

# Eine Untersuchung zu sprachlichen Formulierungen der Items in deutschen Persönlichkeitsfragebogen\*

**Franz-Josef Löhr & Alois Angleitner** Universität Bielefeld

## Zusammenfassung

Der sprachlichen Formulierung von Fragebogenitems wurde bisher in der Persönlichkeitsforschung relativ wenig Beachtung geschenkt. Formulierungsprobleme wurden nur selten systematisch behandelt. Die vorliegende Arbeit untersucht systematisch 1840 Items aus 16 deutschsprachigen Persönlichkeitsfragebogen. Die erhobenen Merkmale umfassen syntaktische Charakteristika wie Verneinungen oder Passivgebrauch, Antwortformate und Itemlänge. Die Ergebnisse zeigen, daß sich in einer Gegenüberstellung ganzer Fragebogen die Itemformulierungen des Gießen-Tests (BECKMANN & RICHTER 1972) als relativ homogen erweisen, während CATTELLS 16 PF eine Fülle unterschiedlicher Itemformulierungen enthält. Items aus Fragebogen für Kinder und Jugendliche sind im allgemeinen einfacher formuliert als Items aus Fragebogen für Erwachsene. Analog zum Grundgedanken einer Multi-trait-multi-method-Matrix wurden Subskalen mit jeweils übereinstimmenden Meßintentionen zusammengefaßt und der Gesamtmenge aller Subskalen gegenübergestellt. Es zeigte sich, daß bestimmte sprachliche Merkmale in einigen dieser Subskalengruppen besonders häufig verwendet wurden. Darüber hinaus sind einige Subskalengruppen in ihrer syntaktischen Struktur homogener als nicht zusammengehörige Skalen, was allerdings auf Itemüberlappungen zurückgeführt werden kann. Bei parallelen Subskalen (16 PF und EYSENCKS EPI) findet sich eine schwache Tendenz zu homogeneren Formulierungen. Die Bedeutung der sprachlichen Formulierung für die Konstruktion und Bewertung von Persönlichkeitsfragebogen und das weitere Vorgehen bei der Validierung des vorgelegten Ansatzes werden diskutiert.

## Abstract

The problem of the formulation of questionnaire items has received little attention in personality research to date; it has seldom been treated systemically. In the present study 1840 items from 16 German-Language personality inventories were analysed. The aspects studied include syntactic features such as negation and the use of the passive voice as well as response formats and item length. Comparisons of entire inventories reveal that the Gießen Test (BECKMANN & RICHTER 1972) is relatively homogeneous, whereas CATTELL'S 16 PF contains a considerable variety of item formulations. In general, items from inventories for children and adolescents are formulated more simply than are items from inventories for adults. In analogy to the principle underlying multitrait-multimethod matrices, sets of subscales measuring similar constructs were compared with the set of all subscales. Certain linguistic features were found to occur with particular frequency in some of these sets of subscales. Moreover, some sets of subscales are more homogeneous than the set of all subscales with respect to the syntactic structures they exhibit; this may, however, be due to item overlap. Parallel subscales in the 16 PF and EYSENCK'S EPI exhibit a weak tendency toward homogeneous formulations. The importance of item formulation in the construction and evaluation of personality inventories is discussed and further steps toward the validation of the present approach are suggested.

\* Dieser Artikel entstand im Rahmen des Forschungsprojekts „Persönlichkeitsfragebogen“ mit finanzieller Unterstützung der DFG.

In der frühen Periode der Fragebogenentwicklung (etwa WOODWORTH 1917) sah man die Antwort eines Probanden (Pb) auf ein Fragebogenitem ausschließlich als Reaktion auf den Inhalt der Items an. Die Itemformulierung als solche wurde nicht als relevante Variable erkannt und untersucht.

In der Phase der empirisch orientierten Fragebogenkonstruktion (internale Strategie wie Faktorenanalyse und Itemanalyse; externale Strategie wie Kriteriengruppenvergleich) wurde das Ankreuzen vorgegebener Antworten durch den Pb unabhängig vom Antwortprozeß und von spezifischen Iteminhalten ausschließlich nach teststatistischen Kriterien bewertet.

Der hier vertretene Ansatz geht im Gegensatz dazu – auch im Gegensatz zu den Annahmen der klassischen Testtheorie – von der Annahme bedeutsamer vermittelnder Prozesse bei der Fragebogenbeantwortung aus. Zur Klärung der Frage

*„Was passiert eigentlich bei einem Menschen, wenn er ein Fragebogenitem nach einem vorgegebenen Antwortschema beantworten soll, und wie kommt er schließlich zu seiner Antwort?“*

sollen zunächst die wesentlichen Determinanten des Beantwortungsvorgangs ermittelt werden; hierzu gehören neben person- und situationsbezogenen Merkmalen auch Eigenschaften der vorgegebenen Stimuli, also der Items selbst. Als Itemeigenschaften sind u. a. die sprachlichen Formulierungen der Items einschließlich der vorgegebenen Antwortmöglichkeiten anzusehen (*formal-syntaktische Merkmale*).

Für die Konstruktion von Fragebogen (Fbn) werden kaum explizite Regeln zur sprachlichen Formulierung von Items angegeben. Die meisten Fragebogenautoren scheinen zwar die allgemeine Auffassung zu teilen, daß Items möglichst einfach zu formulieren seien, insbesondere, wenn man Fbn für Kinder konstruiert. Indessen ist den Beschreibungen der Konstruktionsphasen zu entnehmen, daß explizite Regeln für die Formulierungen wohl nicht verwendet wurden. Statt dessen scheinen viele Autoren das Problem auf die Pbn zu verschieben, indem sie ein ad-

äquates Itemverständnis nur Personen mit mindestens durchschnittlicher Intelligenz zubilligen.

Formal-syntaktische Merkmale werden jedoch nicht nur in der Literatur zur Konstruktion von Fbn weitgehend vernachlässigt; sie finden auch kaum Berücksichtigung bei der Beurteilung der Brauchbarkeit bestehender Fragebogen, da sich solche Einschätzungen überwiegend am Kanon der klassischen Testtheorie orientieren.

In dieser Arbeit soll nun erstmals für deutschsprachige Fragebogen eine Analyse formal-syntaktischer Itemmerkmale vorgestellt werden. Dabei wird zunächst eine umfassende, reliable *Beschreibung* derartiger Merkmale an gängigen Fragebogen für Erwachsene sowie für Kinder und Jugendliche versucht. Diese Beschreibung soll in erster Linie Aufschluß über die Variabilität vs. Homogenität formal-syntaktischer Merkmale geben, wobei das *Gesamt der Fragebogen*, *einzelne Fragebogen* und *einzelne Skalen* als *Beschreibungsebenen* zu unterscheiden sind. Schließlich soll ansatzweise untersucht werden, ob sich *übereinstimmende inhaltliche Meßintentionen* in *homogenen sprachlichen Formulierungen* entsprechender Items niederschlagen.

## 1. Behandlung formal-syntaktischer Itemmerkmale in der Literatur

Als theoretischer Leitfaden zur Erfassung syntaktischer Strukturen bietet sich die generative Transformationsgrammatik von CHOMSKY an. Prinzipiell halten wir eine syntaktische Analyse ausgewählter Fb-Items mit den Mitteln der Transformationsgrammatik durchaus für sinnvoll, zumal psycholinguistische Untersuchungen die Relevanz von Transformationen zu belegen scheinen (vgl. HÖRMANN 1967, Kap. 13).

Jedoch ist der Versuch einer transformationellen Analyse in unserem Rahmen aufgrund der großen Anzahl von Items und mangelnder Kompetenz unsererseits aussichtslos.

Im übrigen ist die Anwendung linguistisch fundierter Theorien auch in den wenigen uns bekannten einschlägigen Untersuchungen aus dem

Bereich der Fragebogenforschung bisher nicht verwirklicht. Die Mehrzahl dieser Untersuchungen belegt die Wichtigkeit syntaktischer Strukturmerkmale und Antwortformate für das Zustandekommen von Fragebogenergebnissen. Bei JONES (1968) wird dies für verschieden stark differenzierte Antwortmöglichkeiten nachgewiesen, bei DUBOIS & BURNES (1975) für das Fragezeichen „?“ als unentschiedene Mittelkategorie, deren Wahl durch völlig verschiedene Überlegungen motiviert sein kann (z. B. Mehrdeutigkeit des Iteminhalts; Item trifft auf den Pb nicht zu). PETERSON & PETERSON (1976) fanden heraus, daß Negationen in der Formulierung beträchtliche Verwirrung stiften können, besonders dann, wenn die Aussage eines negativen Items nicht zutrifft, dieses also korrekterweise mit „Nein“ zu beantworten wäre.

Lediglich bei WIGGINS & GOLDBERG (1965) fanden sich keine praktisch bedeutsamen Korrelationen ( $r > .30$ ) zwischen den Merkmalen „Negation“, „Komplexität der Satzstruktur“, „Satzlänge“ sowie „relative zeitliche Häufigkeit“ einerseits und zahlreichen statistischen Kennwerten sowie „sozialer Erwünschtheit“ andererseits (als Material dienten die 566 Items des MMPI).

Dagegen konnten MICKLIN & DURBIN (1966) in einem komplexen Design die Wirkung der syntaktischen Variablen „einfacher vs. komplexer Satzbau“, „aktive vs. passive Verbform“ sowie „affirmative vs. negative Formulierung“ bei den Items von Einstellungsskalen nachweisen. Diese Autoren untersuchten Paralleltests unter dem Gesichtspunkt, ob Nicht-Äquivalenz der Satzstrukturen in den Items jeweils zweier Parallelformen mit dem Reliabilitätskoeffizienten für diese Parallelformen korreliert. Dazu stellten sie zunächst eine umfangreiche Liste von Merkmalen zusammen, aus der die drei genannten Variablen für die empirische Untersuchung ausgewählt wurden. Diese Variablen wurden dichotom erfaßt und die Items anschließend einer der  $2^3 = 8$  Ausprägungskombinationen zugeordnet. Anhand der prozentualen Anteile der Items mit verschiedenen Merkmalskombinationen an den Skalen berechneten MICKLIN & DURBIN für jedes der untersuchten Parallelskalenpaare einen

Index der Satzstrukturdifferenz und korrelierten diesen mit der Paralleltestreliabilität. Bei Konstanthaltung des Einflusses der Testlänge auf die Reliabilität (Partialkorrelation) erhielten sie einen Koeffizienten von  $r = -.39$ , d. h. je größer die Unterschiede zwischen den Satzstrukturen der Items paralleler Einstellungsskalen sind, desto geringer ist die Antwortkonsistenz der Pbn.

## 2. Beschreibung der untersuchten formal-syntaktischen Itemmerkmale

Die Auswahl der zu erfassenden formal-syntaktischen Merkmale orientierte sich an der von MICKLIN & DURBIN (1969) zusammengestellten Liste. Einige Merkmale aus dieser Liste wurden jedoch nicht berücksichtigt, namentlich solche mit stark semantischer Akzentuierung, wie z. B. eine Aufteilung in transitive vs. intransitive Verben. Ferner wurde zusätzlich versucht, die in den Fbn vorgegebenen Antwortformate sowie die physikalische Länge der Items zu erfassen. Im einzelnen wurden folgende Merkmale untersucht:

- (1) *Zahl der Wörter*: Anzahl der einzelnen Wörter pro Item.
- (2) *Zahl der Buchstaben*: Anzahl der Buchstaben, aus denen das gedruckte Item besteht.
- (3) *Gliederung in Satzeinheiten*: Anzahl der Satzeinheiten mit eigenem Prädikat, d. h. Hauptsätze, Nebensätze und Infinitivkonstruktionen mit Nebensatzfunktion. Dieses Merkmal soll einen Anhaltspunkt für die Komplexität des Satzbaus liefern und stellt eine erweiterte Fassung des dichotomen Merkmals „einfach vs. komplex“ von MICKLIN & DURBIN (1969) dar.
- (4) *Aktiv-Passiv*: Anzahl der im Passiv formulierten Verben pro Item.
- (5) *Negation*: Anzahl der in einem Item vorkommenden syntaktischen Negationen. Im Gegensatz zu der Dichotomie „affirmativ vs. negativ“ nach MICKLIN & DURBIN (1969) erlaubt diese Merkmalsdefinition eine Unterscheidung zwischen einfachen und doppelten Verneinungen.

- (6) *Zeitbezug*: Klassifizierung des Hauptsatzprädikats nach Gegenwart, Vergangenheit und Zukunft.
- (7) *Modus*: Differenzierung des Hauptsatzprädikats nach den Modi „Indikativ“ vs. „Konjunktiv“.
- (8) *Frageform vs. Aussageform*: Diese dichotome Unterscheidung bezieht sich auf die Form des Hauptsatzprädikats eines Items.
- (9) *Persönliche vs. unpersönliche Formulierung*: Dieses Merkmal erfaßt das persönliche Angesprochensein des Pb im Item, d. h. das Vorkommen eines entsprechenden Personalpronomens (Sie, ich).
- (10) *Formale vs. inhaltlich spezifizierte Antwortalternativen*: Es wird zwischen Antworten unterschieden, die lediglich Zustimmung zu bzw. Ablehnung einer inhaltlich abgeschlossenen Feststellung/Frage verlangen (ja – nein usw.), und Antworten, die noch eine inhaltliche Spezifizierung der Itemaussage enthalten (z. B. oft – manchmal – nie) sowie Mischformen.
- (11) *Formulierung der Antwort*: Nähere Spezifizierung der unter (10) vorgenommenen Grobeinteilung, und zwar bei „formalen Antwortalternativen“ Angabe der im einzelnen vorgegebenen Antworten (ja – nein, richtig – falsch, stimmt – stimmt nicht usw.), bei „inhaltlich spezifizierten Antwortalternativen“ Angaben darüber, ob neben der semantischen Verbindung zwischen Itemaussage und Antwort auch eine grammatische Einbindung besteht, d. h. ob das Item erst in Verbindung mit der Antwort einen vollständigen Satz ergibt.
- (12) *Anzahl der vorgegebenen Antwortalternativen*.

Die funktionale Bedeutung dieser Merkmale könnte etwa folgendermaßen umschrieben werden: Die physikalische Länge, die Komplexität des Satzgefüges, Verneinungen und Passivformulierungen sowie die Ausgestaltung der Antwortformate dürften vor allem die Verständlichkeit der Items beeinflussen, während die Satzform (Aussage/Frage), der Zeitbezug, der Modus sowie das persönliche Angesprochensein be-

stimmte Fokussierungen im Sinne des mehr oder weniger direkten Verlangens nach Auskunft, der Aktualität, der Verbindlichkeit und der persönlichen Betroffenheit mit sich bringen dürften.

Die funktionale Bedeutung der formal-syntaktischen Merkmale kann hier allerdings nur den Charakter plausibler Hypothesen über ihre Relevanz haben, sie ist nicht Gegenstand der Untersuchung.

### 3. Formal-syntaktische Merkmale der Items aus zehn deutschsprachigen Fragebogen für Erwachsene

#### 3.1 Durchführung der Datenerhebung

Gegenstand der Untersuchung waren die Items der folgenden 10 Fragebogen für Erwachsene:

- Gießen-Test – „GT“ (BECKMANN & RICHTER 1972)
- Maudsley-Personality-Inventory – „MPI“ (EYSENCK 1959)
- Maudsley-Persönlichkeitsfragebogen – „MMQ“ (EYSENCK 1964)
- Eysenck-Persönlichkeitsinventar Form A – „EPI-A“ (EGGERT 1974)
- Eysenck-Persönlichkeitsinventar Form B – „EPI-B“ (EGGERT 1974)
- Freiburger Persönlichkeitsinventar – „FPI“ (FAHRENBERG & SELG 1970)
- Persönlichkeits-Interessen-Test – „PIT“ (MITTENECKER & TOMAN 1951)
- Sixteen Personality Factor Questionnaire Form A – „16 PF-A“ (CATTELL & EBER 1962)
- Sixteen Personality Factor Questionnaire Form B – „16 PF-B“ (CATTELL & EBER 1962)
- Minnesota Multiphasic Personality Inventory – „MMPI-Saarbrücken“ (SPREEN 1963).

Diese Fbn enthalten insgesamt 1624 Items.

Zur Auszählung der Buchstaben und Wörter pro Item wurde ein Computerprogramm geschrieben. Als Eingabedaten dienten die im Wortlaut verlochten Items (vgl. ANGLEITNER 1976).

Im Gießen-Test wurden die Antworten jeweils mitgezählt, bei den Fbn 16 PF-A und 16 PF-B wurde bis zum ersten abschließenden Satzzeichen („“, „!“, „?“) gezählt, so daß auch hier die Antworten mitgezählt wurden, sofern sie als Satzteile oder Nebensatz Bestandteil des „eigentlichen“ Items waren.

Alle isoliert stehenden Zeichen (Satzzeichen, einstellige Ziffern, Klammern und vereinzelte Buchstaben) wurden automatisch von der Zählung ausgenommen.

Die verbliebenen 10 Merkmale (Nr. 3 – 12 der obigen Aufstellung) sollten anhand einer Codierungsvorschrift durch zwei Beurteiler geratet werden. Die Codierungsvorschrift sollte so präzise sein, daß möglichst kein Spielraum für subjektive Bewertungen blieb. Zu diesem Zweck wurden zunächst für die Kategorien vorläufige Codierungsregeln erstellt. Dann wurden aus jedem Fragebogen zwei Items nach Zufall ausgewählt und codiert. Die Codierung wurde von den beiden Beurteilern begründet; anschließend erhielten die Codierungsregeln durch weitere Präzisierung die Form, in der sie zur Einstufung aller Items verwendet wurden<sup>1</sup>. Für die Beurteilung der Erwachsenen-Fbn benötigten die beiden Beurteiler<sup>2</sup> jeweils ca. 50 Stunden.

### 3.2 Ergebnisse

#### 3.2.1 Güteprüfung der Ratingergebnisse

Nach Abschluß der Erhebung wurden die Beurteilungen von den beiden Ratern einzeln miteinander verglichen. Dabei wurden Fehler korrigiert; soweit es sich nicht um offensichtliche Fehler eines Beurteilers aufgrund von Konzentrationsschwankungen handelte, wurden die Abweichungen diskutiert und gegebenenfalls die

<sup>1</sup> Die Instruktion zur Einstufung der Items ist als Anhang in einem ausführlichen Forschungsbericht enthalten (ANGLEITNER, HAMM, LOHMANN & LÖHR 1978).

<sup>2</sup> Die Einstufung wurde durch den Erstautor und durch Frau Anette Lohmann vorgenommen, der wir an dieser Stelle unseren Dank aussprechen möchten.

Codierungsregeln in Form eines Nachtrags präzisiert. In 13 Fällen konnten die Beurteiler sich nicht einigen; hier wurde durch zwei zusätzliche Rater, in einigen Fällen schließlich durch Münzwurf entschieden.

Die Zahl der nicht übereinstimmenden Einschätzungen betrug insgesamt 205 (152 „Flüchtigkeitsfehler“ sowie 53 Fehler zu Lasten des Codiersystems). Bezogen auf 10258 nicht-redundante Codierungen (d. h. Codierungen, in denen innerhalb eines Fb bei einem bestimmten Merkmal Varianz auftrat) sind dies 2%. Ferner wurde aus den Fehleranteilen der beiden Beurteiler die Wahrscheinlichkeit geschätzt, daß beide übereinstimmend falsch codiert hatten (was dann unentdeckt bliebe). Auf die absolute Häufigkeit umgerechnet ergab sich ein Schätzwert von 0.41.

Zusammenfassend kann festgestellt werden, daß das Ergebnis der formal-syntaktischen Einstufung der Fragebogenitems sehr zuverlässig ist. Die Einstufungen sind anhand der ausführlichen Zuordnungsvorschrift überprüfbar. Die „schwachen Stellen“ in den Daten – d. h. die 13 Einstufungen, die mit Hilfe zusätzlicher Rater bzw. durch Zufall entschieden wurden – sind im einzelnen bekannt. Es ist zu erwarten, daß die beiden Rater bei weniger als einem von 10258 Urteilen übereinstimmend falsch entschieden haben.

#### 3.2.2 Ergebnisse auf Fragebogen-Ebene

Die wichtigsten Informationen über die formal-syntaktischen Merkmale der einzelnen Fbn sowie des Gesamt-Itempools sind in den Tabellen 1 und 2 aufgelistet. Da „2 Verben im Passiv“ und „Hauptsatzprädikat in der Zukunft“ jeweils nur zweimal vorkamen, wurden diese Ausprägungen mit „1 Verb im Passiv“ und „Gegenwart“ zusammengefaßt. Ein infinites Hauptsatzprädikat wurde wie ein Konjunktiv behandelt.

Das „durchschnittliche“ Item eines Fragebogens für Erwachsene (letzte Zeile der Tab. 1 und Tab. 2) besteht aus 12 Wörtern mit zusammen 65 Buchstaben; es enthält zwei Prädikate, besteht also aus einem Satzgefüge.

Tabelle 1: Maße der Itemlänge in 10 Fragebogen für Erwachsene

Fragebogen	Zahl der Wörter		Zahl der Buchstaben		Gliederung in Sätze		N (Items)
	M	SD	M	SD	M	SD	
GT	13.5	2.7	71.6	17.0	2.3	0.68	40
MPI	9.7	3.4	56.7	20.0	1.5	0.68	48
MMQ	10.7	4.9	56.9	25.9	1.9	1.10	56
EPI-A	9.8	4.4	54.9	21.8	1.7	0.77	57
EPI-B	10.1	4.9	53.2	25.0	1.6	0.75	57
FPI	10.9	3.8	57.2	20.7	1.7	0.71	212
PIT	11.1	3.9	62.4	24.5	1.7	0.74	214
16 PF-A	15.9	6.0	88.9	34.9	2.4	1.30	187
16 PF-B	16.1	6.0	92.0	33.7	2.3	1.30	187
MMPI	10.6	5.1	56.2	25.9	1.8	0.91	566
Gesamt	12.0	5.4	65.3	30.3	1.9	1.0	1624

M = Mittelwert

SD = Standardabweichung

Tabelle 2: Syntaktische Merkmale der Items in 10 Fragebogen für Erwachsene

Fragebogen	Prozentsatz von Items mit							N (Items)
	1 Ver- neinung	2 Ver- neinungen	Passiv	Frage- satz	Kon- junktiv	Prädikat in der Vergan- genheit	Proband nicht an- gesprochen	
GT	7.5	0	2.5	0	0	0	0	40
MPI	18.8	0	6.3	100	4.2	2.1	0	48
MMQ	26.8	1.8	0	0	1.8	8.9	0	56
EPI-A	19.3	1.8	5.3	100	7.0	1.8	0	57
EPI-B	14.0	3.5	1.8	100	5.3	5.3	0	57
FPI	22.2	0.5	3.3	0	3.8	5.7	1.9	212
PIT	17.3	0.5	0	0	44.9	7.5	14.0	214
16 PF-A	20.9	2.1	5.3	2.7	16.0	4.3	12.3	187
16 PF-B	24.1	3.7	5.9	3.2	17.6	5.9	15.5	187
MMPI	27.6	2.1	3.9	0	6.9	18.7	4.8	566
Gesamt	22.8	1.8	3.6	10.7	13.3	10.0	7.0	1624

23% der Items enthalten eine einfache Verneinung, knapp 2% eine doppelte, knapp 4% eine Passivkonstruktion. Gut 10% der Items sind als Frage, die übrigen als Aussage formuliert, wobei dies – mit Ausnahme des 16 PF – durch die Fragebogenzugehörigkeit bestimmt ist. Das Hauptsatzprädikat steht bei 13% der Items im Konjunktiv, bei 10% in der Vergangenheit. In 7% der Items wird der Proband nicht persönlich angesprochen.

Die vorgegebenen Antwortalternativen lauten im MMQ, MPI, EPI-A und EPI-B „ja – nein“, im PIT und FPI „stimmt – stimmt nicht“, im MMPI „richtig – falsch“. In den Fbn MPI und PIT ist außerdem ein Fragezeichen „?“ als unentschiedene Mittelkategorie vorgesehen. Der Gießen-Test bietet zu jedem Item zwei inhaltlich ausformulierte Alternativen an, die jeweils grammatisch an einen gemeinsamen Vorsatz angebunden sind; zwischen den beiden Alternati-

Tabelle 3: Antwortformate im 16 PF (absolute und relative Häufigkeiten)

	formale Antwortalternativen			inhaltlich spezifizierte Antwortalternativen		Mischformen
Form A	112 (59.9%)			29 (15.5%)		46 (24.6%)
Form B	110 (58.8%)			29 (15.5%)		48 (25.7%)
	„ja – nein“	„stimmt – stimmt nicht“	Sonstige	gramm. Verbindung mit dem Item		
				vorhanden	nicht vorhanden	
Form A	78 (41.7%)	32 (17.1%)	2 (1.1%)	9 (4.8%)	20 (10.7%)	–
Form B	61 (32.6%)	49 (26.2%)	0	12 (6.4%)	17 (9.1%)	–

ven kann auf einer 7-stufigen Skala eine graduelle Abstufung vorgenommen werden. Über die vielfältigen Antwortformate des 16 PF gibt Tabelle 3 Aufschluß. Für jedes Item stehen 3 Antwortmöglichkeiten zur Verfügung, wobei die „formalen Antwortalternativen“ und fast alle „Mischformen“ (z. B. 16 PF-A, Item 51: „Ich wäre lieber (a) Förster, (b) unsicher, (c) Lehrer“) ein „unsicher“ oder „dazwischen“ als Mittelkategorie enthalten. Tabelle 3 enthält im oberen Teil eine Aufschlüsselung der 16 PF-Antwortformate nach formalen vs. inhaltlich spezifizierten Antwortalternativen sowie Mischformen; im unteren Teil sind die ersten beiden Hauptkategorien weiter aufgliedert.

Will man einzelne Fragebogen zusammenfassend charakterisieren, so erweisen sich die Itemformulierungen des Gießen-Tests – bezogen auf die untersuchten Merkmale – als ziemlich homogen. Auch bei den Längenmaßen besitzt der GT trotz hoher Mittelwerte die geringsten Streuungen.

Auf der anderen Seite sind die Items des 16 PF (beide Formen) deutlich heterogener als die der übrigen Fbn, wenngleich sie nicht in allen Merkmalen die größte Variationsbreite aufweisen. Bei den verbleibenden Fbn sind lediglich Einzelwerte hervorzuheben: So zeichnet sich etwa der PIT durch hohe Prozentsätze von Items im Konjunktiv (44.9%) – bedingt durch die Interessenska-

len – sowie unpersönlich formulierte Items aus (14.0%), während der MMPI verhältnismäßig viele Items enthält (18.7%), die auf Vergangenes abzielen.

Für die Variablen „Zahl der Wörter“, „Zahl der Buchstaben“ und „Gliederung in Satzeinheiten“, die sämtlich mit der Länge zusammenhängen, wurden Interkorrelationen berechnet.

Zwischen Wort- und Buchstabenzahlen besteht ein so hoher Zusammenhang ( $r = .94$ ), daß diese beiden Merkmale als Indikatoren der Itemlänge fast austauschbar sind. Der Zusammenhang der beiden sozusagen „physikalischen“ Längenmaße mit der Zahl der Satzeinheiten, die als Längenmaß der Satzstruktur bezeichnet werden könnte, ist dagegen weniger stark ausgeprägt (Wörter  $\times$  Sätze:  $r = .77$ ; Buchstaben  $\times$  Sätze:  $r = .67$ ).

### 3.2.3 Ergebnisse auf Skalenebene

#### Ausprägung einzelner formal-syntaktischer Merkmale der Skalen

Bei der Auswertung der formal-syntaktischen Merkmale der Items auf Skalenebene wurden nur noch die Merkmale „Gliederung in Satzeinheiten“, „Verneinung“, „Passiv“, „Modus“, „Zeitbezug“ und „personenbezogene Formulierung“ berücksichtigt. Für die genannten Merk-

male sind grundsätzlich in allen Skalen verschiedene Ausprägungsgrade möglich, d. h. die Skalen sind miteinander vergleichbar, auch wenn sie zu verschiedenen Fragebogen gehören. Bei den Antwortformaten und der Formulierung der Items in Aussage- und Frageform ist dagegen keine Vergleichbarkeit gegeben, weil diese Merkmale fast durchgängig durch die Fragebogenzugehörigkeit determiniert sind (Ausnahme: 16 PF). Als Maß der Itemlänge wurde das Merkmal „Gliederung in Satzeinheiten“ verwendet, da es – im Gegensatz zur Anzahl der Wörter und Buchstaben – einen Aspekt der syntaktischen Struktur repräsentiert und somit am ehesten mit den übrigen Merkmalen vergleichbar ist.

Um einen Überblick über die syntaktischen Charakterisierungen der einzelnen Skalen zu gewinnen, wurden zunächst die Kennwerte der 6 ausgewählten Merkmale (Mittelwerte für „Satz-einheiten“, Prozentsätze für die übrigen Merkmale) analysiert. Die wichtigsten Charakteristiken der Häufigkeitsverteilungen über diese Skalenkennwerte sind in Tabelle 4 wiedergegeben.

Es ist hervorzuheben, daß die syntaktischen Charakteristika aller Skalen mit gleichem Gewicht, also unabhängig von der Itemzahl, in die einzelnen Verteilungen eingehen, d. h. jede Skala (insgesamt sind es 91) wird hier als eine Einheit behandelt. Als Folge dieser veränderten Gewichtung weichen die Mittelwerte der Skalenverteilungen geringfügig von den in Tab. 1 und 2 aufgeführten Gesamtwerten ab; lediglich die

Merkmale „Konjunktiv“, „Vergangenheit“ und „Pb nicht angesprochen“ sind um mehr als einen Prozentpunkt verändert.

Die beträchtlich hohe Standardabweichung bei den Variablen „Konjunktiv“ (27%), „Proband nicht angesprochen“ (17%) und „einfache Verneinung“ (13%) resultiert aus den hohen Maximumswerten dieser Variablen; so gibt es Skalen, die ausschließlich unpersönliche Items beinhalten (100%, wie z. B. die Intelligenzskala B des 16 PF) oder Skalen, die in hohem Maße im Konjunktiv formuliert sind (95.2%). Der maximale Prozentsatz einer Skala bezüglich der Variablen „einfache Verneinung“ liegt mit 46.6% ebenfalls noch sehr hoch.

Eine Beschreibung einzelner Skalen kann hier aus Platzgründen nur exemplarisch vorgenommen werden (vollständige Angaben in ANGLEITNER et al. 1978). In Anlehnung an Überlegungen, die der Zusammenstellung von Multitrait-Multimethod-Matrizen (CAMPBELL & FISKE 1959) zugrunde liegen, wollen wir einige Skalen herausgreifen, welche dieselbe Bezeichnung tragen und somit gleiche oder ähnliche Konstrukte zu messen vorgeben. Es sind dies 6 Skalen aus dem Neurotizismusbereich (Tab. 5), 6 Skalen aus dem Extraversionsbereich (Tab. 6) und 4 Lügenskalen (Tab. 7).

Für alle drei Konstruktbereiche gilt, daß die jeweils konstrukt-spezifischen Durchschnittswerte (es handelt sich wieder um ungewichtete Mittelwerte) nur unwesentlich, d. h. um weniger als 1 Standardabweichung, von den entsprechenden

**Tabelle 4:** Beschreibung syntaktischer Merkmale von 91 Skalen aus 10 Erwachsenen-Fragebogen: Charakteristika der Häufigkeitsverteilungen über Skalenmittelwerte bzw. Skalenprozentsätze

Merkmal	Verteilung basiert auf	Charakteristika der Verteilung			
		M <sup>+</sup>	SD	Max	Min
Satz-einheiten	Mittelwerten	1.99	.52	3.38	1.00
einfache Verneinung	Prozentsätzen	20.52	13.04	46.67	0.00
doppelte Verneinung	Prozentsätzen	1.59	3.47	15.38	0.00
Passiv	Prozentsätzen	3.81	5.65	30.00	0.00
Konjunktiv	Prozentsätzen	18.64	27.37	95.24	0.00
Vergangenheit	Prozentsätzen	6.95	9.05	33.33	0.00
Pb nicht angesprochen	Prozentsätzen	8.41	16.92	100.00	0.00

M<sup>+</sup> = Mittelwert; SD = Standardabweichung; Max = Maximum; Min = Minimum

Tabelle 5: Syntaktische Merkmale von 6 *Neurotizismusskalen* der Fragebogen für Erwachsene

	MPI	MMQ	EPI-A	EPI-B	FPI	PIT(D)	Mittelwert der 6 Skalen
Itemzahl	24	38	24	24	24	20	25.67
Satzeinheiten (M)	1.63	1.82	1.46	1.42	1.63	1.80	1.63
einfache Verneinung (%)	29.17	28.95	12.50	25.00	25.00	15.00	22.60
doppelte Verneinung (%)	0.0	0.0	0.0	4.17	0.0	0.0	0.70
Passiv (%)	8.33	0.0	8.33	0.0	12.50	0.0	4.86
Konjunktiv (%)	0.0	0.0	4.17	0.0	0.0	0.0	0.70
Vergangenheit (%)	4.17	7.89	0.0	4.17	8.33	10.00	5.76
Pb nicht angesprochen (%)	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Tabelle 6: Syntaktische Merkmale von 6 *Extraversionsskalen* der Fragebogen für Erwachsene

	MPI	EPI-A	EPI-B	FPI	PIT(C)	MMPI(SI)	Mittelwert der 6 Skalen
Itemzahl	24	24	24	24	15	70	30.17
Satzeinheiten (M)	1.33	1.75	1.79	1.58	1.67	1.93	1.68
einfache Verneinung (%)	8.33	16.67	4.17	20.83	26.67	27.14	17.30
doppelte Verneinung (%)	0.0	4.17	4.17	0.0	0.0	2.86	1.87
Passiv (%)	4.17	0.0	4.17	4.17	0.0	1.43	2.32
Konjunktiv (%)	8.33	8.33	12.50	25.00	6.67	12.86	12.28
Vergangenheit (%)	0.0	0.0	0.0	8.33	20.00	14.29	7.10
Pb nicht angesprochen (%)	0.0	0.0	0.0	0.0	13.33	4.29	2.94

Tabelle 7: Syntaktische Merkmale von 4 *Lügenskalen* der Fragebogen für Erwachsene

	MMQ	EPI-A	EPI-B	MMPI	Mittelwert der 4 Skalen
Itemzahl	18	9	9	15	12.75
Satzeinheiten (M)	2.17	2.22	1.33	1.73	1.86
einfache Verneinung (%)	22.22	44.44	11.11	33.33	27.78
doppelte Verneinung (%)	5.56	0.0	0.0	6.67	3.06
Passiv (%)	0.0	11.11	0.0	6.67	4.45
Konjunktiv (%)	5.56	11.11	0.0	20.00	9.17
Vergangenheit (%)	11.11	11.11	22.22	0.0	11.11
Pb nicht angesprochen (%)	0.0	0.0	0.0	0.0	0.0

Durchschnittswerten für alle Skalen (vgl. Tab. 4) abweichen (die Verteilungen wurden zuvor flächentransformiert, um ihre Schiefe auszugleichen). Eine Analyse der Homogenität bzw. Heterogenität gleichbenannter Skalen soll im folgenden unter dem Aspekt der „syntaktischen Ähnlichkeit“ versucht werden.

#### *Syntaktische Ähnlichkeit der Skalen*

In Anlehnung an das oben beschriebene Verfahren von MICKLIN & DURBIN (1969) wurde für jedes Skalenpaar ein *Index der syntaktischen Strukturdivergenz* berechnet. Allerdings konnten dieser Berechnung – abweichend von MICKLIN

& DURBIN – nicht die Merkmalskonfigurationen der einzelnen Items zugrunde gelegt werden, da die größere Variablenzahl in unserer Untersuchung sowie die größere Kategorienzahl der Merkmale „Gliederung in Satzeinheiten“ (maximal 8!) und „Zahl der Verneinungen“ (3) die Anzahl der möglichen Konfigurationen beträchtlich erhöht und damit einen Vergleich der Skalen untereinander außerordentlich aufwendig gemacht hätte.

Statt dessen wurde auf die Prozentsätze der Merkmalsausprägungen *pro Skala* zurückgegriffen, wie sie in Tab. 5–7 beispielhaft dargestellt sind; auch beim Merkmal „Satzeinheiten“ wurden die jeweiligen Ausprägungen zugrunde gelegt, die in den Tab. 4–7 im einzelnen nicht aufgeführt sind. Für jedes Skalenpaar *i, j* wurde dann zunächst pro Merkmal *m* eine absolute Prozentsatzdifferenz über die möglichen Merkmalsausprägungen *k* hinweg berechnet:

$$d(i, j)_m = \frac{\sum_{c=1}^k |P_{imc} - P_{jmc}|}{2} \quad (1)$$

Dieser Schritt ist in Abb. 1 für die Differenz der Lügenskalen des EPI-A und des MMPI im Merkmal „Zahl der Verneinungen“ dargestellt.

Die Prozentanteile der 3 möglichen Kategorien an den beiden Skalen sind durch schraffierte Rechtecksäulen dargestellt, deren Fläche für eine Skala zusammen jeweils 100% entspricht. Die Prozentsatzdifferenz entspricht der Summe der einfach schraffierten „überstehenden“ Flächenanteile, die zwecks Normierung auf eine maximale Differenz von 100% durch 2 dividiert wird.

Die so errechneten absoluten Prozentsatzdifferenzen wurden abschließend über die 6 Merkmale pro Skalenpaar gemittelt:

$$D(i, j) = \frac{\sum_{m=1}^6 d(i, j)_m}{6} \quad (2)$$

In Anlehnung an SCHLOSSER (1976) kann das Maß *D* auch als „mittlere absolute Profildifferenz“ bezeichnet werden, wenn man sich die Merkmalsprozentätze der Skalen in ein Profil eingezeichnet denkt.

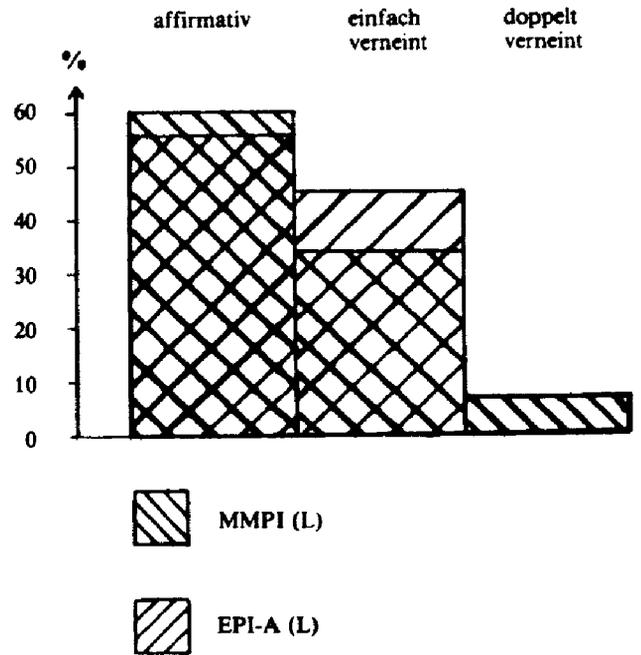


Abbildung 1: Prozentsatzdifferenz zweier Skalen, dargestellt am Beispiel des Merkmals „Zahl der Verneinungen“ der Lügenskalen des EPI-A und des MMPI. Doppelt schraffierte Flächenanteile stellen gemeinsame Prozentsätze dar, darüber hinausgehende einfach schraffierte Flächenanteile Prozentsatzdifferenzen.

In unserem Fall wurden aus 91 Skalen 4095 Skalenpaare gebildet, deren Differenzwerte summarisch als Verteilung beschrieben werden können:

$M = 17.7$ ,  $SD = 7.8$ ;  $Min = 2.4$ ;  $Max = 50.1$ .

Aus der Gesamtheit der möglichen Skalenpaare sollen nun zunächst die bereits oben erwähnten Skalen aus den Bereichen Neurotizismus, Extraversion und Lügenskalen herausgegriffen werden (Tab. 8–10).

Die gleichbenannten Skalen sind sich in ihrer syntaktischen Struktur tendenziell ähnlicher als der Durchschnitt aller Skalen, wobei diese Tendenz bei den Neurotizismusskalen am stärksten, bei den Extraversionsskalen schwächer ausgeprägt ist, während die Lügenskalen einander kaum ähnlicher sind, als für zwei beliebig herausgegriffene Skalen der 10 Fragebogen zu erwarten wäre. Eine zufallskritische Prüfung der Unterschiede ist nicht möglich, da die Werte der Profildifferenzenmatrix – analog einer Korrelationsmatrix – wechselseitig abhängig sind und

**Tabelle 8:** Unterschiede im formal-syntaktischen Aufbau von 6 *Neurotizismusskalen*: Mittlere absolute Profildifferenzen für jedes Skalenpaar

	MPI	MMQ	EPI-A	EPI-B	FPI
MMQ	3.21				
EPI-A	5.56	8.70			
EPI-B	4.17	4.09	6.25		
FPI	4.86	5.63	7.64	6.94	
PIT	7.08	5.22	7.92	7.78	6.67

M = 6.11; SD = 1.55; Min = 3.21; Max = 8.70

**Tabelle 9:** Unterschiede im formal-syntaktischen Aufbau von 6 *Extraversionsskalen*: Mittlere absolute Profildifferenzen für jedes Skalenpaar

	MPI	EPI-A	EPI-B	FPI	PIT
EPI-A	8.33				
EPI-B	5.56	6.25			
FPI	8.33	9.72	8.33		
PIT	16.53	10.28	16.53	15.83	
MMPI	13.54	7.56	8.89	9.25	8.73

M = 10.24; SD = 3.49; Min = 5.56; Max = 16.53

**Tabelle 10:** Unterschiede im formal-syntaktischen Aufbau von 4 *Lügenskalen*: Mittlere absolute Profildifferenzen für jedes Skalenpaar

	MMQ	EPI-A	EPI-B
EPI-A	11.11		
EPI-B	10.19	18.52	
MMPI	10.00	14.07	16.30

M = 13.37; SD = 3.22; Min = 10.00; Max = 18.52

die Verteilungskennwerte daher nur zu deskriptiven Zwecken benutzt werden können.

Bemerkenswert sind ferner die Differenzen zwischen den Parallelskalen des EPI: Die beiden Extraversionsskalen der Formen A und B liegen unter dem konstruktsspezifischen Durchschnitt, die Neurotizismusskalen entsprechen ihm fast, während die Lügenskalen die größte konstruktsspezifische Differenz aufweisen.

Bezieht man die von der Fragebogenkonstruktion her behauptete Parallelität von Skalen in die Analyse ein, so ist es sinnvoll, die Parallelskalen

des EPI und des 16 PF unabhängig von ihrem Inhalt auf ihre syntaktische Ähnlichkeit bzw. Unähnlichkeit hin zu untersuchen. Für die drei Parallelskalenpaare des EPI ergibt sich ein Mittelwert der Profildifferenzen von  $M = 10.34$ , für die 16 PF-Skalen ergeben sich folgende Verteilungskennwerte:

M = 12.92; SD = 4.32; Min = 2.56;  
Max = 21.79; N = 16.

Faßt man EPI und 16 PF zusammen, so resultiert für 19 Parallelskalen ein Mittelwert von  $M = 12.51$  bei einer Standardabweichung von  $SD = 4.58$ . Die in der Fragebogenbatterie enthaltenen Parallelskalen sind also einander syntaktisch etwas ähnlicher, als dies für ein beliebiges Skalenpaar zu erwarten wäre. Exklusive Ähnlichkeit findet sich allerdings nur bei den Intelligenzskalen (B) des 16 PF ( $D = 2.56$ ), die beide zu allen anderen Skalen wesentlich höhere Differenzen (D-Werte zwischen 30 und 50) aufweisen.

#### 4. Formal-syntaktische Merkmale der Items aus acht deutschsprachigen Fragebogen für Kinder und Jugendliche

##### 4.1 Durchführung der Datenerhebung

Folgende Fragebogen wurden in die Untersuchung einbezogen:

Fragebogen für Schüler – „FS 5–10“ (GÄRTNER-HARNACH 1973)

Kinder-Angst-Test – „KAT“ (THURNER & TEWES 1969)

Persönlichkeitsfragebogen zur Erfassung von Introversion und Neurotizismus bei Kindern und Jugendlichen, Form A – „INK(A)“ (NISCHAN 1974)

Angstfragebogen für Schüler – „AFS“ (WIECZERKOWSKI et al. 1974)

Hamburger Neurotizismus- und Extraversionsskala für Kinder und Jugendliche I und II – „HANES I“, „HANES 2“ (BUGGLE & BAUMGÄRTEL 1972)

Eysenck-Persönlichkeitsinventar Form A – „EPI-A“ (EGGERT 1974)

Eysenck-Persönlichkeitsinventar Form B – „EPI-B“ (EGGERT 1974)

Datenerhebung und deskriptive Auswertung auf der Fb-Ebene beziehen den EPI allerdings nicht mit ein (siehe dazu die Angaben zu den Erwachsenen-Fbn). Jedoch scheint es uns sinnvoll, den EPI in einem Vergleich des syntaktischen Aufbaus einzelner Skalen der Fragebogen für Kinder und Jugendliche zu berücksichtigen, da die Altersgruppierung um 14 Jahre die weitest umfänglichste aller im deutschen EPI-Manual aufgeführten Normgruppen bildet.

Die o.a. Fbn enthalten einschließlich EPI 330 Items; ohne den EPI sind es 216 Items.

Die Fragebogen für Kinder und Jugendliche (KJ-Fbn) wurden nach derselben Instruktion eingeschätzt wie die Erwachsenen-Fragebogen. Jedoch wurde die Zahl der Wörter von Hand ausgezählt, die Zahl der Buchstaben pro Item wurde nicht bestimmt. Da die Beurteiler-Übereinstimmung bei den Erwachsenen-Fragebogen sich als sehr hoch erwiesen hatte und zudem die KJ-Fragebogen im syntaktischen Aufbau ihrer Items weniger komplex schienen, wurden die Ratings nur von einem Rater durchgeführt und anschließend stichprobenweise überprüft. Bei der Überprüfung fand sich lediglich ein falsches Urteil.

## 4.2 Ergebnisse

### 4.2.1 Ergebnisse auf Fragebogenebene

In Tabelle 11 sind die wichtigsten Informationen über die formal-syntaktischen Merkmale der KJ-Fragebogen sowie des Gesamt-Itempools dieser Fragebogen (ohne EPI) mitgeteilt.

Zunächst sollen die Gesamtwerte für die Items der KJ-Fragebogen (letzte Spalte in Tab. 11) denen der Erwachsenenfragebogen (vgl. Tab. 1 und 2) gegenübergestellt und die Differenz zu-fallskritisch geprüft werden<sup>3</sup>:

<sup>3</sup> Bei der Variable „Wortzahl“ wurde ein t-Test für unabhängige Stichproben (zweiseitig) durchgeführt, bei den übrigen Merkmalen  $\chi^2$ -Tests für  $2 \times 2$ - bzw.  $2 \times 3$ -Häufigkeitstabellen (letzteres für „Verneinungen“ und „Sätze“, wobei „3 und mehr Sätze“ zusammengefaßt wurden).

Die Items der KJ-Fragebogen enthalten weniger Wörter ( $t = 3.92$ ;  $p < .001$ ,  $df > 216$ ) und bestehen aus weniger Satzeinheiten ( $\chi^2 = 6.02$ ;  $p < .05$ ,  $df = 2$ ). Die relative Häufigkeit einfacher und doppelter Verneinungen ist nur unwesentlich verringert ( $\chi^2 = .866$ ;  $p > .5$ ,  $df = 2$ ); das Passiv wird dagegen in den KJ-Fragebogen sogar häufiger gebraucht als in den Fragebogen für Erwachsene ( $\chi^2 = 16.1$ ;  $p < .001$ ,  $df = 1$ ), ein Ergebnis, das der Erwartung widerspricht. Andererseits sind Items im Konjunktiv, in Vergangenheitsform und ohne direktes Ansprechen des Pb in den KJ-Fragebogen deutlich seltener; die Unterschiede lassen sich statistisch sichern, und zwar auch dann, wenn man im Falle des Konjunktivs die atypischen Interessensskalen des PIT von der Berechnung ausnimmt (jeweils  $\chi^2 > 8$ ;  $p < .005$ ,  $df = 1$ ).

Die Korrelation zwischen Wort- und Satzzahl beträgt für die 216 Items der KJ-Fragebogen  $r = .80$  und ist mit der entsprechenden Korrelation in den Erwachsenenfragebogen ( $r = .77$ ) vergleichbar.

Betrachtet man die einzelnen Fragebogen für sich, so zeichnet sich der FS5 – 10 durch relativ lange, gegliederte Sätze, einfache Verneinungen und Passivformulierungen aus. Letzteres trifft auch für den KAT und den AFS zu. Beim AFS fallen darüber hinaus zahlreiche Verneinungen und einige unpersönliche Formulierungen auf, die ansonsten in den KJ-Fragebogen nicht anzutreffen sind.

Für die Items sind 2 formale Antwortalternativen vorgesehen, und zwar „stimmt – stimmt nicht“ im AFS, „ja – nein“ in den übrigen Fragebogen.

### 4.2.2 Ergebnisse auf Skalenebene

#### *Ausprägung einzelner formal-syntaktischer Merkmale der Skalen*

Die Auswertung für einzelne Skalen der KJ-Fragebogen erfolgte unter den Randbedingungen (Beschränkung der Merkmale auf 6, ungewichtete Mittelwertbildung), wie sie oben unter 3.2.3 für die Erwachsenenfragebogen erläutert sind. Die Prozentsatzdarstellung wurde trotz

**Tabelle 11:** Formal-syntaktische Merkmale bei sechs Fragebogen für Kinder und Jugendliche (ohne EPI)

Fb	FS 5 – 10	KAT	INK(A)	AFS	HANES 1	HANES 2	alle Fbn
Anzahl der Wörter M	14.11	11.74	8.71	10.74	9.56	9.41	10.64
SD	5.12	4.20	3.51	4.61	3.15	3.43	4.49
Anzahl der Satzeinheiten M	2.37	1.79	1.49	1.78	1.56	1.56	1.76
SD	.87	.69	.67	.73	.60	.70	.78
Verneinung einfach (%)	28.90	5.30	12.20	32.00	16.70	15.60	20.40
Verneinung doppelt (%)	–	–	2.40	–	2.80	3.10	1.40
Passiv (%)	18.40	15.80	2.40	16.00	2.80	3.10	9.70
Konjunktiv (%)	–	–	2.40	4.00	5.60	3.10	2.80
Vergangenheit (%)	5.30	–	–	6.00	–	9.40	3.70
Pb nicht angesprochen (%)	–	–	–	6.00	–	–	1.40
N (Items)	38	19	41	50	36	32	216

M = Mittelwert; (%) = Prozentsatz; SD = Standardabweichung

teilweise geringer Itemzahl beibehalten, um die Werte verschiedener Skalen vergleichbar zu machen.

Tabelle 12 gibt die Charakteristika der Verteilungen skalenbezogener Merkmalskennwerte für alle 19 Skalen der KJ-Fragebogen wieder, wobei – im Gegensatz zu Tab. 11 – die EPI-Skalen mit einbezogen sind.

Ein Vergleich der Mittelwerte der Skalen aus KJ-Fragebogen (Tab. 12) mit denen aus den Erwachsenen-Fragebogen (Tab. 4) zeigt durchweg dieselben Relationen, wie sie unter 4.2.1 auf der Fragebogenebene dargestellt wurden; weder die auf Skalenebene vorgenommene Umgewich-

tung noch die Hinzunahme der EPI-Skalen haben bedeutsame Verschiebungen bewirkt.

Tab. 13 enthält Angaben zu einzelnen Skalen der Bereiche „Neurotizismus“, „Extraversion“ und „Ängstlichkeit“ sowie „Schulunlust“, „Soziale Erwünschtheit“ und „Lügen“. Zusammen mit den eindimensionalen Ängstlichkeitsfragebogen FS 5 – 10 und KAT (vgl. Tab. 11) und den 6 EPI-Skalen (Tab. 5 – 7) sind damit die 19 Skalen der KJ-Fragebogen vollständig dokumentiert. Die Bereiche „Neurotizismus“ (5), „Extraversion“ (5), „Ängstlichkeit“ (4) und „Lügen-skalen“ (3) sind mit jeweils mehreren Skalen gleicher oder zumindest ähnlicher Meßintention

**Tabelle 12:** Beschreibung der Verteilung syntaktischer Merkmale der Skalen (N = 19) – Fragebogen für Kinder und Jugendliche (einschließlich EPI)

Merkmal	Verteilung basiert auf	Charakteristika der Verteilung			
		M	SD	Max	Min
Satzeinheiten	Mittelwerten	1.67	0.34	2.37	1.10
einfache Verneinung	Prozentsätzen	19.65	14.31	50.00	0.00
doppelte Verneinung	Prozentsätzen	1.42	2.16	5.00	0.00
Passiv	Prozentsätzen	6.97	9.82	40.00	0.00
Konjunktiv	Prozentsätzen	4.81	5.17	12.50	0.00
Vergangenheit	Prozentsätzen	5.15	9.67	30.00	0.00
Pb nicht angesprochen	Prozentsätzen	1.58	6.88	30.00	0.00

M = Mittelwert; SD = Standardabweichung; Max = Maximum; Min = Minimum

**Tabelle 13:** Syntaktische Merkmale von 11 Skalen aus 4 mehrdimensionalen Fragebogen für Kinder und Jugendliche\*

	Neurotizismus			Extraversion		
	INK(A)	HANES 1	HANES 2	INK(A)	HANES 1 E 1	HANES 2 E 2
Itemzahl	22	20	20	19	8	8
Satzeinheiten (M)	1.55	1.55	1.65	1.42	1.75	1.38
einfache Verneinung (%)	22.73	25.00	20.00	0.0	12.50	0.0
doppelte Verneinung (%)	4.55	5.00	5.0	0.0	0.0	0.0
Passiv (%)	4.55	5.00	0.0	0.0	0.0	0.0
Konjunktiv (%)	0.0	0.0	0.0	5.26	12.50	12.5
Vergangenheit (%)	0.0	0.0	0.0	0.0	0.0	0.0
Pb nicht angesprochen (%)	0.0	0.0	0.0	0.0	0.0	0.0

Fortsetzung Tabelle 13

	AFS Manifeste Angst	AFS Prüfungs- angst	AFS Schul- unlust	AFS Soziale Erwünsch- theit	HANES 2 Lügen- skala
Itemzahl	15	15	10	10	12
Satzeinheiten (M)	1.60	2.33	1.90	1.10	1.42
einfache Verneinung (%)	26.67	20.00	40.0	50.0	8.33
doppelte Verneinung (%)	0.0	0.0	0.0	0.0	0.0
Passiv (%)	6.67	40.00	10.0	0.0	8.33
Konjunktiv (%)	6.67	0.0	10.0	0.0	8.33
Vergangenheit (%)	0.0	0.0	0.0	30.0	25.0
Pb nicht angesprochen (%)	0.0	0.0	30.0	0.0	0.0

\* Die Angstsskalen der eindimensionalen Fbn KAT und FS 5 – 10 sind identisch mit den Angaben in Tab. 11; zum EPI vgl. Tab. 5 – 7.

vertreten. Dabei verzeichnen die Extraversionskalen auffällig hohe Werte beim Konjunktiv (M = 10.22%), die Ängstlichkeitsskalen beim Passiv (M = 20.22%) und die Lügenskala bei der Variable „Vergangenheit“ (M = 19.44%).

Die gleichfalls der Validitätskontrolle dienende Skala „Soziale Erwünschtheit“ des AFS enthält wie die Lügenskala relativ viele Items (30%), die auf Vergangenes abzielen, und weist zur Hälfte negativ formulierte Items auf, womit die Lügenskala des EPI-A (4 von 9 verneinte Items) noch übertroffen wird. Die Skala „Schulunlust“ des AFS fällt durch 40% einfach verneinte Items auf und enthält als einzige KJ-Skala Items, in denen der Pb nicht angesprochen wird (30%).

#### *Syntaktische Ähnlichkeit der Skalen*

Die Methode zur Berechnung syntaktischer Ähnlichkeit wurde unter Punkt 3.2.3 ausführlich dargestellt.

Aus der Verteilung der Profildifferenzen für insgesamt 172 Skalenpaare geht hervor, daß sich zwei beliebige Skalen der KJ-Fragebogen im Durchschnitt etwas weniger hinsichtlich der syntaktischen Struktur der Items unterscheiden als in den Erwachsenen-Fragebogen:

M = 12.0, SD = 5.3, Min = 0.6, Max = 30.6.

Gruppen gleich benannter Skalen sind in den Bereichen „Neurotizismus“ (Tab. 14) und „Extra-

**Tabelle 14:** Unterschiede im formal-syntaktischen Aufbau von 5 Neurotizismusskalen der KJ-Fragebogen (D-Maße)

	INK(A)	HANES 1	HANES 2	EPI-A
HANES 1	0.60			
HANES 2	2.95	3.33		
EPI-A	6.50	6.80	6.25	
EPI-B	5.24	5.00	4.31	6.25

M = 4.72; SD = 1.86; Min = 0.6; Max = 6.8

**Tabelle 15:** Unterschiede im formal-syntaktischen Aufbau von 5 Extraversionsskalen der KJ-Fragebogen (D-Maße)

	INK(A)	HANES 1 E 1	HANES 1 E 2	EPI-A
HANES 1, E 1	9.32			
HANES 1, E 2	3.07	6.25		
EPI-A	10.01	2.08	8.33	
EPI-B	7.24	4.86	5.56	6.25

M = 6.30; SD = 2.42; Min = 2.08; Max = 10.01

**Tabelle 16:** Unterschiede im formal-syntaktischen Aufbau von 4 Ängstlichkeitsskalen der KJ-Fragebogen (D-Maße)

	FS 5-10	KAT	AFS-MA
KAT	9.21		
AFS, MA	10.90	8.83	
AFS, PA	8.30	11.52	13.33

M = 10.35; SD = 1.75; Min = 8.30; Max = 13.33

**Tabelle 17:** Unterschiede im formal-syntaktischen Aufbau von 3 Lügenskalen der KJ-Fragebogen (D-Maße)

	HANES 2	EPI-A
EPI-A	15.28	
EPI-B	5.09	18.52

M = 12.96; SD = 5.72; Min = 5.09; Max = 18.52

version“ (Tab. 15) homogener als der Durchschnitt und ähneln einander stärker als die entsprechenden Untergruppen der Erwachsenen-Fragebogen. Für Ängstlichkeits- und Lügenskalen (Tab. 16 und 17) läßt sich dagegen kein nennenswerter Unterschied zum Gesamtdurchschnitt und – bei den Lügenskalen – zu den Erwachsenen-Fragebogen verzeichnen.

Aufgrund der allgemein größeren syntaktischen Ähnlichkeit der KJ-Skalen untereinander weisen die EPI-Parallelskalen sowohl im Kontext der jeweiligen Skalengruppen als auch im allgemeinen Bezugsrahmen keine verringerten Differenzwerte auf.

## 5. Diskussion

In der vorliegenden Untersuchung wurde versucht, die sprachliche Formulierung der Items deutschsprachiger Persönlichkeitsfragebogen empirisch zu erfassen.

Zwischen Fragebogen für Erwachsene und Fragebogen für Kinder und Jugendliche bestehen markante Unterschiede in der Formulierung der Items; die Vermutung einer intuitiven Berücksichtigung sprachlicher Formmerkmale bei der Itemformulierung wird damit (mit Ausnahme des Merkmals „Passivgebrauch“) in bezug auf unterschiedlich sprachgewandte Zielgruppen bestätigt. Die Bevorzugung einzelner Formmerkmale bei der Formulierung von Items mit bestimmten inhaltlichen Meßintentionen tritt besonders bei den KJ-Fragebogen hervor: Der häufige Bezug auf Vergangenes in den Items der Lügenskalen könnte so interpretiert werden, daß das Zugeben geringfügiger, jedoch als allgemein verbreitet angenommener menschlicher Schwächen dem Pb erleichtert werden soll, indem es auf einen früheren Zeitpunkt zurückverlagert wird. Man kann dies jedoch auch so interpretieren, daß der Pb quasi auf Fakten festgelegt werden soll, daß ihm also bei der Überprüfung seiner Glaubwürdigkeit ein Ermessensspielraum („Geben Sie an, was *gewöhnlich* für Sie am ehesten zutrifft.“) nicht in dem Umfang zugestanden wird wie bei Inhaltsskalen.

Der gehäufte Passivgebrauch in den Ängstlichkeitsskalen der KJ-Fragebogen könnte darauf hindeuten, daß Passivformulierungen nicht nur unter Verständlichkeitsaspekten zu sehen sind, sondern auch eine inhaltliche Funktion haben können; d. h. daß ein Konzept wie „Ängstlichkeit“ stärker als andere über ein passives Etwas-mit-sich-geschehen-lassen operationalisiert wird.

Die häufigere Verwendung des Konjunktivs in Extraversionsskalen (KJ-Fragebogen) kann als Indikator für stärkere Berücksichtigung von bloß vorgestellten Verhaltens- bzw. Erlebnisweisen oder auch von abgeschwächten Eigenschaftszuschreibungen („Ich würde mich selbst als lebhaft bezeichnen“) bei der Operationalisierung dieses Konzepts aufgefaßt werden. Eine aus dem Konzeptinhalt abgeleitete Erklärung hierfür gibt es allerdings nicht, während bei den Interessenskalen des PIT die Bevorzugung des Konjunktivs aus der spezifischen Meßintention klar hervorgeht. Gleiches gilt auch für die unpersönlichen Formulierungen der Intelligenzitems des 16 PF, deren „richtige“ Beantwortung ja von der Person und Meinung des Pb unabhängig ist.

Die Extraversions- und Lügenskalen der Erwachsenen-Fragebogen heben sich – im Gegensatz zu den KJ-Fragebogen – nicht deutlich in bestimmten Merkmalen von den übrigen Skalen ab. Dies ist vor allem auf die größere Heterogenität formal-syntaktischer Merkmale in den Erwachsenen-Fragebogen und deren Skalen zurückzuführen. Die teilweise beträchtliche Variabilität formal-syntaktischer Merkmale belegt das Fehlen expliziter Richtlinien zur Formulierung von Fragebogenitems – unseres Erachtens eine Folge des geringen Stellenwertes, der der Zusammenstellung des Itempools zu Beginn der Fragebogenkonstruktion beigemessen wird.

Die Vereinheitlichung formal-syntaktischer Merkmale geht nur soweit, wie dies für eine standardisierte Auswertung zweckmäßig ist: Sie umfaßt die Antwortformate und – durch diese bedingt – die Formulierung des Items als Aussage oder Frage. Die Schematisierung der Antwortformate wird nur beim 16 PF durchbrochen (vgl. Tab. 3). Der 16 PF verlangt durch seine

zahlreichen Formulierings- und Beantwortungsvarianten vom Pb immer wieder Umstellungen auf andere Sprachmuster und damit – zumal einige Formulierungen recht kompliziert geraten sind – hohe sprachliche Flexibilität. Der 16 PF ist jedoch dadurch noch am ehesten in der Lage, Ermüdungserscheinungen und stereotypen Reaktionstendenzen (Acquieszenz) bei der Beantwortung entgegenzuwirken; seine Testkennwerte hinsichtlich Reliabilität und Validität wären vom Gesichtspunkt der sprachlichen Formulierung her am ehesten generalisierbar. Die Vermutung scheint plausibel, daß die mehrfach nachgewiesene geringe interne Konsistenz und die faktorielle Validität der 16 PF-Skalen (z. B. GREIF 1970; BARTUSSEK 1974; zusammenfassend ANGLEITNER 1976) neben inhaltlicher Heterogenität und der Ausbalanciertheit der Verschlüsselungsrichtung in den Skalen (ANGLEITNER 1976) den heterogenen sprachlichen Formulierungen der Items zuzuschreiben sind.

Die Variabilität syntaktischer Merkmale wurde auf Skalenebene speziell unter dem Aspekt der Übereinstimmung der Merkmalsprofile analysiert; das verwendete Distanzmaß zeigt neben unterschiedlichen Profilverläufen auch Unterschiede in der Höhe der Profile an (im Unterschied zu korrelativen Maßen). Eine gesonderte Betrachtung von Skalengruppierungen mit gleichen Meßintentionen scheint zunächst für eine Tendenz zu größerer Ähnlichkeit im sprachlichen Aufbau der Items gleichbenannter Skalen bzw. Parallelformskalen zu sprechen.

Bei den Parallelformen des 16 PF und des EPI ist diese Tendenz allerdings im Durchschnitt relativ schwach ausgeprägt; sie könnte durchaus in fragebogenspezifischen stilistischen Präferenzen begründet sein, ohne daß eigentlich inhaltlich begründete Vereinheitlichungstendenzen im Spiel sind. Eine eindeutige inhaltliche Interpretation ist nur bei der Intelligenzskala des 16 PF möglich: Leistungsaufgaben sind durch ihre immanente logische Struktur auch in ihrer Form viel strikter festgelegt als nicht-leistungsbezogene Persönlichkeitsitems.

Was gleichbenannte Skalen aus verschiedenen Fragebogen betrifft, so ist die sprachliche Homogenität im Neurotizismus- und Extraversions-

bereich offenbar ein Artefakt der Itemüberlappung zwischen den Fragebogen (vgl. ANGLEITNER & LÖHR 1980). Allerdings bringt eine Konfundierung durch Itemüberlappung nicht automatisch hohe syntaktische Ähnlichkeit mit sich, wie sich am Beispiel der Lügenskalen zeigen läßt. Die Itemüberlappung kann ebenso für die oben erwähnte Anhäufung bestimmter sprachlicher Merkmale in den Lügen- und Extraversionskalen der KJ-Fragebogen verantwortlich sein; unseres Erachtens schränkt sie die Interpretierbarkeit solcher sprachlicher Präferenzen jedoch nicht ein: Vorausgesetzt, die mehrfach verwendeten Items wären besonders häufig im Konjunktiv bzw. mit Bezug auf die Vergangenheit formuliert, so bliebe offen, ob sie möglicherweise gerade wegen dieser Eigenschaft als geeignete Operationalisierungen eines bestimmten Konzepts angesehen werden.

Um die Auswirkung sprachlicher Übereinstimmungen und Differenzen auf die Beantwortung von Fragebogen einschätzen zu können, bietet sich der von MICKLIN & DURBIN (1969) publizierte Ansatz der Gegenüberstellung von Satzstrukturdifferenzen einerseits und Korrelationen zwischen den Skalenpunktwerten einer Pbn-Stichprobe andererseits an. Erste Versuche in dieser Richtung auf der Basis von Skaleninterkorrelationen aus eigenen Fragebogenuntersuchungen an 102 Erwachsenen und 153 Schulkindern (ANGLEITNER 1976) ergaben für die 16 PF-Parallelskalen keinen Zusammenhang; für die Skalenpaare aus dem Extraversions- und dem Neurotizismusbereich beider Fragebogenbatterien (vgl. Tab. 8 und 9 sowie Tab. 14 und 15) ergab sich in drei Vergleichen der erwartete negative Zusammenhang zwischen sprachlichen Distanzmaßen und Skalenkorrelationen, während im vierten Fall ein positiver Zusammenhang resultierte. Obgleich diese Versuche noch wenig systematisch sind und ungelöste methodische Probleme beinhalten (Abhängigkeiten in beiden zu korrelierenden Matrizen), lassen die Ergebnisse den Ansatz von MICKLIN & DURBIN lohnenswert erscheinen.

Zum Abschluß soll auf einige methodische Probleme der Untersuchung eingegangen werden: Das verwendete System formal-syntakti-

scher Kategorien ist deskriptiv; es deckt – neben Maßen der Itemlänge und einer Kategorisierung der Antwortformate – in seinem syntaktischen Teil lediglich einige Aspekte der Oberflächenstruktur von Sätzen ab. Dieses Vorgehen besitzt den Nachteil, daß die Elemente der Oberflächenstruktur – zumal wenn sie, wie hier, isoliert betrachtet werden – funktional nicht in jedem Fall eindeutig sind. Dagegen sind Systeme auf der Basis von Tiefenstrukturanalysen – neben der generativen Transformationsgrammatik nach CHOMSKY (1969) wären etwa Valenztheorie oder Kasusgrammatik mit stark semantischer Akzentuierung zu nennen; vgl. ENGELKAMP (1974) – besser in der Lage, bestimmte Fokussierungen in der Satzformulierung adäquat zu erfassen. Da sie jedoch die Oberflächenstruktur auf die wesentlichen Komponenten reduzieren, gehen Unterschiede in der Verständlichkeit einzelner Formulierungen weitgehend verloren. Dies trifft auch für Kategoriensysteme zu, die empirisch über Skalierungen der funktionalen Ähnlichkeit von Satzelementen gewonnen werden (ROHRMANN 1978). Neben einer auf Sprache im allgemeinen gerichteten und daher für die Analyse von Itemformulierungen nur bedingt geeigneten Zielsetzung sind funktionale Kategoriensysteme zudem erheblich aufwendiger als deskriptive; sie stellen hohe Anforderungen an den Codierer und erbringen nur mäßige Zuverlässigkeitswerte.

Abgesehen vom Kategoriensystem, sind alternative Methoden der *Datenauswertung* denkbar. Hier käme eine Zusammenstellung von Merkmalskonfigurationen auf Itemebene in Betracht (MICKLIN & DURBIN 1969), die anschließend zur Berechnung von Distanzmaßen benutzt werden könnten.

Weitere Möglichkeiten multivariater Verarbeitung der Daten wären Konfigurationsfrequenzanalysen sowie Clusteranalysen und MDS der Distanzmatrizen.

Die Variabilität formal-syntaktischer Merkmale in Persönlichkeitsfragebogen sowie sprachlicher Unterschiede zwischen Persönlichkeitskalen ist durch diese Untersuchung belegt; ihre psychologische Relevanz konnte in der Interpretation der Ergebnisse an einigen Punkten ange-

deutet werden, bedarf allerdings der Bestätigung durch weitere Validierungsuntersuchungen. Erst im Kontext weiterer Ansätze einer methodenkritischen Fragebogenforschung wird das Verhältnis von Form und Inhalt, von methodischem Artefakt und gültiger Information genauer zu klären sein. Wo nicht einmal die inhaltlichen Beziehungen zwischen postulierten Eigenschaftskonzepten und ihren Operationalisierungen durch Fragebogenskalen bzw. -items geklärt sind, scheinen wir davon noch weit entfernt zu sein.

## Literatur

- ANGLEITNER, A.: Methodische und theoretische Probleme bei Persönlichkeitsfragebogen mit einer ausführlichen Analyse deutschsprachiger Persönlichkeitsfragebogen. Unveröffentlichte Habil., Bonn 1976. (2 Bde).
- ANGLEITNER, A., HAMM, S., LOHMANN, A., LÖHR, F. J.: Formal-syntaktische Merkmale der Items von Persönlichkeitsfragebögen. Berichte aus dem Psychologischen Institut der Universität Bonn, 1978, Nr. 22.
- ANGLEITNER, A. & LÖHR, F. J.: Itemüberlappungen zwischen Persönlichkeitsfragebogen als Problem für Validitätsschätzungen. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 1980, 1, 127–136.
- BARTUSSEK, D.: Mitteilungen über Reliabilität und faktorielle Validität des deutschen 16 PF-Tests von CATTELL. *Diagnostica* 1974, 20, 49–55.
- BECKMANN, D. & RICHTER, H. E.: *Gießen-Test (GT). Ein Test für Individual- und Gruppendiagnostik*. Bern: Huber, 1972.
- BUGGLE, F. & BAUMGÄRTEL, F.: *Hamburger Neurotizismus- und Extraversionsskala für Kinder und Jugendliche*. HANES, K.J. Handanweisung für die Durchführung und Auswertung. Göttingen: Hogrefe, 1972.
- CAMPBELL, D. T. & FISKE, D. W.: Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* 1959, 56, 81–105.
- CATTELL, R. B. & EBER, H. W.: Specimen set for the Sixteen Personality Factor Questionnaire „16 PF“, Experimental edition der deutschen Form A und B. Champaign, Illinois: Institute for Personality and Ability Testing, 1962.
- CHOMSKY, N.: *Aspekte der Syntax-Theorie*. Frankfurt: Suhrkamp, 1969.
- DUBOIS, B. & BURNS, J. A.: An analysis of the meaning of the question mark response-category in attitude scales. *Educational and Psychological Measurement* 1975, 35, 869–884.
- EGGERT, D.: Eysenck-Persönlichkeits-Inventar. E-P-I. Handanweisung für die Durchführung und Auswertung. Göttingen: Hogrefe, 1974.
- ENGELKAMP, J.: *Psycholinguistik*. München: Fink, 1974.
- EYSENCK, H. J.: Das „Maudsley Personality Inventory“ (MPI). Göttingen: Hogrefe, 1959.
- EYSENCK, H. J.: Maudsley-Persönlichkeitsfragebogen. 2. verbesserte Auflage, Göttingen: Hogrefe, 1964.
- FAHRENBERG, J. & SELG, H.: Das Freiburger Persönlichkeitsinventar, FPI-Handanweisung. Göttingen: Hogrefe, 1970.
- FAHRENBERG, J., SELG, H., HAMPPEL, R.: Das Freiburger Persönlichkeitsinventar, FPI-Handanweisung, 2. stark erweiterte Auflage. Göttingen: Hogrefe, 1973.
- GARTNER-HARNACH, V.: Fragebogen für Schüler FS 5–10. In: INGENKAMP, K. H. (Ed.): *Deutsche Schultests*. Weinheim: Beltz, 1973.
- GREIF, S.: Untersuchungen zur deutschen Übersetzung des 16 PF Fragebogens. *Psychologische Beiträge* 1970, 12, 186–213.
- HÖRMANN, H.: *Psychologie der Sprache*. Heidelberg: Springer, 1967.
- JONES, R. R.: Differences in response consistency and subjects preferences for three personality inventory formats. Reprinted from the Proceedings, 76th Annual Convention, American Psychological Association, 1968.
- MICKLIN, M. & DURBIN, M.: Syntactic dimensions of attitude scaling techniques: sources of variation and bias. *Sociometry* 1969, 32, 194–206.
- MITTENECKER, E. & TOMAN, W.: Der Persönlichkeits-Interessen-Test. *Beihefte zur Wiener Zeitschrift für Philosophie, Psychologie, Pädagogik* 1, Wien 1951.
- NISCHAN, C.: Der INK – Ein Persönlichkeitsfragebogen zur Erfassung von Introversion und Neurotizismus bei Kindern und Jugendlichen von 9–16 Jahren. In: HELM, J., KASIELKE, E., MEHL, J., (Eds.): *Neurosendiagnostik. Beiträge zur Entwicklung klinisch-psychologischer Methoden*. Berlin: VEB Deutscher Verlag der Wissenschaften, 1974, 228–250.
- PETERSON, C. C. & PETERSON, J. L.: Linguistic determinants of the difficulty of true-false test items. *Educational and Psychological Measurement* 1976, 36, 161–169.
- ROHRMANN, B.: Statistische Untersuchungen syntaktischer Textcharakteristika. *Deutsche Sprache* 1978, 6, 110–142.
- SPREEN, O.: MMPI-Saarbrücken. Handbuch zur deutschen Ausgabe des Minnesota Multiphasic Personality Inventory. Bern: Huber, 1963.

- SCHLOSSER, O.: *Einführung in die sozialwissenschaftliche Zusammenhangsanalyse*. Reinbek: Rowohlt, 1976.
- THURNER, F. & TEWES, U.: Der Kinder-Angst-Test (KAT). Ein Fragebogen zur Erfassung des Ängstlichkeitsgrades von Kindern ab 9 Jahren. Göttingen: Hogrefe, 1969.
- WIECZERKOWSKI, W., NICKEL, H., JANOWSKI, A., FITTKAU, B., RAUER, W.: Angstfragebogen für Schüler (AFS). Braunschweig: Westermann, 1974.
- WIGGINS, J. S. & GOLDBERG, L. R.: Interrelationship among MMPI item characteristics. *Educational and Psychological Measurement* 1965, 25, 381 – 397.
- WOODWORTH, R. S.: *Personal Data Sheet*. Chicago: Stoelting, 1917.

Prof. Dr. A. Angleitner  
Dipl. Psych. F. J. Löhr  
Universität Bielefeld  
Fakultät für Psychologie und Sportwissenschaft  
Abteilung Psychologie  
Postfach 8640  
D-4800 Bielefeld 1