

Eckart Leiser

Kritische Untersuchung einer Faktorenanalyse

Bemerkungen zu einem Aufsatz von A.O. Jäger
zur Intelligenzstrukturforschung*

Seit einiger Zeit biete ich, aufbauend auf meinem »Grundkurs Statistik« (Leiser, 1983) eine Veranstaltung mit dem Titel »Kritische Einführung in die multivariaten Verfahren« an. In ihr werden exemplarisch komplexere statistische Modelle behandelt, die mit der Mittelwerts- und Zusammenhangsstruktur größerer Variablenmengen zu tun haben: Hotellings-T-Quadrat als elementarster Fall der multivariaten Varianzanalyse (MANOVA) sowie Faktoren- und Clusteranalyse. Während die Entfaltung der mathematisch-statistischen Modelle selbst dem für meinen Grundkurs entwickelten genetischen und heuristischen Konzept folgt, das ich unlängst genauer expliziert habe (s. Leiser 1987), erfordert ihre kritische Reflexion ein sehr viel differenzierteres und präziseres Eingehen auf die Praxis dieser multivariaten Verfahren. Das hat mehrere Gründe:

- Das Vertrauen auch kritischer Studenten darauf, daß, bei aller Beschränktheit der Verfahren ihre Erkenntnisfunktion betreffend, ihre Anwendung immanent rational und fehlerfrei sein wird — regelt doch schon die erforderliche mathematische und statistische Kompetenz einen auf qualifizierte Anwender beschränkten Zugang.
- Die schwierigere Rekonstruierbarkeit des konkreten technischen Vorgehens aufgrund der Komplexität der entsprechenden Kalküle.
- Die in der Struktur der Verfahren selbst — und nicht nur in den Defiziten ihrer Anwender — liegenden Dispositive zu Verselbständigungen und zu Irrationalismen.

Nehmen wir als Beispiel die Faktorenanalyse. Auch nach einer grundlegenden Kritik der Faktorenanalyse im Hinblick auf ihre Artefaktproblematik (Variablenauswahl, Abbruchkriterium, Rotationsproblematik) wird der Student i. A. davon ausgehen, daß der — zumal wissenschaftliche — Anwender bei allen Mängeln des Instruments das Beste und Rationalste daraus zu machen versucht. Und wenn ein in der »scientific community« anerkannter Anwender trotz aller besagter Mängel mit diesem Instrument herausfindet, daß weltweit und unter praktisch allen Bedingungen immer die gleichen 4 operativen und 3 inhaltsspezifischen Grunddimensionen der Intelligenz gefunden werden können, wird auch der kritische Student von der sich gegen alle Schwierigkeiten des Verfahrens durchsetzenden Durchschlagskraft dieser »Gesetzmäßigkeit« beeindruckt

* Wir fragten unserer Gepflogenheit gemäß bei A.O.Jäger an, ob er auf Leisers Kritik antworten wolle. Er hat aber auf eine Replik verzichtet. — Red.

und verunsichert sein. Wahrscheinlich wird er das Ergebnis als »Rätsel« stehen lassen und seine Kritik auf die beschränkte Relevanz derartiger »Gesetzmäßigkeiten« richten.

Genau so erging es einem Studenten, der eines Tages einen Aufsatz zum Intelligenzstruktur-Konzept von A.O. Jäger (1984) in mein Seminar einbrachte, als Beispiel für die Anwendungspraxis der Faktorenanalyse. Bei aller kritischen Distanz hatten ihn der Aufsatz und seine so prägnanten Schlußfolgerungen zweifellos verunsichert und fasziniert. Und vor mir tat sich plötzlich in dem bis dahin so erfolgreichen Seminar eine breite Kluft auf zwischen den entwickelten grundlegenden Kritik Kategorien und dieser Art entwaffnender und faszinierender Praxis. Ich fand den Aufsatz im Hinblick auf das didaktische, inhaltliche und wissenschaftstheoretische Problem einer Dechiffrierung irrationaler Methodenpraxis so herausfordernd, daß ich ihn seinerzeit einer ausführlichen Analyse unterzogen habe. Daraus ist die im folgenden entwickelte kritische Rekonstruktion entstanden, die mir über den konkreten Kontext hinaus relevant zu sein scheint:

- Einmal stehen diese Rekonstruktion und ihre Frageweise für eine verallgemeinerbare didaktische Strategie, die Studenten bei der Entmystifizierung methodisch komplexer Texte anleiten kann.
- Darüber hinaus steht sie für einen wissenschaftspolitisch relevanten Kritiktyp, der Grundlagenfragen der psychologischen Methodologie mit deren empirischer Realität zu vermitteln versucht. Mehr dazu am Schluß meines Beitrags.

Über die Wirkung dieses Aufsatzes auf einen weiteren konkreten Kontext, nämlich die Arbeit in dem andauernden und großzügig ausgestatteten Projekt »Produktives Denken, intelligentes Verhalten« mache ich mir nach den bisherigen Reaktionen der Betroffenen allerdings keine Illusionen.

Meine kritische Rekonstruktion stützt sich primär auf den erwähnten Aufsatz »Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven« (A.O. Jäger 1984). Bestimmten offengebliebenen Fragen bin ich dann noch in einem früheren Aufsatz des gleichen Autors (A.O. Jäger 1982) nachgegangen, der zusätzliche Details des methodischen Vorgehens enthält. Auf diesen Aufsatz wird zur Vervollständigung meiner kritischen Analyse an einigen Stellen Bezug genommen. In meiner Kritik folge ich schrittweise dem Vorgehen des Autors im Aufsatz von 1984, wobei ich mich auf die für unsere kritischen Fragestellungen relevantesten Schritte beschränke. Aus dem gleichen Grund spare ich die clusteranalytischen Aspekte seiner Untersuchung aus, was nicht heißt, daß diese nicht eine eigene Kritik wert wären.

Unter jedem Schritt werden die sich mir aufdrängenden methodischen, inhaltlichen und wissenschaftstheoretischen Probleme angesprochen. Dieses Vorgehen stellt angesichts der Lückenhaftigkeit der kritisierten Vorlage am ehesten sicher, daß hart und nachvollziehbar am Gegenstand entlang argumentiert wird.

Schritt 1: Der Autor möchte gern die Intelligenzstruktur erforschen, die hinter einem möglichst unverkürzten Spektrum von Intelligenzleistungen steht. Aus 2000 Aufgabentypen in Intelligenztests werden zu diesem Zweck 191 Aufgabenblöcke als repräsentative Stichprobe ausgewählt. Anhand dieser Aufgaben werden 545 Berliner Abiturienten untersucht.

- *Kontextproblem:* Nicht ein theoretisch begründeter Raum relevanter Intelligenzleistungen bildet den Ausgangspunkt des Vorhabens, sondern ein Arsenal von *Intelligenztests*, die wohl mehr für die Traditionen und Vorlieben der Intelligenztest-Konstrukteure stehen als für das relevante Spektrum kognitiver Funktionen. Aus elementaren inhaltlichen Gesichtspunkten nahegelegte Alternative: als erstes müßten einmal die aus genetischen, funktionalen und möglichst auch historischen Bestimmungen der Kognition folgenden notwendigen Intelligenzleistungen bestimmt werden (s. Kognitionskonzept von Piaget oder der materialistischen Psychologie).
- *Stichprobenproblem:* Was bedeutet es, für ein statistisch ohnehin nicht fundiertes Verfahren wie die Hauptkomponentenmethode der Faktorenanalyse nicht nur auf der Vpn- sondern auch auf der Variablenseite Stichproben zu ziehen? (Umfangreiche Simulationsuntersuchungen, die ich seinerzeit durchgeführt habe, sprechen für eine unerwartet starke statistische Labilität der Hauptkomponentenmethode und für deren hochgradige Abhängigkeit von der Variablenauswahl, insbesondere bei Relationen von Variablen- und Vpn-Zahl wie in der vorliegenden Untersuchung. Zur statistischen Problematik der Faktorenanalyse s. auch Lawley u. Maxwell 1971).

Schritt 2: Diese 191 Variable werden einer Faktorenanalyse unterzogen. Diese ergibt »4 hochgradig generelle Leistungsklassen (Faktoren oder Cluster), die durch ihre operative Eigenart gekennzeichnet sind, nämlich: Einfallsreichtum, Verarbeitungskapazität, Gedächtnis und Bearbeitungsgeschwindigkeit.« (1984, 27)

- *Mystifizierungsproblem:* Die Faktorenanalyse »ergibt« 4 Faktoren. In Wirklichkeit ergibt diese Faktorenanalyse im Extremfall, nämlich bei einer unreduzierten Korrelations-(R-)Matrix, 191 Faktoren. Der Autor versucht zwar anscheinend (1982, 202), den Vektorraum durch Kommunalitätsschätzung zu reduzieren, verrät aber an keiner Stelle etwas über den »Anpassungswert« dieses methodisch und mathematisch immer fragwürdigen Versuchs (s. dazu Harman 1967, Ch. 5). (Konkret: Wieviel »reelle« Faktoren bleiben bei diesen Kommunalitäten übrig? Wie gut läßt sich die R-Matrix mit diesen Kommunalitäten reproduzieren?). Das Auswählen von 4 Faktoren aus einer größeren Gesamtzahl von Faktoren ist also ein *Entscheidungsakt*, dessen Kriterien hier im dunkeln bleiben. Die 4 Faktoren werden so quasi ontologisiert. Konkretisiert heißt die Frage, was eigentlich das Abbruchkriterium für die Hauptkomponentenmethode ist. Eine Antwort auf diese Frage fehlt. Egal wie sie lautet, mit *Sicherheit* sind bei diesem 1. Schritt mehr als 4 Faktoren herausgekommen.

- *Rotationsproblem:* Ist eine »Zielrotation« auf bestimmte den Faktorenbezeichnungen entsprechende Markierungs-Variablen hin versucht worden? Ist, wie es klingt, auf Cluster hin rotiert worden — die 4 Faktoren ließen sich anscheinend auch als Cluster darstellen (1984, 27)? Oder hat nur eine automatische Rotation nach dem Varimax-Kriterium stattgefunden?

Was ist aus den im 1. Schritt zusätzlich gefundenen Faktoren bei der Rotation geworden?

- *Semantisches Problem:* Wie ist es zu den 4 Bezeichnungen gekommen? Diese sind ja keine kosmischen Universalien, sondern aus mehr oder weniger rationalen Assoziationen entstandene Sprachschöpfungen. Diesen naiven Umgang mit der semantischen Problematik von Faktorenanalysen hat diese Untersuchung allerdings — das muß der Gerichtigkeit halber betont werden — mit allen mir bekannten Faktorenanalyse-Anwendungen gemeinsam. Diese Naivität beginnt ja mit der Praxis, das semantische Interpretationsproblem als abhängige Variable eines mathematischen Rotationsproblems zu behandeln, konkret mit dem Glauben also, zu jeder Varimax-Lösung ließe sich schon irgendwie eine sinnvolle Faktoreninterpretation finden. Und sie endet bei Vergleichstechniken von Faktormatrizen wie der von Roppert und Fischer (1965), bei denen die semantischen Implikationen einer Faktorenrotation vollständig ausgeblendet sind (s. Leiser 1971).

Schritt 3: A.O. Jäger wundert sich gleichsam wie ein Naturforscher, daß 3 weitere zu den kosmischen Universalien gezählte Faktoren fehlen. Gemeint sind die material-/inhaltsspezifischen Faktoren für sprachgebundenes, zahlengebundenes und anschauungsgebundenes Denken. (Frage zur didaktischen Zuspitzung: warum sollte nicht mit gleichem Recht nach Faktoren für kausales, systemisches, kombinatorisches, kreatives, klassifikatorisches, intuitives usw. Denken gesucht werden?)

Problem: Woran kann das Nicht-Auffinden der 3 erwarteten Faktoren liegen?

- Hypothese 1: daß nur 4 und nicht 7 Faktoren interpretiert worden sind (also an einem Artefakt);
- Hypothese 2: daß nicht genug Faktoren in die Rotation aufgenommen worden sind (also wieder ein Artefakt);
- Hypothese 3: daß es zwar weitere nicht operativ deutbare Faktoren gibt, die aber nicht ohne weiteres mit den 3 gesuchten kosmischen Universalien identifiziert werden können (also kein Artefakt, aber ein Fall von Falsifikation);
- Hypothese 4: daß die material-/inhaltsspezifischen Faktoren sich nur bei ganz speziellen Variablenauswahlen spezieller Autoren durchgesetzt haben, die man jetzt daraufhin mal genauer untersuchen müßte. Das wäre nach der Logik der vorliegenden Untersuchung dann kein Defizit, sondern ein relevantes Ergebnis: bei einer »repräsentativen« Stichprobe von Intelligenzleistungen haben solche »Inhaltsfaktoren« kein Gewicht mehr (noch ein Fall von Falsifikation).

Angesichts dieser möglichen Hypothesen haben wir *unsererseits* Anlaß, uns über A.O. Jägers Vorgehen zu wundern: Ihm geht es ja nicht um eine generative Untersuchung, in der er freizügig und nach eigenem Ermessen Faktoren zusammenstellen und benennen darf, sondern um eine Art »detection problem«, eben das Aufspüren von 3 postulierten Faktoren. In dieser Situation ist es eigentlich zwingend, als *erstes* das Suchinstrument in Frage zu stellen, also möglichen Artefakten und Ungenauigkeiten der Faktorenanalyse nachzugehen, d.h.

- Anpassungsgüte der benutzten Kommunalitätsschätzungen überprüfen
- gegebenenfalls Kommunalitätsschätzungen variieren
- zum Vergleich unreduzierte R-Matrix verarbeiten
- wenn schon die Analyse abgebrochen wird, dann Abbruchkriterium offenlegen und in Frage stellen
- in die Rotation mindestens die postulierte Zahl von Faktoren, also 7, aufnehmen
- außer Varimax-Rotation Zielrotationen auf die 3 Inhaltsfaktoren markierende Variablen vornehmen.

Wenn das alles ausgeschöpft ist und von den postulierten Faktoren immer noch keine Spur gefunden ist, wäre an die Falsifikation der Ausgangshypothese zu denken, theoretisch gestützt auf ein revidiertes Intelligenzstrukturmodell und empirisch gestützt auf eine alternative Zusammenstellung und Interpretation der gefundenen Faktoren. Gleichzeitig wären von der Datenbasis und dem methodischen Vorgehen her die der Ausgangshypothese zugrundeliegenden Untersuchungen zu kritisieren, die Abweichung der Ergebnisse also plausibel zu machen (s. Hypothese 4).

Schritt 4: A.O. Jägers Motiv ist aber offensichtlich weniger die Falsifikation als die Integration vorhandener »Erkenntnis«. Er muß deshalb den 3 Faktoren wie dem »missing link« nachspüren. Das geschieht folgendermaßen: aus den 191 Variablen werden noch einmal 48 ausgewählt, und zwar genau gemäß den nachzuweisenden kosmischen Universalien. Die 4 (bereits nachgewiesenen) Operationsfaktoren und die 3 (noch nachzuweisenden) Inhaltsfaktoren geben ein »bimodales« Schema vor. Für jede der 12 Kombinationen (4x3) operativ-inhaltlich werden 4 möglichst »reine« oder »typische« Variable ausgewählt. Diese Auswahl der Variablenstichprobe ist also a priori durch das Schema gesteuert, das nachgewiesen werden soll.

Inkonsistenzproblem: Die ursprüngliche Fragestellung einer repräsentativen Stichprobe von Intelligenzleistungen wird aus scheinbar rein methodischen Optimierungsgründen zugunsten einer ergebnisgesteuerten Stichprobe aufgegeben. Nennt man so etwas nicht »petitio principii«?

Die neue Fragestellung heißt sinngemäß: Gibt es eine Auswahl von 48 Variablen, in denen möglichst rein die 4 Operationsfaktoren und die 3 Inhaltsfaktoren repräsentiert sind? Da sich das Klassifizieren der Variablen immer schon auf die postulierten Leistungsdimensionen stützt und die Leistungsdimensionen auf die

diese markierenden Variablen, hat das ganze die Struktur eines Zirkels, um dessen Falsifizierbarkeit es ja bekanntermaßen schlecht bestellt ist. Denn: Im Prinzip genügt es ja, 12 *irgendwie* distinkte Variablentypen zu finden, die für die Kombinationen von 4 Varianten *irgendeiner* Modalität A mit 3 Varianten *irgendeiner* Modalität B stehen. Der Rest liegt im Raum der methodischer Objektivierung weitgehend entzogenen Interpretationsarbeit.

Die so verwandelte Fragestellung zielt also eigentlich auf so etwas wie eine spezielle kognitive Kompetenz von Faktorenanalytikern: ein gewisses intuitives topologisches »feeling« für die Konstruktion einer gewünschten Vektoranordnung. Auch das ist sicher ein reizvolles Problem der Intelligenzstrukturtheorie, das man etwas ironisch unter den Arbeitstitel »Kognitionspsychologie der Faktorenanalytiker« stellen könnte, aber eins, das mit dem Ausgangsproblem wirklich nichts mehr zu tun hat.

Schritt 5: Auch in der Faktorenanalyse dieser 48 Variablen setzen sich zwar die Operationsfaktoren, nicht aber die 3 Inhaltsfaktoren durch.

Mystifikationsproblem

Rotationsproblem stellt sich unverändert, wie unter Schritt 2

Semantisches Problem

Das Mystifikations- und Rotationsproblem verschärft sich hier allerdings noch. Denn wenn ich schon die 3 Inhaltsfaktoren als kosmische Universalien postuliere und nicht finde, sollte ich doch hier erst recht einmal den Artefakten und Ungenauigkeiten nachgehen, die im methodischen Vorgehen liegen. Ich sollte also noch hartnäckiger alle unter Schritt 3 genannten Möglichkeiten durchgehen. Und wenn das alles nichts nützt? Dann sollte ich nüchtern prüfen, was daraus eigentlich folgt:

Die Falsifikation des ursprünglichen Intelligenzstrukturmodells (ein theoretisches Ergebnis) oder — durchaus nicht nur ironisch gemeint — das unzureichende topologische »feeling« der beteiligten Faktorenanalytiker (ein mehr psychodiagnostisches Ergebnis)?

In dem nur lückenhaft dokumentierten konkreten Vorgehen (1982, 221) finden sich keinerlei Indizien für eine solche Hartnäckigkeit, dagegen eine Reihe von zusätzlichen Ungereimtheiten. Um diese näher einzugrenzen, ist mangels präziser Dokumentation des Vorgehens eine Art kriminologische Aufklärungsarbeit nötig:

a) Es sind zunächst (mindestens) 8 Faktoren extrahiert worden, Summe der Eigenwerte = $17.57 = 36,6\%$ der Gesamtvarianz (s. Anhang Tab. 7a, 20). In die in der gleichen Tabelle vorgeführte Varimax-Rotation sind dagegen ohne jede Begründung nur noch 4 Faktoren aufgenommen worden. Die aus der Ausgangshypothese sich geradezu aufdrängende Frage, was nämlich — wie viel oder

wenig — die nächsten 3 Faktoren mit den postulierten Inhaltsfaktoren zu tun haben, kann daraus gar nicht beantwortet werden. Eine Vermutung zu diesem paradoxen Sachverhalt, der wir in den weiteren Auswertungsschritten nachgehen müssen, wäre, daß A.O. Jäger irgendein formales »Abbruchkriterium« eingeführt hat:

- Ist es eine Mindestgröße des Eigenwerts?
Sie müßte dann zwischen 1.59 und 1.20 liegen (Tab. 7a; s. Anhang, S.27).
- Ist es ein prozentualer Mindestbeitrag eines Faktors? Der müßte dann zwischen 3.3 % und 2.5 % liegen (1.59 bzw. 1.20/48).
- Ist es ein prozentualer Mindestbeitrag der Faktoren insgesamt? Dieser müßte dann zwischen 29.75 und 32.25 % liegen (14.28 bzw. 15.48/48).

b) Es fehlt jede Begründung, wie es zu dem Ladungslimit von 0.35 gekommen ist, von dem die ganze weitere Interpretation weitgehend abhängt. Ist es a priori festgelegt worden oder ist es post hoc so gewählt worden, daß die vorweg behauptete Zuordnung von Variablen zu Faktoren möglichst ungetrübt reproduziert werden konnte? Als »Schönheitsfehler« bleibt, daß auch mit diesem Limit 9 der Variablen auf der Strecke bleiben. Sind darunter möglicherweise solche — so darf mit Verlaub gefragt werden — die sich bei besten Willen nicht wunschgemäß zuordnen ließen? Andererseits tauchen erstaunlicherweise 3 Variable in der Tabelle auf, deren Ladungen *unter* dem festgelegten Limit liegen. Zu alledem finde ich im Text nicht *ein* erklärendes Wort.

c) Aufgrund der numerischen Verhältnisse in Tab. 7a darf vermutet werden, daß hier einfach die Varianzanteile der Variablen auf den 4 willkürlich ausgewählten Faktoren als »Kommunalitäten« bezeichnet werden. Mit den aus Schätzungen hervorgegangenen Kommunalitäten, von denen weiter oben die Rede ist (1982, 202), hätte das nichts zu tun. Nur so kann ich mir den Widerspruch erklären, daß den »Kommunalitäten« ein Varianzanteil von 29.8 entsprechen soll, aber bereits die ersten 8 Faktoren einen Varianzanteil von 36.6 % abdecken: eine mathematisch eigentlich unmögliche Relation, wo doch die »Kommunalitäten« die maximal ausschöpfbare gemeinsame Varianz bezeichnen. Die »geringen Kommunalitäten«, die der Autor (a.a.O., 203) flugs inhaltlich kommentiert, sind also möglicherweise einfach auf deren ungewöhnliche Definition zurückzuführen.

Schritt 6: A.O. Jäger schließt aus dem Ergebnis weder auf die Falsifikation des Intelligenzstrukturmodells noch auf sein unbefriedigendes topologisches »feeling«, sondern kommt auf eine unerwartete dritte Möglichkeit: Die Schwäche liegt im bisherigen faktorenanalytischen Denkansatz. Sein Vorschlag: von den einzelnen Variablen als empirische Größen, in die in einer bestimmten Mischung Grunddimensionen intellektueller Leistungsfähigkeit eingehen, überzugehen zu Variablensummen, in denen jeweils über die Operationstypen bzw. über die Inhaltstypen hinweg verschiedene Variable zusammengeworfen sind. Durch eine solche Zusammenfassung entstehen also »Kunst«-Variable, die ein-

mal die inhaltlichen Dimensionen, einmal die operativen Dimensionen intellektueller Leistungsfähigkeit in reiner Form zum Vorschein bringen sollen.

Mystifikationsproblem: Schon intuitiv ist klar, daß Dimensionen, die im Raum der Einzel-Variablen nicht vorkommen, auch nicht durch Summenbildung herbeigezaubert werden können. Umgekehrt: Dimensionen, die für den Raum der Summen-Variablen gefunden werden, müssen sich, wenn auch mit vermutlich anderem Gewicht, im Raum der Einzel-Variablen wiederfinden lassen. Das läßt sich mathematisch auch für Psychologie-Studenten recht einfach zeigen.

Problem der Vermischung von Methoden- und Theorieebene: Was A.O. Jäger macht, ist bei Licht betrachtet ein rein methodischer Kunstgriff, der auf das Verstärken bzw. Herausfiltern bestimmter Dimensionen des Vektorraums hinausläuft. Neu entdeckt wird dabei kein einziger Faktor. Der Preis für diesen Kunstgriff ist allerdings, daß einer der wenigen Pluspunkte der Faktorenanalyse, nämlich die heuristisch recht einleuchtenden Grundgleichungen, verloren geht: diese führen ja in einer elementaren Weise die empirische Vielfalt intellektueller Leistungen auf möglichst wenige *unabhängige* Grunddimensionen zurück. Dieser methodische Kunstgriff wird nun aber zu einem neuen theoretischen Ansatz »aufgeplustert«: Ein neuer Faktorenbegriff, nach dem diese nicht mehr objektiv einen bestimmten Zusammenhang, Vpn-Variablenmenge charakterisieren, vielmehr überhaupt erst durch bestimmte Betrachtungsweisen konstituiert werden, also relativ zu wechselnden Betrachtungsweisen sind. Das verdunkelt die ohnehin nur schwer faßbare »Widerspiegelungsfunktion« von Faktorenanalysen noch zusätzlich (zur Widerspiegelungsproblematik von Logik/Mathematik/Statistik s. Leiser 1978 und 1983/Abschnitt A).

Schritt 7: Ergebnis: bei den »operationshomogenen« Summen-Variablen lassen sich wie gewünscht die 4 Operationsfaktoren auffinden, bei den »inhaltshomogenen« Summen-Variablen die 3 postulierten (und in die Variablenauswahl hineingesteckten) Inhaltsfaktoren.

Fazit: Es gibt die 3 material-/inhaltsspezifischen Intelligenz-Faktoren. Diese stecken aber nicht »objektiv« in den Einzel-Variablen, werden vielmehr erst durch eine bestimmte Betrachtungsweise gewonnen, bei der ich das Variablenmaterial explizit auf die postulierten inhaltlichen Modalitäten hin umstrukturierte.

Frage (wieder zur didaktischen Zuspitzung): Und was ist, wenn ich das Variablenmaterial stattdessen auf kausale, systemische, kombinatorische, kreative, klassifikatorische, intuitive usw. Modalitäten hin umstrukturierte? Die Frage ist durchaus inhaltlich gemeint, soll aber darüber hinaus ein Paradox verdeutlichen: Der Modalitätengesichtspunkt, der hier als Regulativ gegen das voreilige Entwickeln alternativer faktorenanalytischer Erklärungsmodelle zur Intelligenzstruktur ins Spiel gebracht wird, also unter dem Integrationsaspekt, würde zuende gedacht zu einer uferlosen Desintegration faktorenanalytischer Modellbildungsprozesse führen:

Indem die direkte Beziehung zwischen empirischen Variablen und Faktoren als ohnehin schwache Objektivitätsbasis der von Artefaktquellen nur so sprudelnden Faktorenanalyse aufgegeben wird, tut sich ein neuer grenzenloser Raum der Beliebigkeit in Form eigenwilliger Modalitätskonstruktionen auf. Forschungslogisch betrachtet: Konnte bisher an *einer* wie auch immer gefundenen Rotationslösung entschieden werden, ob es für einen gegebenen Variablensatz eine plausible Faktorenstruktur gibt oder nicht, so hat der Forscher nun, lasse ich nur einmal einfache Summenbildungen zu, unter dem Modalitätsgesichtspunkt schon bei 10 Variablen 115.974 Möglichkeiten der Umstrukturierung, ehe er aufgeben muß (s. Leiser 1976, 197).

Dieser letzte Schritt ist soweit im Aufsatz dokumentiert, daß sich den verschiedenen bisher nur vermuteten Willkür- und Artefakteffekten genauer nachgehen läßt.

Tabelle 2 (operationshomogene Variablenbündel, s. Anhang): Hier (1984, 28) sind anscheinend zunächst 6 Faktoren extrahiert worden (Summe der Eigenwert = 11.72 = 73.3 % der Gesamtvarianz). In der vorgeführten Varimax-Rotation tauchen dagegen wieder nur noch 4 Faktoren auf. Auch diesmal gibt es dafür zwar keinerlei Begründung, aber immerhin ein erahnbares Interesse: Bei der Rotation von mehr als 4 Faktoren hätte sich die Wirksamkeit des hier propagierten methodischen Kunstgriffs möglicherweise weniger prägnant demonstrieren lassen. Auf die in Schritt 5 gestellte Frage nach irgendeinem das Vorgehen erklärenden formalen Abbruchkriterium gibt es immer noch keine befriedigende Antwort:

- Eine Mindestgröße des Eigenwerts zwischen 1.59 und 1.20 kann es nicht sein, da diesmal noch ein Faktor mit Eigenwert 1.04 aufgenommen worden ist.
- Ein prozentualer Mindestbeitrag zwischen 3.3 % und 2.5 % kann es auch nicht sein, da der letzte Faktor diesmal einen Beitrag von 6.5 % leistet (1.04/16).
- Ein prozentualer Mindestbeitrag der Faktoren insgesamt zwischen 29.75 % und 32.25 % kann es auch nicht sein, da dieser diesmal bei 64.4 % liegt (Tab. 2 unten rechts).

Auch für die Faktorenextraktion bei der ursprünglichen Hauptkomponentenmethode ist keinerlei konsistentes Abbruchkriterium auszumachen. Erstaunlicherweise ist diesmal das Ladungslimit auf 0.40 angehoben worden. Liegt das vielleicht daran — so darf spekuliert werden —, daß bei einem Beibehalten des Limits von 0.35 weitere, die perfekte Zuordnung störende, Ladungen in die Tabelle Eingang gefunden hätten? An *einem* Punkt immerhin kommt jetzt ein wenig Licht ins Dunkel: Die diesmal komplett nachprüfbare Summe der »Kommunalitäten« entspricht exakt der Eigenwertsumme der ersten 4 Faktoren. Hinter den »Kommunalitäten« verbirgt sich also tatsächlich nichts anderes als der Varianzanteil im willkürlich ausgewählten Raum der ersten 4 Faktoren. Es

kommen damit Zweifel auf, ob überhaupt irgendwelche geschätzten Kommunalitäten in die Faktorenanalyse eingegangen sind oder ob mit einer unreduzierten R-Matrix gerechnet worden ist. Im zweiten Fall wäre das Ergebnis, inhaltlich betrachtet, mehr als enttäuschend: Die zur Stützung des Modalitätenschemas gedachte Faktorenmatrix würde in dieser »gereinigten« Form nur noch 58 % der zu erklärenden Varianz abdecken ($9.22 - \text{Summe der Ladungsquadrate} - / 16$)

Tabelle 3 (inhaltshomogene Variablenbündel, s. Anhang): Auch hier (1984, 29) sind zunächst 6 Faktoren mit 81 % Varianzanteil extrahiert worden. Wieder ist im Vergleich zu den o.g. Tabellen kein konsistentes Abbruchkriterium für diese Faktorenextraktion und die darauf folgende Rotation zu erkennen. Nehmen wir nur einmal die Faktorenextraktion:

- Die Eigenwertgröße kann es nicht sein: der letzte liegt jetzt bei 0.53 nach 0.56 (Tab. 7a) und 0.66 (Tab. 2).
- Der prozentuale Mindestbeitrag kann es auch nicht sein: der letzte liegt jetzt bei 4.4 % ($.53/12$) nach 1.2 % ($.56/48$) und 4.1 % ($66/16$).
- Der prozentuale Mindestbeitrag der Faktoren insgesamt kann es auch nicht sein: die Summe liegt jetzt bei 80.8 % ($9.69/12$) nach 36.6 % ($17.57/48$) und 73.3 % ($11.72/16$).

Berechnet man wie bei Tabelle 2 den Varianzanteil der »gereinigten« Faktorenmatrix, so liegt dieser diesmal bei knapp 57 % (6.78 (Summe der Ladungsquadrate)/12). Inhaltlich gewendet: trotz Bündeln, Verstärken und Filtern ist nur gut die Hälfte der Gesamtvarianz in dem eingeführten Modalitätenschema unterzubringen.

Tabelle 5 (»g«-Bündel, s. Anhang): Hier soll, gleichsam zur Kontrolle, demonstriert werden, daß sich durch eine entsprechende Überlagerung der beiden Modalitäten in den Bündeln jede strukturelle Differenzierung aufheben läßt. Dieser Demonstration wird, so scheint es, auch noch der verbleibende Rest an argumentativer Konsistenz geopfert: Um zu zeigen, wie sehr sich in diesem Fall die Struktur auf einen einzigen »Generalfaktor« reduziert, wird hier plötzlich und unvermittelt von der rotierten zur unrotierten Faktorenanalyse-Lösung gewechselt. Die 79.3 % für den Generalfaktor bei »Neutralisieren« beider Modalitäten sind selbstredend so gar nicht interpretierbar. Denn das unrotierte Ergebnis der Hauptkomponentenmethode führt per definitionem *immer* auf einen gewichtigen 1. Faktor. Wenn schon, müßte das Gewicht des 1. Faktors mit den Prozenten der *unrotierten* Lösungen für die operations- bzw. inhaltshomogenen Variablenbündel verglichen werden. Und auch hierzu müßte, will man seriös sein, der Stellenwert des jeweils 1. Faktors im Eigenwertverlauf bewertet werden. Vergleicht man, was naheliegt, den Varianzanteil des 1. Faktors mit dem Anteil, der bei gleichem Gewicht aller Faktoren herauskäme, geht der Vergleich sogar »nach hinten los«:

Tabelle 2: 36 % zu 6.25 % = 5.76-faches Gewicht

Tabelle 3: 46 % zu 8.33 % = 5.52-faches Gewicht

Tabelle 5: 79 % zu 25.00 % = 3.16-faches Gewicht.

Man kann die Frage auch umgekehrt stellen: Wie sähe denn Tabelle 5 aus, wenn auch hier nach Varimax rotiert würde? Würden darin alle 4 Faktoren einbezogen, kämen sicher weitere gewichtige Faktoren heraus, deren Interpretation A.O. Jäger einige Kopfschmerzen bereiten dürfte. Kurz gesagt: Die Schlüsse, die der Autor aus Tabelle 5 zieht, gründen ausschließlich auf einem manipulierenden Umgang mit der Faktorenanalyse und auf nichts sonst.

Ich fasse zusammen:

Die hier behandelten Aufsätze sind ein prägnantes Beispiel für den unkritischen und manipulativen Umgang mit der Faktorenanalyse. Sie sind in dieser Hinsicht so musterhaft, daß ich sie als festen Bestandteil in mein Seminarkonzept für multivariate Verfahren aufgenommen habe. — Uns wird vorgeführt:

- Ein unbekümmerter Umgang mit der statistischen Problematik der Faktorenanalyse, zugespitzt durch das Einführen von Stichproben auf der Variablen­seite.
- Das zirkelhafte Verhältnis von theoretischem Modell und empirischer Begründung (Auswahl der 48 Variablen).
- Das Verschwinden einer stringenten und falsifikationsfähigen Konfrontation von theoretischem Modell und empirischen Daten (Nicht: Sind die postulierten 7 Faktoren die relevanten Dimensionen der Intelligenz? Sondern: Durch welche Variablenauswahl und Methodentricks lassen sich die 7 Faktoren sichtbar machen?).
- Das Manipulieren von Rotationen durch Faktorenauswahl (Beschränkung auf den modellkonformen Raum von 4 Faktoren bei Tabelle 7a). Verstärkt wird dieser Manipulationseffekt übrigens noch — das gilt für alle angesprochenen Tabellen — durch das »Begradigen« von Faktorenmatrizen mit Hilfe von Ladungslimits, also durch eine Technik des Weglassens. Es hätte ja nichts dagegen gesprochen, wäre es nur um den didaktischen Darstellungseffekt hinsichtlich postulierter Ladungsmuster gegangen, jeweils die komplette Faktorenmatrix wiederzugeben und die über dem Limit liegenden Ladungen in jeder Spalte durch Einrahmung oder Fettdruck hervorzuheben.
- Die Beliebigkeit im Umgang mit solchen Ladungslimits (0.35 vs. 0.40) oder mit Bezugspunkten (Hauptkomponenten-Lösung vs. Varimax-Lösung, bereinigte Faktormatrizen vs. unbereinigte Varianzanteile).
- Die Kluft zwischen Faktorenanalyseanwendung und einschlägiger Kompetenz (das Rechnen und inhaltliche Argumentieren mit einem mißverständlichen »Kommunalitäts«-Maß).

- Das Einführen einer zusätzlichen Manipulationsebene (Bündelung von Variablen unter Modalitätsgesichtspunkten). Hier übersteigen die Aufsätze den Charakter eines Musters, enthalten vielmehr eine manipulationstechnische Innovation.

Am Ende meiner Kritik muß ich damit rechnen, daß diese für den Leser zwar nachvollziehbar ist, er sich aber fragt, worin die Relevanz und die verallgemeinerbare Bedeutung einer solchen Kritik bestehen kann. Ich habe also davon auszugehen, daß der pragmatische Anlaß dieser Kritik und meine argumentationstechnischen Gründe die gewählte Kritikebene als legitime Form wissenschaftlicher Auseinandersetzung für den Leser noch nicht hinreichend plausibel machen.

Daher möchte ich abschließend den hier vorgeführten Kritiktyp über seinen praktischen Vollzug hinaus noch einmal kurz vor dem Hintergrund bestimmter zu antizipierender Einwände kennzeichnen und damit seine eingangs behauptete Relevanz verdeutlichen.

Ein erstes Spezifikum: Ich greife punktuell einen bestimmten Text heraus und »spieße« dabei, sozusagen unter der Lupe, bestimmte inhaltliche und methodische Ungereimtheiten auf.

Einwand: Ist ein solches Vorgehen nicht »spitzfindig«? Hat eine seriöse Kritik nicht zumindest das Gesamtwerk des kritisierten Autors, noch besser den Gesamtkontext der inhaltlichen und methodischen Diskussion zur faktorenanalytischen Intelligenzstruktur einzubeziehen?

Meine Antwort: Der Diskussionskontext in den hier vorliegenden (und anderen) Bereichen der Psychologie hat einen Grad der Verfestigung, Verselbständigung und Verästelung erreicht, daß bestimmte elementare Fragen, etwa nach der inhaltlichen und methodischen Rationalität der Grundkategorien, im Rahmen des vorgegebenen Diskurses gar nicht mehr darstellbar sind.

Beispiel: Was bedeutet eigentlich eine faktorenanalytische Strukturkategorie, die unterschiedslos auf die geometrischen Grunddimensionen der Thurstonischen Zylinder und auf die dynamischen und algorithmischen Charakteristiken kognitiver Prozesse angewendet werden kann?

Was bedeutet eigentlich der Glaube, daß sich in den Dimensionen einer Faktorenanalyse, die immer nur ein Verhältnis kennzeichnen können, nämlich zwischen einer halbwegs wohlbestimmten Population und einer mehr oder weniger künstlichen Menge von Testleistungen, so etwas wie populations- und testinvariante, also absolute, Dimensionen »der Intelligenz« offenbaren könnten?

Was bedeutet es, die nicht individuen- sondern kulturspezifische Werkzeuggebene kognitiver Leistungen einfach auszublenden? Konkret: Der Einsatz allgemeiner Werkzeuge, die sich ein Schüler in Form logischer, mathematischer, algorithmischer oder begrifflicher Schemata *angeeignet* hat, wird die »Struktur« der betroffenen Intelligenzleistungen (in welchem Sinn auch immer) kaum unverändert lassen.

All diese Fragen sind im herrschenden Diskurs zur Faktorenanalyse weitgehend verdrängt, und Versuche, sie wieder freizulegen, bleiben oft selbst im Netzwerk dieses Diskurses hängen. Ein Beispiel ist für mich die interessante und materialreiche, aber in den Zuspitzungen und Schlußfolgerungen eher schwache Untersuchung von Lenk (1983) zur Faktorenanalyse. Meine Konsequenz aus diesem Dilemma des Kritikers, das Irrationale solcher Diskurse nicht von außen entlarven zu können, aber es auch nicht, im Diskurs eingefangen, zu reproduzieren: Wir können mit Freud darauf vertrauen, daß das Verdrängte irgendwie zurückkehrt, in den »Nischen« des Diskurses aufzuspüren ist, in Form von Fehlern, Diskontinuitäten und Ungereimtheiten.

Ein anderes Spezifikum: Ich greife mir ein »durchschnittliches«, fehlerbehaftetes Anwendungsbeispiel der Faktorenanalyse heraus, und nicht eines, das sich auf der erreichbaren Höhe der inhaltlichen und methodischen Argumentation befindet.

Einwand: Zumindest die methodischen Fehler in den kritisierten Texten sind doch vermeidbar. Meine Kritik trifft folglich höchstens die Inkompetenz des speziellen Autors und nicht das Konzept der Faktorenanalyse selbst.

Meine Antwort: Der Wissenschaftsbetrieb wird nicht von einer idealen, sondern von der durchschnittlichen Praxis bestimmt. Darüber hinaus: In einem Wissenschaftsbereich wie der Psychologie, wo sich ideologische Funktion und Erkenntnisfunktion so stark überlagern, kann sich geradezu zwingend eine Funktionsteilung etablieren zwischen einem Methodenideal, das der wissenschaftlichen Legitimation dient, und einer irrationalen Praxis, die ein bestimmtes Problemfeld mit Hilfe eines Diskurses ideologisch zu besetzen und zu kontrollieren hat. Auch wenn dem Leser diese Interpretation zu weit geht: die *durchschnittliche* Praxis ist für mich ein wissenschaftstheoretisch und diskurstheoretisch legitimer Untersuchungsgegenstand und noch dazu einer, der in seiner Phänomenologie streng empirisch zugänglich ist.

Ich begeben mich bei diesem Kritiktyp gleichsam — um das für mich immer wieder paradigmatische Bild im Märchen von des Kaisers neuen Kleidern aufzugreifen — bewußt und nicht naiv in die Rolle des kleinen Kindes, das den ganzen aufgeregten Diskurs von außen beguckt und auf die nackten Stellen zeigt.

Warum frage ich, soll ich mich mit den prächtigen Gewändern der faktorenanalytischen Intelligenzforschung beschäftigen, solange mir diese Nacktheit ins Auge springt?

Um abschließend zum didaktischen Aspekt dieser kritischen Rekonstruktion zurückzukehren. — Die Strategie bestand darin:

— den spezifischen Verselbständigungen und Irrationalismen nachzugehen, die nicht primär auf die Person des kritisierten Autors, sondern auf Dispositive des Verfahrens selbst zurückgehen, die im Kontext eines bestimmten Psychologieverständnisses voll »durchschlagen«;

- dabei auf dem auch bei solchen Methoden noch einfachen methodischen Grundschema zu insistieren, das gegen alle Vernebelungen durchzuhalten ist. Oberstes Prinzip ist hierbei, sich durch verwickelte Praktiken nicht einschüchtern zu lassen, diese vielmehr immer wieder auf dieses einfache Grundschema zurückzuzwingen;
- die vorgetragenen Ergebnisse, also die inhaltlichen Aussagen des Autors, immer wieder mit den elementaren kategorialen Bestimmungen des psychologischen Gegenstands zu konfrontieren: Wie ich zeigen wollte, ist es so möglich, den mathematisch-statistischen »deus ex machina«, der aus Formelapparaten inhaltliche Erkenntnisse herauszaubert, zu entmystifizieren. Didaktik und Kritik durchdringen sich also wechselseitig bei diesem Zugang zur psychologischen Methodenlehre: Methodenausbildung zielt als zentrales Moment auf die Entwicklung von Kritikfähigkeit. Und Kritik zielt auf mehr als Entlarvung, nämlich auf eine Präzisierung und Erweiterung des Problem- und Fragehorizonts.

Literaturverzeichnis

- Harmann, H.H.*, 1967. Modern Faktor Analysis. Chicago: The University of Chicago Press.
- Jäger, A.O.*, 1982. Mehrmodale Klassifikation von Intelligenzleistungen: Experimentell kontrollierte Weiterentwicklung eines deskriptiven Intelligenzstrukturmodells. *Diagnostica*, 28, 195-225.
- Jäger, A.O.*, 1984. Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven. *Psychologische Rundschau*, Band XXXV, 21-35.
- Lawley, D., Maxwell, A.E.*, 1971. Factor Analysis as a Statistical Method. London: Butterworths.
- Leiser, E.*, 1971. Eine Methode zum Vergleich zweier Faktormatrizen. *Methods of Information in Medicine*, 10, 117-120.
- Leiser, E.*, 1976. Simultane vs. sequentielle Clusteranalyse. In: H. J. Bochnik/W. Pittrich: Multifaktorielle Probleme in der Medizin. Wiesbaden: Akademische Verlagsgesellschaft, 195-208.
- Leiser, E.*, 1978. Widerspiegelungscharakter von Logik und Mathematik. Frankfurt: Campus.
- Leiser, E.*, 1983 / 2. Aufl. Grundkurs Statistik. Köln: Pahl-Rugenstein.
- Leiser, E.*, 1987. Psychologische oder mathematische Statistik? *Forum Kritische Psychologie* 19, 59-74.
- Lenk, W.*, 1983. Faktorenanalyse: Ein Mythos? Weinheim: Beltz.
- Roppert, J., Fischer, G.*, 1965. Lineare Strukturen in Mathematik Statistik. Wien: Physica-Verlag.

Tabellen-Anhang (Tab. 7a, aus Jäger 1982; Tab. 2, 3 und 5 aus Jäger 1984)

Tabelle 7a: Faktorenanalyse der 48 ungebündelten Einzelvariablen
Auszug der Varimax-Lösungen mit 4 Faktoren — alle Ladungen > .35

Variable	4er Lösung				h ²	Variable	4er Lösung				h ²
	F1	F2	F3	F4			F1	F2	F3	F4	
ZM 1	63				47	XG 1		76		62	
FM 1	59				44	XG 2		70	54	54	
FM 2	54			36	43	SI 1		55		38	
NM 3	54				34	DR 1		35	63	54	
ZZ 1	50				33	RD 1			39	23	
KF 2	47				30	RZ 2			34	16	
GG 1	47				25	WK 1			33	23	
ZM 2	47				28	WG 2			31	16	
PR 1	45				29	ZN 2			42	47	
ST 1	42				29	AW 1				47	
NM 1	40				23	ZN 1			41	46	
KF 1	36				25	GO 2				46	
EF 2		63			44	HI 4				45	
ZK 2		63			51	AW 2				43	
IT 3		62			43	BN 2				37	
ZK 1		62		37	52	WA 2				37	
ZF 1		59			35	GO 1				36	
OJ 1		57			35						
AF 1		50			27						
TN 1		48			24						
AF 2		46			24						
TN 2		46			27						
Faktor	G	E	B	K		Faktor	G	E	B	K	
% Ges. Varianz						% aufg. Varianz					
7.4						24.8					
8.1						27.2					
7.8						26.2					
6.5						21.7					
29.8						100					

Eigenwerte: 8,09— 2,64—1,96—1,59—1,20—0,80—0,73—0,56. Aufgenommen wurden nur Variablen mit Ladungen > .35. Die am Fuß der Tabelle angegebenen Varianzanteile enthalten auch die hier nicht aufgeführten Ladungen < .35. Die Interkorrelationsmatrix kann beim Verfasser angefordert werden.

Tabelle 2: Faktorenanalyse
16 operationshomogener Bündel
Varimax-Lösung, alle Ladungen > .40

Variable	F1	F2	F3	F4	h ²
G 1	79				67
G 2	84				74
G 3	74				62
G 4	72				62
E 1		85			77
E 2		84			74
E 3		78			69
E 4		74			64
K 1			70		58
K 2			78		65
K 3			70		54
K 4			64		44
B 1				77	72
B 2				76	65
B 3				65	65
B 4				67	58

% ges.V. 16.7 17.8 15.0 14.9 64.4
% aufg.V. 26.0 27.6 23.3 23.2 100

Faktor G E K B
Eigenwerte: 5,82—1,96—1,49—1,04—0,75—0,66

Tabelle 3: Faktorenanalyse
12 inhaltshomogener Bündel
Varimax-Lösung, alle Lad. > .40

Variable	F2	F1	F3	h ²
F 1	77			66
F 2	77			66
F 3	71			61
F 4	71			61
N 1		77		72
N 2		80		78
N 3		82		75
N 4		74		62
V 1			76	66
V 2			77	65
V 3			72	62
V 4			67	61

% ges.V. 21.9 23.7 20.7 66.3
% aufg.V. 33.1 35.7 31.2 100

Faktor F N V
Eigenw.: 5,55—1,26—1,15—0,66—0,54—0,53

Tab. 5:
Faktoranalyse
der 4 »g«
Bündel / erster
unrotiert. Fakt.

Variable	F 1
»g« 1	90
»g« 2	90
»g« 3	89
»g« 4	87

% ges. V. 79.3
% aufg. V. 100

Eigenwerte:
3,17—0,33—
0,28—0,22