüsselwort READ überprüft, der Rest der Eingabezeile wird ignoriert!

3.8.2 END INPUT DATA

END INPUT DATA

Die END INPUT DATA - Anweisung wird notwendig, wenn die Anzahl der Fälle nicht bekannt ist (N OF CASES UNKNOWN). Sie muß direkt hinter dem letzten Datenrecord liegen. END INPUT DATA ist die einzige ASIA-Anweisung, die in Großbuchstaben geschrieben sein muß.

END INPUT DATA ist erlaubt, aber nicht notwendig, wenn die Anzahl der Fälle bekannt ist und dem Programm über die N OF CASES-Anweisung mitgeteilt wurde.

3.8.3 FINISH

FINISH

Die FINISH - Anweisung beendet den ASIA-Lauf und gibt die Kontrolle an das Betriebssystem zurück. Kommandos, die hinter FINISH stehen, werden nicht mehr verarbeitet.

ASIA erkennt das Ende des Steuerfiles auch ohne die FINISH-Anweisung; deshalb ist diese Anweisung wahlfrei.

4 Überblick über die Kommando-Syntax

4.1 Beschreibung der Eingabedaten

VARIABLE LIST	varlist
INPUT FORMAT	FREEFIELD
INPUT FORMAT	<FIXED>
N OF CASES	<& n ! <u>UNKNOWN</u> &>

4.2 Erweiterte Beschreibung der Daten

FILE NAME	<name <label> >
VAR LABELS	varnam,varlabel</varnam, varlabel...>
MISSING VALUES	varlist(codes)</varlist (codes)...> ¹
RANGE	varlist(min,max)</varlist (min,max) ...>
VALUE LABELS	beliebiger Text
PRINT FORMATS	beliebiger Text

4.3 Zur Form der Ein- und Ausgabe

EDIT	<beliebiger Text>
RUN NAME	beliebiger Text ²
TASK NAME	beliebiger Text ³

¹ Die "codes" -Angabe kann aus maximal 3 Werten oder einem Wertebereich und max. 1 Wert bestehen

² keine Fortsetzungskarte erlaubt

³ keine Fortsetzungskarte erlaubt

COMMENT	beliebiger Text
PAGESIZE	beliebiger Text
PRINT BACK	<& NO ! CONTROL ! FORMAT &>
NUMBERED	<& <u>YES</u> ! NO &>
READ INPUT DATA	
END INPUT DATA	
FINISH	

4.4 Daten-Selektion

<*>ONLY IF	varlist(codes)
<*>NOT IF	varlist(codes)
<*>AND IF	varlist(codes)
<*>AND NOT	varlist(codes)
<*>BUT NOT	varlist(codes)

Die "codes"-Angabe entspricht der bei MISSING VALUES

4.5 System-Dateien und Output

SAVE FILE	<name <label> >
SAVE MATRIX	<name <label> >
OLD FILE	<name>
OLD MATRIX	<name>

```

WRITE SCORES      (format)
                   elemente </...>/
                   <MISCODE = & BLANK! * !value & />
                   <FILE      = & 6 ! 7 & />
                   <RECID     = & <*CARDNO(c1<-c2>)> !
                               <*CASENO(c1<-c2>) !
                               <varnam(c1<-c2><,f>)&>>

DEFINE MATRIX     <VARIABLES=varlist / >

CLUSTER          <VARIABLES=varlist / >
                   <SCALE (name label)=varlist />
                   <INVERT(slist)=varlist />
                   <RKCRIT(slist)=value />
                   <DKCRIT(slist)=value />
                   <ALFAD(slist)=value />
                   <MXSIZE(slist)=value />
                   <METHOD(slist)=&CORRMAX ! ALFAMAX
                               ! ANALYZE&/>
                   <MAXCL=value />
                   <LIST /> <MATRIX /> <SUMMARY />
                   <SINGLE /><DIAGRAM(slist)=
                   varlist /> <TURN>

```

5 Literatur

- Bührer, M.:** ASIA-Benutzerhandbuch. SOWI Berichte Nr. 40, 1986.
- Bührer, Michael and Kohr, H.-U.:** ICA Item-Cluster Analysis. A Rationale, an Algorithm, and Computer Program to Improve and Facilitate Construction and Analysis of Scales. In Proceedings of the 26th Conference of the MILITARY TESTING ASSOCIATION, Vol. I, 1985, S. 355-361.
- Cronbach, L.:** Essentials of Psychological Testing. New York 1961.
- Gulliksen, H.:** Theory of Mental Tests. New York 1962.
- Kohr, H.-U.:** Die Item-Kovarianzmatrix als Basis der Testanalyse, Psychologische Beiträge, 20, 1978.
- Magnusson, D.:** Testtheorie. Wien 1975.
- Schmolck, P.:** Methoden der Reliabilitätsschätzung. Universität der Bundeswehr München, Fb Päd, unveröffentl. Manuskript. München 1984.

6 Stichwortverzeichnis

	Seite
?NOMATCH	36
Ablauf-Kontrolle	71
ALFAD	10, 63
ALFAMAX	64
ALL	40
ANALYZE	64
AND IF	54
AND NOT	54
Beschreibung der Eingabedaten	30, 37
BLANK	41
BUT NOT	54
CLUSTER	18, 61
COMMENT	47
CORRMAX	64
Datenselektion	49
DEFINE MATRIX	18, 59
DIAGRAM	65
DKRIT	10, 63
EDIT	45
Einfachzuordnung	7
END INPUT DATA	73
FILE	71
FILE NAME	38
FINISH	20, 73
FIXED	33
Form der Ein- und Ausgabe	44
Format der Steueranweisungen	24
Formatelemente	33
FREEFIELD	32
HIGHEST	41
INPUT FORMAT	17, 31
INVERT	9, 63
Invertierung	8
Itemanalyse	7
Kernvariablen	7
Kommando-Syntax	74

LIST	12, 65
LOWEST	41
MATRIX	12, 65
MAXCL	12, 64
Mehrfachzuordnung	7
METHOD	11, 64
MISSING VALUES	17, 40
Missing-Data-Verarbeitung	16
MXSIZE	11, 63
N OF CASES	17, 37
NOT IF	18, 50
Notation	28
NUMBERED	48
OLD FILE	17, 58
OLD MATRIX	20, 58
ONLY IF	18, 50
PAGESIZE	49
Permanente/temporäre Selektion	49
PRINT BACK	47
PRINT FORMATS	44
Prozeduren	61
RANGE	17, 43
READ INPUT DATA	72
RECID	70
Reihenfolge der Steueranweisungen	22
RKRIT	10, 63
RUN NAME	46
SAVE FILE	17, 19, 56
SAVE MATRIX	20, 57
SCALE	9, 63
SINGLE	10, 65
SUMMARY	12, 65
Systemdateien	55
Systemfile	19, 55
TASK NAME	46
TURN	65, 67
UNKNOWN	37

VALUE LABELS	40
VAR LABELS	17, 39
VARIABLE LIST	17, 30
VARIABLES	9, 59, 61
WRITE SCORES	19, 68
Zentrale Formeln	12

III. ITAMIS

(Itemanalyse-System mit Missing-Data Behandlung)

1 Vorbemerkung

Bei ITAMIS handelt es sich um ein in FORTRAN IV geschriebenes, 1974 bis 1976 entwickeltes Programm, das hinsichtlich seiner Benutzeroberfläche in keiner Weise heute üblichen Anforderungen genügt. Die entsprechende Modifikation des Programms wäre sehr aufwendig gewesen und hätte sich u. E. nicht gelohnt, da sich die meisten Standardprobleme auch (und meist einfacher bzw. besser) mit ASIA lösen lassen. Wir haben deshalb von umfangreichen Änderungen gänzlich abgesehen. Dies hat zur Folge, daß ITAMIS die Benutzer, die die Zeiten der Eingabe von Benutzerprogrammen über Lochkarten nur noch vom Hörensagen kennen, merkwürdig und antiquiert anmutet. So ist in der Eingabebeschriftung von "Karten"¹ bzw. "Steuerkarten" die Rede - heute würde man von Records sprechen. Die Steuerparameter für ITAMIS sind also über "Karten" einzugeben, und ITAMIS erwartet vom Benutzer, daß die programmsteuernden Informationen rechtsbündig genau in den "Spalten" (Recordpositionen) stehen, die dafür laut Eingabebeschriftung vorgesehen sind.

Zur Pflege von Bequemlichkeit und Nostalgie haben wir die Terminologie der "Lochkartenzeiten" beibehalten und hoffen, daß menügewohnte Benutzer dadurch nicht verschreckt werden. In der Konsequenz dieser Entscheidung finden sich daher in der Ein-

¹ Zur Erinnerung an die "alten Zeiten": eine Lochkarte hatte 80 "Spalten", entspricht also einem Record mit 80 Characters/ASCII-Zeichen

gabebeschreibung auch veraltete Begriffe wie "Parameterkarte", "Textkarte", "Formatkarte", "Zuordnungskarte", "Schablone" etc. ITAMIS benutzt diese Records zur Programmablaufsteuerung und zur Zuweisung von Parametern. Allgemein wird hierfür der Begriff "Steuerkarten" in der folgenden Programmbeschreibung verwendet. Die Beschreibung der Programmeingabe folgt genau der in der letzten ausführlichen Darstellung von ITAMIS² wiedergegebenen; einige Parameter sind allerdings für die PC-Version überflüssig geworden. Der Einfachheit halber werden diese in der Eingabebeschreibung nicht mehr erwähnt (daraus erklärt sich, daß in einigen Parameterfeldern numerische Vorgaben gefordert werden, die etwas eigenwillig anmuten mögen).

Abschließend sei darauf hingewiesen, daß in der folgenden Kurzbeschreibung von ITAMIS nur die Informationen wiedergegeben sind, die man zur sinnvollen Anwendung des Programms dringend benötigt. Weitere finden sich für den interessierten Benutzer bei Kohr 1976 bzw. 1978.

2 Leistungen des Programmsystems

2.1 Allgemeine Hinweise

Die folgende Darstellung veranschaulicht die wesentlichen Funktionen von ITAMIS. Diese werden im Abschnitt 2.2. besprochen. In der Eingabebeschreibung wird - sofern dies zum Verständnis erforderlich - auf diese Funktionen verwiesen.

Die Abfolge, in der die Funktionen beschrieben werden, entspricht nur z.T. dem tatsächlichen Programmablauf. Über diesen informiert die ausführliche Programmbeschreibung (Kohr 1976/1978). Auf Literatur zu Formeln, die zur Bestimmung von Koeffizienten verwendet wurden, wird nur verwiesen, wenn es sich um weniger gebräuchliche Koeffizienten handelt. Korrelationskoeffizienten werden grundsätzlich anhand der Formel für den Produkt-Moment-Koeffizienten nach PEARSON/BRAVAIS bestimmt. Gleiches gilt auch für die Berechnung von Trennschärfekoeffizienten³.

Für das Programmsystem gelten folgende Beschränkungen:

Zahl der Items	: maximal 500;
Zahl der Skalen	: maximal 30;
Zahl der Items pro Skala	: maximal 40;
Zahl der Probanden	: maximal 9999;
Zahl der Einleseformatkarten:	maximal 9;

- jedes Item kann nur einer Skala zugeordnet werden;

³ Literatur: DIEHL/KOHR: Deskriptive Statistik, Kap. 16 "Itemanalyse nach dem Konzept der klassischen Testtheorie", Frankfurt: Verlag Klotz, 1977, S. 376-419)

- das Programm kann folgende Codes verarbeiten: BLANK, 0 ... 9. Demzufolge können Items mit maximal zehnstufigen Reaktionsschemata verarbeitet werden; nichtnumerische Lochungen oder Lochungen mit einem Zahlenwert größer 9 werden vom System als MISSING-DATA behandelt. Diesen wird programmintern der Zahlenwert 10 zugeordnet (negative Zahlen werden wie zweistellige Zahlen behandelt). Das System diagnostiziert probandenweise die Zahl der MISSING-DATA: die Diagnose kann zwar unterdrückt werden, die MISSING-DATA Verarbeitung bleibt davon jedoch unberührt;
- wenn die Daten von einem File gelesen werden sollen, muß diesem die logische Nummer 1 zugeordnet werden;
- Skalenwerte können auf ein File ausgegeben werden. Das System benutzt für das Ausgabe-File die logische Nummer 2.

2.2 Die wesentlichen Funktionen des Programmsystems

2.2.1 Datenmanagement

	K ⁴
F 1. Selektion von Datenvektoren nach maximal drei Kriterien vor Analysebeginn.	
F 2. Zuordnung von Items zu wenigstens einer und höchstens dreißig Skalen. Jedes Item kann nur einer Skala zugeordnet werden.	SK

⁴ In dieser Spalte sind die im Programmausdruck verwendeten, dort nicht näher erläuterten Kurzbezeichnungen verzeichnet. Diese Kurzbezeichnungen sind in der Kopfzeile der Ergebnistabellen enthalten, die von ITAMIS ausgegeben werden.

F 3. Erleichterung der Dateneingabe. Zum Einlesen der Daten können einfache Formate verwendet werden, da Items, die nicht zugeordnet wurden, aus der Analyse ausgeschlossen werden.	
F 4. Invertierung von Beantwortungen. Für jedes Item kann eine unterschiedliche minimale und maximale Skalierung angegeben werden.	I U O
F 5. Dichotomisierung von Beantwortungen. Für jedes Item kann ein unterschiedliches Dichotomisierungskriterium angegeben werden.	D
F 6. Dichotomisierung und Invertierung. Die Daten können zunächst dichotomisiert und anschließend invertiert werden.	
F 7. Verarbeitung von Mehrfachwahlantworten. Für jedes Item wird die richtige Lösung angegeben. Auf diese Weise werden die Daten nach 1 (richtige Lösung) und 0 (falsche Lösung) transformiert.	R
F 8. Bestimmung von Skalenwerten durch Summierung. Es können maximal dreißig Skalenwerte für jede Person bestimmt werden (vgl. F. 2.). Außerdem kann ein Gesamtpunktwert (Summe aller Beantwortungen einer Person) berechnet werden	
F 9. Ausgabe der Skalenwerte. Die Skalenwerte können - mit oder ohne Gesamtpunktwert - auf folgende Weise ausgegeben werden: 1) Drucken 2) Ausgabe auf File mit der logischen Nummer 2 3) Ausgabe auf File und Drucken	

<p>F 10. Möglichkeiten zur Identifikation ausgegebener Skalenwerte.</p> <ol style="list-style-type: none"> 1) Ein Identifikationsfeld (LABEL; max 80 beliebige ASCII-Zeichen) kann von den Datenkarten übernommen werden 2) Ein zusätzliches Identifikationsfeld (ZUSATZ-LABEL; max. 20 beliebige ASCII-Zeichen) kann vorgesehen werden 	
<p>F 11. Definition fehlender Daten.</p> <p>Für jedes Item kann definiert werden, welche der unten genannten Lochkombinationen durch das Programmsystem als fehlende Daten (MISSING DATA; MD) behandelt werden sollen. Diese Definition kann für jedes Item unterschiedlich sein.</p> <ol style="list-style-type: none"> 1) Es liegen keine MD vor. In diesem Fall werden BLANK und 0 gleich behandelt und gehen mit dem Wert 0 in die Berechnung ein 2) Nur BLANKS gelten als MD 3) Sowohl BLANKS als auch eine der Lochungen 0, bzw. 1, ..., bzw. 9 gelten als MD 4) BLANKS und Nullen sowie eine der Lochungen 1, bzw. 2, ..., bzw. 9 gelten als MD <p>Als MD definierte Werte werden vom Datenmanagement ausgeschlossen</p>	MDC

2.2.2 Itemanalyse

	K
<p>F 12. Itemmittelwerte.</p> <p>Für jedes Item werden die Mittelwerte bestimmt.</p> <ol style="list-style-type: none"> 1. Summe aller von MD verschiedenen Beantwortungen bei diesem Item geteilt durch deren Anzahl (N1) 2. Summe der Beantwortungen derjenigen Probanden, die bei allen Items der Skala, welcher das Item zugeordnet wurde, eine von MD verschiedene Antwort gegeben haben, geteilt durch deren Anzahl (N2) 3. Summe der Beantwortungen derjenigen Probanden, welche bei allen Items eine von MD verschiedene Antwort gegeben haben, geteilt durch deren Anzahl (N3). 	<p>MITT1</p> <p>MITT2</p> <p>MITT3</p>
<p>F 13. Standardabweichungen der Items.</p> <p>Für jedes Item werden Standardabweichungen bestimmt (Quadratwurzel aus: Summe der quadrierten Abweichungen vom jeweiligen Itemmittelwert, geteilt durch die entsprechende Probandenzahl - 1). Die jeweilige Berechnungsstrategie ist unter F 12. beschrieben.</p>	<p>STAND1</p> <p>STAND2</p> <p>STAND3</p>
<p>F 14. Skalenmittelwerte und Gesamtmittelwert.</p> <p>In deren Berechnung gehen nur die von Blank/Missing verschiedenen Skalenwerte ein.</p>	
<p>F 15. Standardabweichungen der Skalen und des Gesamtpunktwertes.</p> <p>Quadratwurzel aus: Summe der quadrierten Abweichungen vom Skalenmittelwert, geteilt durch die Probandenzahl - 1. Es gehen nur die von Blank/Missing verschiedenen Skalenwerte ein.</p>	

<p>F 16. Häufigkeitsverteilung der Antwortalternativen.</p> <p>Für jedes Item wird die Besetzung jeder Antwortalternative (vor und nach Datenmanagement; F4 - F7) wiedergegeben.</p> <p>Für jedes Item wird in der Spalte NMDI die Anzahl der MD angegeben. Als MD gelten:</p> <ul style="list-style-type: none"> - BLANKS und Überschreitungen der oberen Grenze des Skalierungsbereichs - ggf. zusätzlich die als MD definierten Lochungen (vgl. F 11.3 und F 11.4) 	NMDI
<p>F 17. Trennschärfekoeffizienten.</p> <p>Für jedes Item werden drei Koeffizienten bestimmt (Produkt-Moment-MD-Korrelationskoeffizienten):</p> <ol style="list-style-type: none"> 1. Trennschärfe des Items in bezug zur Skala, der das Item zugeordnet wurde. In die Berechnung gehen folgende Größen ein: N_2 (F12.)⁵, MITT2 (F12.), STAND2 (F13.), Skalenmittelwert (F14.), Standardabweichung der Skala (F15.) 2. Part-whole korrigierte Trennschärfe des Items in bezug zur Skala, der das Item zugeordnet wurde. Die unter 1. genannten Größen gehen in die Berechnung ein 3. Part-whole korrigierte Trennschärfe des Items in bezug zum Gesamtpunktwert. Die folgenden Größen gehen in die Berechnung ein: N_3 (F12.), MITT3 (F12.), STAND3 (F13.), Gesamtmittelwert (F14.), Standardabweichung des Gesamtpunktwertes (F 15.) 	R-SKA RCSKA RCGES

⁵ bestimmt nach dem unter F 12 beschriebenen Verfahren

<p>F 18. Koeffizienten zur Schätzung der Reliabilität (Interne Konsistenz)</p> <ol style="list-style-type: none"> 1. CRONBACH ALPHA bzw. KUDER-RICHARDSON, Formel 20. In die Berechnung gehen folgende Größen ein: Anzahl der Items, Standardabweichung der Items (STAND2, F13.), Standardabweichung der Skala (F 15.). <p>Die folgenden Koeffizienten werden nach odd-even split-half Methode bestimmt.</p> <ol style="list-style-type: none"> 2. Produkt-Moment-Korrelation der Testhälften 3. SPEARMAN-BROWN 4. FLANAGAN KONSISTENZ- 5. KRISTOF SCHÄTZUNGEN <p>In die Berechnung werden nur die splits der Skalenwerte einbezogen, die frei von MD sind</p>	
<p>F 19. Durchschnittliche Trennschärfekoeffizienten.</p> <p>Für jede Skala werden die durchschnittlichen Trennschärfekoeffizienten angegeben. Zu deren Berechnung werden die Koeffizienten nach FISHER's z transformiert, gemittelt und rücktransformiert.</p> <ol style="list-style-type: none"> 1. Durchschnittliche Trennschärfe der Items in bezug zur Skala, der das Item zugeordnet wurde. 2. Durchschnittliche, part-whole korrigierte Trennschärfe des Items in bezug zur Skala, der das Item zugeordnet wurde. Berechnungsgrundlage: RCSKA (F17.) 3. Durchschnittliche, part-whole korrigierte Trennschärfe des Items in bezug zum Gesamtpunktwert. Berechnungsgrundlage: RCGES (F17.) 	

<p>F 20. Interkorrelation der Skalenwerte und durchschnittliche Interkorrelation</p> <p>Auf Wunsch kann die Interkorrelation der Skalenwerte und deren Korrelation zum Gesamtpunktwert bestimmt werden. Dabei handelt es sich um MD-Korrelationen, die anhand der unter F8. beschriebenen Werte bestimmt werden.</p> <p>Zur Bestimmung der durchschnittlichen Interkorrelation werden die Koeffizienten über die Hauptdiagonalen der R-Matrix nach FISHER transformiert, summiert und gemittelt. Der Durchschnittswert wird dann in einen Korrelationskoeffizienten zurücktransformiert.</p>	
<p>F 21. (Part-whole korrigierte) Trennschärfe der Items in bezug auf alle Skalen (Item-Skalen-Korrelation).</p> <p>Für jedes Item kann die (korrigierte) Trennschärfe in bezug auf alle Skalen bestimmt werden. Diese Koeffizienten sind insbesondere bei der Konstruktion von Tests bzw. Fragebögen mit mehreren Skalen zur Beurteilung der a priori getroffenen Zuordnung von Items zu Skalen und zur eventuellen Modifikation der Zuordnung nützlich (vgl. KOHR, 1972: Beschreibung der Zuordnungsalgorithmen). Es handelt sich hierbei um MD-Korrelationskoeffizienten.</p>	

<p>F 22. Interkorrelation der Items jeder Skala und durchschnittliche Interkorrelation</p> <p>Auf Wunsch kann die Matrix der Iteminterkorrelationen für jede Skala bestimmt werden. Es handelt sich hierbei um MD-Korrelationen. Zur Bestimmung der durchschnittlichen Interkorrelation werden die Koeffizienten über der Hauptdiagonalen der R-Matrix nach FISHER transformiert, summiert und gemittelt. Der Durchschnittswert wird dann in einen Korrelationskoeffizienten zurücktransformiert (vgl. F28.).</p>	
<p>F 23. Bestimmungen der primären Häufigkeitsverteilung für die Skalenwerte</p> <p>Für jede Skala werden alle auftretenden Skalenwerte, deren Häufigkeit (absolut, prozentual und kumuliert) sowie die den Skalenwerten entsprechenden z-Werte und Prozentränge angegeben.</p>	

2.2.3 Hilfsfunktionen

	K
<p>F 24. Ausdrucksteuerung.</p> <p>Der Zeilenvorschub kann "von außen" gesteuert werden, z.B. um Tabellen auf jeweils eine neue Seite drucken zu lassen oder um dies zu verhindern.</p>	

<p>F 25. Steuerung der Ausgabe von Skalenwerten.</p> <p>Damit der Benutzer die für sein Problem sinnvollste Art der Datenausgabe (vgl. F 9.) wählen kann, sind die Formate für die Ausgabesteuerung selbst zu schreiben.</p>	
<p>F 26. Skalenkennzeichnungskarten.</p> <p>Die Spalten 1 - 40 dieser Karten (pro Skala eine Karte) werden auf dem Ausdruck zur Beschriftung der Skalen verwendet.</p>	
<p>F 27. Itemkennzeichnungskarten.</p> <p>Mithilfe dieser Karten (pro Item eine Karte) können die Tabellen der Ergebnisse der Itemanalyse beschriftet werden. Im Ausdruck erscheinen jeweils die ersten dreißig Zeichen.</p>	
<p>F 28. Steuerung der Berechnung von Iteminterkorrelationen bzw. (korrigierten) Trennschärfen der Items einer Skala in bezug auf alle Skalen.</p> <p>Die Ausführung beider Funktionen (vgl. F21. und F22.) ist relativ aufwendig (hohe Rechenzeit wegen der MD-Korrelationen; hoher I/O-Bedarf, da der gesamte Datensatz pro Skala, für die eine Funktion angefordert wird, erneut von einem Hilfsfile eingelesen werden muß).</p> <p>Mit Hilfe der Korrelationsschablone kann deshalb die Anforderung dieser beiden Funktionen gesteuert werden. Wegen des hohen Zeitbedarfs sollten die Funktionen nur angefordert werden, wenn sie tatsächlich benötigt werden.</p>	

<p>F 29. Datenmanagement-Summary</p> <p>Vor Beginn der Analyse wird für jedes Item folgendes angegeben:</p> <ol style="list-style-type: none"> 1) die Itemnummer 2) die Nummer der Skala, der das Item zugeordnet wurde 3) ggf. das Dichotomisierungskriterium 4) "JA", falls Invertierung angefordert wurde, sonst "NEIN" 5) ggf. der Code, der als richtige Lösung gewertet werden soll 6) die für das Item vorgenommene MISSING-DATA Definition 	
<p>F 30. MISSING-DATA Diagnose</p> <p>Wenn die Angabe der Zahl der MD pro Proband nicht erforderlich ist, kann diese durch Lochung der Ziffer 1 in Spalte 78 der Parameterkarte unterdrückt werden. Bei großen Stichproben läßt sich hierdurch der Umfang des Ausdrucks u.U. erheblich reduzieren.</p>	
<p>F 31. Ersetzung fehlender Daten</p> <p>Auf Wunsch der ITAMIS-Benutzer wurden die Möglichkeiten zur Verarbeitung von BLANKS/MD erweitert: wenn dies für sinnvoll erachtet wird, können BLANKS/MD nunmehr vom System ersetzt werden. Dazu ist in der Parameterkarte (Spalten 79-80) anzugeben, wieviel Prozent BLANKS/MD pro Skala maximal tolerabel sind (dieser Prozentsatz gilt also für alle Skalen). Wird ein Prozentsatz größer als 50 angegeben, so bricht ITAMIS die Bearbeitung ab. Sind die Spalten 79 u. 80 BLANK oder 0, so arbeitet das Programmsystem ohne Ersetzung von BLANKS/MD.</p>	

Beschreibung des Ersetzungsalgorithmus:

Auf der Parameterkarte wird angegeben, bei welchem Prozentanteil von BLANKS/MD eine Ersetzung für noch vertretbar gehalten wird. Aus diesem Prozentsatz errechnet ITAMIS die Absolutzahl der pro Skala tolerablen (d.h. ersetzbaren) BLANKS/MD. Ist die Anzahl der bei einem Probanden in einer Skala registrierten BLANKS/MD kleiner oder gleich der errechneten Anzahl tolerabler BLANKS/MD, so werden diese durch die gerundete durchschnittliche Reaktion des Probanden auf alle Items dieser Skala ersetzt. Die Ersetzung erfolgt nach dem Datenmanagement. Liegen mehr BLANKS/MD vor, als tolerabel ist, so werden die betreffenden Skalensummenwerte mit BLANK besetzt und aus den Berechnungen ausgeschlossen. Sind Summenwerte zu drucken oder auszugeben, so werden diese ebenfalls mit BLANK besetzt; die problemlose Weiterverarbeitung mit den üblichen Statistikprogrammen ist damit gewährleistet.

2.2.4 Sonstiges

- Fremd-Trennschärfen (vgl. DIEHL/KOHR, 1977, S. 401-404) werden ohne Korrektur bestimmt;
- unter der Matrix der Fremd-Trennschärfen und der Matrix der Skaleninterkorrelationen werden ggf. die Skalennamen gelistet;
- bei den Item-Antwortverteilungen werden ausgegeben:
 - + die ersten 48 Zeichen des Itemtextes;
 - + unter *NVP* die Zahl der Probanden mit Itemwerten, die von BLANK/MD verschieden sind;

- + die Häufigkeit der Antwortalternativen vor dem Datenmanagement (Zeile 1) und nach dem Datenmanagement (Zeile 2) sowie entsprechende Prozentangaben (Prozentuierungsbasis: NVP).

2.2.5 Anwendung des Systems auf "pathologische Daten":

Die Erfahrungen mit der Benutzung von ITAMIS haben gezeigt, daß das System im Regelfall Eingabefehler sinnvoll diagnostiziert und auch mit relativ pathologischen Datensätzen (fehlende Itemvarianzen, etc.) "fertig" wird. Eine Einschränkung, die durch die komplexe MISSING-DATA Behandlung bedingt ist, muß jedoch erwähnt werden: Wenn sehr viele und bezüglich der erfaßten Merkmalsbereiche heterogene Skalen in eine Analyse einbezogen werden und wenn darüber hinaus die Probandenzahl relativ gering ist, kann der Fall eintreten, daß sich für keinen einzigen Probanden ein Gesamtpunktwert bestimmen läßt, weil bei jedem Probanden in der einen oder anderen Skala Daten fehlen. In diesem, freilich seltenen, Extremfall läßt sich bei Anforderung einer Itemanalyse u.U. ein Programmabbruch von ITAMIS selbst nicht mehr abfangen, da keine Möglichkeit zur Berechnung von Item-Gesamt-Korrelationen besteht (die Bildung von Skalensummenwerten ist hiervon nicht betroffen; Skalensummenwerte können auch in diesem Fall berechnet und ausgegeben werden). Vorschlag zur Abhilfe: Man eliminiert die Skalen, in denen die Probanden extrem viele MISSING-DATA aufweisen. Um welche Skalen es sich handelt, läßt sich leicht an den ausgedruckten Skalensummenwerten erkennen, die in diesem Fall BLANK sind.

2.3 Fehlerdiagnostik

ITAMIS untersucht die Eingaben auf Fehler und Unstimmigkeiten. Wenn Fehler auftreten, wird auf die Ursache des Fehlers hingewiesen und der Lauf abgebrochen. Folgende Eingabeelemente werden geprüft:

1. TEXTKARTE: Enthält diese Karte in den Spalten 1 bis 6 eine von TEXTKA verschiedene Zeichenkombination, so bricht das Programm mit einem Hinweis ab und sucht nach einer weiteren Textkarte (prüft also, ob ein weiteres Problem zu verarbeiten ist).
2. PARAMETERKARTE: Enthält diese Karte in den Spalten 1-6 eine von PROBLM verschiedene Zahlenkombination, so erfolgt ein Abbruch mit Diagnostik.
3. Werden unerlaubte Parameter oder - im Rahmen des Absehbaren - unzulässige Parameterkombinationen verwendet, so erfolgt ein Abbruch mit entsprechendem Hinweis. Übersteigt die Item- und/oder Skalenzahl den zulässigen Bereich, so wird - falls dies möglich ist - nur die davon betroffene Funktion ausgelassen und weiter gerechnet.
4. FORMATKARTEN: ITAMIS nimmt keine Prüfung der Formatkarten vor; entsprechende Fehlerdiagnosen und Hinweise werden durch das Betriebssystem des Rechners ausgegeben.
5. SCHABLONEN: Für die einzelnen Funktionen des DATENMANAGEMENT (vgl. 2.2.1) sind Schablonen erforderlich. Die formale Prüfung der Schablonen erfolgt durch ITAMIS auf folgende Weise:
 - a. Zuerst wird kontrolliert, ob die SCHABLONEN-ANKÜNDIGUNGSKARTE vorhanden ist. Fehlt diese oder enthält sie nicht die erforderlichen

Zeichen, so bricht ITAMIS mit einer Diagnose ab.

- b. Dann wird geprüft, ob die ZUORDNUNGSKARTEN für diese Schablonen in den Spalten 1-6 ZUWEIS enthalten. Ist dies nicht der Fall, so wird mit einer Diagnostik abgebrochen.
 - c. Es wird kontrolliert, ob der Wert, der zugewiesen werden soll, den zulässigen Bereich überschreitet (ggf. Hinweis und Abbruch).
 - d. Wenn irrtümlich Mehrfachzuordnung angefordert wird (z.B. soll ein Item zugleich zu Skala 1 und 3 gehören), so wird darauf hingewiesen und abgebrochen.
 - e. Die Nummer der Items, die keiner Skala zugewiesen wurden (vgl. F 3.), werden angegeben.
6. Bei der Berechnung der Itemkennwerte wird geprüft, ob Standardabweichungen von 0.0 vorkommen. In diesem Fall werden die Trennschärfekoeffizienten 0.0 gesetzt. Nach Hinweis im Ausdruck wird weiter gerechnet.
 7. Bei der Bestimmung von Korrelationsmatrizen wird geprüft, ob die Zahl der Kreuzprodukte zur Berechnung eines Koeffizienten ausreichend ist. Ist das nicht der Fall, so wird die Korrelationsberechnung mit einem Hinweis beendet.
 8. Wenn SKALEN- und/oder ITEMKENNZEICHNUNGSKARTEN angefordert werden, so wird zunächst geprüft, ob die entsprechenden ANKÜNDIGUNGSKARTEN vorhanden sind. Fehlen diese (bzw. enthalten sie unzulässige Zeichen), dann bricht das Programm mit Hinweis ab. Stimmt lediglich die Anzahl der Bezeichnungskarten nicht, so arbeitet ITAMIS mit einem Hinweis auf diesen Fehler weiter.

9. Für jedes ITEM muß der Skalierungsbereich (kleinster Wert U und größter Wert O) angegeben werden. Es wird geprüft, ob eine Beantwortung über der oberen Grenze O liegt. Ist das der Fall, so wird die Beantwortung wie BLANK behandelt.

3 Eingabebeschreibung

3.1 Anordnung und Reihenfolge der Steuerkarten für ITAMIS: Schematische Übersicht

Die folgende Übersicht zeigt die Anordnung der Steuerkarten, sowie deren Reihenfolge. Die nicht-obligatorischen Steuerkarten, die ITAMIS nur bei speziellen Anforderungen des Benutzers erwartet, sind eingeklammert. Sie entfallen, wenn entsprechende Angaben auf der Parmeterkarte nicht gemacht wurden.

1. Textkarte
2. Parameterkarte
3. Einleseformatkarte(n)
4. (Zusatzlabelkarte)
(Karte mit dem Zusatzlabel)
6. (Druckformatkarte)
(Karte mit Druckformat)
7. Subskalenschablone
Zuweis-Karten
Ende-Karte

8. (Invertierungsschablone)
(Zuweis-Karten)
(Ende-Karte)
9. (Richtig-Falsch-Schablone)
(Zuweis-Karten)
(Ende-Karte)
10. (Dichotomisierungsschablone)
(Zuweis-Karten)
(Ende-Karte)
11. (Minimax-Schablone)
(Zuweis-Karten)
(Ende-Karte)
12. (Missing-Data-Schablone)
(Zuweis-Karten)
(Ende-Karte)
13. (Korrelationsschablone)
(Zuweis-Karten)
(Ende-Karte)

14. (Subskalenkennzeichnungskarten)
(Erste Kennzeichnungskarte)
15. (Itemkennzeichnungskarten)
(Kennzeichnungskarte für Item1, Skala 1)
(Kennzeichnungskarte für Item2, Skala 1)
16. Finish-Karte (bzw. Textkarte des folgenden Problems)

3.2 Beschreibung der Steuerkarten

3.2.1 Textkarte

1- 6: TEXTKA
7-78: Beliebiger Text

3.2.2 Parameterkarte

Der Benutzer teilt ITAMIS seine Wünsche über bestimmte Lochungen in der Parameterkarte mit. Nachfolgend wird angegeben, welche Funktion durch welche Lochung in Gang gesetzt wird. Benötigt man eine Funktion nicht, so läßt man die betreffende Spalte in der Parameterkarte leer (BLANK).

1- 6: (PARA)1⁶ PROBLM
7-10: (NS) Anzahl der Probanden

⁶ In dieser Spalte sind die in ITAMIS verwendeten Parameternamen verzeichnet.

- 11-13: (NI) Anzahl der Items (genauer: Anzahl der Elemente des Datenvektors)⁷
- 14-15: (NSU) Anzahl der Skalen (min. 1, max. 30)
- 16-17: (LAB) Anzahl der Labelstellen (min. 1, max. 80)
- 18-19: (MD) Definition fehlender Daten:
- 2 : Keine spezielle Definition.
Null und blank werden als 0 behandelt (keine MISSING DATA)
 - 1 : Nur BLANKS werden als MD behandelt
 - 0 - 9: BLANKS und eine Lochung im Bereich von 0 bis 9 werden als MD behandelt
 - 11 - 19: BLANK und NULL, sowie eine Lochung im Bereich von 1 bis 9 werden als MD behandelt. Locht man z.B. in Spalte 18-19 eine 11, so gelten 0, BLANK und 1 als MD.
 - 10: Über MISSING-DATA-SCHABLONE wird für jedes Item gesondert eine Definition vorgenommen (die MD-De-

⁷ Max. 500 Itemelemente; max. 40 Items pro Skala.

- 20-21: (LAZU) Anzahl der Stellen des Zusatzlabels. Das Programm erwartet dann eine Zusatzlabelkarte.
- 22-23: (ISU) 1: Drucken ohne Gesamtpunkt-
wert
4: Drucken mit Gesamtpunkt-
wert
7: Skalenwerte werden ohne Gesamtpunkt-
wert geschrieben
8: Skalenwerte werden mit Gesamtpunkt-
wert geschrieben
9: Skalenwerte werden ohne Gesamtpunkt-
wert geschrieben und gedruckt
10: Skalenwerte werden mit Gesamtpunkt-
wert geschrieben und gedruckt
- Hinweise:
ITAMIS erwartet eine Druckformatkarte bei 1, 7, 8, 9 und 10. Die Daten werden im Druckformat auf File geschrieben.
- 24-25: (NDI) 1-9: Dichotomisierungskriterium. Ist eine Lochung größer als die angegebene Zahl oder ihr gleich (.GE.), dann wird diese nach 1 transformiert, ist sie kleiner, so wird 0 zugewiesen.
10: Über Dichotomisierungsschablone wird für jedes Item gesondert ein Kriterium angegeben.
- definition ist nicht für alle Items gleich).

- 26: (IFO) BLANK: Es wird eine Einleseformatkarte benötigt.
1 - 9: Anzahl der notwendigen Einleseformkarten.
- 27: (IVOR) BLANK: Normaler Vorschub auf dem Schnelldrucker
1: Tabellenbeginn am Anfang einer Seite (größerer Papierverbrauch)
- 28: (ITS) 1: Zur Bezeichnung der Skalen werden Skalenbezeichnungskarten benutzt (max. 80 Zeichen)
- 29: (ITI) 1: Zur Bezeichnung der Items werden Itembezeichnungskarten benutzt (max. 72 Zeichen)⁸
- 30: (MI) Untere Grenze der Antwortkodierung - (U) für alle Items
- 31: (MA) Obere Grenze der Antwortkodierung - (O) für alle Items
- Beispiel: Bei jedem Item können die Probanden eine von fünf Antwortalternativen auf einer Skala ankreuzen, die von "völlig zutreffend (1)" bis "völlig unzutreffend (5)" reicht. In diesem Falle wäre U = 1 und O = 5.

⁸ davon werden die ersten 48 Zeichen ausgegeben

- Sind die Spalten 30 und 31 BLANK, so erwartet ITAMIS eine Minimax-Schablone. Mit Hilfe dieser Schablone kann U und O für jedes Item gesondert angegeben werden. Dies ist z.B. erforderlich, wenn die Antwortkodierungen nicht für alle Items gleich sind.
- 32: (NIV) 1: Zur Umkehrung der Antwortkodierungen wird eine Invertierungsschablone benötigt.
- 33: (NRF) 1: Zur Angabe der Kodierungen richtiger Lösungen eines Tests mit Mehrfachwahl-Antworten wird eine Richtig-Falsch-Schablone benötigt.
- 35: (ITOP) 1: Es werden nur Summenwerte bestimmt und auf Drucker oder File 2 ausgegeben. Itemanalyse etc. entfällt (IV, IK, ISK und IR werden 0 gesetzt).
- 36: (IV) 1: Bestimmung der primären Häufigkeitsverteilung der Skalenwerte für jede Skala.
- 37: (IK) 1: Interkorrelation der Skalenwerte und deren Korrelation zum Gesamtpunktwert.
- 38: (ISK) 1: Item-Skalen-Korrelationen werden für j e d e Skala bestimmt.

- 2: Über Korrelationsschablone wird ITAMIS mitgeteilt, für welche Skalen diese Koeffizienten zu bestimmen sind.
- 39: (IR)
- 1: Interkorrelationen der Items werden für j e d e Skala bestimmt.
 - 2: Über Korrelationsschablone wird ITAMIS mitgeteilt, für welche Skalen diese Koeffizienten bestimmt werden sollen.
- 40: (ITA)
- 1: Nach dem Lesen der Daten von File 1 erfolgt ein REWIND. Dadurch kann der Datensatz erneut gelesen werden (Anwendungsbeispiel: ein Datensatz soll getrennt für weibliche und männliche Probanden analysiert werden).
- 78: (NOMI)
- 1: Die Anzahl der BLANKS/MD die bei einem Probanden aufgetreten sind, werden n i c h t ausgegeben (bei 0 oder BLANK erfolgt eine probandenweise BLANK/MD-Diagnose).
- 79-80: (PMAXB)
- Prozentsatz der BLANKS/MD, die pro Skala tolerabel sind (vgl. Beschreibung des Algorithmus zur Ersetzung fehlender Daten).
- Treten bei einem Probanden PMAXB Prozent oder weniger BLANKS/MD in einer Skala auf,

so werden diese durch die mittlere Reaktion (gerundet) des Probanden auf alle beantworteten Items dieser Skala ersetzt. Bei mehr als PMAXB Prozent BLANKS/MD setzt ITAMIS den Skalensummenwert BLANK und verarbeitet diesen als MD. PMAXB muß als ganze Zahl angegeben werden; die Zahl darf nicht größer als 50 sein. Ist PMAXB BLANK oder 0, so erfolgt keine Ersetzung von BLANKS/MD. Da die für den Ersetzungsalgorithmus erforderlichen Berechnungen recht aufwendig sind, sollte man die Ersetzung nur benutzen, wenn dies unerlässlich erscheint. Ferner ist zu prüfen, ob sich die Ersetzung überhaupt plausibel begründen läßt.

Beispiel für die Arbeitsweise des Ersetzungsalgorithmus:

Angenommen, es liegen 2 Skalen vor. Skala 1 hat 8 Items, Skala 2 enthält 6 Items. In den Spalten 79-80 der Parameterkarte ist 20 gelocht, d.h. maximal 20 Prozent BLANKS/MD können ersetzt werden. Bei Skala 1 dürfen demzufolge $(20/100) \cdot 8 = 1.6 \approx 2$, bei Skala 2 dürfen $(20/100) \cdot 6 = 1.2 \approx 1$ auftretende BLANKS/MD ersetzt werden.

Ein Proband habe in diesen beiden Skalen folgende Werte:

Skala 1: 12B22B22 (B steht hier für BLANK)
 Skala 2: 4B33B4

Bei Skala 1 ergibt sich eine mittlere Reaktion von $11/6 = 2$ (gerundet). Nach Ersetzung wird der Skalensummenwert somit 15. Skala 2 enthält hingegen mehr BLANKS als tolerabel ist; eine Ersetzung kann demzufolge nicht vorgenommen werden. Der Summenwert wird deshalb von ITAMIS mit BLANK besetzt und aus den Berechnungen ausgeschlossen (Strategie der "echten" MISSING-DATA Verarbeitung).

Wird der Ersetzungsalgorithmus angefordert, so gibt ITAMIS für jede Skala die Anzahl der BLANKS/MD an, die ersetzt werden können.

Der Prozentsatz tolerabler BLANKS/MD (PMAXB) gilt für a l l e Skalen - unterschiedliche Angaben für einzelne Skalen sind nicht möglich.

3.2.3 Einleseformatkarte(n)

Je nach angeforderten Funktionen sind verschiedene Fälle bei der Anfertigung der Formatkarten zu unterscheiden. Im Normalfall (Fall 1 bzw. 2) ist das Eingabeformat einfach zu erstellen.

Nachfolgend werden verschiedene Fälle näher erläutert. Dabei ist zu beachten, daß ITAMIS zwischen dem Datenvektor und dem Labelvektor unterscheidet. Der Labelvektor enthält die Spalten (Min. 1, Max. 80), aus denen das Identifikationsfeld für die ggf. auszugebenden Skalenwerte zusammengesetzt wird. Da dessen Aufbau einfach und unproblematisch ist, wird an dieser Stelle nicht näher darauf eingegangen.

Der Datenvektor kann bestehen aus:

1. Itemelementen, d.h. aus den Items, die in der Itemanalyse verarbeitet werden (Fall 1, Normalfall);

2. Itemelementen und Dummyelementen, d.h. Spalten, die zwar eingelesen, aber nicht in der Itemanalyse berücksichtigt werden (Fall 2, Normalfall mit Eingabeerleichterung durch Ausschluß von Items, die nicht mit Hilfe der Subskalenschablone einer Skala zugeordnet wurden).

3.2.3.1 Eingabefälle

Nachfolgend werden die einzelnen Fälle anhand von Beispielen verdeutlicht und die Anfertigung der Einleseformatkarte gezeigt.

Im Zusammenhang mit den vom Normalfall abweichenden Anforderungen des Benutzers an ITAMIS muß an dieser Stelle kurz auf das Konzept der Schablonen und Zuweis-Karten eingegangen werden. Wie oben ausgeführt, besteht der Datenvektor für ITAMIS aus einzelnen Elementen, die vom ersten bis zum letzten Element fortlaufend numeriert werden. Über Schablonen und Zuweis-Karten (diese sind der wichtige Bestandteil der Schablonen) wird ITAMIS mitgeteilt, welches Element eines Datenvektors ein Itemelement, das in der Itemanalyse verarbeitet werden soll, darstellt. Dazu gibt man auf der Zuweis-Karte einfach die Nummer des Elements im Datenvektor an. Daraus folgt: Wenn Eingabeerleichterung angefordert wird, dann muß man bei allen Schablonen auf den Zuweis-Karten die Nummer des Elements im Datenvektor anstelle der Itemnummer angeben; nur im Fall 1 sind Itemnummer und Nummer des Elements im Datenvektor identisch. Für die Anfertigung der Zuweis-Karten ergibt sich folgende einfache Verfahrensweise:

- Numeriere alle Elemente des Datenvektors fortlaufend;
- kreuze die Elemente an, die in die Itemanalyse einbezogen werden sollen (Itemelemente);

- trage die Nummern der Itemelemente auf den Zuweis-Karten ein.

Unter den nachfolgend aufgeführten Beispielen wird die jeweils angemessene Zuweis-Karte (nach der allgemeinen Formulierung des Eingabeformats) angeführt; es handelt sich jeweils um die Zuweis-Karten, die die Items einer Skala zuordnen. Der Einfachheit halber ist dabei davon auszugehen, daß alle Items einer Skala zugeordnet werden sollen.

FALL 1: Normalfall

NI=20 Items auf einer Lochkarte. Der Label ist fünfstellig (LAB=5). Die Items sind von folgenden Spalten der Lochkarte zu lesen:

1-3, 7, 10-14, 20, 22, 32-40

Der Label soll von den Spalten 74 und 77-80 der Karte übernommen werden. Dazu ist folgende Formatkarte anzufertigen:

(3I1,3X,I1,2X,5I1,5X,I1,1X,I1,9X,9I1,33X,A1,2X,4A1)

Allgemein formuliert: I1-Formate für alle NI Items; A1-Formate für die LAB Labelstellen

ZUWEISO1001,002,003,004,005,006,007,008,009,010,011
ZUWEISO1012,013,014,015,016,017,018,019,020

FALL 2: Normalfall mit Eingaberleichterung

Dieser Fall tritt ein, wenn von der Möglichkeit der Vereinfachung der Eingabe durch Vermeidung komplizierter Formate Gebrauch gemacht werden soll (nur ein Teil der Items wird einer oder mehreren Skalen über die Subskalenschablone zugeordnet). Hier kann der Benutzer zwei Strategien wählen, die an dem unter Fall 1 eingeführten Beispiel erläutert werden.

Strategie 1: In diesem Beispiel sind die 20 Items, die in die Analyse einbezogen werden sollen, in den Spalten 1-40 abgelocht. Man liest nun alle 40 Spalten ein und ordnet nur die 20 oben genannten Items über Subskalenschablone zu.

Bitte beachten: In diesem Fall ist NI=40, da 40 Spalten gelesen werden.

Folgende Formatkarte wäre anzufertigen:

(40I1, 33X, A1, 2X, 4A1)

Der Vergleich mit Fall 1 macht deutlich, daß sich die Eingabe wesentlich vereinfacht. Dies gilt umso mehr, wenn große Itemzahlen vorliegen und die Items im Datenvektor nicht unmittelbar hintereinander stehen.

Allgemein formuliert: I1-Formate für die Item- und Dummyelemente; A1-Formate für die LAB Labelstellen.

ZUWEISO1001,002,003,007,010,011,012,013,014,020,022
ZUWEISO1032,033,034,035,036,037,038,039,040

Strategie 2: Da nur die über die Subskalenschablone zugewiesenen Items für die Analyse berücksichtigt werden, kann man das Format alternativ auch so schreiben:

(3I1,I3,I1,I2,5I1,I5,I1,I1,I1,I9,9I1,33X,A1,2X 4A1)

In diesem Fall wäre NI=25 (zwanzig "wirkliche" Items plus fünf Dummyelemente, die nicht in die Analyse eingehen; die Formatangaben für diese sind unterstrichen).

Allgemein formuliert: Mischung von I1-Formaten für die Itemelemente und I-Formaten für die Dummyelemente (NI=Anzahl der "wirklichen" Items + Anzahl

der Dummyelemente); A1-Formate für die LAB Labelstellen.

ZUWEISO1001,002,003,005,007,008,009,010,011,013,
ZUWEISO1015,017,018,019,020,021,022,023,024,025

Die Strategien 1 und 2 können auch kombiniert werden.

3.2.3.2 Allgemeine Gesichtspunkte, die bei der Anfertigung der Einleseformatkarte(n) zu beachten sind

Im System ITAMIS bedeutet NI (vgl. Parameterkarte, Spalte 11-13) die Anzahl der Elemente eines Datenvektors (Anzahl der Itemelemente + ggf. Anzahl der Dummyelemente); nur im Fall 1 ist folglich NI gleich der Anzahl der Items. Bei der Erstellung der Formatkarten zur Steuerung des Einlesens der Daten muß deshalb folgendes berücksichtigt werden:

Es müssen stets NI Formatangaben im I-Format und - für den Labelvektor - LAB Formatangaben im A1-Format geschrieben werden.

- Für die Itemelemente immer I1-Formate benutzen;
- für die Labellemente stets A1-Formate verwenden;
- für die Dummyelemente in der Regel ebenfalls I1-Formate benutzen; nachfolgend wird davon ausgegangen, daß auch für die Dummyelemente I1-Formate angegeben werden;
- maximal 9 Einleseformatkarten sind zulässig;
- die Formatkarten können von Spalte 1-80 beschrieben werden.

3.2.3.3 Allgemeine Gesichtspunkte, die bei der Anfertigung von Schablonen zu beachten sind

Für die folgenden Karten und/oder Schablonen, die ITAMIS je nach Wunsch des Benutzers erwartet, gilt:

- ITAMIS unterscheidet grundsätzlich zwischen Ankündigungs- und Informationskarten.
- Ankündigungskarten teilen ITAMIS mit, welche Information die unmittelbar nachfolgenden Karten enthalten.
- Informationskarten enthalten die für ITAMIS wesentliche Information. Diese Abfolge (zuerst Ankündigungskarte, dann Informationskarte bzw. - bei Schablonen - Informationskarten) wurde vorgesehen, damit eine möglichst eingehende Prüfung der Eingabe durch ITAMIS vorgenommen werden kann. Enthält eine erwartete Ankündigungskarte nicht die korrekte Information, so bricht ITAMIS mit einem entsprechenden Hinweis ab.
- Ankündigungs- und Informationskarten werden nur erwartet, wenn der Benutzer auf der Parameterkarte entsprechende Angaben gemacht hat. Im Gegensatz zur Text-, Parameter- und Einleseformatkarte sind diese Karten also nicht obligatorisch. Um dies zu verdeutlichen, sind die Überschriften, unter denen die Karten im folgenden Text näher erläutert werden, in Klammern gesetzt.

3.2.4 Zusatzlabelkarte

Diese Karte wird erwartet, wenn in Spalte 20-21 der Parameterkarte eine von 0 oder BLANK verschiedene Zahl K steht.

Ankündigungskarte: Spalte 1-16: Zusatzlabelkarte
 Informationskarte: enthält in den Spalten 1-K die Zeichen, die bei der Ausgabe der Skalenwerte zu deren Kennzeichnung zusätzlich zum Label, der von den Datenkarten übernommen wird, eingebracht werden sollen. Die Verwendung einer Zusatzlabelkarte ist nur sinnvoll, wenn die Skalenwerte auf File geschrieben werden sollen.

3.2.5 Druckformatkarte

Diese Karte wird erwartet, wenn Spalte 22-23 der Parameterkarte eine der folgenden Zahlen enthält:

1,3,4,6,7,8,9,10

Ankündigungskarte: Spalte 1-16: Druckformatkarte

Informationskarte: Enthält in den Spalten 1-80 die Formate zur Steuerung der Ausgabe der Skalenwerte (I-Formate) und des Labels (Al-Formate). Falls ein Zusatzlabel eingebracht werden soll, müssen für diesen ebenfalls Al-Formate vorgesehen werden. Der Zusatzlabel muß nach den Skalenwerten und vor dem Label ausgegeben werden. Wichtig: Die Formate zur Ausgabe der Skalenwerte sollten so gewählt werden, daß der Vorschub des Druckers richtig gesteuert wird. Um hier sicher zu gehen, sollte man für einstellige Skalenwerte I2-Formate und für zweistellige Skalenwerte I3-Formate verwenden (max.: I4).

Die unterschiedlichen Parameter haben folgende Bedeutung:

- 1: Skalenwerte werden ohne Gesamtpunktwert gedruckt.
- 4: Skalenwerte werden mit Gesamtpunktwert gedruckt.

- 7: Skalenwerte werden ohne Gesamtpunktwert auf File mit der logischen Nummer 2 geschrieben.
- 8: Skalenwerte werden mit Gesamtpunktwert auf File mit der logischen Nummer 2 geschrieben.
- 9: Skalenwerte werden ohne Gesamtpunktwert auf File mit der logischen Nummer 2 geschrieben und auf dem Drucker ausgegeben.
- 10: Skalenwerte werden mit Gesamtpunktwert auf File mit der logischen Nummer 2 geschrieben und auf dem Drucker ausgegeben.

3.2.6 Hinweise zur Anfertigung von Schablonen

ITAMIS kennt insgesamt 7 Schablonen mit unterschiedlicher Wirkung, aber prinzipiell gleichem Aufbau. Alle Schablonen bestehen aus:

- Einer Ankündigungskarte;
- einer oder mehreren (max. 1000) Zuweis-Karten;
- einer Ende-Karte.

Über die Zuweis-Karten, die die für ITAMIS wesentliche Information enthalten, wird jedem Item nach Maßgabe der auf der Parameterkarte durch den Benutzer spezifizierten Anforderungen eine Reihe von Werten zugewiesen, welche den Ablauf der Analyse und die Verarbeitung der Daten steuern.

Der Aufbau der Zuweis-Karten ist einfach; sie enthalten in Spalte 1-6 die Zeichenfolge "ZUWEIS". Danach steht in Spalte 7-8 die Zahl, die den Items zur Steuerung der Analyse zugeordnet werden soll. Dann werden die Itemnummern bzw. die den Itemnummern entsprechenden Nummern der Elemente des Datenvektors fortlaufend abgelocht (dreistellig, rechts-

bündig, durch Komma getrennt). Auf eine Zuweis-Karte passen 18 Item- bzw. Elementnummern. Es können beliebig viele Zuweis-Karten verwendet werden. Die Reihenfolge, in der die Item- bzw. Elementnummern abgelocht werden, ist beliebig. Es ist nicht erforderlich, daß eine Zuweis-Karte vollständig ausgenutzt wird, sie muß jedoch mindestens eine Item- bzw. Elementnummer enthalten. Einem Item darf nur eine Zahl zugeordnet werden. Mehrfach Zuordnungen werden von ITAMIS diagnostiziert und als Fehler behandelt. Den Abschluß einer Schablone bildet eine Ende-Karte (Spalte 1-4: ENDE).

Nachfolgend werden die einzelnen Schablonen beschrieben und durch kurze Beispiele erläutert. Zur Vereinfachung wird dabei davon ausgegangen, daß auf den Zuweis-Karten Itemnummern eingetragen werden (Normalfall: Fall 1). Es wird hier nochmals daran erinnert, daß in anderen Fällen nicht Itemnummern, sondern Nummern der Elemente im Datenvektor einzutragen sind. Zu beachten sind hierbei die ausführlichen Erläuterungen und Beispiele im Abschnitt zur Beschreibung der Einleseformatkarte (Abschnitt III.3).

3.2.7 Subskalenschablone

Diese Schablone wird von ITAMIS in jedem Fall erwartet (auch dann, wenn nur eine Skala vorliegt). Über die Schablone wird ITAMIS mitgeteilt, welches Item zu welcher Skala gehört. Die Skalen werden fortlaufend numeriert. Maximal 30 Skalen sind möglich. Es ist nicht erforderlich, daß alle Items einer Skala zugeordnet werden.

Ankündigungskarte: Spalte 1-18: SUBSKALENSCHABLONE

Zuweis-Karten : Spalte 1-6 : ZUWEIS
Spalte 7-8 : Nummer der Skala,

der das Item zugeordnet werden soll.

Spalte 9-80: Dreistellige Itemnummern, jeweils durch Komma getrennt.

Beispiel: 20 Items; die ersten 10 Items sollen Skala 2, die nächsten 10 Skala 1 zugeordnet werden. Folgende Zuweis-Karten wären anzufertigen:

ZUWEIS01011,012,013,014,015,016,017,018,019,020
ZUWEIS02001,002,003,004,005,006,007,008,009,010

Ende-Karte : Spalte 1-4 : ENDE

3.2.8 Invertierungsschablone

ITAMIS erwartet diese Schablone, wenn die Spalte 32 der Parameterkarte eine 1 enthält. Über die Schablone wird ITAMIS mitgeteilt, bei welchen Items die Beantwortungen umzukehren sind.

Beispiel: Der Reaktionsmodus für ein Item ist fünfstufig und kodiert mit 1, 2, ... 5. Ein Proband hat bei einem Item den Wert 4. Nach Invertierung erhält er den Wert 2 (minimale + maximale Kodierung - Lochung = 1+5-4=2).

Ankündigungskarte: Spalte 1-22:
INVERTIERUNGSSCHABLONE

Zuweis-Karten : Spalte 1-6: ZUWEIS
Spalte 8: 1
Spalte 9-80: Dreistellige Itemnummer, jeweils durch Komma getrennt.

Beispiel: Ein Fragebogen enthält insgesamt 30 Items. Bei folgenden Items sind die Beantwortungen umzukehren: 1, 7, 14, 27, 28, 29. Folgende Zuweis-Karte wäre anzufertigen:

ZUWEISO1007,001,027,029,028,014

(Die Reihenfolge, in der die Itemnummern angegeben werden, ist beliebig)

Ende-Karte : Spalte 1-4: ENDE

3.2.9 Richtig-Falsch-Schablone

ITAMIS erwartet diese Schablone, wenn die Spalte 33 der Parameterkarte eine 1 enthält. Über diese Schablone wird ITAMIS mitgeteilt, welche Beantwortung bei Mehrfachwahl-Antworten als richtige Lösung verarbeitet werden soll. Weist ein Proband bei einem Item die der richtigen Lösung entsprechende Kodierung auf (zulässiger Bereich: 1, ... , 9; 0 nicht erlaubt), so wird der Wert 1 zugeordnet, andernfalls der Wert 0. Beispiel: Der Subtest eines Leistungstests hat 10 Items. Ein Proband muß bei jedem Item unter vier Antwortalternativen (kodiert mit 1, 2, 3, 4) die richtige angeben. Folgende Kodierungen repräsentieren die richtigen Lösungen: 1, 4, 3, 2, 1, 1, 3, 2, 4, 4. In diesem Fall wird über die Zuweis-Karten den Items die richtige Lösung zugeordnet (s.u.).

Ankündigungskarte: Spalte 1-24:
RICHTIG-FALSCH-SCHABLONE

Zuweis-Karten : Spalte 1-6 : ZUWEIS
Spalte 8 : Kodierung der richtigen Lösung
Spalte 9-80: Dreistellige Itemnummern, jeweils

durch Komma getrennt.

Beispiel: Es werden die Zuweis-Karten für das oben erwähnte Beispiel gezeigt:

ZUWEISO1001,005,006
ZUWEISO2004,008
ZUWEISO3003,007
ZUWEISO4002,009,010

Ende-Karte : Spalte 1-4: ENDE

3.2.10 Dichotomisierungsschablone

Diese Schablone wird von ITAMIS erwartet, wenn die Spalten 24-25 der Parameterkarte mit 10 gelocht sind. Über die Schablone wird ITAMIS mitgeteilt, nach welchem Kriterium die Beantwortungen dichotomisiert (d.h. in den Antwortmodus 0 und 1 umgewandelt) werden sollen.

Ankündigungskarte: Spalte 1-26:
DICHOTOMISIERUNGSSCHABLONE

Zuweis-Karten : Spalte 1-6:
ZUWEIS

Spalte 8:
Dichotomisierungskriterium.
Ist eine Kodierung größer als dieser Wert oder gleich diesem Wert (.GE.), dann wird die Beantwortung in 1 umkodiert; allen Kodierungen, die kleiner als das Dichotomisierungskriterium sind, wird 0 zugeordnet.

Spalte 9-80:
Dreistellige Itemnummer, jeweils
durch Komma getrennt.

Beispiel: 20 Items; für die ersten 10 Items sollen alle Kodierungen, die größer oder gleich 4 sind, in 1 umkodiert werden. Allen Kodierungen die kleiner als 4 sind, soll 0 zugeordnet werden. Für die Items 11-20 soll 3 als Dichotomisierungskriterium verwendet werden. Folgende Zuweis-Karten wären anzufertigen:

ZUWEIS04001,002,003,004,005,006,007,008,009,010
ZUWEIS03011,012,013,014,015,016,017,018,019,020

Ende-Karte : Spalte 1-4:
 ENDE

3.2.11. Minimax-Schablone

ITAMIS erwartet diese Schablone, wenn die Spalten 30 und 31 der Parameterkarte 0 oder BLANK sind. Über diese Schablone wird ITAMIS mitgeteilt, in welchem Bereich der Antwortmodus bei jedem Item liegt. Dazu wird die untere Grenze U und die obere Grenze O angegeben. Beispiel: Bei einem Item können die Probanden ihre Beantwortung auf eine Skala angeben, die von "völlig unzutreffend = 1 bis völlig zutreffend = 5" reicht. Für dieses Item wäre U = 1 und O = 5. ITAMIS prüft, ob die Kodierung die obere Grenze überschreitet. Ist das der Fall, so wird die Kodierung als fehlendes Datum betrachtet.

Ankündigungskarte: Spalte 1-17: MINIMAX-SCHABLONE

Zuweis-Karte : Spalte 1-16: ZUWEIS
 Spalte 7: Untere Grenze des
 Beantwortungsmodus
 (U)

Spalte 8: Obere Grenze des
 Beantwortungsmodus
 (O)
Spalte 9-80: Dreistellige Item-
 nummern, jeweils
 durch Komma ge-
 trennt.

Beispiel: 10 Items; die Items 1, 3, 5, 7, 10 haben einen dreistufigen Beantwortungsmodus (0, 1, 2), die restlichen Items einen vierstufigen (1, 2, 3, 4). Folgende Zuweis-Karten wären anzufertigen:

ZUWEIS02001,003,005,007,010
ZUWEIS14002,004,006,008,009

Ende-Karte : Spalte 1-4: ENDE

3.2.12 Missing-Data-Schablone

Diese Schablone wird erwartet, wenn die Spalten 18-19 der Parameterkarte 10 enthalten. Diese Lochung dient zur Definition fehlender Daten (Missing-Data) für jedes Item. Folgende Definitionen sind möglich:

- 2: 0 und BLANK werden als 0 behandelt.
Keine MD-Definiton.
- 1: Nur BLANKS werden als MD behandelt.
- 0 bis 9: BLANKS und die hier angegebene Lochung werden als MD behandelt.
Beispiel: Locht man eine 3, so wird BLANK und 3 als MD gewertet.
- 11 bis 19: BLANK und 0 sowie die um 10 verminderte angegebene Lochung werden als MD behandelt. Locht man z.B. 11, so gelten BLANK, 0 und 1 als MD.

Bei den MD-Definitionen ist zu beachten, daß ITAMIS prüft, ob eine Lochung über der festgesetzten obe-

ren Grenze der Kodierung liegt (vgl. 11.). Ist das der Fall, so wird diese Lochung ebenfalls als MD betrachtet.

Vorsicht: Durch eine MD-Definition kann der Skalierungsbereich eines Items verändert werden! Beispiel: Bei einem Item liegt ein fünfstufiger Beantwortungsmodus vor. Die Beantwortungen der Probanden wurden mit 0, 1, ..., 4 kodiert. Über die Missing-Data-Schablone wird nun 0 als MD definiert. Dadurch ist faktisch der Skalierungsbereich eingeschränkt worden; die untere Grenze U ist nunmehr 1, die obere Grenze O bleibt unverändert. Bei Invertierung ergibt sich daraus folgendes Problem: Der Inversionsalgorithmus bildet zur Umkehrung einer Beantwortung B die Größe B^* nach folgender Formel: $B^* = U+O-B$. Wenn nun $B=4$ ist, so folgt daraus bei $U=0$: $B^* = 0+4-4 = 0$, d.h. ein Wert, der als MD definiert wurde. Dieser Widerspruch läßt sich dadurch vermeiden, daß man $U=1$ setzt. Dann gilt nämlich: $B^* = 1+4-4 = 1$.

ITAMIS gibt auf solche eventuell versehentlich vorgenommenen Eingriffe in den Skalierungsbereich von Items, die durch MD-Definition entstanden sind, einen Hinweis.

Ankündigungskarte: Spalte 1-22:
MISSING-DATA-SCHABLONE

Zuweis-Karten : Spalte 1-6 : ZUWEIS
Spalte 7-8 : MD-Definition
Spalte 9-80: Dreistellige Itemnummern, jeweils durch Komma getrennt.

Beispiel: 10 Items; bei Item 1 und 10 sollen 0 und BLANK als Missing-Data behandelt werden (keine Missing-Data-Definition). Bei Item 2, 3 und 4 gelten

nur BLANKS als MD. Für Item 5, 6 und 7 gilt BLANK und 0 als MD. Bei Item 8 und 9 soll BLANK, 0 und 1 als MD gelten. Folgende Zuweis-Karten wären anzufertigen:

ZUWEIS-2001,010
ZUWEIS-1002,003,004
ZUWEIS00005,006,007
ZUWEIS11008,009

Ende Karte : Spalte 1-4 : ENDE

3.2.13 Korrelationsschablone

Diese Schablone wird erwartet, wenn Spalte 38 und/oder Spalte 39 der Parameterkarte eine 2 enthält. Mit Hilfe dieser Schablone wird die Berechnung von Iteminterkorrelationen (F22) und part-whole korrigierten Trennschärfen der Items einer Skala zu allen Skalen (F21) gesteuert (siehe auch F28).

Ankündigungskarte: Spalte 1-21:
KORRELATIONSSCHABLONE

Zuweis-Karten : Spalte 8:
1, nur Item-Skalen-Korrelationen werden angefordert.
2, nur Iteminterkorrelationen sind zu bestimmen.
3, beide Funktionen werden für die Skalen angefordert, deren Nummern (dreistellig) in den Spalten 9-80 stehen.
4, wenn für die folgenden Skalen keine Analysen erfolgen sollen; damit entfällt für diese Skalen die Itemanalyse, die Bestimmung der primären Häu-

figkeitsverteilung, die Berechnung der Item-Skalen-Korrelationen und die Bestimmung der Iteminterkorrelationen, Summenbildung findet jedoch statt.

Wird einer Skala keiner dieser Werte zugeordnet, so setzt ITAMIS den Wert 0 ein und führt eine Itemanalyse für die Items der betreffenden Skala durch. Item-Skalen-Korrelationen und Iteminterkorrelationen entfallen somit.

Spalte 9-80:

Dreistellige Skalennummern, jeweils durch Komma getrennt. Die Anzahl der Zuweis-Karten ist beliebig.

Beispiel: 10 Skalen; für die Skalen 1 und 7 soll nur die part-whole korrigierte Trennschärfe der Items zu allen 10 Skalen bestimmt werden; für die Skalen 3 und 4 ist die Iteminterkorrelation zu berechnen; für die Skalen 9 und 10 werden beide Funktionen angefordert. Folgende Zuweis-Karten wären anzufertigen:

ZUWEISO1001,007
ZUWEISO2003,004
ZUWEISO3009,010

Ende-Karte : Spalte 1-4: ENDE

3.2.14 Subskalenkennzeichnungskarten

Diese Karten, die zur Beschriftung der Skalen im Ausdruck dienen, werden von ITAMIS erwartet, wenn Spalte 28 der Parameterkarte eine 1 enthält.

Ankündigungskarte: Spalte 1-29:

SUBSKALENKENNZEICHNUNGSKARTEN

Nach der Ankündigungskarte folgen die Kennzeichnungskarten. Spalte 1-80 dieser Karten wird mit der Überschrift der Tabellen ausgegeben (z.T. werden nur die ersten 28 Zeichen ausgedruckt).

3.2.15 Itemkennzeichnungskarten

Diese Karten dienen der Beschriftung im Ausdruck. Sie werden erwartet, wenn Spalte 29 der Parameterkarte eine 1 enthält. Die Karten müssen "skalenweise" angeordnet werden (Item 1, Skala 1, Item 2, Skala 1, ..., Item 1, Skala 2, Item 2, Skala 2, usf.).

Ankündigungskarte: Spalte 1-24:

ITEMKENNZEICHNUNGSKARTEN

Nach der Ankündigungskarte folgen die Kennzeichnungskarten. Spalte 1-72 wird zur Beschriftung der Tabellen ausgegeben, die die wesentlichsten Itemkennwerte enthalten (Mittelwert, Standardabweichung, Trennschärfekoeffizienten. In einigen Tabellen werden ebenfalls (vgl. 14) nur die ersten 28 Zeichen ausgegeben).

3.2.16 Finish-Karte

Die Finish-Karte beendet den Programmablauf. Soll ein weiteres Problem bearbeitet werden, so folgt anstelle der Finish-Karte die Textkarte für dieses Problem, dann die Parameterkarte, usf.

Finish-Karte : Spalte 1-6 : FINISH

4 Eingabebeispiel

Im folgenden wird die Anwendung von ITAMIS an einem konstruierten, zur Demonstration des wesentlichen reduzierten Beispiel gezeigt. Die Anfertigung der notwendigen Steuerkarten wird erläutert.

Es soll eine relativ einfache itemanalytische Aufgabenstellung besprochen werden. Zur Lösung genügen wenige Steuerkarten.

Ein 40 Items umfassender Fragebogen zur Beurteilung von Hochschulveranstaltungen wurde von Studenten in der Mitte und am Ende des Semesters ausgefüllt. Der Fragebogen enthält vier Subskalen mit jeweils 10 Items vom LIKERT-Typ. Der vorgegebene Reaktionsmodus ist vierstufig, es können Kodierungen von 1-4 auftreten. Bei einigen Items ist Invertierung erforderlich.

Mit Hilfe der Itemanalyse soll geprüft werden, ob sich Itemkennwerte, Skalenkennwerte, Interkorrelationen, Koeffizienten zur Schätzung der internen Konsistenz etc. für die Daten von Erst- und Zweidurchführung wesentlich unterscheiden.

Ein Problem der Analyse besteht darin, daß nicht alle Studenten, die den Bogen in der Mitte des Semesters bearbeitet haben, bei Semesterende noch erreichbar sind ("Teilnehmerschwund"). Für einige Veranstaltungsteilnehmer ist folglich der zweite Datenvektor leer. Deshalb ist Missing-Data Verarbeitung erforderlich. Wir nehmen an, die Daten einer solchen Untersuchung seien wie folgt abgelocht:

Plan für die Ablochung der Daten

Spalte: Karte:

1-40	1	40 Items des Fragebogens (Erstdurchführung). Zu den vier Skalen gehören die folgenden Items: Skala 1: 11, 13, 15, 17, 22, 25, 29, 32, 38, 40 Skala 2: 2, 3, 4, 6, 8, 12, 19, 20, 31, 37 Skala 3: 1, 5, 7, 10, 14, 21, 26, 27, 33, 36 Skala 4: 9, 16, 18, 23, 24, 28, 30, 34, 35, 39
77	1	1 = Kartennummer; zugleich Kennzeichnung für den bei Semestermitte erhobenen Datenvektor
78-80	1	Probandennummer
1-40	2	40 Items des Fragebogens (Zweitdurchführung). Da es sich um denselben Fragebogen handelt, werden bei den Skalen die laufenden Itemnummern angegeben (der gesamte Datenvektor enthält 2*40=80 Items). Skala 1: 51, 53, 55, 57, 62, 65, 69, 72, 78, 80 Skala 2: 42, 43, 44, 46, 48, 52, 59, 60, 71, 77 Skala 3: 41, 45, 47, 50, 54, 61, 66, 67, 73, 76 Skala 4: 49, 56, 58, 63, 64, 68, 70, 74, 75, 79

Für folgende Items ist Invertierung erforderlich (Erst- und Zweitedurchführung, laufende Nummern):

1, 2, 3, 11, 15, 16, 17, 18, 19, 22, 24, 28, 29, 30, 34, 37, 38, 39, 40 (Erstdurchführung)

41, 42, 43, 51, 55, 56, 57, 58, 59, 62, 64, 68, 69, 70, 74, 77, 78, 79, 80 (Zweitedurchführung)

77 2 2 = Kartennummer, zugleich Kennzeichnung für den bei Semesterende erhobenen Datenvektor

78-80 2 Probandennummer

Notwendige Steuerkarten:

Nachfolgend werden lediglich die Steuerkarten ohne nähere Erläuterungen wiedergegeben. Der Testdatensatz enthält die Datenvektoren für 10 Meßwertträger (20 Karten).

A) Textkarte

/1-6:TEXTKA/8-27:Testlauf für Itamis/

B) Parameterkarte

/1-6:PROBLM/9-10:10/12-13:80/15:8/17:3/18-19:-1/23:1/30:1/31:4/32:1/36:1/37:1/38:2/39:2/

Durch diese Parameterkarte werden folgende Anforderungen spezifiziert:

- Für alle Items gelten BLANKS als MISSING DATA;
- die Summenwerte sind ohne Gesamtpunktwert auszugeben;
- bei einigen Items ist Invertierung erforderlich;
- der Label ist dreistellig;
- Interkorrelation der Skalen soll berechnet werden;
- die primäre Häufigkeitsverteilung der Summenwerte soll bestimmt werden;
- für Skala 2 und 3 sollen nur die Item-Skalen-Korrelationen bestimmt werden;
- für die Skalen 1 und 4 sind zusätzlich die Iteminterkorrelationen auszugeben;
- die Skalen 5 und 8 sind völlig aus der Analyse auszuklammern.

C) Einleseformatkarte

(40I1/40I1,37X,3A1)

D) Ankündigungskarte für Druckformatkarte

/1-16:DRUCKFORMATKARTE/

Karte mit dem Druckformat

(8I3, 10X, 3A1)

E) Ankündigungskarte für Subskalenschablone

/1-18:SUBSKALENSCHABLONE/

Zuweis-Karten

ZUWEISO1011,013,015,017,022,025,029,032,038,040
 ZUWEISO2002,003,004,006,008,012,019,020,031,037
 ZUWEISO3001,005,007,010,014,021,026,027,033,036
 ZUWEISO4009,016,018,023,024,028,030,034,035,039
 ZUWEISO5051,053,055,057,062,065,069,072,078,080

ZUWEISO6042,043,044,046,048,052,059,060,071,077
 ZUWEISO7041,045,047,050,054,061,066,067,073,076
 ZUWEISO8049,056,058,063,064,068,070,074,075,079

Ende-Karte

/1-4:ENDE/

F) Ankündigungskarte für Invertierungsschablone

/1-22:INVERTIERUNGSSCHABLONE/

Zuweis-Karten

ZUWEISO1001,002,003,011,015,016,017,018,019,022
 ZUWEISO1024,028,029,030,034,037,038,039,040,041
 ZUWEISO1042,043,051,055,056,057,058,059,062,064
 ZUWEISO1068,069,070,074,077,078,079,080

Ende-Karte

/1-4:ENDE/

G) Ankündigungskarte für Korrelationsschablone

/1-21:KORRELATIONSSCHABLONE/

Zuweis-Karten

ZUWEISO1002,003
 ZUWEISO3001,004
 ZUWEISO4005,008

Ende-Karte

/1-4:ENDE/

Hinter dieser Karte folgen die zwanzig Datenkarten für dieses Beispiel (pro Meßwertträger zwei Kar-

ten). Der Programmablauf wird durch die FINISH-KARTE (/1-6:FINISH/) abgeschlossen, falls nicht ein weiteres Problem bearbeitet werden soll. In diesem Fall liegt hinter der letzten Datenkarte des ersten Problems die TEXTKARTE zur Kennzeichnung des folgenden Problems.

5 Literatur

- Cronbach, L.J.:** Essentials of Psychological Testing; New York, Evanston, London, 1970
- Diehl, J.M. und Kohr, H.-U.:** Deskriptive Statistik; Frankfurt, Verlag Klotz, 1977
- Fisher, R.A. and Yates, F.:** Statistical Tables for Biological, Agricultural and Medical Research; London, 1963
- Guilford, J.P.:** Psychometric Methods; New York, 1954
- Kohr, H.-U.:** DATANA - Ein Programmsystem zur Daten- und Itemanalyse nach dem Konzept der klassischen Testtheorie; Berichte aus dem Psychologischen Institut der Universität Giessen, Nr. 4, 1972
- Lienert, G.A.:** Testaufbau und Testanalyse; Berlin, Weinheim, 1967

6 Stichwortverzeichnis

	Seite
Beschränkungen	83
D	85
Datenvektor	108
Dichotomisierungsschablone	119
Druckformatkarte	114
Einleseformatkarte(n)	108
Ersetzungsalgorithmus	94
Fehlerdiagnostik	96
Finish-Karte	126
Funktionen des Programmsystems	84
I	85
IFO	104
IK	105
Invertierungsschablone	117
IR	106
ISK	105
ISU	103
ITA	106
Itemkennzeichnungskarten	125
ITI	104
ITOP	105
ITS	104
IV	105
IVOR	104
Karten	81
Korrelationsschablone	123
LAB	102
Labelvektor	108
LAZU	103
MA	104
MD	102
MDC	86
MI	104
Minimax-Schablone	120
Missing-Data-Schablone	121
MITT1	87

MITT2	87
MITT3	87
NDI	103
NI	102
NIV	105
NMDI	88
NOMI	106
NRF	105
NS	101
NSU	102
O	85
PARA	101
Parameterkarte	101
Pathologische Daten	95
PMAXB	106
R	85
R-SKA	88
RCGES	88
RCSKA	88
Records	81
Reihenfolge der Steuerkarten	99
Richtig-Falsch-Schablone	118
SK	84
STAND1	87
STAND2	87
STAND3	87
Steuerkarten	81
Subskalenkennzeichnungskarten	125
Subskalenschablone	116
Textkarte	101
U	85
Zusatzlabelkarte	113

IV Allgemeine Hinweise zu den Programmen

1 Voraussetzungen an die Hardware

Um mit den PC-Versionen von ASIA und ITAMIS arbeiten zu können, müssen folgende Hardwarevoraussetzungen gegeben sein:

IBM-Personal Computer (oder kompatibler Rechner) mit 80386-Prozessor oder höher. Ein mathematischer Koprozessor (bei 80386) wird dringend empfohlen.

Diskettenlaufwerk 1,2 oder 1,44 MB, zusätzlich wird eine Festplatte empfohlen.

MS-DOS Version 3.1 oder höher.

Arbeitsspeicher: 640 KB oder mehr.

Zum Ausdruck der Ergebnislisten sollte ein Drucker zur Verfügung stehen, der die Möglichkeit bietet, Endlospapier zu verarbeiten.

An den Grafikadapter des Rechners werden keine besonderen Anforderungen gestellt.

2 Start der Programme

Obwohl ASIA und ITAMIS auch von Diskette aus lauffähig sind, sollten beide Programme wegen der wesentlich höheren Arbeitsgeschwindigkeit grundsätzlich von der Festplatte des Rechners gestartet werden. Dazu wird auf der Festplatte ein Arbeitsverzeichnis (z.B. ASIA, ITA, ARBEIT o.ä.) angelegt und es werden alle Files dort hineinkopiert, die sich im Verzeichnis ASIA bzw. ITAMIS auf der über sandten Diskette befinden. Auf jeden Fall müssen aber die Files ASIA.EXE bzw. ITAMIS.EXE und DOSXMSF.EXE im Arbeitsverzeichnis (oder zumindest im Suchpfad des Rechners) stehen. Ebenfalls muß ein Steuerfile (siehe unten) vorhanden sein. Das jeweilige Ergebnisfile wird erst bei der Programmausführung erzeugt, muß also nicht etwa schon vorher vom Benutzer angelegt worden sein.

2.1 ASIA

ASIA wird durch die folgende Eingabezeilen am Systemprompt gestartet:

```
ASIA FILE1 FILE2[RETURN]
```

- FILE1 ist der Name des entsprechenden Steuerfiles (z.B. STEUER.FIL) mit den ASIA-Anweisungen und den Rohdaten oder ein bereits vorher erzeugtes ASIA-Systemfile.
- FILE2 ist der Name eines Files, das ASIA während der Programmausführung erzeugt und in dem die Ergebnisse der Rechenarbeit in Form einer Liste abgelegt werden (z.B. ERGEBNIS.LIS).

Für das auf der übersandten Diskette befindliche erste Beispiel würde die Eingabezeile wie folgt aussehen:

```
ASIA BEISP1 LISTE[RETURN]
```

2.2 Beispiel für ein ASIA-Steuerfile

Um Ihnen einen Eindruck zu vermitteln, wie ein Steuerfile für ASIA aussehen kann, haben wir nachfolgend das erste Beispiel (BEISP1) abgedruckt, das sie auf der übersandten Diskette finden – sofern Sie diese wie unter Nr. 4 beschrieben anfordern.

```
RUN NAME ASIA-TESTLAUF;
VARIABLE LIST ID EIG1 TO EIG29
INPUT FORMAT (A4,29F2.0)
MISSING VALUES EIG1 TO EIG29(BLANK)
RANGE EIG1 TO EIG29(1,7)
VAR LABELS EIG1 AUTORITAER/EIG2 PARTNERSCHAFTLICH/
EIG3 ENGSTIRNIG/EIG4 FAEHIG, GUT AUSGEBILDET/
EIG5 EHRGEIZIG /EIG6 OBRIGKEITSHOERIG/
EIG7 KRITISCH,UNABHAENGIG/EIG8 AM WOHL D.ANDEREN INTERESS/
EIG9 STANDESBEWUSST/EIG10 GEBILDET/
EIG11 WICHTIG F.UNSER LAND/EIG12 AM GEHALT INTERESSIERT/
EIG13 AUFOFFERUNGSBEREIT/EIG14 INTELLIGENT/
EIG15 ERFUELLT AUFGABEN GUT/EIG16 POL. UNZUVERLAESSIG/
EIG17 IDEALISTISCH/EIG18 AN LEISTG INTERESSIERT/
EIG19 GEWISSENLOS/EIG20 EINGEBILDET/
EIG21 POL.EINFLUSSREICH/EIG22 REFORMFREUDIG/
EIG23 GEHALT ZU HOCH/EIG24 VORBILD F.DIE JUGEND/
EIG25 KARRIEREORIENTIERT/EIG26 PATRIOTISCH/
EIG27 SELBSTBEWUSST/EIG28 HALTEN ZUSAMMEN/
EIG29 AUF DISZIPLIN BEDACHT/ID IDENTIFIKATION
COMMENT *****
* DIE FRAGE LAUTETE: *
* SIE FINDEN IM FOLGENDEN EINE LISTE VON EIGENSCHAFTEN. *
* KRINGELN SIE BITTE FUER JEDE EIGENSCHAFT DEN PUNKT UNTER *
* DER DANEBENSTEHENDEN SKALA, DER ANGIBT, WELCHER TEIL DER *
* OFFIZIERE DER BUNDESWEHR DIE EIGENSCHAFT IHRER MEINUNG NACH *
* BESITZT. JE WEITER SIE AUF DIESER "SKALA" NACH RECHTS GEHEN, *
* DESTO GROESSER IST DER ANTEIL DER OFFIZIERE DER BUNDESWEHR, *
* DER DIE EIGENSCHAFT IHRER MEINUNG NACH BESITZT. *
*****
VALUE LABELS EIG1 TO EIG29 (1) KAUM EINER (2) WENIGE
(3) WENIGER ALS DIE HAELFTE (4) ETWA DIE HAELFTE
(5) MEHR ALS DIE HAELFTE (6) VIELE (7) FAST ALLE
TASK NAME ERSTER LAUF MIT DEFAULT-WERTEN
CLUSTER VARS=EIG1 TO EIG29 /SINGLE
```

```
READ INPUT DATA
1001 3 5 6 6 2 6 5 2 4 4 4 4 5 6 5 6 6 7
1001 4 4 2 6 6 5 5 5 6 6 7 3 5 6 6 2 6 5 2 4 4 4 4 5 6 5 6 6 7
1002 4 5 3 6 5 5 4 5 6 5 7 3 5 6 5 1 5 6 1 4 5 5 4 5 5 5 6 6 7
1003 4 4 2 6 5 5 4 4 5 6 6 3 4 6 6 1 3 5 2 4 4 4 5 5 5 5 6 6 6
1004 5 6 2 7 6 4 4 4 6 6 7 4 5 6 6 1 5 5 1 3 4 5 4 5 6 6 6 6 7
1005 6 5 2 7 6 5 5 5 5 6 7 3 5 7 7 7 7 1 2 4 4 3 5 5 6 6 6 7
END INPUT DATA
TASK NAME ZWEITER LAUF MIT SKALEN-RUEMPFEN
CLUSTER VARIABLES=EIG1 TO EIG29/RKRIT=.40/SUMMARY/
SCALE(IDOL)=EIG24/SCALE(GELD)=EIG12,EIG23
FINISH
```

2.3 ITAMIS

ITAMIS wird durch die folgende Eingabezeile gestartet:

```
ITAMIS FILE1 FILE2 FILE3 FILE4 FILE5[RETURN]
```

- FILE1 ist der Name des Steuerfiles (z.B. STEUER.ITA), das die ITAMIS-Steuerkarten enthält und im Anschluß daran die Rohdaten.
- FILE2 wird bei der Programmausführung erzeugt und enthält die vom Programm errechneten Ergebnisse (z.B. LISTE).
- FILE3, FILE4 und FILE5 sind die Namen dreier Files, die ITAMIS zur Programmausführung benötigt (z.B. 1, 2 und 3 oder A, B und C). Die Files werden automatisch bei der Ausführung des Programms erzeugt und sind nach Beendigung der Arbeit leer.

Für das Beispiel auf Diskette lautet die Eingabezeile:

```
ITAMIS BEISP LISTE 1 2 3[RETURN]
```


4 Bezug der Programme

Die Programme können von Interessenten bei folgender Adresse bezogen werden:

Sozialwissenschaftliches Institut der Bundeswehr
Winzererstraße 52
80797 München

Bitte fügen Sie Ihrer Anforderung einen ausreichend frankierten und mit der Versandadresse versehenen Rückumschlag sowie eine formatierte Leerdiskette (HD, 3,5 Zoll) bei.

Trotz umfangreicher Tests können weder die Autoren noch das Sozialwissenschaftliche Institut der Bundeswehr Garantien für eine fehlerfreie Funktion übernehmen.

Rückfragen unter 089/12003-352 (Telefax) oder 089/12003-356 (Telefon).