

- HORNKE, L., Verfahren zur Mitteilung von Korrelationen. Psychol. Beiträge, 1973, 15, 87-105.
- JÄGER, R., Biographische Daten und Differentielle Diagnostizierbarkeit. Drei Studien zur Verwendung von Moderatoren und biographischen Daten im Zusammenhang mit Leistungskriterien in Schule und Hochschule. Dissertation (unveröffl.) Mannheim 1973.
- JÄGER, R., Korrigierte Schätzung des Korrelationskoeffizienten nach OLKIN & PRATT für $n = 8$ bis 40. Biometr. Z. 1974, 16, 115-124.
- JÄGER, R., Moderatoransatz als vereinheitlichendes Prinzip in der psychologischen Forschung. Möglichkeiten und Grenzen. Arch. Psychol. 1974a, 126, 97-113.
- JÄGER, R., Probleme der differentiellen Diagnostizierbarkeit. In: TACK, H. (Ed.): Bericht vom 29. Kongreß der DGfP 1974b (im Druck).
- JÄGER, R., Transformationen bei Zensuren. Eine kritische Stellungnahme. 1974c (zur Veröffl. vorbereitet).
- JÄGER, R., Methoden zur Mittelung von Korrelationen. Psychol. Beitr. 1974d, 16, 417-427.
- JÄGER, R., Ähnlichkeit und Konsequenzen von Suppression und Multicollinearität. Psychol. Beitr. 1974e (im Druck).
- JÄGER, R., Bemerkungen zu: „Die Standardskala der kritischen Differenz“, Diagnostica 1974f, 4, 165-168.
- KLEITER, E., Über Theorie und Modell kategorialer Fehler des Lehrerurteils. Psychol. Beitr. 1973, 15, 185-229.
- LIENERT, G. A., Über die Anwendung von Variablen-Transformationen in der Psychologie. Biometr. Z. 1962, 4, 145-181.
- MAASSEN, S., Das ideale Korrelationsmaß. Technische Mitteilungen. Hüttenwerke Rheinhausen, 1957, 319ff.
- OLKIN, I. & W. PRATT, Unbiased estimation of certain correlation coefficients. An. mathem. Stat. 1958, 29, 201-211.
- OLKIN, I., Correlation analysis: two or more bivariate normal distribution. In: STANLEY, J. (Ed.): Improving experimental designs and statistical analysis. Chicago, 1967.
- SACHS, L., Statistische Auswertungsmethoden. Berlin, Heidelberg, N.Y., 1973³.
- WEBER, E., Biometrische Bearbeitung multipler Regressionen unter besonderer Berücksichtigung der Auswahl der Transformation und der Linearkombinationen von Variablen. Statistische Hefte, 1968, 3, 228-251.
- WINER, J. B., Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Anschrift des Verfassers:

Dr. REINHOLD JÄGER, Otto-Selz-Institut für Psychologie und Erziehungswissenschaft der Universität Mannheim (WH), 68 Mannheim 1, Schloß

Sammelreferat

Die Anwendung der Faktorenanalyse. Darstellung und Kritik der Praxis einer Methode

Von

Helmut Lukesch und Gernot D. Kleiter

(Angenommen am 8. Oktober 1974)

Résumés

Es werden an Hand einer Stichprobe von fünf deutschsprachigen Zeitschriften die Anwendungen der Faktorenanalyse (FA) im Zeitraum zwischen 1953 und 1972 untersucht. Zur Beurteilung der Arbeiten werden dabei Kriterien herangezogen, die auf empirische Arbeiten im allgemeinen angewandt werden können und solche, welche für die FA allein zutreffend sind. Aufgrund dieser Kritik wird die Frage bearbeitet, warum faktorenanalytischen Untersuchungen nicht die theoretische Bedeutsamkeit zukommt, die man an Hand der zahlreichen Veröffentlichungen, in denen diese Methode verwendet wird, vermuten könnte.

The application of factor analysis to empirical data as shown in five German journals between 1953 and 1972 is investigated. For judging the articles criteria were used which are valid for empirical studies in general and others which are appropriate for factoranalytical studies alone. By a lot of precarious remarks the question was answered, why factoranalytical investigations do not possess the theoretical significance, which could be assumed by the lot of publications which make use of this method.

1. Einleitung

1.1. Zielsetzung

Grundsätzliche Arbeiten zur Faktorenanalyse (FA) in der Psychologie behandeln diese Methode meist unter anspruchsvollen theoretischen Gesichtspunkten. Die Beiträge befassen sich insbesondere mit (1) der Erweiterung und Verallgemeinerung der Ausgangssituation, auf welche die FA anwendbar ist, (2) mit der Bereitstellung neuer Lösungsmöglichkeiten und (3) mit der Untersuchung von Eigenschaften der Algorithmen selbst. Beispiele hierfür sind Erweiterungen auf mehrdimensionale Datenmatrizen (BARTUSSEK, 1972;

FISCHER, 1967; TUCKER, 1964, 1966), Rotationsverfahren zur Bestimmung der Ähnlichkeit von Faktoren aus verschiedenen Analysen (WITTE, 1972; GEBHARDT, 1967; FISCHER und ROPPERT, 1965), die Propagierung wenig bekannter Rotationstechniken (PIAGGIO, 1972a, b) und Arbeiten zur Konvergenz von Iterationstechniken oder zum Einfluß der Rechengenauigkeit. Problem-Input, Problem-Output und Eigenschaften der Verarbeitungsalgorithmen liefern auch Ansätze zu grundsätzlicher Rahmenkritik in Bezug auf die Adäquatheit und Sinnhaftigkeit der Verwendung der FA zur Bearbeitung psychologischer Fragestellungen (KEMPF, 1972; KALVERAM, 1970 a, b, c, d; Münchner Symposion über Faktorenanalyse, 1967).

In Ergänzung zu diesen weitgehend theoretischen Untersuchungen soll in den nachfolgenden Ausführungen versucht werden, eine Darstellung der tatsächlichen Anwendung der FA zu geben. Die Bestandsaufnahme wird das faktorenanalytische Datenmaterial und seine Gewinnung darstellen, die gebräuchlichen Analyse- und Verrechnungstechniken umreißen, auf Kompromisse hinweisen, die bei der mehr oder weniger konsequenten Überprüfung und Berücksichtigung von Adäquatheitsbedingungen eingegangen werden und schließlich auch versuchen, einen Überblick zu den Fragestellungen zu geben, die mit der FA bearbeitet werden. Es ist Ziel dieses Beitrages, aus der Anwendung der FA eine allgemeine Kritik abzuleiten, mit der die grundsätzliche methodische Kritik bezüglich der mathematischen Grundlagen und theoretischen Tragweite der FA abgerundet werden kann.

1.2. Vorgehen

Die Untersuchung beschränkt sich auf die deutsche Zeitschriftenliteratur. Von fünf Zeitschriften wurden die Jahrgänge zwischen 1953 und 1972 auf solche Artikel durchgesehen, in denen die FA theoretisch abgehandelt oder auf Daten angewandt wird. Neben drei vorwiegend „experimentell“ orientierten Zeitschriften (Zeitschrift für experimentelle und angewandte Psychologie, Diagnostica, Psychologische Rundschau) wurden auch zwei „klinische“ Zeitschriften (Zeitschrift für Klinische Psychologie und Psychotherapie, vorm. Jahrbuch für Psychologie und Psychotherapie; Praxis der Kinderpsychologie und Kinderpsychiatrie) in die Stichprobe aufgenommen¹. Die Auswahl war dabei nicht unbedingt von Gesichtspunkten der Repräsentativität geleitet, sondern es wurden solche Zeitschriften ausgewählt, die leicht zugänglich waren. Die „Psychologischen Beiträge“ wurden nicht mit aufgenommen, da diese Zeitschrift erst ab 1967 in Jahrgangsbänden erscheint. Insgesamt wurden 136 „faktorenanalytische“ Zeitschriftenartikel bearbeitet.

¹ Für die Durchsicht eines Teiles der genannten Zeitschriften sei an dieser Stelle Herrn stud. phil. FRANZ MATER herzlich gedankt.

Die Kriterien, nach denen die einzelnen Arbeiten besprochen werden, ergeben sich zum einen aus den allgemeinen Grundsätzen empirischen Arbeitens, etwa der Beurteilung von Daten- und Probandenstichproben, zum anderen aus Prinzipien der FA selbst (MULAIK, 1972; ÜBERLA, 1968, S. 355 ff.; PAWLIK, 1968, S. 264 ff.; CATTELL, 1966; EYSENCK, 1953; GUILFORD, 1952).

1.3. Die Entwicklung in der deutschsprachigen Literatur

Während der ersten fünf Jahre des untersuchten Zeitraumes wird von der FA in deutschsprachigen Zeitschriften von ausländischen Autoren (CATTELL und SANDERS, 1954; EYSENCK, 1954; EKMAN, 1954) oder nur in Ausnahmefällen Gebrauch gemacht (WEBER, 1953; HOFSTÄTTER, 1956; ROTH, 1957). Auch in den nächsten fünf Jahren gehört die Anwendung der FA zu den Ausnahmefällen möglicher statistischer Bearbeitung empirischen Datenmaterials. Im Jahre 1964 sind in den untersuchten Zeitschriften zum ersten Mal zehn solche Artikel zu finden (vgl. Tab. 1). Die Beliebtheit der FA —

Tabelle 1. Überblick über die Anzahl der Zeitschriftenartikel, in denen die Faktorenanalyse methodisch angewandt oder theoretisch abgehandelt wurde

Zeitschrift	ZEAP ¹	DIAG ²	PR ³	PKK ⁴	JPP ⁵	Summe
Jahr						
1953	1	0	0	0	0	1
1954	3	0	0	0	0	3
1955	0	0	0	0	0	0
1956	1	0	0	0	0	1
1957	1	0	0	0	0	1
1958	2	1	0	0	1	4
1959	2	2	0	0	0	4
1960	0	1	0	0	0	1
1961	4	0	1	0	0	5
1962	1	1	1	0	0	3
1963	5	2	0	0	0	7
1964	7	3	0	0	0	10
1965	7	2	0	0	0	9
1966	4	1	0	2	0	7
1967	5	0	0	0	0	5
1968	8	7	1	1	1	18
1969	6	8	3	0	0	17
1970	4	5	1	1	0	11
1971	5	5	1	0	0	11
1972	9	5	3	1	0	18
Summe	75	43	11	5	2	136

¹ Zeitschrift für experimentelle und angewandte Psychologie

² Diagnostica

³ Psychologische Rundschau

⁴ Praxis der Kinderpsychologie und Kinderpsychiatrie

⁵ Jahrbuch für Psychologie und Psychotherapie (1953—1971) und

Zeitschrift für Klinische Psychologie und Psychotherapie (1971—1972)

gemessen an der Häufigkeit ihrer Verwendung — hat sich in den letzten Jahren aber auf einem relativ konstantem Niveau eingependelt. Ab 1968 erscheinen jährlich mindestens elf solcher Arbeiten in den untersuchten Zeitschriften. In den Jahren 1968 und 1972 findet man mit jeweils 18 Publikationen Höhepunkte.

In den 50er Jahren fehlte es offensichtlich an leicht zugänglicher deutschsprachiger Literatur (LIENERT, 1959; HOFSTÄTTER, 1938, 1953), an leicht zugänglichen Computern und — Psychologen. Die hinlänglich bekannte Situation führte zu einem „cultural-lag“ nicht nur im Bereich dieser Methode. Die leicht lesbaren Lehrbücher im anglo-amerikanischen Raum (z. B. THURSTONE, 1947; FRUCHTER, 1954) oder das Standardwerk von HARMAN (1960) fanden erst ihr Pendant in den Publikationen von ÜBERLA (1968) und PAWLIK (1968). Es ist in gewissem Sinne bezeichnend, daß ausgerechnet in den Fachkreisen, in denen an der Schließung der Lücke am intensivsten gearbeitet wurde — und dies dürfte bis zur Erstellung der Computerprogramme gereicht haben — die forcierteste Kritik an der Methode formuliert wurde (Münchener Symposion über Faktorenanalyse, 1967). Ebenso bezeichnend ist es — unabhängig von jeder Bewertung — daß eben diese Kritik die wachsende Popularität der FA nicht einschränken konnte; die Durchführung einer FA ist durch den Aufbau von Rechenzentren oder fachbereichseigener Anlagen und mit der Verfügbarkeit an software zu einer Bagatelle geworden. Als ein Facit ergibt sich: Es gibt nun im deutschen Sprachraum auch die praktischen Beispiele, auf welche die Kritik anwendbar ist, in genügend großer Zahl.

Neben dem zeitlichen Trend fällt der Unterschied in der Häufigkeit von faktorenanalytischen Beiträgen zwischen den „experimentell“ und den „klinisch“² ausgerichteten Zeitschriften ins Auge. In der Praxis wissenschaftlichen Forschens haben sich hier offenbar Rituale ausgebildet, die sich in der Bevorzugung bestimmter Auswertungstechniken dokumentieren. Dies kann wohl kaum mit dem Gegenstandsbereich selbst zusammenhängen, da die Verhältnisse bei amerikanischen Zeitschriften genau umgekehrt sind: Im „Journal of Clinical Psychology“ erschienen 1971 insgesamt 19 Arbeiten, in denen die FA wesentlich angewandt wurde, 1972 waren es 18 Arbeiten; im „Journal of Experimental Psychology“ findet man in diesem Zeitraum überhaupt keine FA! Bei aller Unterschiedlichkeit in den spezifischen Eigenheiten dieser Zeitschriften — hier sind methodische Akzente gegeben, die im deutschen und im anglo-amerikanischen Raum genau umgekehrt liegen. Wenn es sich hier um eine Art „Phasenverschiebung“ handelt, dann könnte es sein, daß die klinischen Psychologen im deutschsprachigen Raum die FA doch noch

² Die zur Abgrenzung verwendete Nomenklatur sollte nicht dahingehend interpretiert werden, daß die Einteilung theoretisch begründbar, konsistent durchführbar oder etwa praktisch fruchtbar sei.

lernen müssen, während die Experimentalpsychologen sie schon bald wieder vergessen könnten.

2. Die Beurteilung des Datenmaterials

Die legitime Verwendung statistischer Methoden ist an die Erfüllung von einschränkenden Bedingungen gebunden. Das faktische Datenmaterial muß nämlich ein Modell der entsprechenden mathematischen Theorie sein. Komplexe statistische Methoden machen in der Regel strengere Voraussetzungen und Annahmen notwendig. „In fast allen Fällen haben wir es mit einer Superposition von statistischen Hypothesen verschiedener Allgemeinheitsstufen zu tun“ (STEGMÜLLER, 1973, S. 74). Es ist nicht immer klar, welche Annahmen in der Kette der „Oberhypothesen“ (STEGMÜLLER, 1973) stillschweigend als gültig vorausgesetzt werden können und welche zu prüfen sind und mit welchem Grad der Strenge. Wesentlich ist wohl, daß die Interpretation invariant gegenüber möglichen Verletzungen von ungeprüften Oberhypothesen oder gegenüber nachweislich verletzten Oberhypothesen sein muß. Dieses Eindeutigkeitskriterium der Interpretation kann etwa durch die Robustheit des entsprechenden Verfahrens gewährleistet sein oder durch einen allerdings schwer kontrollierbaren Rückzug auf schwächere Interpretationsansprüche mit geringerem empirischen Gehalt versucht werden. „The use of a complicated statistical procedure like factor analysis does not permit one to forget about the usual safeguards that should surround scientific observations. Statistical operations do not compensate for carelessness in making observations. Rather, they presuppose careful observations.“ (GUILFORD, 1952, S. 26).

2.1. Merkmalsstichprobe

Die FA wird u. a. mit dem Ziel eingesetzt, eine schwer überschaubare Menge von Variablen zu strukturieren. Die Strukturierung erfolgt aufgrund der Spezifikationsgleichung und führt — je nach den verwendeten Zwischenschritt-kriterien — zu einer in der Regel deutlich reduzierten Zahl „neuer“ Variablen. Die „neuen“ Variablen oder Faktoren sind unabhängig voneinander oder stehen in bekannter Beziehung zueinander.

Drei Dinge gilt es bei dieser Zielsetzung zu klären: (1) Zunächst wurde in obiger Formulierung nur ein Anspruch auf eine kognitive Erleichterung der Informationsverarbeitung seitens des Wissenschaftlers geltend gemacht; es bleibt aber eine jeweils zu überprüfende Frage, ob sich die Strukturierung nur auf diesen kognitiven Aspekt bezieht oder ob der Anspruch erhoben wird, grundlegende Strukturen des Sachbereichs aufzufinden. (2) In Analogie zu varianzanalytischen Modellen mit fixen Effekten und mit Zufallseffekten lassen sich zwei weitere Fälle unterscheiden: Der Variablensatz kann alle theoretisch relevanten Variablen umfassen oder lediglich eine Auswahl davon. Es stellt sich also die Frage: Soll die Strukturierung nur in Bezug auf die tatsächlich

vorliegenden Variablen vorgenommen werden oder soll darüber hinaus auf ein Variablenuniversum generalisiert werden? „... one may question whether the principal aim of factor analysis is to make inferences about the factors in a universe of tests. Many factor-analytic studies seem concerned not with discovering all the latent factors in a domain of tests but rather with isolating just those latent factors which are common to certain distinct tests“ (MULAIK, 1972, S. 216). (3) Bei einem „common-factor“-Modell ist die Zahl der Faktoren größer als die des ursprünglichen Variablensatzes (KALVERAM, 1970, S. 112f.); bei r erhaltenen Faktoren und bei p ursprünglichen Variablen werden in der Spezifikationsgleichung $p + r$ Faktoren angenommen. Zu einer kognitiven Erleichterung oder zu einer Vereinfachung im Sachbereich kann es also nur kommen, wenn man die spezifischen Faktoren als irrelevant ansieht. Man ist dann nicht an allen Faktoren interessiert, sondern eben nur an den gemeinsamen. Darüber hinaus weist CATTELL darauf hin (1966, S. 201), daß auch die Zahl der gemeinsamen Faktoren durchaus größer sein kann als die Zahl der untersuchten Variablen: Wenn in einem 50 Variablen-Problem 12 Faktoren wirksam sind, dann sind bei einer Reduktion auf 10 Variable noch immer alle 12 Faktoren wirksam.

Je nach der Beantwortung der ersten beiden Fragen ergeben sich vier mögliche Positionen. Den generellsten Anspruch erhebt dabei jene Position, die Strukturen des Sachbereiches finden und diese generalisieren will. Sie kommt etwa in der heroischen Intention CATTELLS (1950) zum Ausdruck, zu Beginn der Bildung einer Persönlichkeitstheorie alle möglichen Variablen des Merkmalbereiches zu verwenden. Dies ist unmöglich, da grundsätzlich unendlich viele Variablen an einem Gegenstand feststellbar sind. So etwas wie eine Zufallsauswahl scheint aber auch unmöglich, werden doch ganz „unsinnige“ Variablen nicht mit in das Ausgangsmaterial aufgenommen, sondern eher doch solche, von denen man mehr oder weniger begründet erwartet, daß sie sich sinnvoll in den „Struktur-Zusammenhang“ mit anderen Variablen einordnen lassen. KALLINA (1967, S. 84), der auf diesen Auswahlgesichtspunkt hinweist, gibt das Beispiel, daß wohl kaum die Länge zwischen linkem Schulterblatt und rechtem Knie zur Ermittlung körperlicher Konstitutionsfaktoren Berücksichtigung finden dürfte. Gegen eine Zufallsauswahl spricht zudem die fragliche Invarianz faktorenanalytischer Ergebnisse; auf diesen Punkt wird noch weiter unten einzugehen sein.

Die Zusammenstellung des Ausgangsmaterials für eine FA kann also nicht hypothesenfrei vorgenommen werden. Werden diese Hypothesen nicht explizit genannt, so sind Ergebnisse von FAs manipulierbar, da für die Wahl der Ausgangsvariablen keine spezifischen Restriktionen vorliegen. CATTELL (1966, S. 231) empfiehlt daher wesentlich die Einschließung von Markierungsvariablen. „... the factor analyst who plans skillfully will always carefully choose a minimum of two good marker variables for each common factor relevant to his research . . . far too many factor analyses — roughly two-

thirds of all those published between 1950 and 1960 — are practically lost to science because they operate in ‘private universes’ . . . they have no mean, generally, of tying the new construction intelligently into the fabric of existing scientific knowledge.“ Wegen der weiten Streuung der untersuchten Gegenstandsbereiche und wegen der nicht standardisierten sprachlichen Diktion bei der theoretischen Begründung für diese oder jene Variablenauswahl ist es nicht möglich, die Frage nach der Repräsentativität der Variablenstichproben durch einfache Auszählungen zu beantworten. Allerdings sei darauf hingewiesen, daß die Zahl der Variablen, die bei den untersuchten Arbeiten verwendet wird, außerordentlich stark schwankt und zwischen 4 (SEYDEL, 1972) und 101 (SEITZ und BRÄTTH, 1970) liegt.

2.1.1. Skalenniveau der Daten

Die Zulässigkeit eines statistischen Verfahrens scheint u.a. vom Skalenniveau der verrechneten Daten abzuhängen. Wörtlich genommen ist diese Formulierung allerdings unrichtig; so können etwa ohne weiteres Varianzanalysen mit Ordinaldaten durchgeführt werden und Intervallskalenniveau garantiert nicht unbedingt die Invarianz entsprechender Ergebnisse von parametrischen Verfahren (ANDERSON, 1961). Ausschlaggebend ist vielmehr die Bedeutung der empirischen Aussagen; sie muß unter den das Skalenniveau definierenden zulässigen Transformationen invariant bleiben. BREDEKAMP vermerkt zu Recht (1972, S. 123), „... daß die Anwendbarkeit eines statistischen Verfahrens nicht vom Skalenniveau, sondern vom Zutreffen statistischer Annahmen abhängt“.

Trotz gewisser Spitzfindigkeiten bei der Zuordnung von Skalenniveau und statistischer Verfahren kann vom pragmatischen Gesichtspunkt dieser Arbeit von folgender Feststellung ausgegangen werden: Es ist nur dann sinnvoll, eine FA durchzuführen, wenn zumindest angenommen werden kann, daß die einzelnen untersuchten Relationssysteme einen Ordnungsgrad aufweisen, der auf einer Intervallskala abbildbar ist. Faßt man die einzelnen empirischen Relationssysteme zusammengenommen als ein multivariates auf, so wäre entsprechend der Ordnungsgrad des Gesamtsystems zu untersuchen. Die eben getroffene Feststellung kann nun nachweislich erfüllt sein, sie kann aber auch ungeprüft sein und dann besteht die Möglichkeit, daß sie nicht erfüllt ist. Auch dann, wenn die Annahme, daß der Ordnungsgrad des gesamten empirischen Relationssystems dem einer Intervallskala entspricht, vielleicht zutrifft, kann aus praktischen und meßtechnischen Gründen der Fall eintreten, daß nur die Ordinalinformation des Relationssystems ausgeschöpft wird.

Die Frage, welches Skalenniveau im einzelnen Fall vorliegt, ist jedoch fast nie entscheidbar, da den verwendeten Verfahren keine Meßtheorie zugrunde liegt. „Die Frage, ob die Werte von Persönlichkeits- oder Intelligenztests, ob ratings usw. auf einer Intervallskala liegen, kann nicht einfach mit ‚ja‘ beantwortet werden, denn die Transformationseigenschaften dieser Skalen

sind unbekannt“ (BREDEKAMP, 1972, S. 132). Ähnlich äußern sich KRANTZ *et al.* (1971, S. 33): „... in scaling aptitude, intelligence, or social attitudes, test scores or numerical ratings are usually interpreted as measures of the attribute in question. But in the absence of a well-defined homomorphism between an empirical and a numerical relational structure, it is far from clear how to interpret such numbers.“

Obwohl die weitaus überwiegende Zahl der Fälle hinsichtlich des Skalenniveaus nicht eindeutig beurteilt werden können, wurden von uns versucht, das faktorenanalytische Ausgangsmaterial in fünf Gruppen einzuteilen³. Daß die entsprechende Einteilung angreifbar ist, ist uns bewußt.

(1) Zu einer ersten Gruppe lassen sich solche Arbeiten zusammenfassen, in denen Alternativdaten (zumeist Testitems) als Variablen verwendet werden (SCHNEIDER und MINKMAR, 1972; HEINERTH, 1972; BOTTENBERG und WEHNER, 1972; WENDELER, 1971; SEITZ und BRÄTH, 1970; LÜER *et al.*, 1970; FISCH und SCHMALT, 1970; WIENDIECK und LÜCK, 1969; WENDELER, 1969; BUGGLE *et al.*, 1968; LÜER *et al.*, 1966; HARTMANN und ENGELMANN, 1966; MEYER und GOLLE, 1966; FAHRENBERG, 1965; HEDINGER, 1965). Die faktorenanalytische Untersuchung von Tests auf Itemebene ist erst seit kürzerer Zeit gebräuchlich, macht aber bereits 11% aller herangezogenen Arbeiten aus. Insofern es das Ziel solcher Ansätze ist, homogene Subskalen zu gewinnen, kann auf die Kritik verwiesen werden, die von den Vertretern der Rasch-Analyse vorgebracht wurde (FRICKE, 1972; FISCHER, 1968). Technisch ergebnisich gewisse Schwierigkeiten aus dem Umstand, daß die verwendeten Vierfelderkoeffizienten bei nicht-symmetrischen Randbesetzungen – und diese liegen insbesondere bei Items mit stark differierenden Schwierigkeitsindices vor – nicht auf die Grenzen -1 und $+1$ normiert sind (GUILFORD, 1965). In einzelnen vergleichenden Analysen wurde allerdings versucht, die Robustheit der FA gegenüber der Verwendung von Korrelationsmatrizen, die auf Alternativdaten basieren, nachzuweisen (ÜBERLA, 1971, S. 284; ANKER *et al.*, 1963).

(2) In einer zweiten Gruppe von Anwendungsfällen lassen sich solche Arbeiten zusammenfassen, in denen von einzelnen Subtests oder von ganzen Testbatterien ausgegangen wird (SEYDEL, 1972; HOBI und KLÄR, 1972; JOHN und KEIL, 1972; STEINGRÜBER, 1971; SEITZ, 1971; KRÜSKEMPER und KRÜSKEMPER, 1971; CATTELL *et al.*, 1971; ALTHOFF, 1971; WENDELER, 1970; SIDDIQUI und KEIL, 1970; SASSENSCHIEDT und BUGGLE, 1970; STEINHAGEN, 1970; KNABE, 1969; SEITZ und LÖSER, 1969; MIKUS, 1969; WEHNER und BOTTENBERG, 1969; FÜRNRATT, 1969c; TIMM, 1968; BALZERT, 1968; FÜRNRATT, 1968a; BEYME und FAHRENBERG, 1968; BÄUMLER *et al.*, 1968; COHEN und WITTEMANN, 1967; SELG und LISCHKE, 1966; HERRMANN und STAPF, 1966;

³ Einige Artikel werden bisweilen doppelt angeführt; dies geschieht dann, wenn mehrere FAs mit verschiedenen Datensätzen durchgeführt wurden.

KATZENBERGER, 1965; SCHEIER und CATTELL, 1965; HUNZIKER, 1964; JÄGER und TODT, 1964; JÄGER, 1964; BÄUMLER, 1964; LIENERT, 1963; LIENERT und FABER, 1963; EYFERTH, 1963; KEREKJARTO und SCHMIDT, 1962; KRAAK, 1961; SCHAEDELI, 1961; FISCHER und BUTSCH, 1961; RIEGEL, 1960; SEIFERT, 1959; LIENERT und LIENHÖFT, 1959; FISCHER, 1958; LIENERT, 1958a, b; ROTH, 1957; WEBER, 1953). Meist steht hier die Frage einer faktoriellen Validierung im Vordergrund: Laden weniger gut untersuchte und gut untersuchte Markierungsvariable gemeinsame Faktoren? In diesen Fällen, knapp 37% aller untersuchter Arbeiten, liegen Daten wohl vorwiegend auf Intervallskalenniveau vor.

(3) Bei 20,5% der bearbeiteten Publikationen wurden die Ausgangsdaten über semantische Differentiale gewonnen. Die für r Vpn für p Objekte auf q Skalen gewonnenen Rohdaten lassen sich zunächst als ein Kubus darstellen. Mit der von TUCKER (1966, 1964) entwickelten dreidimensionalen FA lassen sich die Daten unreduziert in dieser Form bearbeiten und man erhält Versuchspersonen-Faktoren, Objekt-Faktoren und Skalen-Faktoren sowie eine „Core-Matrix“ für die Beziehungen zwischen diesen drei Faktor-Typen. Zumeist werden die Rohdaten jedoch zweidimensional einer konventionellen FA zugeführt. Dabei kann entweder (a) die dritte Dimension „herausgemittelt“ werden oder (b) die dritte Dimension kann gleichsam auseinandergefaltet werden. Insgesamt resultieren nicht weniger als 10 mögliche Fälle für FAs bei semantischen Differentialen.

Bei der folgenden Aufzählung werden mit N die Zahl der Datenpaare bezeichnet, aufgrund derer die Korrelationen berechnet werden. An der ersten Stelle bei der untenstehenden Liste der Möglichkeiten steht jeweils der Bereich, für den die Faktoren gewonnen werden sollen.

- (a) Dreidimensionale Tucker-Analysen
- (b) Skalen \times Objekte, $N = p$
- (c) Objekte \times Skalen, $N = q$
- (d) Skalen \times Vpn, $N = r$
- (e) Vpn \times Skalen, $N = q$
- (f) Objekte \times Vpn, $N = r$
- (g) Vpn \times Objekte, $N = p$
- (h) Skalen \times Objekte \times Vpn, $N = p \times r$
- (i) Objekte \times Skalen \times Vpn, $N = q \times r$
- (j) Vpn \times Skalen \times Objekte, $N = q \times p$.

Damit sind die Möglichkeiten aber noch nicht erschöpft. In einer Arbeit (MIKUS, 1969) wurden Korrelationsmatrizen für jede Vp berechnet und aus der Vereinigung dieser über gemittelte FISHER'sche z -Werte eine FA gerechnet. Dieses Vorgehen scheint erstmals von MIRON und OSGOOD (1966) praktiziert worden zu sein: „... scale correlation matrices are obtained for each of the concept replications on each subject and then averaged over the N subjects“

(1966, S. 793). Ein weiterer Extremfall liegt vor, wenn pro Individuum eine FA vorgenommen wird. So wurde von WARE verfahren (1958, zit. nach VANNOY, 1965), und zwar verwendete er den durch den ersten Faktor aufgeklärten Varianzanteil als Indikator für die kognitive Komplexität der einzelnen Vpn. Eine andere Variante dieser Methode wurde von WARE 1963 angewandt (zit. nach MIRON und OSGOOD, 1966, S. 793): Zunächst wurde für jede Vp eine Korrelationsmatrix berechnet, anschließend wurde für jeweils zwei Vpn die Ähnlichkeit ihrer Korrelationsmatrizen bestimmt. Dazu wurden die Querproduktschritte der Korrelationskoeffizienten in den entsprechenden Zeilen gebildet. Die resultierende $V_{pn} \times V_{pn}$ -Matrix wurde dann einer FA unterzogen. Die Berechnung von FAs für jeweils einzelne Vpn konnte in den untersuchten Publikationen ebenfalls einmal gefunden werden (ERTEL, 1965b).

Sonderfälle der Möglichkeiten (d) und (e) ergeben sich auch dann, wenn nur ein Objekt beurteilt werden soll, d. h. daß dann in diesem Fall diese dritte Dimension nicht herausgemittelt werden muß. Dies ist z. B. bei Arbeiten über Selbstbeurteilung mittels Rating-Skalen der Fall (JOHN und KEIL, 1972; TIMM, 1971). Ob damit aber alle Möglichkeiten erschöpft sind, bleibt noch eine offene Frage. Wer weiß, welche Varianten an dem Kubus noch zu entdecken sind!

Die von OSGOOD ursprünglich angewandte Technik entspricht dem Fall (h) (TANAKA *et al.*, 1963; SUCI, 1960; KUMATA und SCHRAMM, 1956; OSGOOD und SUCI, 1955). MIRON und OSGOOD (1966, S. 792) bezeichnen dieses Verfahren als „stringing out“. Die Verfahren (b) bis (g) werden von MIRON und OSGOOD unter „summation strategy“ angeführt (1966, S. 793). Der Fall (b) findet sich in der amerikanischen Literatur ziemlich häufig (DI VESTA, 1966; OSGOOD, 1964; ELLIOTT und TANNENBAUM, 1963; MIRON, 1961). Es besteht hier allerdings die Gefahr, daß die Zahl der Skalen die der Objekte übersteigt. Bei DI VESTA (1966) waren es 37 Skalen und 20 Objekte. Ein weiterer Extremfall von (d) ist dann gegeben, wenn nur ein Objekt vorliegt bzw. die faktorielle Struktur isolierter Objekte miteinander verglichen werden soll (PROTHRO und KEEHN, 1957). Zuweilen werden mit den Daten mehrere FAs „durchprobiert“, z. B. bei DI VESTA (1966) die Fälle (b) und (h) oder bei JAKOBOVITS und OSGOOD (1967) die Fälle (b), (c) und (h). Sehr selten hingegen findet man auch im amerikanischen Raum den Fall (a) (LEVIN, 1965; Hinweise auf nicht veröffentlichte Arbeiten finden sich bei WIGGINS und FISHBEIN, 1969).

In den untersuchten Arbeiten fehlten die Fälle (a), (e) und (g). Es scheint so zu sein, daß TUCKER-Analysen vor allem wegen fehlender Computerprogramme noch nicht die entsprechende Publizität besitzen; zum anderen ist man an einer Gruppierung von Vpn, wie sie durch die Möglichkeiten (e) und (g) versucht werden könnte, nicht interessiert. Am häufigsten wird von den Möglichkeiten (b) (ROEDER, 1972; BORTZ, 1972; FRANKE und BORTZ, 1972; SCHÖNPFLUG, 1972; SATERDAG *et al.*, 1971; BORTZ, 1971 a, b; GRONER, 1967; ERTEL, 1965 a, b) und (c) Gebrauch gemacht (ROEDER, 1972; BUSZ *et al.*, 1972; SATERDAG *et al.*, 1971; COHEN und SCHÜMER, 1969; GRONER, 1967; TRAXEL,

1961; MERZ, 1958). Auch die Berechnung von Skalen-Faktoren über die Vpn wird bisweilen vorgenommen (JOHN und KEIL, 1972; TIMM, 1971; MÖBUS und AHRENS, 1970; AHRENS und MÖBUS, 1968; JANSEN und HOFFMANN, 1965; HOFSTÄTTER, 1963; RÜSSEL, 1964). Werden Ähnlichkeitsbeurteilungen analysiert, so sind diese Methoden der Möglichkeit (f) zuzuordnen (KENTLER, 1959; EKMAN, 1954). Die Möglichkeit der Auseinanderfaltung der drei Dimensionen und die Berechnung von Skalenfaktoren (h) (ERTEL und SCHINDLER, 1969; HOFSTÄTTER, 1956), Objektfaktoren (i) (HOFSTÄTTER, 1962; MICKO, 1962; HOFSTÄTTER, 1956) und Versuchspersonenfaktoren (j) (BORTZ, 1971 b) wird insgesamt seltener ausgenutzt.

MESSICK (1957) hebt bei der Verwendung von semantischen Differentials in FAs drei wesentliche metrische Annahmen hervor: (1) gleiche Intervalle für jede einzelne Skala, (2) gleiche Intervalle zwischen verschiedenen Skalen und (3) Nullpunkte, die bei jeder Skala an der gleichen Stelle liegen. MESSICK analysierte Teile der von OSGOOD und SUCI (1955) verwendeten Daten mit der Methode der sukzessiven Intervalle. Die Ergebnisse zeigen „... an apparent inequality of intervals within any one of the scales“ und weiter, „... the same categories tend to be too large or too small in similar amounts over all scales“ (1957, S. 202). Die Korrelationen zwischen den empirischen Skalen und den theoretisch errechneten liegen jedoch durchwegs bei .98 und .99. MESSICK schließt daraus, „that the scaling properties implied by the semantic differential procedures have some basis other than mere assumption“ (1957, S. 206). Diese Ergebnisse sind insbesondere dort wesentlich, wo direkt ungemittelte Rohdaten verwendet werden. Aber auch bei der Verarbeitung von Mittelwerten dürften sich konsistente metrische Verzerrungen bemerkbar machen. Dies vor allem, wenn man auf die Information aus den unterschiedlichen Skalenstreuungen verzichtet und die FA aufgrund von Korrelationskoeffizienten und nicht von Querprodukten berechnet (LEVIN, zit. nach MIRON und Osgood, 1966, S. 796).

Abschließend sei bei dieser Gruppe von Anwendungsfällen nochmals betont, daß hier Schätzurteile als Intervalldaten behandelt werden. Diese Voraussetzung wird dabei vor allem dort unter den Tisch fallen gelassen, wo gemittelte Daten einer FA zugrunde gelegt werden. Man könnte allerdings sagen, wenn die FA selbst gegenüber der Verwendung von Alternativdaten robust ist, dann sind auch die vorhandenen metrischen Fehler des semantischen Differentials von untergeordneter Bedeutung.

(4) Eine vierte Gruppe von Arbeiten kann man unterscheiden, und zwar solche, in denen Ratings, die nicht oder nur teilweise über Polaritätenprofile gewonnen wurden, mit in die FA eingegangen sind. Hierher gehören Arbeiten, in denen Verhaltensratings, Schulnoten, Kategorisierungen aufgrund von Schätzurteilen und dergleichen mehr zusammen mit anderen Variablen verwendet werden. Diese Beurteilungen dürften hier wohl nicht über das Ordinalniveau hinausgehen. Bisweilen werden diese Beurteilungen mit solchen Varia-

blen verarbeitet, bei denen das Skalenniveau weniger zweifelhaft auf dem Intervallniveau liegt; etwa dann, wenn Anmutungsqualitäten von Handschriften und objektiv gemessene Größen (wie Schriftgröße oder Schreibdruck) oder wenn Verhaltensratings und Ergebnisse aus geeichten Persönlichkeits-tests zusammen untersucht werden. Bei diesen Arbeiten, die etwa 18% der Untersuchungen ausmachen, wird meist nicht thematisiert, daß die Daten wenigstens zum Teil auf einem niedrigeren Skalenniveau liegen, als üblicherweise bei einer FA vorausgesetzt wird (BOTTENBERG, 1972; BOTTENBERG und WEHNER, 1972; ANGERMAIER, 1972; ALTHOFF, 1971; HAAS und LOEWER, 1971; TIMM, 1971; SEITZ, 1971; OTTE, 1970; BOTTENBERG, 1970; BRANDSTÄTTER, 1969; KOPFMANN, 1969; SEITZ und LÖSER, 1969; MAXEINER, 1968; FÜRNRATT 1968b, c; BOTTENBERG, 1968; KLEBELSBERG *et al.*, 1968; LOCKOWANDT, 1968; LÜER und NEUFELDT, 1968; TAUSCH *et al.*, 1967; HELM, 1967; TAUSCH *et al.*, 1966; FAHRENBERG und CONRAD, 1965; SEIFERT, 1964; KLEBELSBERG, 1963).

(5) Zu einer letzten Gruppe, die ca. 11% der Publikationen ausmacht, lassen sich solche Arbeiten vereinigen, in denen Daten auf Absolutskalenniveau (Häufigkeiten, physikalische Messungen) erhoben wurden oder diese nur mit Intervallskalendaten vermischt sind (KLIPPSTEIN, 1972; BORTZ, 1971b; LAUX und FRÖHLICH, 1970; FÜRNRATT, 1969a, c; WIECZERKOWSKI *et al.*, 1968; CARL, 1968; FAHRENBERG und MYRTEK, 1966; KRISTOF, 1964; KALLINA, 1964; FISCHER, 1964; PAWLIK, 1963; KRISTOF, 1963; DIRKS, 1961; MERZ, 1961). Bei diesen Fällen ergeben sich vom Skalenniveau der Daten keine Einwände gegen eine faktoranalytische Weiterverarbeitung.

2.1.2. Verteilungsannahmen

Es kann kein Zweifel darüber bestehen, daß die Anwendung der FA die Annahme einer multivariaten Normalverteilung in der Referenzpopulation des Datenkörpers voraussetzt. Eine solche Annahme ist in so gut wie allen Fällen multivariater Korrelationstechniken im Zusammenhang mit inferenzstatistischen Fragen notwendig, da Stichprobenverteilungen nur für diesen Fall entwickelt wurden. „In multivariate analysis, especially in those branches which deal with components, factors and canonical correlations, exact inferences from sample to parent offer some exceedingly difficult problems in distribution theory . . . Practically everything that is known about exact distributions in the multivariate case depends on the assumption that the parent distribution is multivariate normal. This assumption will be made throughout“ (KENDALL, 1965, S. 86). Bisweilen wird auch explizit auf diese Annahme hingewiesen, wenn bestimmte Extraktionsmethoden besprochen werden (LINN, 1968, S. 38).

Bereits auf der Ebene einfacher Korrelationskoeffizienten kommt es bei der Überprüfung von Verteilungsannahmen zu gewissen Mißverständnissen. Zunächst sollten Korrelation und Regression klar getrennt werden. Wird ein Korrelationskoeffizient in inferenzstatistischem Kontext behandelt, so ist die Annahme einer bivariaten Normalverteilung notwendig. Dies trifft nicht für

die Aufstellung einer Regressionsbeziehung zu. Nur wenn der Korrelationskoeffizient als rein beschreibende Statistik verwendet wird — wie das auch bei einem Mittelwert oder einer Prozentangabe vorkommen kann — sind keine besonderen Annahmen erforderlich (BINDER, 1970); dieser Fall dürfte jedoch sehr selten sein. Trifft die Annahme einer bivariaten Normalverteilung zu, so impliziert dies, daß (1) die beiden Randverteilungen von X und Y normalverteilt sind und weiter, daß (2) bei einem gegebenen X -Wert die bedingte Verteilung von Y normal und bei gegebenem Y -Wert die bedingte Verteilung von X normal ist und schließlich (3), daß die Regressionsbeziehung zwischen X und Y linear ist (HAYS, 1963, S. 528). „ . . . under the bivariate or multivariate normal assumption, the absence of a linear relation is the absence of any systematic relation“ (HAYS, 1963, S. 565). Eine bivariate Normalverteilung impliziert übrigens auch die selten diskutierte Eigenschaft der Homoscedasdität (GARRETT, 1962, S. 163). Es ist klar, daß als Methode der Wahl für einen Vor-test bezüglich der Verteilungsannahme nur ein solcher Test in Frage kommt, der prüft, ob eine multivariate Normalverteilung vorliegt.

Tests, die für die wissenschaftliche Praxis empfohlen — allerdings dort nicht allzuoft angewandt werden — untersuchen lediglich notwendige aber nicht hinreichende Bedingungen (BAUMANN, 1971, S. 67) für die Annahme einer multivariaten Normalverteilung, nämlich die Normalität der Randverteilungen und die Linearität der Regressionsbeziehungen. Neben der Tatsache, daß solche Überprüfungen nicht schlüssig sind, werden solche Tests mit den einzelnen Variablen bzw. Variablenpaaren isoliert durchgeführt. Die dadurch „in Serie geschalteten“ Irrtumswahrscheinlichkeiten erschweren das Gesamturteil in Bezug auf den ganzen Datenkörper und sind analog zur Durchführung einzelner t -Tests anstelle einer multifaktoriellen (u. U. multivariaten) Varianzanalyse. Weiter ist zu vermerken, daß bei der Untersuchung der Randverteilungen die Forschungshypothese mit einer scharfen Nullhypothese identifiziert wird. Es ist bekannt, daß dieses Vorgehen nicht unproblematisch ist (BREDENKAMP, 1972; KLEITER, 1969; WILSON, MILLER und LOWER, 1967; EDWARDS, 1965; WILSON und MILLER, 1964; BINDER, 1963; GRANT, 1962). Dies führt dazu, daß die Beurteilung unter welchen Umständen eine Anpassung als gut oder als weniger gut anzusehen ist, eher nach groben Faustregeln bestimmt wird (LIENERT, 1973, S. 164). Grundsätzliche Überlegungen, wie sie etwa von BREDENKAMP (1972) bezüglich der Akzeptierung von Nullhypothesen diskutiert werden, wären sicherlich auch für diesen Spezialfall angebracht. Dabei können Entscheidungskriterien nur im Zusammenhang mit Robustheitseigenschaften der sich anschließenden Verfahren entwickelt werden.

In den untersuchten Zeitschriftenartikeln wird bisweilen angegeben, daß die Randverteilungen annähernd „normalverteilt“ waren; meist wird jedoch nicht erwähnt, wie dies geprüft wurde (MIKUS, 1969; COHEN und WITTEMANN, 1967; SEIFERT, 1964; SELG und LISCHKE, 1966; LIENERT und FABER, 1963; LIENERT, 1963). Sind nicht alle Variablen normalverteilt, so wird entweder (1)

trotzdem mit der üblichen Produkt-Moment-Korrelation gerechnet (STEINGRÜBER, 1971; BOTTENBERG, 1970; SCHAEDEL, 1961) oder (2) es werden T-Transformationen vorgenommen (TIMM, 1971; ALTHOFF, 1971; BOTTENBERG, 1968; ERTEL, 1965a, b; BÄUMLER, 1964) bzw. (3) Flächentransformationen durchgeführt (LAUX und FRÖHLICH, 1970; FÜRNRATT, 1969b; FAHRENBURG und MYRTEK, 1966). Bei T-Transformationen bleiben die ursprünglichen Relationen zwischen den Zahlenwerten erhalten, d. h. die Verteilungen werden dadurch nicht normalisiert. Dies ist nur dann der Fall, wenn zugleich Flächentransformationen vorgenommen werden. Transformationen, wie sie von ●TTE (1970) erwähnt werden, haben allerdings den Nachteil, daß die Ergebnisse nur schwer zu interpretieren sind, da die Relationen der Ursprungsdaten verändert werden.

Auch die Linearität der Regressionsbeziehung wird selten überprüft. Geschehen ist dies nur in den Arbeiten von ALTHOFF (1971) sowie bei SELG und LISCHKE (1966). In einem Artikel wird angegeben, daß nicht alle Beziehungen linear seien (FAHRENBURG und MYRTEK, 1966), in einem weiteren, daß solche Variablen, die nicht-lineare Beziehungen aufwiesen, von der weiteren Verarbeitung ausgeschlossen wurden (SASSENSCHIEDT und BUGGLE, 1970).

Auf die Frage nach der Homoscedasizität wurde in keinem der Artikel eingegangen, so daß auch hier auf eine weitere Diskussion verzichtet werden muß.

2.1.3. Zusammenfassung

Zu der Praxis der Variablenauswahl kann man sagen, daß diese wesentlich verbessert werden könnte, wollte man etwa den Ratschlägen PAWLIKS (1968, S. 270f.) folgen. Er meint, daß der Zusammenstellung des Ausgangsmaterials einer FA eine möglichst vollständige Aufzählung der Variablen des Merkmalbereiches vorangehen sollte. Dies käme auch dem Explizieren von gehegten Hypothesen nahe. Diese Variablen oder zumindest eine repräsentative Auswahl davon, die bereichert durch Markierungsvariable aus früheren Untersuchungen (CATTELL und WARBURTON, 1967, S. 6; SCHEIER und CATTELL, 1965) werden sollten, könnten dann einer FA unterzogen werden. Falls man nur an der Struktur der ursprünglichen Variablen interessiert ist und begründet vermuten kann, daß die durch Markierungsvariable mitberücksichtigten Faktoren nichts mit den vermutlich sich neu ergebenden Faktoren zu tun haben, könnten diese Markierungsvariablen nicht in die analytischen Rotationskriterien mit aufgenommen werden (TIMM, 1971).

Nach der Praxis der vorgenommenen Analysen scheint es aber eher so zu sein, daß immer dann, wenn ein anfallendes oder vorgefundenes Datenmaterial miteinander „korrelierbar“ erscheint, auch mit einer gewissen Zwangsläufigkeit einer FA gerechnet wird. Dabei muß festgestellt werden, daß die Voraussetzungen für eine sachgerechte Behandlung des Datenmaterials, sowohl in bezug auf das Skalenniveau als auch in Hinblick auf die notwendige Vertei-

lungsform der Daten, wenn schon nicht ausdrücklich verletzt, so doch ungeprüft waren.

2.2. Probandenstichproben

Neben der Auswahl der Merkmale ist die Auswahl der Probanden von wesentlicher Bedeutung für die Ausgangsdaten einer FA. Die Generalisierung faktorenanalytischer Ergebnisse auf ein Personenkollektiv, das größer ist als das der tatsächlich untersuchten Stichprobe, scheint von der Zielsetzung wohl aller Arbeiten her selbstverständlich. Trotzdem wurde nicht zuletzt in der theoretischen Entwicklung der FA die Frage des im Zusammenhang mit der Stichprobenauswahl der Vpn stehenden „sampling error“ im Vergleich zum im Zusammenhang mit der Merkmalsauswahl stehenden „measurement error“ vernachlässigt (LINN, 1968).

Zwei Fragenbereiche scheinen bei der Untersuchung der Anwendungspraxis einer Forschungsmethode in diesem Zusammenhang wichtig: (1) Einmal die Frage nach der Vielfalt und Reichhaltigkeit der Anwendungspraxis dieser Methode auf psychologisch interessante Populationen; (a) für welche Grundgesamtheit kann die entsprechende Stichprobe überhaupt repräsentativ sein und (b) gewährleistet das Auswahlverfahren die Repräsentativität in ausreichendem Maße? (2) Die grundsätzlichere und theoretische Frage ist aber die, in wie weit faktorenanalytische Ergebnisse (a) invariant sind gegenüber verschiedenen Stichproben gleicher Grundgesamtheiten und (b) unter welchen Voraussetzungen Vergleiche faktorenanalytischer Ergebnisse von Stichproben verschiedener Grundgesamtheiten sinnvoll sind.

2.2.1. Zur Repräsentativität der Probandenstichproben

Eine oft notwendige Einschränkung des Generalisierungsbereiches ist, daß die Aussagen – vor allem wenn es um soziologische Fragen geht, aber auch etwa um Aussagen über Intelligenzstrukturen, erfaßt mit den üblichen Intelligenztests – eine nur auf unseren Kulturkreis beschränkte Gültigkeit besitzen. Dies ergibt sich oft auch schon daraus, daß interkulturelle Vergleichsuntersuchungen mit denselben Untersuchungsmethoden nur selten möglich sind. Vermutlich nur dieser Einschränkung sind die Untersuchungen unterworfen, in denen das „Probandenmaterial“ von Eichstichproben zur Verfügung stand (KLIPPSTEIN, 1972; BUGGLE *et al.*, 1968; JÄGER und TODT, 1964; JÄGER, 1964; KEREKJARTO und SCHMIDT, 1962; LIENERT, 1963; RIEGEL, 1960). In diesen Fällen ist sowohl die zugehörige Grundgesamtheit für die psychologische Theorienbildung relevant und ebenso die Stichprobengewinnung unproblematisch. Schwieriger ist die Frage nach der adäquaten Stichprobengewinnung bei den Arbeiten zu beantworten, wo nicht nach expliziten Kriterien vorgegangen wurde, sondern wo die Repräsentativität durch den großen Stichprobenumfang gewährleistet werden soll (SEYDEL, 1972; ANGERMAIER, 1972; CATTELL *et al.*, 1971; HAAS und LOEWER, 1971; ALTHOFF, 1971; KOPF-

MANN, 1969; FÜRNRATT, 1969a; WENDELER, 1969; COHEN und WITTEMANN, 1967; FAHREBERG, 1965; LIENERT und FABER, 1963).

Da man leider nicht alle Fragestellungen mit Rattenversuchen beantworten kann, greift man oft auf das „Material“ zurück, das am leichtesten zugänglich ist, und das sind — nach der Anzahl der Arbeiten zu schließen, in denen so verfahren wird — sehr oft Studenten und Schüler. Aber selbst für diese Referenzpopulationen sind die herangezogenen Untersuchungen, mit Ausnahme der oben angeführten Fälle, nicht repräsentativ. Eine weitere Beschränkung des Gültigkeitsbereiches ist auch dadurch gegeben, daß die untersuchten Probanden Psychologiestudenten waren (BUSZ *et al.*, 1972; BORTZ, 1971a, b; MÖBUS und AHRENS, 1970; FISCH und SCHMALT, 1970; CARL, 1968; AHRENS und MÖBUS, 1968; LÜER und NEUFELDT, 1968; FAHREBERG und MYRTEK, 1966; SELG und LISCHKE, 1966; ERTEL, 1965a, b); diese Liste dürfte allerdings nicht vollständig sein, da vermutet werden kann, daß in einigen Untersuchungen, in denen die globale Etikettierung „Studenten“ verwendet wurde, Psychologiestudenten gemeint waren. Bei allen diesen Fällen kann man in begründeter Weise vermuten, daß die leichtere Zugänglichkeit bei der Datenerhebung mit einer Einschränkung des Generalisierungsbereiches der Untersuchungen erkauft wurde.

Aus dem Gesagten ergibt sich ferner, daß Untersuchungen mit „Erwachsenen“ nur selten zu finden sind (SEYDEL, 1972; KRÜSKEMPER und KRÜSKEMPER, 1971; HOBI und KLÄR, 1972; ALTHOFF, 1971; OTTE, 1970; SASSENSCHIEDT und BUGGLE, 1970; STEINHAGEN, 1970; FÜRNRATT, 1969b; 1968a, c; KLEBELSBERG *et al.*, 1968; MEYER und GOLLE, 1966; FAHREBERG, 1965; RÜSSEL, 1964; KALLINA, 1964; KRISTOF, 1963; KLEBELSBERG, 1963; RIEGEL, 1960). Diese Untersuchungen machen nur ca. 14% der bearbeiteten Artikel aus. Als Schlußfolgerung aus dieser Anlage der Untersuchungen wird die Ansicht nahegelegt, daß man viel weiß über Verhaltensweisen von Psychologiestudenten (in den ersten Semestern), weniger über die von Studenten und Schülern im allgemeinen und am allerwenigsten über die von Erwachsenen.

2.2.2. Die Verwendung von inhomogenen Stichproben

Aus verschiedenen Untersuchungen ist bekannt, daß eine Faktorenlösung auch von den die Stichprobe definierenden Merkmalen abhängt. Im Intelligenzbereich spielt etwa die Zusammensetzung der Stichprobe nach Lebensalter, Leistungsfähigkeit oder etwa Neurotizismus eine erhebliche Rolle (WEWETZER, 1958; LIENERT, 1961; COHEN und WITTEMANN, 1967). Die Alterszusammensetzung wird für eine FA belangvoll, wenn man altersinhomogene Stichproben verwendet (JOHN und KEIL, 1972; HOBI und KLÄR, 1972; KRÜSKEMPER und KRÜSKEMPER, 1971; HAAS und LOEWER, 1971; OTTE, 1970; SASSENSCHIEDT und BUGGLE, 1970; FÜRNRATT, 1968a, b, c; EYFERTH, 1963; MICKO, 1962; MERZ, 1958). Leider sind die Angaben zur Beschreibung der Stichproben (Mittelwerte, Streuungen, Variationsbreite) nicht immer vollständig, so

daß die Altershomogenität bei einigen Stichproben nicht beurteilt werden kann. Auf diesen Punkt weist bereits GUILFORD hin (1952, S. 32): „Too often, samples are taken from populations that happen to be most available. What is worse, different populations are thrown together without questioning what effect this may have upon the intercorrelations.“ Zieht man solche Effekte nicht schon von vornherein in Betracht, etwa durch Verwendung von Markierungsvariablen für diese Einflüsse, so kann es sein, daß die FA schon im Ansatz fehl am Platze ist. In diesem Zusammenhang sei auch auf die Modoratorentechnik (BARTUSSEK, 1970) und für den Fall von Alternativdaten auf die Konfigurationsfrequenzanalyse (LIENERT, 1971) verwiesen.

Noch weitere Effekte können diskutiert werden; und zwar ist für den bivariaten Fall allgemein bekannt, daß die Homogenisierung von Stichproben zu einer Verkleinerung der entsprechenden Korrelationen führt. Korrelationskoeffizienten sind gegenüber einer Varianzreduktion nicht invariant; hierfür wurden auch entsprechende Korrekturformeln entwickelt (GUILFORD, 1954, S. 392). Anders ist es mit Regressionskoeffizienten, sie sind gegenüber einer durch Teilstichprobenbildung bedingten kleineren Varianz invariant. Weniger bekannt sind analoge Effekte für den multivariaten Fall und speziell für faktorenanalytische Ergebnisse. Die Kenntnis solcher Effekte ist jedoch dann wichtig, wenn aus Gesamtstichproben Teilstichproben gebildet werden, FAs gerechnet und die entsprechenden Ergebnisse anschließend vergleichend interpretiert werden (SEYDEL, 1972; HOBI und KLÄR, 1972; CATTELL *et al.*, 1971; BOTTENBERG, 1970; OTTE, 1970; STEINHAGEN, 1970; SEITZ und BRÄTH, 1970; BALZERT, 1968; SELG und LISCHKE, 1966; MEYER und GOLLE, 1966; LIENERT und FABER, 1963; KEREKJARTO und SCHMIDT, 1962; RIEGEL, 1960; EYSENCK, 1953; WEBER, 1953). CATTELL *et al.* (1972) halten allerdings Untersuchungen von Stichproben, die sich aus Normalen und klinischen Krankheitsgruppen zusammensetzen, für wünschenswert und führen auch eine entsprechende Untersuchung durch. Dabei geht es ihnen insbesondere darum zu zeigen, daß die Faktorenlösungen solcher gemischter Gruppen nicht wesentlich verschieden sind von denen, die bei ausschließlich normalen Probanden resultieren. Die Effekte scheinen darüber hinaus auch in all den Fällen interessant zu sein, in denen verschiedene Stichproben miteinander verglichen werden, deren Merkmale aber unterschiedliche Varianz aufweisen. Sie sind selbstverständlich ebenso dann wichtig, wenn aufgrund des Stichprobenauswahlverfahrens die Randbereiche einer Variable „gestutzt“ sind. Die Invarianz faktorenanalytischer Ergebnisse bei selektierten Stichproben wurde erst in jüngerer Zeit eingehend behandelt. Eine Übersicht dazu findet man bei MULAİK (1972). Für den Vergleich mehrerer FAs weist MULAİK (1972, S. 356) neben der Verwendung verschiedener Rotations- und Extraktionsmethoden auf folgende fünf häufig anzutreffende Fehler hin:

(1) Die entsprechenden Faktorenmuster werden aufgrund der Basis von Korrelationsmatrizen und nicht auf der Basis von Varianz — Kovarianz-

Matrizen bestimmt. Durch die Verwendung von Korrelationen entsteht nämlich eine Normierung der Varianz jeder Stichprobe auf Eins; wenn die Variablen jedoch von Stichprobe zu Stichprobe eine unterschiedliche Varianz aufweisen, so wird damit pro Stichprobe eine unterschiedliche Metrik eingeführt. Bei Verwendung einer unterschiedlichen Metrik müssen sich die Faktorenmuster jedoch unterscheiden. Auf diesen Punkt weist auch CATTELL hin: „... the two analyses to be compared should be made in terms of cross-products, not correlations...“ (1966, S. 190).

(2) In den zu vergleichenden Stichproben werden jeweils orthogonale Faktorenlösungen bestimmt. Dabei wird die Interkorrelation zwischen den Faktoren zwangsweise mit Null festgelegt. Es läßt sich jedoch zeigen, daß „if a set of factors associated with a factor-pattern matrix F are mutually uncorrelated in the parent population, the same factors are likely to be correlated in the sub-population“ (MULAIK, 1972, S. 348). Ähnlich spricht sich auch CATTELL aus (1966, S. 211). Will man also die Invarianz des Faktorenmusters F erhalten, so darf man nicht die Interkorrelationen der Faktoren zwangsweise gleich Null festlegen. Das Faktorenmuster kann nur dann invariant sein, wenn mindestens eine von zwei zu vergleichenden Faktorenlösungen nicht-orthogonal ist. Setzt man die Faktorenkorrelationen zwangsweise gleich Null an, so entspricht die Lösung nicht einem „tatsächlich vorhandenem“ Faktorenmuster, sondern der zu diesem korrespondierenden Faktorenstruktur. Die Faktorenstruktur ist jedoch nachweislich invariant. (Unter Faktorenmuster werden hierbei die Gewichte der Linearkombinationen verstanden, unter Faktorenstruktur die Korrelationen zwischen beobachteten Variablen und Faktoren.)

(3) Es werden zu wenig Faktoren extrahiert, um die gemeinsamen Faktoren genügend zu stabilisieren.

(4) In den zu vergleichenden Analysen sind die Anteile der Variablen an den spezifischen Faktoren nicht von den Selektionsmerkmalen unabhängig.

(5) Die FAs wurden zunächst nicht so transformiert, daß sie einander möglichst ähnlich sind.

Abschließend kann man zu diesem Punkt sagen, daß schon PAWLIK vor der Anwendung von sog. „Großklumpenstichproben“ gewarnt hat, d. h. vor solchen Personenkollektiven, die aus Stichproben verschiedener Populationen (z. B. auch von Altersgruppen) unrepräsentativ zusammengesetzt sind. „Es gibt dann keine natürliche Population, auf welche die erhaltenen Korrelationen und Faktoren verallgemeinert werden können.“ Daß solche Stichproben aber auch aus dem Grund der leichten Zugänglichkeit dennoch untersucht werden und die Ergebnisse auch einer faktorenanalytischen Bearbeitung und Interpretation unterzogen werden, mag u. a. zu der von ORLIK (1967b, S. 87) beklagten „erschreckenden Unverbindlichkeit“ der Resultate faktorenanalytischer Forschung mit beigetragen haben.

2.2.3. Zur Relation zwischen Variablen- und Probandenzahl

Nach PAWLIK (1968, S. 278) sollte die Anzahl der Versuchspersonen das Dreifache der Anzahl der Variablen betragen. Gegen diese Faustregel finden sich bisweilen Verstöße, z. B. wenn über Polaritätenprofile korreliert wird (s. o.). Dieser Fehler wird z. T. aber dadurch überdeckt, daß viele Skalen zur Beurteilung wenigen Versuchspersonen vorgelegt werden oder diese viele Objekte einzuschätzen haben. Verstöße kann man auch finden, wenn viele Variable verwendet werden (SEITZ und BRÄTH, 1970; LOCKOWANDT, 1968; KALLINA, 1964) oder schließlich auch ohne weitere Begründungsmöglichkeit, wenn man von der Schwierigkeit der Datenerhebung überhaupt absieht (BORTZ, 1972; KRÜSKEMPER und KRÜSKEMPER, 1971; FAHRENBERG und MYRTEK, 1966; FAHRENBERG und CONRAD, 1965; FISCHER, 1964). Der krasseste Fall stammt dabei aus einer Untersuchung aus dem klinischen Bereich (KRÜSKEMPER und KRÜSKEMPER, 1971), bei der 11 Variable, aber nur 14 Probanden verwendet werden.

Die angegebene Faustregel läßt sich intuitiv verständlich machen, wenn man daran denkt, daß bei einer FA die Korrelationskoeffizienten als fixe Größen behandelt werden, während sie aber eigentlich als Zufallsvariable angesehen werden müßten, die mit einer gewissen Wahrscheinlichkeit im Vertrauensintervall schwanken (diese Tatsache dürfte es auch verhindert haben, daß die Verteilung der Faktorenladungen bekannt ist). Bekanntlich nimmt die Breite des Vertrauensintervalls mit zunehmender Stichprobengröße ab und daher die Begründung dieser Forderung. Daß durch die notwendige Voraussetzung fixer Koeffizienten für die meisten Extraktionsmethoden auch der theoretische Geltungsbereich dieser Methode eingeschränkt wird, sei noch am Rande erwähnt. Als Folge davon ergibt sich, daß auch für die Beurteilung der Bedeutsamkeit von Faktorenladungen nur Faustregeln angegeben werden können, die Faktorenladungen aber nicht auf ihre Signifikanz hin überprüft werden können.

3. Beurteilung der Güte von Faktorenanalysen

Eine kritische Beurteilung einer experimentellen Untersuchung setzt bekanntlich voraus, daß die Daten bzw. die aus ihnen abgeleiteten Maße hinreichend genau referiert werden. Für faktorenanalytische Arbeiten wäre es ebenfalls von Vorteil, wenn zumindest gewisse Standards in der Berichterstattung allgemein eingehalten würden (z. B. Angabe der Korrelationsmatrix, rotierte Ladungsmatrix, Berechnung der Varianzaufklärung absolut und in Anteilen der Kommunalität, Analysetechnik, usw.). Bisweilen werden die Korrelations- und Ladungsmatrizen, wenn die Anzahl der Variablen sehr groß oder sehr viele FAs für eine Publikation gerechnet wurden, für einen Zeitschriftenartikel zu viel Raum einnehmen; die Angaben zur Varianzaufklärung (und zwar sowohl absolut als auch bezogen auf die Kommunalität) müßten aber immer referiert werden. Die Kette der Schlüsse und Interpreta-

tionen eines Autors können eben nur dann nachvollzogen werden, wenn nicht nur diese selbst, sondern auch das Material, auf welches die Untersuchung aufbaute, hinlänglich genau wiedergegeben wird.

3.1. Voraussetzungen für eine FA

3.1.1. *Signifikanzprüfung der Korrelationsmatrix*

Bevor eine FA gerechnet wird, sollte sichergestellt sein, daß sich die Korrelationsmatrix signifikant von einer Nullmatrix unterscheidet; Prüfmethode zu diesem Punkt sind bei ÜBERLA (1968, S. 131f.) zu finden. Erst wenn diese Voraussetzung gegeben ist, kann man eine FA sinnvollerweise durchführen⁴. Enthält außerdem die Korrelationsmatrix nur niedrige Korrelationen, dann ist von vornherein mit einer großen Anzahl von Faktoren zu rechnen, die dabei nur wenig Varianz aufklären.

In der Praxis der Untersuchungsberichte findet man vollständige Korrelationsmatrizen nur in ca. 30% der Fälle angegeben (ANGERMAIER, 1972; SEYDEL, 1972; HOBI und KLÄR, 1972; FRANKE und BORTZ, 1972; SEITZ, 1971; STEINGRÜBER, 1971; KRÜSKEMPER und KRÜSKEMPER, 1971; OTTE, 1970; LAUX und FRÖHLICH, 1970; SASSENSCHIEDT und BUGGLE, 1970; STEINHAGEN, 1970; FÜRNRATT, 1969b; KOPFMANN, 1969; KNABE, 1969; BÄUMLER *et al.*, 1968; BALZERT, 1968; TIMM, 1968; CARL, 1968; AHRENS und MÖBUS, 1968; COHEN und WITTEMANN, 1967; TAUSCH *et al.*, 1966; SELG und LISCHKE, 1966; HARTMANN und ENGELMANN, 1966; KATZENBERGER, 1965; JÄGER und TODT, 1964; JÄGER, 1964; KALLINA, 1964; KRISTOF, 1964; HUNZIKER, 1964; FISCHER, 1964; SEIFERT, 1964; EYFERTH, 1963; KENTLER, 1959; LIENERT, 1963; KRISTOF, 1963; HOFSTÄTTER, 1962; MICKO, 1962; KRAAK, 1961; MERZ, 1961; FISCHER, 1958; MERZ, 1958; ROTH, 1957; WEBER, 1953). Bei der Verwendung sehr vieler Variablen ist das Fehlen dieser Angaben zwar nicht weiter verwunderlich, es erschwert aber doch die Überprüfung dieses Punktes.

Die ausdrückliche Erwähnung der Prüfung der Korrelationsmatrix auf ihre Signifikanz gegenüber einer Nullmatrix findet man nur bei ANGERMAIER (1972) und FÜRNRATT (1969c). Zumindest eine graphische Kontrolle der Verteilung der Korrelationskoeffizienten der Matrix ist zu fordern, vor allem bei solchen Matrizen, die sowohl positive wie auch negative Koeffizienten enthalten.

3.1.2. *Art der Kommunalitätsschätzung*

Bevor eine FA durchgeführt werden kann, erfolgt (außer bei der Bestimmung von Hauptkomponenten) eine Schätzung der Kommunalitäten, d.h. des Anteiles der Varianz jeder einzelnen Variablen, die durch die gemeinsamen Faktoren aufgeklärt werden soll. Die Ermittlung dieser Koeffizienten in der

⁴ CATTELL (1966, S. 200) führt dazu ein recht schönes Beispiel an, das Faktorenladungen auf vier Variablen zeigt; die korrespondierende Korrelationsmatrix weist nur Null-Korrelationen auf!

Hauptdiagonale ist scheinbar ein noch nicht befriedigend gelöstes Problem. Wie KALVERAM ausgeführt hat (1970d, S. 105), sind die Kommunalitäten auch theoretisch nicht eindeutig zu ermitteln und praktisch angewandte iterative Schätzverfahren, die von verschiedenen Ausgangsschätzungen ausgehen, konvergieren zu numerisch verschiedenen Werten, die aber unter gewissen Umständen ihrerseits die ursprüngliche Korrelationsmatrix wieder richtig reproduzieren können.

Auf welche Art nun die Kommunalitätsschätzung erfolgt, wird in den untersuchten Artikeln ebenfalls nur selten angegeben. Heutigen Standards entsprechend, wird der quadrierte multiple Korrelationskoeffizient (SMC) als Verfahren der Wahl angesehen (ANGERMAIER, 1972; LAUX und FRÖHLICH, 1970; BÄUMLER *et al.*, 1968; BALZERT, 1968; LOCKOWANDT, 1968; SELG und LISCHKE, 1966). Man findet aber auch Arbeiten, in denen das Iterationsverfahren von einer Eins in der Diagonale seinen Ausgangspunkt nimmt (WENDELER, 1969); vor allem in früheren Arbeiten wurde dazu auch die höchste Zeilenkorrelation verwendet (COHEN und WITTEMANN, 1967; HEDINGER, 1965; FISCHER, 1964; KRISTOF, 1964; 1963; RIEGEL, 1960). Schlüsse über die verwendeten Schätzverfahren könnte man auch noch auf Grund der Angabe der verschiedenen Rechenzentren ziehen, in denen die Arbeiten gerechnet wurden, aber dies dürfte inhaltlich nicht mehr Aufschlüsse bringen. Jedenfalls muß man feststellen, daß man sich bei der Anwendung der FA über die theoretischen Bedenken oder sogar die eventuelle theoretische Unlösbarkeit des Kommunalitätenproblems hinwegsetzen muß.

3.1.3. *Kommunalität und Reliabilität der Variablen*

Bekanntlich bildet nach dem Modell der FA die Reliabilität der Variablen eine obere Grenze für die zu extrahierende gemeinsame Varianz (ÜBERLA, 1968, S. 58); d.h. die Reliabilität einer Variablen entspricht der Kommunalität, wenn keine spezifischen Faktoren mehr zur systematischen Variabilität der Variablen beitragen. Aus diesem Grund empfiehlt auch ÜBERLA (1968, S. 358) nach einer FA die geschätzten und die rückerrechneten Kommunalitäten mit Reliabilitätsschätzungen der Variablen zu vergleichen. Dadurch soll sichergestellt werden, daß nicht zu viel an systematischer Varianz einer Variablen extrahiert wird, d.h. um nicht Fehlervarianz und systematische Varianz zu konfundieren. Übersteigt aber die rückerrechnete Kommunalität den quadrierten Reliabilitätskoeffizienten, dann sind offensichtlich die Modellvoraussetzungen der FA verletzt.

MULAİK (1972) schlägt vor, das Kommunalitätenproblem auf ein Testuniversum zu beziehen. Unter einem Testuniversum versteht er die Menge aller logisch möglichen Tests, die konzipiert werden können, um Verhaltens-eigenheiten in einem bestimmten Verhaltensbereich zu messen. Es sollen dabei die folgenden Annahmen gelten: (1) Die Zahl der Dimensionen wahrer Test-

werte ist fix und endlich, (2) für alle Tests gilt, daß es mindestens einen zweiten Test im Universum gibt, so daß die wahren Testwerte sich nur um einen unendlich kleinen Betrag unterscheiden und (3) für jeden Test gibt es eine unbegrenzte Zahl anderer Tests mit unterschiedlicher Fehlervarianz. Für ein solches Universum läßt sich zeigen, daß die untere und die obere Grenze für die Kommunalität zusammenfallen und mit der Reliabilität identisch sind; zur Bestimmung davon können multiple Korrelationen verwendet werden. Kommunalitäten, die auf der Basis der Stichprobe über multiple Korrelationen mit anderen Variablen bestimmt werden, sind dann Minimalschätzungen. Durch die Hinzunahme weiterer Variablen kann eine multiple Korrelation nur erhöht, nicht jedoch vermindert werden. Die multiple Korrelation kann im Testuniversum nur größer oder höchstens gleich hoch sein als bei der Merkmalsstichprobe. Aus der Unterschätzung der Kommunalitäten, die danach bei vielen FAs vorliegen müßte, würde dann wieder folgen, daß die Zahl der extrahierten Faktoren zu klein ist.

Entsprechende Vergleiche zwischen Reliabilität und Kommunalität sind jedoch nur bei 11% der Arbeiten zu finden (JOHN und KEIL, 1972; SEYDEL, 1972; STEINGRÜBER, 1971; SIDDIQI und KEIL, 1970; BOTTENBERG, 1970; WEHNER und BOTTENBERG, 1969; BALZERT, 1968; HUNZIKER, 1964; FISCHER, 1964; KRISTOF, 1964; LIENERT, 1963; SCHAEDELI, 1961; MERZ, 1958; WEBER, 1953). Bisweilen könnte man auch als vorsichtige Schätzung der Reliabilität Koeffizienten der Auswertungs- oder Übereinstimmungsobjektivität heranziehen (TIMM, 1971; OTTE, 1970; BRANDSTÄTTER, 1969; FÜRNRATT, 1969b), denn die Objektivität geht ja voll in die Reliabilität mit ein. Bei diesen Vergleichen findet man aber immer, daß die Modellvoraussetzungen der FA (wenn auch nicht bei allen, so doch bei einigen Variablen) auch tatsächlich verletzt sind, d.h. daß immer mehr systematische Varianz extrahiert wurde als aufgrund der Reliabilitätsschätzung zulässig wäre. Bedauerlich bleibt aber, daß nicht für alle Untersuchungen solche Reliabilitätsbestimmungen (z. B. aus anderen Arbeiten) für einen Vergleich herangezogen werden. Es scheint so zu sein, daß dies für Einzeluntersuchungen ein zu aufwendiges Unternehmen sei und man daher eher in Kauf nimmt, eine der Modellvoraussetzungen der FA zu verletzen.

3.2. Bemerkungen zur Analysetechnik

3.2.1. Faktorenextraktion

Als Methode der Wahl hat die Hauptachsenanalyse die früher gebräuchliche Centroidmethode abgelöst, welche vorher das einzig praktikable Verfahren gewesen zu sein scheint. Erfahrungen mit anderen Varianten der Faktorenextraktion, wie etwa eine Maximum-Likelihood-Schätzung von Faktoren, konnte nur einmal gefunden werden (STEINHAGEN, 1970). Diese Lösungen können aber in allen Fällen durch geeignete Rotation miteinander in Deckung

gebracht werden; dies gilt allerdings nur unter der Voraussetzung, daß die gleichen Kommunalitätsschätzungen verwendet wurden.

3.2.2. Kriterien zur Bestimmung der Faktorenzahl

Legt man sich nicht von vornherein darauf fest, Hauptkomponenten zu berechnen und damit so viele Faktoren wie vorhandene Variable zu bestimmen, so muß man sich auf ein Kriterium einigen, nach welchem die Bestimmung der Faktorenanzahl zu erfolgen hat. Soweit zu diesem Punkt Angaben vorhanden waren und dies war bei 25% der Arbeiten der Fall, werden als Abbruchkriterien der Scree-Test von CATTELL verwendet (ANGERMAIER, 1972; BOTTENBERG und WEHNER, 1972; HAAS und LOEWER, 1971; CATTELL *et al.*, 1971; MÖBUS und AHRENS, 1970; BOTTENBERG, 1970; WEHNER und BOTTENBERG, 1969); ferner die Anzahl der Eigenwerte, die größer als Eins sind (ROEDER, 1972; SEYDEL, 1972; HOBI und KLÄR, 1972; BORTZ, 1971b; TIMM, 1971; OTTE, 1970; WENDELER, 1970; LAUX und FRÖHLICH, 1970; WIENDIECK und LÜCK, 1969; BRANDSTÄTTER, 1969; FÜRNRATT, 1969b; KEIL und SADER, 1968; FISCHER, 1964); in einem Fall wurden so viele Faktoren bestimmt als Eigenwerte vorhanden waren, die größer als Null waren (BALZERT, 1968); schließlich wird auch die Beurteilung der Restkorrelationen oder die des Anteiles der extrahierten Varianz als Abbruchkriterium verwendet (KLIPPSTEIN, 1972; LOCKOWANDT, 1968; BÄUMLER, 1964; KRISTOF, 1964; MERZ, 1958). Angegeben wird bisweilen auch das Kriterium von TUCKER (PAWLK, 1968, S. 168) (BOTTENBERG, 1968; EYFERTH, 1963; KEREKJARTO und SCHMIDT, 1962; RIEGEL, 1960). Andere Möglichkeiten (Guttman-Kriterium, Lawley-Test, Fürntratt-Kriterium) kommen nur sporadisch oder gar nicht zur Anwendung.

Diese Kriterien besitzen zwar alle eine gewisse Plausibilität, aber dieser steht auf der anderen Seite eine gewisse Willkür des Autors (oder zumindest des Programmierers) gegenüber, der darüber entscheidet, welches er wirklich verwenden will. Eine zwingende Notwendigkeit, ein bestimmtes Kriterium als verbindlich anzusehen, besteht aber nicht. Eine Konsequenz davon ist wieder, daß Untersuchungen zum selben Problem in dieser Hinsicht nicht gut vergleichbar sind. Ist aber schon für die Bestimmung der Anzahl der zu extrahierenden Faktoren kein allgemein akzeptiertes Kriterium vorhanden, so wird sich das auch bei der Interpretation der einzelnen FAs auswirken. Das bedeutet, daß die Anzahl der für interpretierbar gehaltenen Faktoren zum Großteil der Willkür des Untersuchers anheim gestellt ist.

3.2.3. Rotationsmethoden

Hier ist ein Trend zur Beschränkung auf ein einziges analytisches Rotationskriterium festzustellen: Mit nur wenigen Ausnahmen wird das Varimax-Kriterium KAISERS (ÜBERLA, 1968, S. 209) verwendet. Diese Konzentration auf nur ein Rotationskriterium ist bedauerlich, denn z. B. wird durch dieses

Verfahren nur dann eine Einfachstruktur erreicht, wenn dies durch eine orthogonale Transformation überhaupt möglich ist. CATTELL (1966, S. 186) nimmt an, daß dies eigentlich nie der Fall sei und Orthogonalität und Einfachstruktur von da her als Kontradiktionen zu betrachten seien. Er kritisiert analytische Kriterien insbesondere in Hinblick auf die Replizierbarkeit faktorenanalytischer Ergebnisse. „And it remains true that three-fourth of publicated factor analyses are demonstrably nowhere near simple structure to the field they are intended to clarify. What these casual and unworkmanlike studies have actually done in the last decade is to create an atmosphere of pointlessness and disillusionment by accumulating a junk heap in which factors can rarely be matched from any one research to another“ (1966, S. 186).

Das Varimax-Verfahren verleitet wegen seiner Praktikabilität zur Anwendung; der Untersucher braucht sich hier um das vorprogrammierte Rotationsverfahren nicht zu kümmern und bekommt ein fertiges Ergebnis ins Haus geliefert. Auf eine interessante Variante sei jedoch verwiesen: Man kann z. B. so vorgehen, daß man nicht alle Variablen in das analytische Rotationskriterium einbezieht, sondern die Rotation von nur bestimmten Variablen durchführt und anschließend daran die Ladung der übrigen Variablen auf den rotierenden Faktoren bestimmt (TIMM, 1971). Dieses Verfahren ist z. B. dann von Interesse, wenn einige Variablen als Markierungsvariable mit in die Analyse aufgenommen wurden, um vielleicht einen Interpretationsansatz für die gefundenen Faktoren zu bekommen, oder wenn man abgeleitete Maße in eine Analyse aufnehmen will, z. B. Zwischensummen von Testitems, die man aber bei der Bestimmung der Stellung der Rotation ausschalten muß, will man nicht Artefakte produzieren.

Andere analytische Rotationskriterien kommen nur selten zur Anwendung, das Oblimin-biquartimin-Kriterium wurde nur von JÄGER und TODT (1964) und JÄGER (1964) verwendet, das Equimax-Verfahren von KATZENBERGER (1965). Besonders bedauerlich ist dabei, daß schiefwinkelige Rotationsmethoden, von denen ja angenommen wird, daß sie gewissermaßen eine „natürlichere“ oder zumindest besser interpretierbare Lösung bringen, dadurch ins Hintertreffen geraten (KEREKJARTO und SCHMIDT, 1962; KRAAK, 1961; CATTELL und SAUNDERS, 1954). Manchmal findet man auch Arbeiten, welche mit dem Hinweis versehen sind, daß sich die schiefwinkeligen Lösungen z. B. von einer Varimax-Lösung nicht wesentlich unterscheiden und daß man daher nicht auf sie einzugehen brauche (FÜRNRATT, 1969b; 1968b, c; MAXEINER, 1968). Die analytischen Rotationsverfahren haben wegen ihrer leichteren Anwendbarkeit auch die älteren graphischen oder kriteriumsorientierten Rotationsmethoden verdrängt (HUNZIKER, 1964; SEIFERT, 1964; PAWLIK, 1963; LIENERT, 1963; MERZ, 1961; 1958; ROTH, 1957). Diese Verfahren weisen aber bisweilen eine leichtere Interpretierbarkeit auf und erlauben auch eher eine Einordnung in den Zusammenhang mit anderen Untersuchungen (SCHEIER und CATTELL, 1965; HEDINGER, 1965; EYFERTH, 1963; SCHAEDELI, 1961; WEBER, 1953).

3.3. Varianzaufklärung durch die FAs

Daß die Beurteilung der Ergebnisse einer FA u. a. von der Vollständigkeit der Berichterstattung abhängt, wurde bereits erwähnt. Zur Vollständigkeit gehört aber auch, daß angegeben wird, welcher Anteil der Variabilität der Daten durch eine FA (u. zw. nach der Rotation) aufgeklärt wird. In einem zweiten Schritt ist es erst dann von Wichtigkeit anzugeben, welchen Anteil der Varianz an der Gesamtkommunalität durch die einzelnen Faktoren abgedeckt wird, um damit zu einer Beurteilung der relativen Bedeutsamkeit und Wichtigkeit einzelner Faktoren zu kommen.

3.3.1. Aufklärung der Gesamtvarianz

Wieviel Varianz insgesamt durch die FA aufgeklärt wurde, ist nur in rund einem Drittel der Arbeiten angegeben. Da die Aufklärung der Gesamtvarianz aber auch ein Kriterium dafür ist, ob der Einsatz einer FA überhaupt sinnvoll war, könnten die eher spärlichen Angaben zu diesem Punkt zu der Vermutung Anlaß geben, daß verhindert werden soll, eine solche Frage überhaupt aufkommen zu lassen.

Ergänzt man diese Angaben – dies ist natürlich nur dann möglich, wenn das Ladungsmuster vollständig angegeben wird oder zumindest die Kommunalitäten berechnet worden sind – so zeigt sich, daß die Analysen, in denen mehr als 80% der Gesamtvarianz durch die gemeinsamen Faktoren aufgeklärt wird nur etwa 7,5% der Untersuchungen ausmachen. Dieses Kriterium wurde gewählt, weil es ja auch ein allgemeines Abbruchkriterium für eine FA darstellt (BUSZ *et al.*, 1972; FRANKE und BORTZ, 1972; BORTZ, 1972; BORTZ, 1971b; KOPFMANN, 1969; COHEN und SCHÜMER, 1969; TAUSCH *et al.*, 1967; HOFSTÄTTER, 1962; KENTLER, 1959; MERZ, 1958; HOFSTÄTTER, 1956). In dem Gros der Untersuchungen liegt der aufgeklärte Varianzbetrag zwischen 50 und 80%; aber auch die Fälle, in denen weniger als 30% der Varianz durch gemeinsame Faktoren abgedeckt werden kann, sind nicht allzu selten anzutreffen (SCHNEIDER und MINKMAR, 1972; ANGERMAIER, 1972; BOTTENBERG und WEHNER, 1972; WENDELER, 1971; FISCH und SCHMALT, 1970; WENDELER, 1969; SEITZ und LÖSER, 1969; MEYER und GOLLE, 1966; HARTMANN und ENGELMANN, 1966; FAHRENBERG, 1965; JÄGER und TODT, 1964; LIENERT und FABER, 1963; FISCHER und BUTSCH, 1961).

3.3.2. Größe der Kommunalitäten

Die Kommunalitäten geben den Anteil der Varianz der einzelnen Variablen wieder, der durch gemeinsame Faktoren aufgeklärt werden konnte. Es ergänzt das vorher gefundene Ergebnis, wenn man feststellen kann, daß Werte zwischen .02 und .25 keine Seltenheit sind. Die mitgeführten Variablen lassen sich also nicht immer befriedigend in dem gemeinsamen Faktorenraum darstellen. Durch diese beiden Tatsachen – geringe Aufklärung der Gesamtvarianz und der Varianz der einzelnen Variablen – wird das Argument, das

KALVERAM (1970d, S. 112) gegen die FA ins Spiel bringt, unterstützt: Es gelingt nämlich in vielen Fällen durch eine FA nicht, eine ökonomischere Beschreibung der Ausgangsdaten zu liefern. Wenn ein solcher Anspruch erhoben wird, dann wird allzuoft dabei vergessen, daß willkürlich die sehr beträchtliche spezifische Varianz der Variablen unter den Tisch fallen gelassen wird.

4. Anwendungsaspekte von FAs

Als letztes soll überprüft werden, ob die Ansprüche, die mit dem Einsatz von FAs einhergehen, auch tatsächlich erfüllt werden können. Es sollen hier also Gesichtspunkte besprochen werden, ob und in welcher Weise die Ergebnisse faktorenanalytischer Forschung theoretisch bedeutsam oder zumindest praktisch relevant sein können. Einschränkend muß dabei gesagt werden, daß hier kein vollständiges Bild aller möglichen Anwendungen von FAs gegeben werden kann.

4.1. Anwendungen in der Psychologischen Diagnostik

Es wäre interessant, mehr über die Art der Urteilsbildung von Psychologen bei Verwendung von diagnostischen Tests zu wissen, über die Berücksichtigung verschiedener Testkriterien, die Art ihrer Gewichtung und ihrer kognitiven Verarbeitung. Eine wesentliche Beurteilungsdimension dürfte der „Informationsgehalt“ eines Tests sein; dabei meint man mit „Informationsgehalt“ in einer inkorrekten Anlehnung an informationstheoretisches Wissen, indem „stochastisch unabhängig“ und „orthogonal“ nicht unterschieden werden, insbesondere die Zahl unabhängiger Dimensionen, über die in einem Test Aussagen gemacht werden können. Reliabilität und Validität bilden bei diesem Prozeß der Urteilsbildung vielleicht gewisse einschränkende Minimalbedingungen.

4.1.1. Die Verwendung der FA zur Itemselektion

Die FA scheint diesem Prozeß der Urteilsbildung entgegenzukommen. Sie findet häufig bei der Testkonstruktion Anwendung und liefert Ergebnisse, die sich relativ leicht in die eben beschriebene Denkweise integrieren lassen. Aus einem Item-Pool sollen Items so zusammengestellt werden, daß jeweils eine Gruppe von Items ein und dieselbe Verhaltensdimension mißt. Die Gesamtheit der Items soll ein breites Band von Verhaltensdimensionen umfassen, die einzelnen Dimensionen sollen aber nach Möglichkeit gegenseitig orthogonal sein. Das übliche Verfahren dabei ist, daß aus der Iteminterkorrelationsmatrix Faktoren berechnet werden und die Items, die auf einem Faktor „substantiell“ laden, zu Subskalen zusammengefaßt werden, die somit definitionsgemäß wenig miteinander korrelieren sollen. Diese so gebildeten Subskalen können anschließend den üblichen Prozeduren der Testkonstruktion unterworfen werden (SCHNEIDER und MINKMAR, 1972; HEINERTH, 1972; FISCH und

SCHMALT, 1971; WENDELER, 1971; LÜER *et al.*, 1970; SEITZ und BRÄTH, 1970; WENDELER, 1969; WIENDIECK und LÜCK, 1969; FÜRNRATT, 1968a, b; BUGGLE *et al.*, 1968; LÜER *et al.*, 1966; MEYER und GOLLE, 1966).

Wesentlich dafür scheint eine implizite Annahme zu sein: Die erhaltenen Subskalen sollen Faktoren darstellen, die Punktwerte auf den Subskalen somit Faktorenwerte! Dies kann aus mindestens drei Gründen nicht zutreffen: (1) Der Punktwert auf einer Subskala entsteht in der Regel durch einfaches Auszählen von Itembeantwortungen, d.h. durch Gleichgewichtung. (2) In der Regel geht ein Item nur in einen Subtest ein, d.h. es werden „non-overlapping“ Faktoren angenommen. (3) Die Normierung der Subtests in der Eichstichprobe, die in der Regel nach einer FA vorgenommen wird, transformiert die ursprünglichen Subskalen. Dazu ein Beispiel: In der sicherlich im Vergleich zu vielen anderen Tests als sorgfältig anzusehenden Konstruktion des FPI wurde u.a. eine FA mit zunächst 240 Items durchgeführt (FAHRENBERG und SELG, 1970). Nach Extraktionen von neun Faktoren nach der Hauptkomponenten-Methode — die Zahl der Faktoren wurde insbesondere durch den Scree-Test bestimmt — wurde eine orthogonale Rotation nach dem Varimax-Kriterium durchgeführt. Die Kommunalitäten der Items sind sehr niedrig; die höchste beträgt .41. Die Faktoren klären 27% der Gesamtvarianz auf. Wie werden nun die neun Hauptskalen für das Testprofil erstellt? Ein Rohwert ergibt sich aus der Summe der Items, die jeweils ihre höchste Ladung bei dem entsprechenden Faktor aufweisen. Die Skala 9 (Offenheit) setzt sich aus 12 Items zusammen, deren höchste Ladung .44 und deren niedrigste .23 ist. Auch wenn CATTELL (1966, S. 195) meint, wir sollten nicht überrascht sein, „that important factors never have loadings above 0.3 to 0.4“, erscheinen uns die Ladungen im vorliegenden Kontext niedrig. Die Summenbildung allerdings soll die Bildung von Faktorwerten ersetzen. Dies wird zwar nirgends explizit erwähnt, aber warum sollte man sonst Punktwerte für die zunächst faktoranalytischen neun Skalen ermitteln? Neben der Gleichgewichtung der Items müßte zudem ein „non-overlapping“ Modell (CATTELL, 1966, S. 214) angenommen werden, da jedes Item nur in einen Subtest miteingeht. In der Folge ergibt sich auch, daß trotz orthogonaler Rotation die durch einfache Summenbildung und Gleichgewichtung entstandenen 9 Skalen beiinsgesamt 36 möglichen nicht weniger als 30 signifikante Interkorrelationen aufweisen (ähnliche Ergebnisse für faktoranalytisch konstruierte Tests konnten auch BARTUSSEK *et al.* (1972) nachweisen). Die Interkorrelationsmatrix der 9 Subtests wurde erneut einer FA unterzogen. Warum werden aber nicht die multiplen Regressionskoeffizienten angegeben, mit denen die Faktorwerte bestimmt werden könnten? Gewiß, die Testauswertung ist dann nicht mehr durch einfaches Auszählen von Itembeantwortungen möglich. Da das Testmanual aber sowieso eine Computer-Auswertung als Option anbietet, wäre eine differenzierte Auswertung auch für den „Testkonsumenten“ möglich, der nicht von dieser Möglichkeit Gebrauch machen will.

Bei FAs auf Itemniveau ist die erklärte Gesamtvarianz meist außerordentlich gering. Man findet Prozentsätze bis hinan zu 20%. „While it is common in factoring intercorrelations of scales with high reliabilities or communalities to account for 70–80% variance, the present figures should be interpreted with reference to the relatively lower reliabilities and intercorrelations characteristic of true–false items“ (LUNNEBORG, 1972, S. 314). Diese Begründung ist sicherlich richtig, doch sollte man aus den Umständen die Konsequenzen ziehen, sonst ist die FA eine unbeschränkte Methode, die man immer anwenden kann und bei der immer etwas herauskommt. Eine nach einer FA auf Itemniveau vorgenommene Summenbildung widerspricht natürlich auch offen den Modellannahmen einer FA. Es sei in diesem Zusammenhang aber auf das Rasch-Modell verwiesen, mit dem ausdrücklich nachgeprüft wird, ob Summenwerte erschöpfende Statistiken darstellen (FRICKE, 1972, S. 49).

Eine ähnliche Zielsetzung von FAs liegt auch dann vor, wenn untersucht wird, ob ein diagnostisches Verfahren für verschiedene Populationen dieselbe Struktur besitzt. Eine methodische Variante dieses Vorgehens ist, daß für verschiedene Stichproben FAs gerechnet werden (BUGGLE *et al.*, 1968) und diese dann durch Ähnlichkeitsrotation zur maximalen Übereinstimmung gebracht werden (FISCHER und ROPPERT, 1965; KRISTOF, 1964). Auch hier besteht bisweilen das Ziel darin, Items zu selektieren, die in verschiedenen Stichproben nicht dieselbe Struktur aufweisen (SEITZ und BRÄTH, 1970; MEYER und GOLLE, 1966). Ein möglicher Einwand gegen ein solches Vorgehen ist, daß die FA keine „harten“ Kriterien liefert, an Hand derer man überprüfen könnte, inwieweit Items nicht dieselbe Struktur aufweisen. Die FA läßt sich daher in die Reihe der „weichen“ Verfahren (FRICKE, 1972, S. 38; FISCHER, 1968, S. 207) der sogenannten klassischen Testtheorie einordnen.

4.1.2. Aufklärung der faktoriellen Struktur von Testmaßen

In dem Zusammenhang mit der Testkonstruktion und insbesondere der Testvalidierung scheint es auch eine legitime Frage zu sein, die Bedeutung der Durchführung von FAs zu überprüfen, die über Testmaße (z. B. Items, Subskalen, Indikatoren aus dem Rorschach-Test oder dem TAT, Kriteriumsmaßen, usw.) gerechnet werden, und die nicht dem Zweck dienen, Items zu selektieren oder zu gruppieren. Folgen solche praktischen Konsequenzen nicht nach, so besteht das Resultat solcher Analysen zumeist in der Feststellung, daß ein Testverfahren vorliege, das im Sinne der FA als „violdimensional“ zu bezeichnen sei (HEINERTH, 1972; SCHNEIDER und MINKMAR, 1972; JOHN und KEIL, 1972; BOTTENBERG und WEHNER, 1972; SEYDEL, 1972; ALTHOFF, 1971; OTTE, 1970; WEHNER und BOTTENBERG, 1969; BOTTENBERG, 1968; BALZERT, 1968; BÄUMLER *et al.*, 1968; LOCKOWANDT, 1968; KLEBELSBERG *et al.*, 1968; SELG und LISCHKE, 1966; FAHRENBERG, 1965; BÄUMLER, 1964; SEIFERT, 1964; KALLINA, 1964; FISCHER, 1964; KLEBELSBERG, 1963; EYFERTH, 1963; KEREKJARTO und SCHMIDT, 1962; FISCHER und BUTSCH, 1961; LIENERT, 1958a, b).

Der Feststellung der „Mehrdimensionalität“ eines Verfahrens kommt aber nur dann Bedeutung zu, wenn nun auch die Punktwerte der Vpn auf der Faktorebene berechnet werden. Werden für die Vpn aber keine Faktorwerte berechnet, dann handelt es sich um eine weiter nicht anwendbare Feststellung, denn die Befunde können nicht auf Faktorebene diskutiert werden, wie dies durch die „Interpretation“ der Faktoren nahegelegt wird. Hier wird also dasselbe wiederholt, was sich bereits auf Itemebene abgespielt hat. KALVERAM beschreibt dies (1970d, S. 110), wenn er kritisiert, daß die FA durch ihre Modellvoraussetzungen bestimmte Größen definiert, die auf einer Ebene liegen, die nicht direkt der Messung zugänglich ist. Nachdem diese Größen aber definiert sind, wird darauf vergessen, die Ausstattung der Vpn mit diesen Faktoren festzustellen. In diesem Sinn sind die vorliegenden Analysen also unvollständig, man könnte aber auch sagen „sinnlos“.

4.1.3. FA und empirische Validierung von Testverfahren

Erfüllt ein Test grundlegende teststatistische Kriterien, so stellt sich die Frage nach der „empirischen Validität“ (HÖRMANN, 1964) dieses Verfahrens. Ein beliebtes Vorgehen dabei ist, korrelative Beziehungen mit Außenkriterien zu suchen. Diese können selbst wieder andere Testverfahren sein oder auch andere Validierungsmaße; LIENERT (1967) spricht dabei einmal von „innerer“ und das andere Mal von „äußerer“ Validität. Diese korrelativen Beziehungen werden dann wieder faktoranalysiert (BOTTENBERG und WEHNER, 1972; HAAS und LOEWER, 1971; SEITZ, 1971; SASSENSCHIEDT und BUGGLE, 1970; FISCH und SCHMALT, 1970; SEITZ und LÖSER, 1969; FÜRNTTRATT, 1968a; BOTTENBERG, 1968; TIMM, 1968; FAHRENBERG und MYRTEK, 1966; FAHRENBERG und CONRAD, 1965; FISCHER, 1964; BÄUMLER, 1964; FISCHER und BUTSCH, 1961; FISCHER, 1958; LIENERT, 1958a, b; CATTELL und SAUNDERS, 1954). Im Prinzip analog ist dabei das Verfahren, wenn überhaupt Beziehungen zwischen zwei Datenmengen mittels einer FA untersucht werden (BORTZ, 1971b; KOPFMANN, 1969; LÜER und NEUFELDT, 1968; KALLINA, 1964).

Die Begründung für ein solches Vorgehen sucht man darin zu finden, daß man die Stellung eines oder mehrerer Tests im diagnostischen Meßfeld abklären will, daß ein Überblick über das nomologische Netzwerk – allerdings nur korrelativer Beziehungen – gefunden werden soll oder einfacher, daß man abklären will, ob der gerade interessierende Test mit anderen Tests oder Außenkriterien gemeinsame Varianz besitzt. Es ist dabei allerdings zu fragen, ob sich solche Probleme nicht eher durch Inspektion der ursprünglichen Korrelationsmatrix beantworten ließen oder indem partielle, multiple oder kanonische Korrelationen gerechnet werden. Stellt man aber darüber hinaus den Anspruch, daß noch mehr erfaßt wird, etwa hypothetische Einflußgrößen im Sinne von hypothetischen Konstrukten oder intervenierenden Variablen, so kann man dem – ähnlich wie im vorigen Punkt – mit KALVERAM (1970d, S. 110) entgegenhalten, daß „die zahlreichen Veröffentlichungen der

Ergebnisse von Faktorenanalysen tatsächlich nur Mitteilungen von Eichkoeffizienten ganz spezieller Meßinstrumente (sind), nämlich der in einer bestimmten Situation an bestimmten Probanden angewendeten Testbatterie; wobei jedoch nach erfolgter Eichung auf die Durchführung der Messungen selbst verzichtet wurde . . .“.

4.2. FA und psychologische Theorienbildung

4.2.1. Veränderungen der „Struktur“ von Teilsystemen der Persönlichkeit

Unter der Annahme, daß es möglich sei, durch Verwendung von FAs zu den „Bausteinen“ der menschlichen Persönlichkeit vorzustoßen, ist es nur folgerichtig, Verschiedenheiten faktorenanalytischer Untersuchungen als Änderung der Zusammensetzung dieser Bausteine zu interpretieren. Unwillkürlich wird man dabei an das Beispiel der Chemie erinnert, wobei man sich die chemischen Elemente als Analoga zu den Faktoren vorstellen könnte. Allerdings trägt dieser Vergleich nicht allzu weit, denn er reicht nur bis zu einer sehr äußerlichen Ähnlichkeit; Analogien zu irgendwelchen Verbindungsgesetzlichkeiten sind offenbar schon unplausibel. Untersuchungen, welche diesem Denkschema in etwa folgen, befassen sich beispielsweise mit Veränderungen der Intelligenzstruktur in Abhängigkeit von anderen Bedingungen, z. B. des Neurotizismus (HOBİ und KLÄR, 1972; COHEN und WITTEMANN, 1967; LIENERT, 1963), des Lebensalters (STEINHAGEN, 1970; LIENERT und FABER, 1963; RIEGEL, 1960), dem Grad der Intro- oder Extraversion (BALZERT, 1968) oder auch des Intelligenzniveaus selbst (JÄGER, 1964; LIENERT und FABER, 1963). Solche Fragestellungen lassen sich natürlich auch für andere Teilbereiche der menschlichen Persönlichkeit aufstellen, wie dies z. B. von SIDDIQI und KEIL (1970) getan wird, wenn diese die Altersdifferenzierung des Neugierverhaltens mit faktoranalytischen Methoden untersuchen.

Ist man allerdings nicht geneigt, der oben angedeuteten Interpretation faktorenanalytischer Ergebnisse zu folgen, dann erhalten diese Untersuchungen ein anderes Gewicht. Als Kritik und Erklärung läßt sich nämlich dann anführen, daß die Ergebnisse z. B. auf die verschiedenen Stichprobenzusammensetzungen und Stichprobenhomogenisierungen zurückzuführen sind. Wird aber behauptet, daß eine Interpretation auf einer solchen reistischen Basis nie angestrebt wurde, so bleibt zu fragen, welchem Zweck diese Analysen überhaupt dienen oder welchen Erkenntniswert sie haben. Offenbar sind die „sichersten“ Aussagen aufgrund der Korrelationsmatrix allein zu machen, während hingegen die Zahl der Faktoren unmittelbar auf die mittlere Größe der Korrelationskoeffizienten zurückzuführen ist.

4.2.2. Identifikation von Intelligenzfaktoren

Will man sich gegen den Vorwurf schützen, daß FAs nur Artefakte produzieren, so ist die Reproduzierbarkeit von Ergebnissen als eine Minimalforderung zu stellen. Dabei ist noch immer zu bedenken, daß sich auch Artefakte

reproduzieren lassen können. Die Kontinuität in der Verfolgung einzelner Forschungsrichtungen könnte dabei als Prüfstein für die Fruchtbarkeit einzelner Ansätze angesehen werden. In diesem Zusammenhang sollte der methodische Hinweis von ÜBERLA (1968, S. 358), die FA als hypothesentestende Methode einzusetzen, indem man z. B. das hypothetisch erwartete Faktorenmuster vor Durchführung einer FA aufstellt, erwähnt werden.

Einer der wenigen Bereiche, in welchem die Anwendung von FAs Tradition besitzt, ist die Intelligenzforschung. Hierbei sind die Arbeiten zu nennen, die der Bestätigung der Intelligenzfaktoren von MEILI dienen sollten (HEDINGER, 1965; HUNZIKER, 1964; SCHAEDELI, 1961); ferner die Arbeiten, die sich an dem Intelligenzmodell von THURSTONE orientieren (JÄGER und TODT, 1964; JÄGER, 1964; FISCHER, 1958) oder schließlich dem von WECHSLER (SEYDEL, 1972; EYFERTH, 1963). Daneben finden sich noch einige Untersuchungen, die der Identifikation von einzelnen Intelligenzfaktoren gewidmet sind (KRAAK, 1961; ROTH, 1957; WEBER, 1953). Die gezeigten Ergebnisse lassen sich allerdings schwer einer Würdigung unterziehen, denn die zu anderen Untersuchungen gezogenen Parallelen liegen auf einer rein verbalen Ebene. Stringentere Vergleiche, welche beispielsweise auf der Verwendung analytischer Ähnlichkeitsrotationen beruhen könnten, sind in diesem Bereich leider nicht zu finden. Es kommt bei diesen Arbeiten außerdem mit ins Spiel, daß die verschiedenen Analysen mit unterschiedlichen Variablensätzen durchgeführt worden sind und daß verschieden viele Faktoren für bedeutsam erachtet wurden. Damit wurde aber auf der Interpretationsebene der scheinbar so exakte Ansatz völlig aufgegeben, zugunsten von intuitiven Plausibilitätsüberlegungen.

4.3. Zusammenfassung

Bei den untersuchten Anwendungsfällen faktorenanalytischen Vorgehens konnte eine Reihe zu kritisierender Punkte aufgewiesen werden. Diese betreffen (1) einmal Probleme, die allgemein mit empirischen Arbeiten verbunden sind (z. B. unrepräsentative Probandenstichproben oder ungeklärtes Datenniveau), (2) zum anderen aber Punkte, die für die FA spezifisch sind (z. B. Zusammensetzung der Variablenstichprobe, Kommunalitätsschätzungen und Vergleiche mit der Reliabilität der Daten). Für die Nachvollziehbarkeit faktorenanalytischer Ergebnisse stellte es sich ebenfalls als sehr nachteilhaft heraus, daß einzelne Analysen unvollständig berichtet werden. Dadurch geht auch die Kontinuität der Forschung, die sich solcher Methoden bedient, verloren.

Die Anwendungsfälle zeigten ebenfalls, daß die Möglichkeiten, die in dem faktorenanalytischen Ansatz stecken, rein methodisch gesehen bei weitem nicht ausgenutzt werden. Aus „ökonomischen“ Gründen, die aber keineswegs durch die Sache selbst bedingt sind, findet eine Beschränkung auf wenige gängige Analyse- und Rotationsmethoden statt. Vor allem auf schiefwinkelige Rotationsverfahren oder auf kriteriumsorientierte, denen aber aus theoretischen

schen Überlegungen der Vorzug gegeben werden muß und die sich auch in Modelluntersuchungen bewährt haben, wird fast vollständig verzichtet. Die Interpretation faktorenanalytischer Ergebnisse wird wieder vorwiegend von Plausibilitätsüberlegungen geleitet, die ihrerseits mit den analytischen Kriterien, welche gerade den Exaktheitsanspruch der FA sichern soll, nicht im Einklang stehen. Damit wird aber auf der Interpretationsebene der mit der Verwendung mathematischer Modelle einhergehende Gültigkeitsanspruch wieder aufgegeben.

Diese Punkte der Kritik stellen dabei aber noch nicht den faktoranalytischen Ansatz selbst in Frage. Man könnte allerdings zu der Überzeugung kommen, daß die in der Spezifikationsgleichung zum Ausdruck kommende Annahme einer additiven Merkmalswirkung zu simpel ist, so daß die Hoffnung durch ein solches Modell zu einer adäquaten Abbildung der Wirklichkeit zu kommen, von vornherein ausgesprochen gering ist. Hinzu kommt auch noch, daß die üblichen FAs (ausgenommen sind Tucker-Analysen) zumeist Querschnittsuntersuchungen darstellen, d.h. sie lassen Entwicklungsgesetzmäßigkeiten von vornherein außer acht. Gerade aber auch diese Einschränkung auf statische Zustandsbeschreibungen engt die Brauchbarkeit dieses Ansatzes erheblich ein.

5. Schlußbemerkung

MULAIK (1972, S. 9) unterscheidet zwei Anwendungsformen der FA: (1) die blinde FA, bei welcher der Faktoranalytiker gleichsam automatisch auf die latenten Schlüsselvariablen vorstößt und (2) die hypothesentestende FA, bei der durch prokruste Transformationen spezifische Strukturhypothesen überprüft werden. Die erste Anwendungsform, die in den 50er und 60er Jahren dominierte (im deutschen Sprachraum auch noch in den nachfolgenden Jahren), dürfte relativ wenig zur Theorienbildung beigetragen haben. „In this period factor analyses was frequently applied agnostically as regards structural theory to all sorts of data, from personality-rating variables, Rorschach-test-scoring-variables, physiological variables, semantic differential variables, and biographical-information variables . . . In all these applications the hope was that factor analyses could bring order and meaning to the many relationships between variables“ (MULAIK, 1972, S. 9). Der FA käme hier ein kreatives Moment zu, insofern sie nicht „nur“ Hypothesen prüft, sondern sogar zu neuen „Begriffen“ führt: „It (die FA) is thus a means of *creating* concepts, not merely of employing them or checking their fit to new data“ (CATTELL, 1966, S. 174). Der FA wird hier also nicht nur eine beschreibende Funktion zugesprochen, sondern darüber hinaus eine erklärende. „Obviously, the central aim of factor analyses is to represent or explain observed vocariational relations . . .“ (CATTELL, 1966, S. 174). Sie wird zur „direktesten, universellst anwendbaren und repräsentativsten“ wissenschaftlichen Methode (CATTELL, 1966, S. 175). MULAIK hingegen meint, man könne zwar über die erklärende Funktion der

FA diskutieren, in den Einzelbereichen hat die blinde FA jedoch nicht zu einer sinnvollen Erklärung der den beobachteten Variablen zugrunde liegenden Strukturen geführt. Beispiele sind etwa der Rorschach-Test, Ratings von Persönlichkeitseigenschaften, bei der die FA zu einer Vermischung von Rater- und Ratee-Prozessen führt, oder biographische Daten, aus deren Analyse keine brauchbare Theorie über Lebenslaufdaten hervorgeht.

Im Zusammenhang mit psychometrischen Testfragen schuf SUTCLIFFE (1965) die Bezeichnung „platonian true score“. Sie läßt sich sofort auf die oben angesprochene reistische Interpretation von Faktoren übertragen. Die Intention mit der FA auf die wirklich existenten wesentlichen Grunddimensionen eines Sachbereiches zu stoßen, kommt in den dazu geplanten Plasmodienstudien recht gut zum Ausdruck. Sieht man aber, mit welcher Mühe etwa Variable zu Kraftfahrzeug-Faktoren (BAUMANN, 1971) führen oder zu welcher absurden Faktoren Gewichtskorrelationen verschiedener chemischer Zusammensetzung führen (JONES, 1960; zit. nach MULAIK, 1972, S. 364), so kann man sich die Konsequenzen eines entsprechenden methodischen Vorgehens in weitgehend unerforschten Bereichen vorstellen. Die faktorenanalytischen Methoden zur Überprüfung expliziter Strukturhypothesen wurden erst während der letzten 10 Jahre entwickelt. Ernst-Untersuchungen liegen noch kaum vor. Das Arbeiten mit ihnen dürfte weniger wegen des mathematischen Aufwandes schwieriger werden (wenn einmal das Stadium des Programmwurfes für die Rechenanlagen vorbei ist) als wegen der notwendigen Explizitheit der zu formulierenden Hypothesen. Der teilweise tautologische Charakter faktorenanalytischer Ergebnisse dürfte hier jedoch wegfallen.

Abschließend soll zu den untersuchten Arbeiten noch angemerkt werden, daß der reistische Anspruch der FA mit ein Grund für die extensive Anwendung dieser Methode gewesen sein dürfte. Damit soll nicht unmittelbar etwas über die theoretische Bedeutsamkeit dieser Methode gesagt sein, sondern über die ihrer Anwendungen. EYSENCK schrieb dazu bereits 1953 (S. 108): „ . . . factor analyses is sometimes used as a last resort to try and rescue worthless data accumulated at random from the fate such data so richly deserve“. Wollte man böswillig sein, so könnte man auch sagen, daß es sich bei der Durchführung von FAs um Rituale handelt, die sich im Wissenschaftsbetrieb vor etlichen Jahren ausgebildet haben und die nur mehr der Befriedigung des mit solchen Ritualen verbundenen Wiederholungszwanges dienen. So kommt es auch, daß bisweilen der einzige Schluß, den man aus der Berichterstattung einer FA ziehen kann, darin besteht, daß der Autor Zugang zu einem Rechenzentrum hatte, an welchem ein Programm für die Durchführung einer FA vorhanden war.

„Die Psychologie hat (leider!) keine Patentrezepte“ (CARUSO, 1973, S. 16), es sei aber doch erwähnt, daß eine Reihe von Eigenheiten der faktorenanalytischen Methode eine gewisse Verwandtschaft zu einer empirischen Phänomenologie nahelegen: ihr explorativer Charakter, die Abneigung gegen expli-

zite Hypothesen, die „Unvoreingenommenheit“ bei der Erfassung eines breiten Variablenspektrums, das Suchen hinter den „Oberflächenvariablen“ oder die reistische Interpretation der „wesentlichen“ Faktoren. Der Unterschied besteht eigentlich nur in der verwendeten Sprache, die einmal verbal das andere Mal aber numerisch und mathematisch ist; in beiden Fällen ist aber eine ganze Menge nicht nachvollziehbarer Subjektivität enthalten.

6. Literaturverzeichnis

- AHRENS, H. J., C. MÖBUS, Zur Verwendung von Einstellungsmessungen bei der Prognose von Wahlentscheidungen. *Z. exper. angew. Psychol.* 1968, 15, 543–563.
- ALTHOFF, K., Die Voraussage des Berufserfolges von Regierungsinspektoren und Revierförstern – 2 Bewährungskontrollen. *Diagnostica* 1971, 17, 145–180.
- ANDERSON, N. H., Scales and statistics: Parametric and nonparametric. *Psychol. Bull.* 1961, 58, 305–316.
- ANGERMAIER, M., Faktorenstruktur eines Fragebogens von R. Tausch zur Beurteilung der Vorlesungstätigkeit. *Psychol. Rdsch.* 1972, 23, 183–194.
- ANKER, J. M., J. C. TOWNSEND, J. P. O'CONNOR, A multivariate analysis of decision making and related measures. *J. Psychol.* 1963, 55, 211–221.
- BALZERT, C., Untersuchungen über Zusammenhänge zwischen Extraversion/Introversion und der Faktorenstruktur der Intelligenz. *Z. exper. angew. Psychol.* 1968, 15, 195–238.
- BARTUSSEK, D., Zur Interpretation der Kernmatrix in der dreimodalen Faktorenanalyse von R. L. TUCKER. *Arbeiten aus dem Psychol. Inst. d. Univ. Hamburg*, Nr. 20, Hamburg 1972.
- BARTUSSEK, D., Eine Methode zur Bestimmung von Moderatoreffekten. *Diagnostica* 1970, 16, 57–76.
- BARTUSSEK, D., G. WEISE, B. HEINZE, Reliabilität und faktorielle Validität des deutschen 16 PF-Tests von CATTELL mit einer ausführlichen Analyse der Items. *Arbeiten aus dem Psychol. Inst. d. Univ. Hamburg*, Nr. 19, Hamburg 1972.
- BÄUMLER, G., Zur Faktorenstruktur der Paulitestleistung unter Berücksichtigung des sogenannten numerischen Faktors. *Diagnostica* 1964, 10, 107–119.
- BÄUMLER, G., W. SEITZ, G. KNOP, Die faktorielle Struktur der „Nicht verbalen Intelligenztestreihe“ von Snijders und Snijders-Omen bei Taubstummen. *Diagnostica* 1968, 14, 87–93.
- BAUMANN, U., *Psychologische Taxometrie*. Bern 1971 (Huber).
- BEYME, F., C. J. FAHRENBERG, Zur deutschen Bearbeitung des Anxiety-Tests von R. B. CATTELL. *Diagnostica* 1968, 14, 39–44.
- BINDER, A., Considerations of the place of assumptions in correlational analysis. In: E. F. HEERMANN; L. A. BRASKAMP (Hrsg.): *Readings in statistics for the behavioral sciences*. New Jersey (Prentice Hall) 1970 (338–349).
- BINDER, A., Further considerations of testing the null hypotheses and the strategy and tactics of investigating theoretical models. *Psychol. Rev.* 1963, 70, 107–115.
- BORTZ, J., Beiträge zur Anwendung der Psychologie auf den Städtebau. *Z. exper. angew. Psychol.* 1972, 19, 226–281.
- BORTZ, J., Möglichkeiten einer exakten Kennzeichnung der Sprechstimme. *Diagnostica* 1971a, 17, 3–14.
- BORTZ, J., Ein multivariater Ansatz zur Erfassung der Möglichkeiten einer diagnostischen Verwertung der Sprechstimme. *Z. exper. angew. Psychol.* 1971b, 18, 15–50.
- BOTTENBERG, E. H., Elementare und komplexe Geschlechtsunterschiede im Rorschachtest (Ro 30). *Diagnostica* 1972, 18, 180–184.
- BOTTENBERG, E. H., Exploration der Struktur des Rorschachtests (Ro 30). *Z. exper. angew. Psychol.* 1970, 17, 553–591.
- BOTTENBERG, E. H., Zur Diagnose von „Psychisch gestörtem Verhalten“ und „Hirnorganischer Schädigung“ mit Hilfe des Rorschachtests (Ro 30). *Praxis Kinderpsychol. Kinderpsychiatr.* 1968, 17, 82–87.
- BOTTENBERG, E. H., E. G. WEHNER, Suggestibilität: II. Einige persönlichkeits- und leistungsdiagnostische Korrelate des Würzburger Suggestibilitäts-Tests (WST). *Praxis Kinderpsychol. Kinderpsychiatr.* 1972, 21, 282–288.
- BRANDSTÄTTER, H., Zur Diagnose der Persönlichkeitsintegration aus der Handschrift. *Psychol. Rdsch.* 1969, 20, 159–172.
- BREDENKAMP, J., *Der Signifikanztest in der psychologischen Forschung*. Frankfurt a. M. 1972 (Akad. Vlg.ges.).
- BUGGLE, F., K. GERLICHER, F. BAUMGÄRTEL, Analyse einer deutschen Übersetzung des Junior Eysenck Personality Inventory (JEPI). *Diagnostica* 1968, 14, 3–18.
- BUSZ, M., R. COHEN, U. POSER, A. SCHÜMER, R. SCHÜMER, C. SONNENFELD, Die soziale Bewertung von 880 Eigenschaftsbegriffen sowie die Analyse der Ähnlichkeitsbeziehungen zwischen einigen dieser Begriffe. *Z. exper. angew. Psychol.* 1972, 19, 282–308.
- CARL, W., Eine Untersuchung zur Faktorenstruktur von Antworttendenzen („response sets“) bei Antwortskalen unterschiedlicher Stufenzahl. *Z. exper. angew. Psychol.* 1968, 15, 419–434.
- CARUSO, I. A., Kirche bunt, St. Pöltener Kirchenzeitung, 31. Sonntag im Jahreskreis, 28. Jahrgang, 1973, 16.
- CATTELL, R. B., The meaning and strategic use of factor analysis. In: R. B. CATTELL (Hrsg.): *Handbook of multivariate experimental psychology*. Chicago 1966 (McNally), 174–243.
- CATTELL, R. B., *Personality: a systematic, theoretical, and factual study*. New York 1950 (McGraw Hill).
- CATTELL, R. B., D. R. SAUNDERS, Beiträge zur Faktorenanalyse der Persönlichkeit. *Z. exper. angew. Psychol.* 1954, 2, 325–357.
- CATTELL, R. B., L. R. SCHMIDT, A. BJERSTEDT, Clinical diagnosis by the objective-analytic personality batteries. *J. Clin. Psychol.* 1972, 28, Special Monograph Supplement, 240–312.
- CATTELL, R. B., W. SEITZ, A. RAUSCHE, Zur Konsistenz der Persönlichkeitsstruktur von Kindern, gemessen durch Questionnaire-(Q-) Daten. *Z. exper. angew. Psychol.* 1971, 18, 513–524.
- CATTELL, R. B., F. W. WARBURTON, *Objective personality and motivation tests*. Chicago–London 1967 (Univ. of Illinois Press).
- COHEN, R., R. SCHÜMER, Die diagnostische Verarbeitung widersprüchlicher Informationen in Abhängigkeit vom Ausmaß der Widerspruchs. *Diagnostica* 1969, 15, 3–13.
- COHEN, R., K. WITTEMANN, Eine Untersuchung zur Abhängigkeit der Intelligenzstruktur vom Neurotizismus. *Z. exper. angew. Psychol.* 1967, 14, 71–88.
- DIRKS, H., Über den Standortsozialpsychologischer Forschung innerhalb der Betriebspsychologie. *Psychol. Rdsch.* 1961, 12, 168–187.
- DI VESTA, F. J., A developmental study of the semantic structures of children. *J. Verbal Learning Verbal Behavior* 1966, 5, 249–259.
- EDWARDS, W., Tactical note on the relation between scientific and statistical hypotheses. *Psychol. Bull.* 1965, 63, 400–402.

- EKMAN, G., Eine neue Methode zur Erlebnisanalyse. *Z. exper. angew. Psychol.* 1954, 2, 167–174.
- ELLIOTT, J. L., P. H. TANNENBAUM, Factor-structure of semantic differential responses to visual forms and prediction of factor scores from structural characteristics of the stimulus shapes. *J. Psychol.* 1963, 76, 589–597.
- ERTEL, S., Standardisierung eines Eindruckdifferentials. *Z. exper. angew. Psychol.* 1965a, 12, 22–58.
- ERTEL, S., Weitere Untersuchungen zur Standardisierung eines Eindruckdifferentials. *Z. exper. angew. Psychol.* 1965b, 12, 177–208.
- ERTEL, S., U. SCHINDLER, Intelligenzleistung und Testverhaltensdifferential. *Diagnostica* 1969, 15, 75–98.
- EYFERTH, K., Über die Abhängigkeit der Urteile über olfaktorische Qualitäten von der Empfindung der Reizintensität. *Z. exper. angew. Psychol.* 1965, 12, 209–222.
- EYFERTH, K., Gültigkeit und faktorielle Struktur des Hamburg-Wechsler-Intelligenztests für Kinder (HAWIK) bei seiner Anwendung auf schwachsinnige Erwachsene. *Diagnostica* 1963, 9, 11–26.
- EYFERTH, K., P. B. BALTES, Über Normierungseffekte in einer Faktorenanalyse von Fragebogendaten. *Z. exper. angew. Psychol.* 1969, 16, 38–51.
- EYSENCK, H. J., Das „Maudsly Personality Inventory“ als Bestimmer der neurotischen Tendenz und Extraversion. *Z. exper. angew. Psychol.* 1959, 6, 167–190.
- EYSENCK, H. J., The logical basis of factor analysis. *Amer. Psychologist* 1953, 8, 105–114.
- FAHRENBERG, J., Ein itemanalysierter Fragebogen funktionell-körperlicher Beschwerden (VELA). *Diagnostica* 1965, 11, 141–153.
- FAHRENBERG, J., W. CONRAD, Eine explorative Faktorenanalyse graphometrischer und psychometrischer Daten. *Z. exper. angew. Psychol.* 1965, 12, 223–238.
- FAHRENBERG, J., M. MYRTEK, Ein kritischer Beitrag zur psychophysiologischen Persönlichkeitsforschung. *Z. exper. angew. Psychol.* 1966, 13, 222–247.
- FAHRENBERG, J., H. SELG, Das Freiburger Persönlichkeitsinventar. FPI – Handanweisung. Göttingen 1970 (Hogrefe).
- FISCH, R., H.-D. SCHMALT, Vergleich von TAT und Fragebogendaten. *Z. exper. angew. Psychol.* 1970, 17, 608–634.
- FISCHER, G. H. (Hrsg.), *Psychologische Testtheorie*. Bern–Stuttgart 1968 (Huber).
- FISCHER, G. H., Zu Tuckers Methode der Faktorenanalyse von Lerndaten. *Psychol. Beitr.* 1967, 10, 136–146.
- FISCHER, G. H., Zum Problem der Interpretation faktorenanalytischer Ergebnisse. *Psychol. Beitr.* 1967, 10, 122–135.
- FISCHER, G. H., Zur faktoriellen Struktur der Handschrift. *Z. exper. angew. Psychol.* 1964, 11, 254–280.
- FISCHER, G. H., J. ROPPERT, Ein Verfahren zur Transformationsanalyse faktorenanalytischer Ergebnisse. In: G. H. FISCHER, J. ROPPERT (Hrsg.): *Lineare Strukturen in Mathematik und Statistik*. Wien–Würzburg 1965 (Physica).
- FISCHER, H., Ein Vergleich zwischen dem IST von Amthauer und dem PMA von Thurstone. *Diagnostica* 1958, 4, 25–32.
- FISCHER, H., C. BUTSCH, Musikalische Begabung und Intelligenz. *Z. exper. angew. Psychol.* 1961, 8, 508–518.
- FRANKE, J., J. BORTZ, Beiträge zur Anwendung der Psychologie auf den Städtebau. *Z. exper. angew. Psychol.* 1972, 19, 76–108.
- FRICKE, R., *Über Meßmodelle in der Schulleistungsdiagnostik*. Düsseldorf 1972 (Schwann).
- FRUCHTER, B., *Introduction to factor analysis*. Princeton–New York 1954 (Van Nostrand).

- FÜRNRATT, E., Ein Test zur Messung der kritischen Urteilsfähigkeit. *Diagnostica* 1968a, 14, 19–34.
- FÜRNRATT, E., Ein Fragebogen zur Erfassung emotionaler Reaktionen in sozialen Stress-Situationen. *Diagnostica* 1968b, 14, 121–130.
- FÜRNRATT, E., Die Messung auf menschliches Verhalten bezogener Wertvorstellungen. *Diagnostica* 1968c, 14, 145–155.
- FÜRNRATT, E., Faktorenanalyse inhaltlicher Motive in van Lennep-Bildgeschichten Jugendlicher. *Psychol. Rdsch.* 1969a, 20, 79–102.
- FÜRNRATT, E., Antworttendenzen in Fragebögen: 1. Bejahungs- und Varianztendenz. *Psychol. Rdsch.* 1969b, 20, 1–18.
- FÜRNRATT, E., Hauptachsenanalyse und Varimax-Rotation von Thurstones Primary Mental Abilities – Testbatterie. *Diagnostica* 1969c, 15, 161–177.
- FÜRNRATT, E., Zur Bestimmung der Anzahl interpretierbarer Faktoren in Faktorenanalysen psychologischer Daten. *Diagnostica* 1969d, 15, 62–75.
- GARRETT, H. E., *Statistics in psychology and education*. New York 1962 (McKay).
- GEHARDT, F., Über Ähnlichkeit von Faktorenmatrizen. *Psychol. Beitr.* 1967, 10, 591–599.
- GRANT, D. A., Testing the null hypotheses and the strategy and tactics of investigating theoretical models. *Psychol. Rev.* 1962, 69, 54–61.
- GRÖNER, R., Dimensionen der Wahrnehmung von Gesichtern. *Z. exper. angew. Psychol.* 1967, 14, 133–154.
- GUILFORD, J. P., When not to factor analyze. *Psychol. Bull.* 1952, 49, 26–37.
- GUILFORD, J. P., The correlation of an item with a composite of the remaining items in a test. *Educ. Psychol. Measmt.* 1953, 13, 87–93.
- GUILFORD, J. P., *Psychometric methods*. New York 1954 (McGraw-Hill).
- GUILFORD, J. P., The minimal Phi coefficient and the maximal Phi. *Educ. Psychol. Measmt.* 1965, 25, 3–8.
- HAAS, R., H. D. LOEWER, Das motivationspsychologische Signierungssystem nach Graf Hoyos in seiner Anwendung beim Vier-Bilder-Test nach van Lennep: Objektivität, klinische und faktorielle Validität. *Diagnostica* 1971, 17, 132–139.
- HARMAN, H. H., *Modern factor analysis*. Chicago 1960 (Univ. of Chicago Press).
- HARTMANN, K., W. ENGELMANN, Eine faktorenanalytische Untersuchung von Labilitätskriterien „erziehungsschwieriger“ männlicher Minderjähriger. *Praxis Kinderpsychol. Kinderpsychiatr.* 1966, 15, 19–23.
- HAYS, W. L., *Statistics*. London–New York 1963 (Holt).
- HEDINGER, U. K., Die Faktorenstruktur komplexer Denkaufgaben. *Z. exper. angew. Psychol.* 1965, 12, 337–403.
- HEINERTH, K., Prüfungsangst von Studenten. *Psychol. Rdsch.* 1972, 23, 80–90.
- HELM, J., Studien zur Korrelationsanalyse von Variablen der Leistungsmotivation. *Z. exper. angew. Psychol.* 1967, 14, 600–615.
- HERRMANN, T., K.-H. STAPF, Untersuchungen zum Erlernen von Wahlreaktionen, Teil II: Zur Theorie leistungsrelevanten Lernverhaltens. *Z. exper. angew. Psychol.* 1966, 13, 539–580.
- HERRMANN, T., A. STAPF, H. W. KROHNE, Die Marburger Skalen zur Erfassung des elterlichen Erziehungsstils. *Diagnostica* 1971, 17, 118–131.
- HÖBI, V., A. KLÄR, Zur Frage der Struktur und des Altersabbaues der Intelligenz bei Toxikomanen. *Diagnostica* 1972, 18, 159–179.
- HÖRMANN, H., *Aussagemöglichkeiten psychologischer Diagnostik*. Göttingen 1964 (Hogrefe).
- HOFSTÄTTER, P. R., Über Faktorenanalyse. *Arch. ges. Psychol.* 1938, 100, 223–279.
- HOFSTÄTTER, P. R., *Quantitative Methoden der Psychologie*. München 1953 (Barth).

- HOFSTÄTTER, P. R., Dimensionendes mimischen Ausdrucks. *Z. exper. angew. Psychol.* 1956, 3, 505–529.
- HOFSTÄTTER, P. R., Theoretische Grundlagen der Verhaltenssteuerung. *Psychol. Rdsch.* 1962, 13, 163–179.
- HOFSTÄTTER, P. R., Über sprachliche Bestimmungsleistungen: das Problem des grammatikalischen Geschlechts von Sonne und Mond. *Z. exper. angew. Psychol.* 1963, 10, 91–108.
- HUNZIKER, H. W., Plastizität als Faktor der Spannungsüberwindung in Denkaufgaben. *Z. exper. angew. Psychol.* 1964, 11, 185–237.
- JÄGER, A. O., Faktorenanalyse der WIT-Leistungen bei überdurchschnittlich intelligenten Probanden. *Diagnostica* 1964, 10, 45–53.
- JÄGER, A. O., E. TOBT, Zur Faktorenstruktur des Wilde-Intelligenztests (WIT) bei Siebzehnjährigen; Faktorenanalyse der WIT-Langform. *Diagnostica* 1964, 10, 3–14.
- JAKOBOVITS, L. A., C. E. OSGOOD, Connotations of twenty psychological journals to their professional readers. *Amer. Psychologist* 1967, 22, 792–800.
- JANSEN, G., H. HOFFMANN, Lärmbedingte Änderungen der Feinmotorik und Lästigkeitsempfindungen in Abhängigkeit von bestimmten Persönlichkeitsdimensionen. *Z. exper. angew. Psychol.* 1965, 12, 594–613.
- JOHN, D., W. KEIL, Selbsteinschätzung und Verhaltensbeurteilung. *Psychol. Rdsch.* 1972, 23, 10–29.
- KALLINA, H., Das Unbehagen in der Faktorenanalyse. *Psychol. Beitr.* 1967, 10, 81–86.
- KALLINA, H., Validitätsuntersuchung und Faktorenanalyse verkehrspsychologischer diagnostischer Methoden. *Z. exper. angew. Psychol.* 1964, 11, 56–70.
- KALVERAM, K. T., Probleme der Selektion in der Faktorenanalyse. III. Die „Invarianz“ von Faktorenlösungen unter Selektion. *Arch. Psychol.* 1970a, 122, 223–230.
- KALVERAM, K. T., Probleme der Selektion in der Faktorenanalyse. II. Kritik an einschlägigen Arbeiten. *Arch. Psychol.* 1970b, 122, 215–222.
- KALVERAM, K. T., Probleme der Selektion in der Faktorenanalyse. I. Die theoretische Behandlung der Selektion. *Arch. Psychol.* 1970c, 122, 199–204.
- KALVERAM, K. T., Über Faktorenanalyse. Kritik eines theoretischen Konzepts und seine mathematische Neuformulierung. *Arch. Psychol.* 1970d, 122, 92–118.
- KATZENBERGER, L., Dimensionen des Gedächtnisses. *Z. exper. angew. Psychol.* 1965, 12, 451–492.
- KEIL, W., M. SADER, Situationsspezifische Einflüsse auf die Leistungsmotivationsmessung nach Heckhausen. *Z. exper. angew. Psychol.* 1968, 15, 100–121.
- KEMPF, W. F., Zur Bewertung der Faktorenanalyse als psychologische Methode. *Psychol. Beitr.* 1972, 14, 610–625.
- KENDALL, M. G., *A course in multivariate analysis.* London 1965 (Griffin).
- KENTLER, H., Zur Problematik der Profilmethode. *Diagnostica* 1959, 5, 5–18.
- KEREKJARTO, M., G. SCHMIDT, Faktorenanalysen des Hamburg-Wechsler-Intelligenztests für Kinder (HAWIK). *Diagnostica* 1962, 8, 95–110.
- KLEBELSBERG, D., Ergebnis einer Verhaltensanalyse des Kraftfahrers. *Z. exper. angew. Psychol.* 1963, 10, 597–603.
- KLEBELSBERG, D., B. BIEHL, G. DERFLINGER, G. FISCHER, J. FUHRMANN, U. SEYDEL, Zweite Verhaltensanalyse von Kraftfahrern und Validitätsuntersuchung verkehrspsychologischer Untersuchungsmethoden. *Z. exper. angew. Psychol.* 1968, 15, 601–618.
- KLEITER, G., Krise der Signifikanztests in der Psychologie. *Jb. Psychol. Psychother.* 1969, 17, 144–163.
- KLIPPSTEIN, E., Eine Analyse der Rosenzweig P-F-Test-Situationen (Form für Kinder). *Z. exper. angew. Psychol.* 1972, 19, 444–459.

- KNABE, G., Multidimensionale experimentelle Analysen des Legasthenie-Syndroms. *Z. exper. angew. Psychol.* 1969, 16, 570–612.
- KOPFMANN, G., Zusammenhang zwischen Abiturnoten von Physikstudenten und ihre Leistung an der TH München. *Z. exper. angew. Psychol.* 1969, 16, 434–451.
- KRAAK, B., Räumliches Vorstellen als Voraussetzung technischen Denkens. *Z. exper. angew. Psychol.* 1961, 8, 78–109.
- KRANTZ, D. H., R. D. LUCE, P. SUPPES, A. TVERSKY, *Foundations of measurement.* Bd. 1. New York 1971 (Academic Press).
- KRISTOF, W., Die beste orthogonale Transformation zur gegenseitigen Überführung zweier Faktorenmatrizen. *Diagnostica* 1964, 10, 87–90.
- KRISTOF, W., Die faktorielle Struktur einiger geometrisch-optischer Täuschungen. *Z. exper. angew. Psychol.* 1964, 11, 71–94.
- KRISTOF, W., Eine Faktorenanalyse geometrisch-optischer Täuschungen. *Z. exper. angew. Psychol.* 1963, 10, 583–596.
- KRÜSKEMPER, G., L. KRÜSKEMPER, Schilddrüsenfunktion und geistige Leistungsfähigkeit Erwachsener, gemessen mit dem Hamburg-Wechsler-Intelligenztest. *Z. exper. angew. Psychol.* 1971, 18, 411–433.
- KUMATA, H., W. SCHRAMM, A pilot study of cross-cultural meaning. *Publ. Opin. Quart.* 1956, 20, 229–238.
- LAUX, L., W. D. FRÖHLICH, Serielles Wahrnehmen, Aktualgenese, Informationsintegration und Orientierungsreaktion II: Dimensionsanalyse von Maßen des Hautwiderstandes (GHR). *Z. exper. angew. Psychol.* 1970, 17, 266–276.
- LEVIN, J., Three-mode factor analysis. *Psychol. Bull.* 1965, 64, 442–452.
- LIENERT, G. A., Ein Formlegetest als Prüfmittel der praktischen Intelligenz. *Z. exper. angew. Psychol.* 1958a, 5, 82–107.
- LIENERT, G. A., Ein Test zur Beurteilung des mechanisch-technischen Verständnisses. *Z. exper. angew. Psychol.* 1958b, 5, 605–620.
- LIENERT, G. A., Prinzip und Methode der multiplen Faktorenanalyse demonstriert an einem Beispiel. *Biometr. Z.* 1959, 1, 88–141.
- LIENERT, G. A., Überprüfung und genetische Interpretation der Differenzierungshypothese von Wewetzer. *Vita Humana* 1961, 4, 112–124.
- LIENERT, G. A., Die Faktorenstruktur der Intelligenz als Funktion des Neurotizismus. *Z. exper. angew. Psychol.* 1963, 10, 140–159.
- LIENERT, G. A., Testaufbau und Testanalyse. Weinheim—Berlin 1967 (Beltz).
- LIENERT, G. A., Die Konfigurationsfrequenzanalyse. *Z. Klin. Psychol. Psychother.* 1971, 19, 99–115.
- LIENERT, G. A., Verteilungsfreie Methoden in der Biostatistik. Meisenheim 1973 (Hain).
- LIENERT, G. A., C. FABER, Über die Faktorenstruktur des HAWIK auf verschiedenen Alters- und Intelligenzniveaus. *Diagnostica* 1963, 9, 3–11.
- LIENERT, G. A., R. LIENHÖFT, Versuch einer Standardisierung der Drahtbiegeprobe. *Z. exper. angew. Psychol.* 1959, 6, 792–815.
- LINN, R. L., A monte carlo approach to the number of factors problem. *Psychometrika* 1968, 33, 37–71.
- LOCKOWANDT, O., Faktorenanalytische Validierung der Handschrift mit besonderer Berücksichtigung projektiver Methoden. *Z. exper. angew. Psychol.* 1968, 15, 487–530.
- LÜER, G., R. COHEN, D. EGGERT, Zur Erfassung der motorischen Begabung bei minderbegabten Kindern durch eine Hamburger Version der Lincoln-Oseretzky Motor Development Scale. *Praxis Kinderpsychol. Kinderpsychiatr.* 1970, 19, 20–25.

- LÜER, G., R. COHEN, W. NAUCK, Eine Kurzform der „Vineland Social Maturity Scale“ für minderbegabte Kinder. *Praxis Kinderpsychol. Kinderpsychiatr.* 1966, 15, 101–105.
- LÜER, G., B. NEUFELDT, Über den Zusammenhang zwischen Maßen der galvanischen Hautreaktion und Beurteilungen von Reizen durch Versuchspersonen. *Z. exper. angew. Psychol.* 1968, 15, 619–648.
- LUNNEBORG, P. W., Dimensionality of MF. *J. Clin. Psychol.* 1972, 18, 313–317.
- MAXEINER, J., Eine Untersuchung zur Validität des Satzergänzungstestes zur Erfassung der Aggressivität. *Diagnostica* 1968, 14, 94–98.
- MEILL, R., Die Zwiespältigkeit faktorieller Untersuchungen der Intelligenzstruktur. *Z. exper. angew. Psychol.* 1969, 16, 278–294.
- MERZ, F., Über die individuelle Inferenzneigung. *Z. exper. angew. Psychol.* 1961, 8, 381–392.
- MERZ, F., Über die soziale Determination von „Eigenschaften“. *Jb. Psychol. Psychother.* 1958, 5, 293–309.
- MESSICK, S. J., Metric properties of the semantic differential. *Educ. Psychol. Measmt.* 1957, 17, 200–206.
- MEYER, A. E., R. GOLLE, Zur Validierung des Brengelmann Fragebogens E–N–NR an klinischen Stichproben. *Diagnostica* 1966, 12, 93–105.
- MICKO, H. C., Die Bestimmung subjektiver Ähnlichkeiten mit dem semantischen Differential. *Z. exper. angew. Psychol.* 1962, 9, 242–280.
- MIKUS, P., Persönlichkeitskorrelate der Stimmungslabilität. *Diagnostica* 1969, 15, 14–22.
- MIRON, M. S., A cross-linguistic investigation of phonetic symbolism. *J. Abnorm. Soc. Psychol.* 1961, 62, 623–630.
- MIRON, M. S., C. E. OSGOOD, Language behavior: The multivariate structure of qualification. In: CATTELL, R. B. (Hrsg.), *Handbook of multivariate experimental psychology*. Chicago 1966 (McNally), 790–819.
- MÖBUS, C., H.-J. AHRENS, Multivariate Weiterentwicklung eines Modells zur Prognose von Wahlentscheidungen. *Z. exper. angew. Psychol.* 1970, 17, 386–413. *Münchener Symposium über Faktorenanalyse. Psychol. Beitr.* 1967, 10, 81–153.
- MULAIK, S. A., *The foundations of factor analysis*. New York 1972 (McGraw-Hill).
- ORLIK, P., Eine Technik zur erwartungsgetreuen Skalierung psychologischer Merkmalsräume auf Grund von Polaritätenprofilen. *Z. exper. angew. Psychol.* 1967a, 14, 616–650.
- ORLIK, P., Das Dilemma der Faktorenanalyse – Zeichen einer Aufbaukrise in der modernen Psychologie. *Psychol. Beitr.* 1967b, 10, 87–98.
- OSGOOD, C. E., G. J. SUCI, Factor analysis of meaning. *J. Exper. Psychol.* 1955, 50, 325–338.
- OTTE, H., Bemerkungen zur Zuverlässigkeit und Validität einiger „Inhaltsskalen“ des Rorschach bei psychosomatischen Stichproben. *Diagnostica* 1970, 16, 3–16.
- PAWLIK, K., *Dimensionen des Verhaltens*. Bern–Stuttgart 1968 (Huber).
- PAWLIK, K., Psychologische Maße der Aktivierung. *Z. exper. angew. Psychol.* 1963, 10, 19–34.
- PIAGGIO, L., Bemerkungen zum Promax-Verfahren für schiefwinkelige Faktorenrotation. *Z. exper. angew. Psychol.* 1972a, 19, 460–475.
- PIAGGIO, L., Modifikation des Promax-Verfahrens. *Z. exper. angew. Psychol.* 1972b, 19, 580–592.
- PIAGGIO, L., Eine Studie über die Anwendung der Methode der Adjektivskalen in der Persönlichkeitsforschung. *Z. exper. angew. Psychol.* 1969, 16, 367–401.
- PROTHRO, E. T., J. D. KEEHN, Stereotypes and semantic space. *J. Soc. Psychol.* 1957, 45, 197–209.

- RIEGEL, R. M., Faktorenanalysen des Hamburger-Wechsler-Intelligenztests für Erwachsene (HAWIE) für die Altersstufen 20–34, 35–39, 50–64 und 65 und älter. *Diagnostica* 1960, 6, 41–66.
- ROEDER, B., Die Bestimmung diskrepanten Antwortverhaltens. *Z. exper. angew. Psychol.* 1972, 19, 593–640.
- ROTH, K., Faktorenanalytische Überprüfung der Theorie der technischen Begabung. *Z. exper. angew. Psychol.* 1957, 4, 58–83.
- RÜSSEL, A., Selbst- und Fremdeinstufung des Befindens bei der Arbeit. *Z. exper. angew. Psychol.* 1964, 11, 95–129.
- SACHS, L., *Statistische Auswertungsmethoden*. Berlin–Heidelberg–New York 1969 (Springer).
- SASSENSCHIEDT, H., F. BUGGLE, Prädiktorvariable praktisch-technischen Verständnisses. *Diagnostica* 1970, 16, 30–41.
- SATERDAG, H., E. APENBURG, R. FISCH, P. ORLIK, Die Selbstcharakterisierung des Faches Psychologie auf dem Hintergrund der klassischen fünf Wissenschaftsdisziplinen. *Psychol. Rdsch.* 1971, 22, 103–113.
- SCHAEDEL, R., Untersuchungen zur Verifikation von Meilis Intelligenzfaktoren. *Z. exper. angew. Psychol.* 1961, 8, 211–264.
- SCHIEFER, I. H., R. B. CATTELL, Bestätigung von objektiven Testfaktoren und Beurteilung ihrer Beziehung zu Fragebogenfaktoren. *Diagnostica* 1965, 11, 95–120.
- SCHNEEWIND, K. A., R. B. CATTELL, Zum Problem der Faktorenidentifikation: Verteilungen und Vertrauensintervalle von Kongruenzkoeffizienten für Persönlichkeitsfaktoren im Bereich objektiv-analytischer Tests. *Psychol. Beitr.* 1970, 12, 214–226.
- SCHNEIDER, J., H. MINKMAR, Deutsche Neukonstruktion einer Konservatismuskala. *Diagnostica* 1972, 18, 37–48.
- SCHÖNPFUG, W., Ein Problem bei der Arbeit mit Kategorienskalen: Welchen Einfluß hat die Zahl der Skalenkategorien? *Z. exper. angew. Psychol.* 1972, 19, 141–171.
- SCHÜMER, A., C. LAEMMERHOLD, R. COHEN, Beschreibung von zwei Ansätzen zur Erfassung der Differenziertheit von Variablenstrukturen (illustriert an einem Beispiel zwischenmenschlicher Beziehungen). *Z. exper. angew. Psychol.* 1972, 19, 507–516.
- SEIFERT, R., Referat über Untersuchungsmethoden der Psychomotorik. *Diagnostica* 1959, 5, 135–155.
- SEIFERT, T., Faktorenanalyse einiger Schriftmerkmale. *Z. exper. angew. Psychol.* 1964, 11, 645–666.
- SEITZ, W., Über die Beziehung von Persönlichkeitsmerkmalen zu Schul- und Intelligenztest-Leistungen bei Volksschülern. *Z. exper. angew. Psychol.* 1971, 18, 307–336.
- SEITZ, W., H. BRÄTH, Empirische Untersuchung eines Fragebogens zur Prüfung kritischer Einstellungen gegenüber der Schule (ausgehend von einer erweiterten Version des „School Inventory“ nach Bell). *Diagnostica* 1970, 16, 172–185.
- SEITZ, W., W. LÖSER, Über die Beziehung von Persönlichkeitsmerkmalen zu Schul- und Intelligenz-Leistungen bei Gymnasial-Schülern. *Z. exper. angew. Psychol.* 1969, 16, 651–679.
- SELG, H., G. LISCHKE, Eine Faktorenanalyse von Aggressionsvariablen. *Z. exper. angew. Psychol.* 1966, 13, 506–526.
- SEYDEL, U., HAWIE-Kurzformen und deren Kreuzvalidierung. *Diagnostica* 1972, 18, 121–136.

- SIDDIQI, J. A., W. KEIL, Neugierverhalten in der Sprache. *Psychol. Rdsch.* 1970, 21, 270–278.
- SIXTL, F., Faktoreninvarianz und Faktoreninterpretation. *Psychol. Beitr.* 1967, 10, 99–111.
- SIXTL, F., Vorläufige Mitteilung: Ein Verfahren zur Bestimmung von Faktorenstrukturen ohne vorherige Berechnung der skalaren Produkte. *Z. exper. angew. Psychol.* 1964, 11, 349–351.
- STEGMÜLLER, W., Personelle und Statistische Wahrscheinlichkeit. Berlin–Heidelberg–New York 1973 (Springer).
- STEINGRÜBER, H. J., Zur Messung der Händigkeit. *Z. exper. angew. Psychol.* 1971, 8, 337–357.
- STEINHAGEN, K., Untersuchungen zur Veränderung von faktoriellen Intelligenzstrukturen im Erwachsenenalter. *Diagnostica* 1970, 16, 149–164.
- SUCI, G. J., A comparison of semantic structures in american southwest culture groups. *J. Abnorm. Soc. Psychol.* 1960, 61, 25–30.
- TANAKA, Y., T. OYAMA, C. E. OSGOOD, A cross-cultural and cross-concept study of the generality of semantic space. *J. Verbal Learning Verbal Behavior* 1963, 2, 392–405.
- TAUSCH, A.-M., R. TAUSCH, B. FITTKAU, Merkmalszusammenhänge der verbalen Interaktion und kritische Überprüfung typologischer Verhaltenskonzepte. *Z. exper. angew. Psychol.* 1967, 14, 522–541.
- TAUSCH, R., H. KÖHLER, B. FITTKAU, Variablen und Zusammenhänge der sozialen Interaktion in der Unterrichtung. *Z. exper. angew. Psychol.* 1966, 13, 354–365.
- THURSTONE, L. L., Multiple factor analysis. Chicago 1947 (Univ. of Chicago Press).
- TIMM, U., Eigenschaftsratings als Validitätskriterien eines Persönlichkeitsfragebogens. *Diagnostica* 1971, 17, 26–52.
- TIMM, U., Reliabilität und Faktorenstruktur von Cattells 16 P.F.-Test bei einer deutschen Stichprobe. *Z. exper. angew. Psychol.* 1968, 15, 354–373.
- TRAXEL, W., Über Dimensionen und Dynamik der Motivierung. *Z. exper. angew. Psychol.* 1961, 8, 418–428.
- TUCKER, L. R., Some mathematical notes on three-mode factor analysis. *Psychometrika* 1966, 31, 279–311.
- TUCKER, L. R., The extension of factor analysis to three-dimensional matrices. In: FREDERIKSEN, N., H. GULLIKSEN (Hrsg.), Contributions to mathematical psychology. New York 1964 (Holt, Rinehart, Winston), 107–109.
- ÜBERLA, K., Faktorenanalyse. Berlin–Heidelberg–New York 1968 (Springer).
- VANNOY, J. S., Generality of cognitive complexity-simplicity as a personality construct. *J. Pers. Soc. Psychol.* 1965, 2, 385–396.
- VUKOVICH, A., Faktorielle Typenbestimmung. *Psychol. Beitr.* 1967, 10, 112–121.
- WEBER, H., Untersuchungen über die Faktorenstruktur numerischer Aufgaben. *Z. exper. angew. Psychol.* 1953, 1, 336–393.
- WEHNER, E. G., E. H. BOTTENBERG, Wiederholungsreliabilität und faktorielle Konstanz des MMPI Saarbrücken. *Diagnostica* 1969, 15, 36–40.
- WENDELER, J., Persönlichkeitseigenschaften und „Lügen“-Tendenz. *Diagnostica* 1971, 17, 73–82.
- WENDELER, J., Vergleich einiger Faktorenanalysen muttersprachlicher Leistungen. *Diagnostica* 1970, 16, 76–94.
- WENDELER, J., Extraversion, neurotische Tendenz und Leistungsmotivation. *Diagnostica* 1969, 15, 22–36.
- WEWETZER, K. H., Zur Differenzierung der Leistungsstrukturen bei verschiedenen Intelligenzgraden. In: WELLEK, A. (Hrsg.), Bericht über den 21. Kongreß der Dt. Ges. f. Psych. in Bonn – 1957. Göttingen 1958 (Hogrefe).

- WIECZKOWSKI, W., H. NICKEL, L. ROSENBERG, Einige Bedingungen der unterschiedlichen Bewertung von Schüleraufsätzen. *Psychol. Rdsch.* 1968, 19, 280–295.
- WIENDIECK, G., H. E. LÜCK, Zur Entwicklung von Kurzformen der Bregelmann-Skalen E, N und NR. *Diagnostica* 1969, 15, 40–43.
- WIGGINS, N., M. FISHBEIN, Dimensions of semantic space: A problem of individual differences. In: SNIDER, J. G., OSGOOD, C. E. (Hrsg.), Semantic differential technique. Chicago 1969 (Aldine), 183–193.
- WILSON, W. R., H. L. MILLER, J. S. LOWER, Much ado about the null hypothesis. *Psychol. Bulletin* 1967, 67, 188–196.
- WILSON, W. R., H. L. MILLER, A note on the inconclusiveness of accepting the null hypothesis. *Psychol. Review* 1964, 71, 238–242.
- WITTE, E. H., Ein Vergleich von Intelligenzstrukturen: Eine Anwendung der bilateralen Transformation und eine Kritik zu der Arbeit von K. Steinhagen (1970) über die Dedifferenzierungshypothese. *Diagnostica* 1972, 18, 18–25.
- ZELLINGER, E., Theorie, Rechenkalkül, phänomenale Erfahrung – Einwände gegen den faktorenanalytischen Formalismus. *Jb. Psychol. Psychother.* 1968, 16, 130–134.

Anschriften der Verfasser:

Dr. phil. HELMUT LUKESCH, Fachbereich Erziehungswissenschaft der Universität Konstanz, D–755 Konstanz, Postfach 7733

Dr. phil. GERNOT D. KLEITER, Psychologisches Institut der Universität Salzburg, Akademiestraße 26, A–5020 Salzburg