

Selbstbeurteilungs-Skalen zur Abschätzung des „subjektiven Befundes“ in psychopathologischen Querschnitt- und Längsschnitt-Untersuchungen* **

D. v. Zerssen

Max-Planck-Institut für Psychiatrie, München

Eingegangen am 18. April 1973

Self-Rating Scales for Evaluation of a Patient's Subjective State in Cross-Sectional and Longitudinal Studies of Psychopathology

Summary. Self-rating scales for the evaluation of a patient's subjective state can be used to document his symptoms. They also yield quantitative scores representing individual degrees of various forms of psychopathology. These scores lend themselves to statistical comparison of groups of patients and controls. Repeated administration of the scales makes it possible to follow the course of symptomatology over a period of time on the basis of these scores. Three different self-rating scales are described in detail. Each of them exists in two parallel forms which correlate closely with each other ($r \approx 0.9$): 1. The unidimensional „Befindlichkeits-Skala“ (mood-scale, B-S and B-S'), each test form consisting of 28 pairs of adjectives. It is particularly suitable for evaluation of the subjective state in longitudinal investigations of patients with affective illnesses. 2. The „Beschwerden-Liste“ (list of symptoms, B-L/B-L'), comprising 24 symptoms listed as nouns in each test form. It was developed mainly for estimation of the degree of general and somatic symptoms. 3. A multidimensional psychopathologic symptom scale for differential evaluation of three dimensions of mental symptoms indicative of paranoid, neurotic and depressive tendencies. In its provisional form this scale (PND-S/PND-S') is composed of 99 statements, most of which relate to the psychopathological dimensions under discussion while others are related to certain aspects of faking. A scale shortened by 30 items in each of the parallel versions is being prepared. The self-rating scales discussed in this paper have been successfully administered in clinical practice and clinical research in Germany and Switzerland. They can also be recommended for the examination of out-patients and for epidemiological field work.

Key words: Self-Rating Scales — Psychopathological Measurement and Documentation — Mood-Scale (B-S) — List of Symptoms (B-L) — Paranoid-Neuroticism-Depression-Scale (PND-S).

Zusammenfassung. Selbsbeurteilungs-Skalen zur Erfassung des „subjektiven Befundes“ liefern — über die Dokumentation des individuellen Beschwerdenbildes

* Nach einem auf der 15. Jahrestagung der Deutschen Gesellschaft für Medizinische Dokumentation und Statistik in der DGD in Frankfurt a.M. im Oktober 1970 gehaltenen Vortrag.

** Herrn Professor Dr. Gerd Peters zum 65. Geburtstag gewidmet.

hinaus — quantitative Werte, durch die sich individuelle Ausprägungsgrade verschiedener Formen subjektiver Gestörtheit kennzeichnen lassen. Diese Werte sind sowohl für den statistischen Vergleich verschiedener Patientengruppen untereinander und/oder mit Kontrollgruppen und — bei wiederholter Anwendung der Skalen beim selben Fall — zur Verlaufskontrolle geeignet. — Drei Untersuchungsinstrumente der bezeichneten Art, die in jeweils zwei hoch ($r \approx 0,9$) miteinander korrelierenden Parallelformen vorliegen, werden im einzelnen besprochen: 1. Die aus je 28 Eigenschaftspaaren bestehende eindimensionale Befindlichkeits-Skala (B-S bzw. B-S'), die besonders zur fortlaufenden Registrierung des subjektiven Befindens im Rahmen von Längsschnittuntersuchungen bei Patienten mit affektiven Störungen geeignet ist. 2. Die aus je 24 — als Substantiva angegebenen — Beschwerden aufgebaute Beschwerden-Liste (B-L bzw. B-L'), die vornehmlich zur Erfassung von somatischen und Allgemeinbeschwerden entwickelt wurde. 3. Eine multidimensionale psychopathologische Skala zur differenzierteren Erfassung von im engeren Sinne psychischen Beschwerden, nämlich solchen paranoider, neurotischer bzw. depressiver Färbung. Diese Skala (PND-S bzw. PND-S') besteht in ihrer vorläufigen Fassung aus je 99 Sätzen, die sich großteils auf die genannten Formen subjektiver Gestörtheit beziehen, zum kleineren Teil auf Verfälschungstendenzen bei der Selbstbeurteilung. Eine um je 30 Sätze gekürzte Fassung wird vorbereitet. Die hier besprochenen Selbstbeurteilungs-Skalen haben sich bei klinischen Routineuntersuchungen sowie im Rahmen von Forschungsprojekten in der Bundesrepublik und der Schweiz bewährt, können aber auch für ambulante Untersuchungen und für epidemiologische Feldstudien empfohlen werden.

Schlüsselwörter: Selbstbeurteilungs-Skalen — Psychopathologische Befunderhebung und -dokumentation — Befindlichkeits-Skala (B-S) — Beschwerden-Liste (B-L) — Paranoid/Neurotizismus/Depressivitäts-Skala (PND-S).

I. Allgemeine Grundsätze

Die Forderung nach Reproduzierbarkeit und damit Kontrollierbarkeit von Ergebnissen wissenschaftlicher Untersuchungen muß auch an die Psychiatrie gestellt werden. Insbesondere die weltweite Entwicklung der psychopharmakologischen Forschung hat dieser Forderung Nachdruck verliehen. Folge ist ein zunehmendes Bemühen darum, klinische Prozeduren strenger zu standardisieren, die Befunde genauer zu dokumentieren und die klinische Befunderhebung durch objektivierende Verfahren zu ergänzen. Zu diesen gehören Selbstbeurteilungs-Fragebögen zur Abschätzung des „subjektiven Befundes“ im Sinne von Jaspers (1965), die den Probanden in vorformulierter Form vorgelegt und von ihnen selbst ausgefüllt werden (vgl. v. Zerssen, 1973). Die Probanden haben dazu die von ihnen als für sie zutreffend angesehenen Formulierungen — einer schriftlichen Anweisung gemäß — zu markieren. Das kann etwa durch Unterstreichen eines jeder einzelnen Formulierung zugeordneten „Ja“ bzw. „Nein“ geschehen oder durch das Ankreuzen einer abgestuften Beantwortungskategorie (z.B. „trifft ausgesprochen zu“ — „trifft überwiegend zu“ — „trifft kaum zu“ — „trifft gar nicht zu“). Die Formulierungen — Eigenschaftswörter, Substantiva oder ganze Sätze in indikativer oder Frageform — sollen möglichst der Sprache des „Durch-

schnittspatienten“ angemessen sein. Sie beziehen sich auf bestimmte Aspekte des „subjektiven Befundes“, also auf Beschwerden, wie sie auch in jeder eingehenden ärztlichen Untersuchung erfragt werden.

Ein leicht ersichtlicher, in erster Linie praktischer Vorteil des Selbstbeurteilungsmodus gegenüber der üblichen Explorationstechnik ist wirtschaftlicher Natur: Für eine gründliche, streng standardisierte, insbesondere auch gleichartig zu verschiedenen Zeitpunkten der Behandlung wiederholte Befragung von Patienten fehlt es den an einer Untersuchung beteiligten Ärzten oft an Zeit. Die globale klinische Einschätzung des jeweiligen Zustandsbildes kann aber Detailinformationen meist nicht vollständig ersetzen. Diese lassen sich jedoch durch Selbstbeurteilungs-Fragebögen der bezeichneten Art relativ einfach gewinnen: Für die jeweilige Fragestellung besonders geeignete Bögen können den Patienten u.U. auch von entsprechend instruierten Pflegepersonen vorgelegt und im allgemeinen von den Kranken ohne fremde Hilfe ausgefüllt werden. Die Auswertung, die ebenfalls vom Pflegepersonal oder von technischem Personal übernommen werden kann, erfolgt nach einem vorgegebenen Auswertungsschlüssel.

Daraus ergibt sich ein weiterer, vor allem für wissenschaftliche Aussagen bedeutungsvoller Vorteil des Selbstbeurteilungsmodus: Die Möglichkeit einer subjektiven Beeinflussung von Untersuchungsergebnissen durch das „Ergebnis-Wunschkbild“ (Hellpach, 1936) des Untersuchers wird dadurch ausgeschaltet. Deshalb sollte auf derartige Untersuchungsverfahren auch dann nicht verzichtet werden, wenn eine eingehende — gegebenenfalls auch fortlaufende — klinische Untersuchung erfolgt; denn auf diese Weise lassen sich die Explorationsergebnisse mittels eines unabhängig von ihnen gewonnenen, aber in wesentlichen Punkten vergleichbaren Kriteriums kontrollieren.

Selbstverständlich beinhaltet der in entsprechenden Fragebögen erfaßbare „subjektive Befund“ nur *einen* Aspekt der Psychopathologie. Der „objektive Befund“ muß, wenn es auf eine umfassende Objektivierung klinischer Eindrücke ankommt, in anderer Weise — z.B. durch Wahrnehmungs- und bzw. oder Leistungs-Tests etc. — geprüft werden. Für viele Fragestellungen ist aber die Information über den „subjektiven Befund“ und seine Veränderungen durchaus relevant und u.U. sogar im wesentlichen ausreichend. Das gilt z.B. für die Verlaufsanalyse endogen-depressiver Verstimmungen, für die eine weitgehende Kongruenz von subjektivem und objektivem Befund kennzeichnend ist (Alsen, 1961).

Bei der Auswertung von Fragebögen faßt man im allgemeinen mehrere der Formulierungen, die zur Beantwortung angeboten werden und die man als „Test-Items“ (engl.) bezeichnet, zu Einheiten zusammen; evtl. bilden sogar sämtliche Test-Items eine solche Einheit. Die Zu-

sammenfassung geschieht entweder hypothetisch nach rein inhaltlichen Gesichtspunkten — etwa nach der Zusammengehörigkeit der durch die Items repräsentierten Beschwerden im Rahmen bekannter klinischer Syndrome — oder empirisch aufgrund statistisch ermittelter Kriterien, nämlich entweder des Trennvermögens der Items beim Vergleich verschiedener klinischer Gruppen (z.B. endogen Depressive versus Schizophrene) oder ihrer Zusammengehörigkeit im Sinne statistischer Syndrome (als Cluster oder als Dimensionen = Faktoren; vgl. Pawlik, 1968). Man kann bei der Erstellung eines Auswertungsschlüssels die genannten Vorgehensweisen auch miteinander verbinden und Einheiten aus solchen Items bilden, die der klinischen Vorhersage (Hypothese) entsprechend empirisch bestimmte klinische Gruppen differenzieren und darüber hinaus einen statistischen Zusammenhang aufweisen.

Eine aus mehreren (evtl. allen) Fragebogen-Items bestehende Einheit bezeichnet man als Skala, weil sie — *cum grano salis* — die Eigenschaften einer Meß-Skala besitzt: Bewertet man die Einzelantworten mit Punkten (etwa „Ja“ = 1; „Nein“ = 0) und addiert die zu einer Einheit gehörigen Punkte zu einem „Test-Score“ auf, so lassen sich die Werte der untersuchten Fälle ihrem Ausprägungsgrad pro Item-Gruppe nach wie auf einer Skala anordnen. Über die Objektivierung des subjektiven Befundes hinaus wird damit auch seine Quantifizierung ermöglicht. Der Vorteil für Gruppenvergleiche (etwa Patienten versus gesunde Kontrollpersonen, psychiatrische versus internistische Fälle, paranoid versus nicht-paranoid Schizophrene), für Verlaufsanalysen (etwa endogen Depressive bei Klinikeintritt, nach einwöchiger, zweiwöchiger, etc. Behandlung, kurz vor und vier Wochen nach Entlassung) und für eine Kombination von beidem (etwa beim Vergleich der Wirkung zweier Thymoleptica auf den Verlauf endogener Depressionen) liegt auf der Hand: Man kann Ausprägungsgrade miteinander vergleichen und dafür die entsprechenden statistischen Verfahren einsetzen (vgl. Eysenck, 1963; Hofstätter u. Wendt, 1967). Auf diese Weise gelingt es manchmal schon bei relativ kleinen Fallzahlen, statistisch gesicherte Unterschiede zwischen Versuchsgruppen aufzudecken bzw. Ergebnisse anderer Untersucher, die ebenfalls mit solchen quantifizierenden Testverfahren gearbeitet haben, zu überprüfen.

Der störende Einfluß von — bewußten oder unbewußten — Verfälschungstendenzen auf das Testergebnis läßt sich bis zu einem gewissen Grade durch die Art der Skalen-Konstruktion (z.B. willkürlicher Wechsel der Schlüsselrichtung bei den einzelnen Items) einschränken. Spezielle Kontroll-Skalen (Parallel-Skalen und Kontroll-Skalen im engeren Sinne, z.B. Skalen zur Erfassung von Tendenzen zur Schönfärberei, auch „Lügen-Skalen“ genannt) ermöglichen es zudem, solche Verfälschungstendenzen

einigermaßen abzuschätzen und bei der Testinterpretation zu berücksichtigen (vgl. Eysenck, 1953; Hoeth u. Gregor, 1964).

Für die Interpretation ist letztlich die Validität eines Tests ausschlaggebend, d.h. das Ausmaß, in dem er tatsächlich erfaßt, was er erfassen soll — z.B. den Grad der Depressivität. Eine wesentliche Bedingung dafür ist die Zuverlässigkeit des Tests, d.h. die Genauigkeit, mit der er bestimmte Werte liefert — was auch immer diese bedeuten mögen. Dazu müssen die in einer Skala zusammengefaßten Bestandteile (z.B. zwei gleichwertige Hälften einer Skala) zu übereinstimmenden Resultaten führen, d.h. hoch miteinander korrelieren (sog. Halbierungszuverlässigkeit); ferner müssen die Testergebnisse bei wiederholter Anwendung unter gleichen Bedingungen miteinander übereinstimmen (sog. Wiederholungszuverlässigkeit) — vorausgesetzt, daß keine Veränderung in der Psychopathologie der untersuchten Fälle eingetreten ist. Da diese Voraussetzung häufig nicht gegeben ist, spielt diese zweite Art der Zuverlässigkeit für Skalen zur Bestimmung aktueller psychischer Störungen nur eine untergeordnete Rolle.

Weitere Einzelheiten über die Eigenschaften psychopathologischer Selbstbeurteilungs-Skalen müssen der Literatur entnommen werden (Anastasi, 1969; Brengelmann, 1963; Eysenck, 1953; Gilbert, 1969; Lienert, 1967; Michel, 1964; Mittenecker, 1964).

II. Eigene Beiträge

In dem Bestreben, die psychopathologische Befunderhebung zu standardisieren und die Befunde soweit wie möglich zu objektivieren und zu quantifizieren, wurden an der Psychiatrischen Abteilung des Max-Planck-Instituts für Psychiatrie in München — neben verschiedenen Formen ärztlicher Untersuchungsbögen, insbesondere denen nach Lorr *u. Mitarb.* (1963) und der AMP (vgl. Angst *et al.*, 1969; ferner Mombour 1972) — auch Selbstbeurteilungs-Skalen zur Bestimmung des subjektiven Befundes erprobt bzw. neu entwickelt. Aus diesen Versuchen ist eine Serie von Verfahren hervorgegangen, die es — als „Testbatterie“ zusammengefaßt — gestattet, ein breites Spektrum von Beschwerden zu erfassen und — bei wiederholter Anwendung — ihre Veränderungen unter der Behandlung objektivierend zu verfolgen (vgl. v. Zerssen, 1973).

Es handelt sich um drei Fragebögen, die jeweils in zwei hoch miteinander korrelierenden Parallelformen vorliegen. Jeder Bogen enthält eine bis drei klinische Skalen. Zusammen mit einem kurzen verbalen Intelligenztest in Papier- und Bleistiftform (dem Untertest „Allgemeines Wissen“ aus dem HAWIE; vgl. Wechsler, 1964) und Test-Skalen zur Erfassung prämorbidier Persönlichkeitszüge (vgl. v. Zerssen *et al.*, 1969, 1970b), werden sie routinemäßig bei allen Patienten unserer Abteilung

angewendet, und zwar kurz nach der Aufnahme und kurz vor der Entlassung, z.T. auch in regelmäßigen Abständen zwischen diesen Zeitpunkten. Im Zusammenhang mit anderen Fragestellungen können sie aber auch unabhängig voneinander eingesetzt werden. Die Kombination mit einem Intelligenztest ist allerdings in jedem Fall zu empfehlen, da bei einem Verbal-IQ unter 80 damit zu rechnen ist, daß die intellektuellen bzw. die motivatorischen Voraussetzungen für eine sinnvolle Selbstbeurteilung fehlen.

Im einzelnen handelt es sich bei den von uns verwendeten Symptom-Fragebögen um eine Befindlichkeits-Skala, eine Beschwerden-Liste und eine mehrdimensionale psychopathologische Symptom-Skala.

1. Die Befindlichkeits-Skala (vgl. v. Zerssen *et al.*, 1970a, 1973) besteht aus 28 jeweils gegensätzlichen Eigenschaftspaaren, die zur Kennzeichnung der aktuellen Gemütsverfassung geeignet sind und sich — laut Testanweisung — auf den *augenblicklichen* Zustand beziehen. Die Skala ist „eindimensional“, d.h. sie „mißt“ nur *einen* — von uns als „Befindlichkeit“ bezeichneten — Aspekt der subjektiven Verfassung, und zwar mit großer Zuverlässigkeit. Die Skala wurde speziell für Verlaufsuntersuchungen an Patienten mit affektiven Psychosen entworfen. Tatsächlich reagieren die Skalenwerte sehr empfindlich auf Schwankungen der Gemütslage, aber in einer keineswegs für phasenhafte Verstimmungen vom depressiven bzw. manischen Typ spezifischen Weise. Deshalb eignet sich das Verfahren ganz allgemein besonders zur Längsschnittanalyse der Befindlichkeit, was z.B. für psychopharmakologische Untersuchungen von grundsätzlicher Bedeutung ist. Für eine Querschnittsanalyse, etwa die differenzierte Erfassung der Beschwerden eines Falles oder der für eine Patientengruppe typischen Beschwerdenkonstellation, ist das Verfahren weniger geeignet.

Die beiden Parallelformen des Tests (B-S und B-S' genannt) korrelieren, wenn man sie unmittelbar nacheinander anwendet, in der Höhe von 0,9 miteinander. Sie können deshalb im Rahmen einer Längsschnittstudie demselben Patienten einzeln in beliebiger Reihenfolge vorgelegt werden. Trägt man die Werte pro Testanwendung gegen die Zeitskala der Untersuchungsreihe ab und verbindet die Markierungspunkte durch einen Linienzug miteinander, so erhält man für jeden Fall eine Verlaufskurve seiner Befindlichkeit, die einer Fieberkurve ähnelt. Man kann danach typische Verlaufsformen von Befindlichkeitsstörungen unterscheiden (Abb. 1 und 2). Die klinische Validität dieser Verlaufsformen konnte — zumindest für endogen Depressive — durch Untersuchungen von Schwarz u. Strian (1972) an unserem Institut eindrucksvoll belegt werden: Über mindestens 12 Zeitpunkte durchgeführte Testungen korrelierten bei insgesamt 30 endogen Depressiven mit einer unabhängig

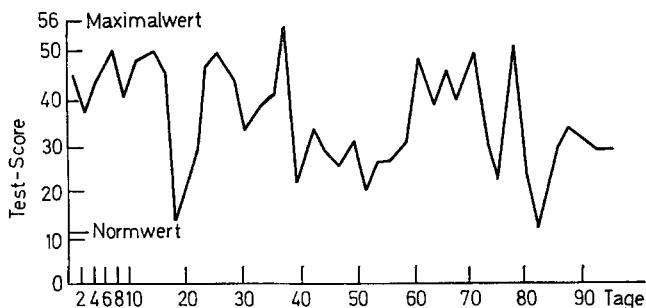


Abb. 1. Verlaufstyp der Befindlichkeit bei einem Fall von neurotischer Depression (nach v. Zerssen *et al.*, 1970 a). Typ einer Fieberkurve bei undulierendem Fieber

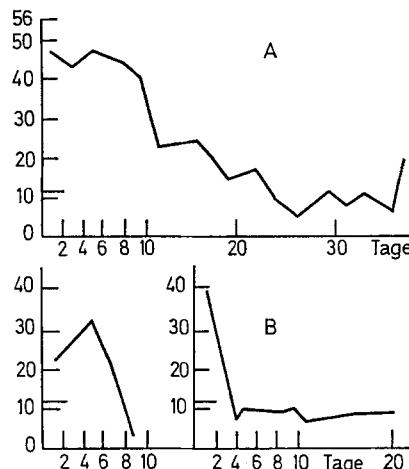


Abb. 2. Verlaufstyp der Befindlichkeit bei zwei Fällen von endogener Depression (nach Zerssen *et al.*, 1970 a)

Fall A: Typ einer Fieberkurve bei lytischer Entfieberung; Fall B (zwei Phasen): Typ einer Fieberkurve bei kritischer Entfieberung

durchgeführten globalen klinischen Schätzung des Verstimmungszustandes (entlang einer 7-Punkt-Skala) im Schnitt in der Höhe von 0,92 (Spearmanscher Rangkoeffizient)!

2. Die Beschwerden-Liste (vgl. Baumann, 1972; v. Kerekjarto *et al.*, 1972; v. Zerssen, 1971 a) dient uns zur Erfassung ubiquitärer Beschwerden, wie sie — außer von psychiatrischen Fällen — auch von körperlich Kranken häufig angegeben werden. In der ursprünglichen Fassung

enthält die Liste 64 Items in „Ja“/„Nein“-Form und ein zusätzliches Item für weibliche Probanden („Regelstörungen“), das im Test-Score nicht berücksichtigt wird. Neuerdings verwenden wir die Liste in einer parallelisierten Kurzform mit 24 Items pro Bogen (B-L und B-L') und einer abgestuften Skalierung pro Item („stark“ — „mäßig“ — „kaum“ — „gar nicht“). Hinzu kommt ein Ergänzungsbogen (B-L°) mit solchen Beschwerden, deren Angabe für den Arzt zwar informativ, für eine zusammenfassende Skalierung aber aus statistischen Gründen nicht geeignet ist und deshalb bei der quantitativen Auswertung nicht berücksichtigt wird. Beide Parallelformen korrelieren in der Höhe von 0,8 bis 0,9 miteinander (Baumann, 1972; v. Zerssen, 1971 a).

Tabelle 1. Statistische Kennwerte der Beschwerden-Liste: Gesamt-Score der ursprünglichen Testfassung (nach v. Zerssen, 1971 a)

Kollektiv	<i>N</i>	\bar{X}	<i>s</i>
Gesunde	92	10,1	7,9
Schizophrene	214	21,7	12,3
Neurotiker	87	23,7	12,0
end. Depressive	181	26,6	10,8

Die Gesamtzahl der angegebenen Beschwerden gibt (in der Neufassung mit abgestufter Skalierung pro Item — von „kaum“ = 1 über „mäßig“ = 2 bis zu „stark“ = 3 — zusammen mit dem jeweiligen Ausprägungsgrad) einen Anhaltspunkt für das Ausmaß subjektiver Beeinträchtigung durch — vorwiegend körperliche — Beschwerden. Läßt sich eine organische Ursache für die angegebenen Beschwerden ausschließen, so kann daraus auf eine entsprechende psychovegetative Labilität bzw. auf starke Klagsamkeit geschlossen werden. Die Skalenwerte sind demgemäß — ebenso wie die der Befindlichkeits-Skala — bei endogen Depressiven und bei Neurotikern im Schnitt besonders stark erhöht, ebenfalls deutlich, wenn auch durchschnittlich nicht ganz so ausgeprägt, bei Schizophrenen (s. Tab.1). In zeitlichem Zusammenhang mit der klinischen Besserung normalisiert sich der Beschwerden-Score, während er konstant bleibt, wenn keine wesentliche Änderung des Befindens eintritt. Er verhält sich also prinzipiell so wie ein sog. Neurotizismus-Score (vgl. Rautenstrauch, 1967). Erwartungsgemäß korreliert er mit dem Score der Neurotizismus-Skalen aus dem MMQ und aus dem MPI nach Eysenck (1953, 1959) in derselben Höhe wie diese untereinander (= 0,6 bis 0,7; vgl. v. Kerekjarto *et al.*, 1972; v. Zerssen, 1971 a) und geht (in einer faktorenanalytischen Untersuchung an klinisch gesunden Studenten) mit der

höchsten Ladung (0,9) in einen Faktor ein, der vornehmlich von den genannten und ähnlichen Skalen konstituiert wird.

Der Vorteil der Beschwerden-Liste gegenüber den üblichen Neurotizismus-Skalen liegt einmal in ihrer einfacheren Form, durch die auch intellektuell wenig differenzierten oder psychisch stark gestörten Patienten die Beantwortung erleichtert wird; denn die Items bestehen nicht — wie bei konkurrierenden Verfahren — aus jeweils verschiedenen formulierten Sätzen, sondern lediglich aus den substantivisch in Ausdrücken der Umgangssprache formulierten Beschwerden, z.B. „Schlaflosigkeit“ statt: „Ich leide unter Schlaflosigkeit“ bzw. „Schwindelgefühl“ statt: „Mir wird gelegentlich schwindlig“. Zum Vergleich: Im MMQ (Eysenck, 1953) müssen für 38 Neurotizismus-Items 295 einzelne Wörter gelesen werden, im VELA (Fahrenberg, 1965) für 56 Items 624 Wörter, in der (ursprünglichen) Beschwerden-Liste dagegen für 64 Items nur 158 Wörter.

Der für den Kliniker entscheidende Vorzug unseres Verfahrens liegt aber darin, daß die Beschwerden nach der Häufigkeit ihres Vorkommens bei Patienten aus Klinik und Praxis ausgewählt worden sind und damit das Spektrum organbezogener oder allgemeiner Beeinträchtigungen des Befindens weitgehend decken. Sie können deshalb dem Arzt u.U. wesentliche Aufschlüsse über das Beschwerdenbild eines Einzelfalles geben. Das faktorenanalytisch definierte Item-Komplexe aus unserer Beschwerden-Liste in einer statistisch nachgewiesenen Beziehung zu bestimmten Formen intermedizinischer, psychosomatischer bzw. psychovegetativer Erkrankungen stehen (v. Kerekjarto *et al.*, 1972), mag zudem für bestimmte klinische Fragestellungen von Interesse sein. Für eine differenzierte Beurteilung des Erscheinungsbildes psychischer Störungen ist die Beschwerden-Liste aber genauso wenig geeignet wie die üblichen Neurotizismus-Fragebögen, da sie — wie diese — hierfür zu wenig Information über einzelne psychische Syndrome enthält. Für deren Erfassung haben wir

3. die „multidimensionale“ Paranoid-Neurotizismus-Depressivitäts-Skala — abgekürzt PND-S bzw. (als Parallelform hierzu) PND-S' — konstruiert. Aus der Bezeichnung geht schon hervor, daß es sich um zwei gleichwertige Fragebögen mit je drei (parallelisierten) klinischen Skalen handelt, die paranoide, neurotische, respektive depressive Tendenzen erfassen sollen. Sie werden durch zwei Kontroll-Skalen ergänzt, eine „Lügen-Skala“, durch die eine Neigung zur Schönfärberei bei der Selbstbeurteilung erfaßt werden soll, und eine weitere Skala, die für die Erfassung einer Tendenz zur Krankheitsverleugnung von uns entwickelt wurde. Diese Tendenz ist bekanntlich bei schizophrenen Patienten häufig so stark ausgeprägt, daß Ergebnisse eines klinischen Selbstbeurteilungs-Verfahrens bei ihnen auch dann weitgehend unbrauchbar sind, wenn die

allgemeine Tendenz zur Schönfärberei im Sinne der „sozialen Erwünschtheit“ (vgl. Edwards, 1959) keineswegs abnorm erhöht ist. Mit Hilfe der Paranoid-Skala unseres Fragebogens soll aber gerade ein wesentlicher Aspekt psychotischer, insbesondere schizophrener Symptomatik erfaßt werden. Das ist auch prinzipiell möglich — nicht zuletzt deshalb, weil bei der Test-Auswertung die Tendenz zur Krankheitsverleugnung mit in Rechnung gestellt werden kann (s.u.).

Die Konstruktion der PND-Skala beruht auf mehrjährigen Vorarbeiten. An ihnen waren außer dem Autor vor allem ein weiterer Psychiater (Häfner), zwei Psychologen (Hüneke und Rey) und mehrere medizinische Doktoranden beteiligt (Beissert, 1967; Lerch, 1967; Schäfer, 1969; Urban, 1965; bezüglich der P- und der D-Skala insbesondere Zimmermann 1971 und bezüglich der N-Skala und der beiden Kontroll-Skalen Volbehr, 1971). Die Dissertationen von Zimmermann (1971) und von Volbehr (1971) enthalten die meisten für das Verständnis der vorliegenden Testformen erforderlichen Informationen. Hier können nur einige Gesichtspunkte kurz referiert werden.

Ausgangspunkt der Skalen-Konstruktion (vgl. Edwards, 1959) war die Überlegung, daß eine Differenzierung verschiedener psychischer Beschwerdenkomplexe in einem objektivierenden Testverfahren nur dann sinnvoll ist, wenn sie einerseits in Übereinstimmung mit der klinischen Syndromatologie psychischer Erkrankungen erfolgt, andererseits die getroffene Differenzierung sich auch vom Testmaterial her — aufgrund statistischer Zusammenhänge zwischen den verschiedenen Test-Items — anbietet. Das veranlaßte eine kombinierte Vorgehensweise bei der Auswahl der Items und ihrer Kombination zu verschiedenen Skalen:

Zunächst wurden nach klinischen Gesichtspunkten Beschwerdenkomplexe definiert und ihnen zugehörige Test-Items zusammengestellt, und zwar aus der Test-Literatur (u. a. Beck *et al.*, 1961; v. Kerekjarto u. Lienert, 1970; Spreen, 1963; Zung, 1965), der psychiatrischen Literatur (u.a. Bleuler, 1969; Bräutigam, 1968; Bumke, 1942; Freedman, Kaplan, 1967; Jaspers, 1965; Kolb, 1968; Kraepelin, 1915; Mayer-Gross *et al.*, 1969; Schneider, 1930; Weitbrecht, 1968) und aus Krankengeschichten. Die Items wurden sodann auf ihr Trennvermögen zwischen — wiederum klinisch definierten — Patientengruppen geprüft, die jeweils geeignet erschienen, einen der in Frage stehenden Beschwerdenkomplexe zu repräsentieren. Schließlich wurden die so aus den ursprünglich hypothetisch gewonnenen Skalen empirisch abgeleiteten Skalen auf ihre faktorielle Homogenität und gegenseitige Unabhängigkeit geprüft; d.h. es wurde korrelationsstatistisch untersucht, ob die in einer Skala zusammengefaßten Items — über alle Patientengruppen hinweg — untereinander ein solches Maß an gemeinsamen Variationstendenzen aufwiesen und ob diese

Tendenzen sich für jede Skala prinzipiell so voneinander unterschieden, daß die Skalen-Konstruktion auch unter rein formalen testtheoretischen Gesichtspunkten gerechtfertigt erschien.

Als Beschwerdenkomplexe, die sich klinischen Syndromen zuordnen lassen, wurden solche gewählt, die jeweils als „*pars pro toto*“ eine der Grundformen aktueller, körperlich nicht begründbarer psychischer Erkrankungen repräsentieren: ängstliche Gespanntheit und emotionelle Labilität als häufigster Ausdruck neurotischer Störungen (vgl. Pawlik, 1968), depressive Züge als häufigster Ausdruck affektiver Psychosen und paranoide Tendenzen als häufigster Ausdruck anderer funktioneller, insbesondere schizophrener Psychosen. Bei dieser Auswahl spielte u.a. die Überlegung mit, daß die genannten Beschwerdenkomplexe in relativ spezifischer Weise auf eine der drei großen Gruppen von Psychopharmaka ansprechen (vgl. Degkwitz, 1967; Haase, 1972; Kalinowsky u. Hippius, 1969), nämlich auf Tranquillizer (ängstliche Gespanntheit), auf Thymoleptica (Depressivität) und auf Neuroleptica (paranoide Tendenzen). Dieser zweite Auslesegesichtspunkt sollte die Eignung der Skalen für psychopharmakologische Untersuchungen erhöhen.

Die Berücksichtigung weiterer Beschwerdenkomplexe hätte den Umfang des Fragebogens in einer für Patienten unzumutbaren Weise verlängert und — nach den von Spreen 1963 veröffentlichten Daten über das „MMPI Saarbrücken“ zu urteilen — voraussichtlich keinen entscheidenden Gewinn an klinisch relevanter Information erbracht. Wichtiger erschien uns die Möglichkeit, Verfälschungstendenzen anhand einer Parallelform zu jeder klinischen Skala und durch spezielle Kontroll-Skalen aufdecken zu können.

Die Items der „Lügen-Skala“ entstammen ausnahmslos bereits existierenden, statistisch ausreichend durchgearbeiteten Skalen (vgl. Ling, 1966; Hummel, 1968). Dagegen wurden die Items der Skala zur Erfassung von Krankheitsverleugnung erstmalig vom Autor formuliert, und zwar als „Allerweltsbeschwerden“, die ehrlicherweise von jedem Probanden bejaht werden müßten (z.B. „Ich war hin und wieder erkältet“). Tatsächlich wird — wie angedeutet — von Schizophrenen, die eine für Patienten dieser Krankheitsgruppe auffallend niedrige Anzahl „paranoider“ Test-Items als für sich zutreffend angeben, auch relativ häufig das Vorliegen solcher banalen Beschwerden bestritten. Die beiden Parallelformen der PND-Skala (PND-S und PND-S') bestehen aus je 99 Items in Satzform, davon 24 „paranoiden“, 20 „neurotischen“ und 16 „depressiven“ Items, welche die vorher genannten Auslesekriterien erfüllt haben. Hinzu kommen je 12 Items für die beiden (faktoriell voneinander unabhängigen!) Kontroll-Skalen, die in den Parallelformen identisch sind, ferner 12 Items, die klinisch informativ sind (z.B. bezüglich Aggres-

sivität oder Suizidalität), aber die genannten Kriterien nicht erfüllen und deshalb in keine der drei Skalen eingehen, sowie 3 Fragen, die sich auf das Verständnis der Testaufgabe und die Motivation zu ihrer Bewältigung beziehen. Die Stellungnahme zu den Items erfolgt durch Ankreuzen einer von drei Antwortkategorien („trifft voll und ganz zu“ — „trifft nur zum Teil zu“ — „trifft gar nicht zu“).

Nach unseren bisherigen Erfahrungen ist das Verfahren für eine differenzierte Erfassung der Psychopathologie des Einzelfalles und bestimmter klinischer Gruppen im Querschnitt wie im Längsschnitt (z.B. bei psychopharmakologischen Untersuchungen) besser geeignet als andere klinische Fragebögen, die entweder ein zu pauschales oder ein zu einseitiges Bild ergeben oder für klinische Zwecke zu umfangreich sind (z.B. das MMPI mit über 550 Items!). Zur Zeit wird aufgrund neuerer Datenanalysen eine um je 30 Items gekürzte Version von PND-S/PND-S' mit einer um eine Beantwortungskategorie pro Item erweiterten Skalierung ausgearbeitet. Wir hoffen, dadurch Anwendbarkeit, Zuverlässigkeit und Validität der Skala noch erhöhen zu können.

Unsere günstigen Erfahrungen mit klinischen Selbstbeurteilungs-Skalen sind sicherlich teilweise der Tatsache zuzuschreiben, daß unsere Patienten ein relativ hohes Maß an therapeutischer Zuwendung erhalten und dadurch eher zur Mitarbeit — auch beim Ausfüllen von Fragebögen — bereit sind, als dies an Häuslern mit einem ungünstigeren Bettenschlüssel der Fall sein mag. Die positive Motivation der Patienten ist natürlich für die Verwendbarkeit von Selbstbeurteilungs-Skalen von entscheidender Bedeutung. Nicht zu unterschätzen ist aber auch der Einfluß, den die Motivation der Untersucher und des gesamten Pflegepersonals der Stationen auf die Mitarbeit von Patienten ausübt. Eine negative Einstellung überträgt sich rasch auf die Patienten und führt unweigerlich zu — womöglich unüberwindbar scheinenden — Schwierigkeiten bei der Anwendung entsprechender Untersuchungsverfahren. Deren Einführung muß daher psychologisch sorgfältig vorbereitet werden. Die mit ihrer Verwendung verbundenen Vorteile sowohl für Querschnitt- wie für Längsschnitt-Untersuchungen rechtfertigen aber u.E. diesen Aufwand.

In den letzten Jahren wird in zunehmendem Umfang auch an anderen — bundesdeutschen wie ausländischen — Institutionen, z.T. sogar in der freien Praxis, mit den von uns entwickelten Untersuchungsinstrumenten (die in englischen, französischen u.a. fremdsprachlichen Fassungen vorliegen) gearbeitet. Dabei sind bisher die eigenen positiven Erfahrungen bezüglich praktischer Anwendbarkeit, formaler Zuverlässigkeit und inhaltlicher (klinischer) Validität der Skalen im wesentlichen bestätigt worden. So hat sich die Befindlichkeit-Skala (B-S/B-S') besonders in

psychopharmakologischen Versuchen an Gesunden und Kranken bewährt (N. Matussek, München; Heimann *u. Mitarb.*, Lausanne; u. a.). Die Beschwerden-Liste (B-L/B-L') war unter anderem bei der Untersuchung körperlich, psychovegetativ und/oder psychosomatisch Kranker von Nutzen (Baumann, 1972; Cohrs, 1967; Jores, v. Kerekjarto, 1967; v. Kerekjarto *et al.*, 1972; Scherf, 1968; Ziegler, 1970). Die Paranoid-Neurotizismus-Depressivitäts-Skala (PND-S/PND-S') erwies sich auch in der Schweiz an einem stationären psychiatrischen Krankengut als brauchbares Untersuchungsinstrument (Meyer von Wohlenschwill, 1973). Die genannten Verfahren kommen darüber hinaus auch für Untersuchungen an ambulanten Patienten und für die epidemiologische Feldforschung in Betracht (v. Zerssen, 1971 b).

Literatur

- Alsen, V.: Das Kernsyndrom der endogenen Depression. *Nervenarzt* **32**, 470—473 (1961).
- Anastasi, A.: *Psychological Testing*. 3rd ed. London: Macmillan 1969.
- Angst, J., Battegay, R., Bente, D., Berner, P., Broeren, W., Cornu, F., Dick, P., Engelmeier, M. P., Heimann, H., Heinrich, K., Helmchen, H., Hippius, H., Pöldinger, W., Schmidlin, P., Schmitt, W., Weis, P.: Das Dokumentationssystem der Arbeitgemeinschaft für Methodik und Dokumentation in der Psychiatrie (AMP). *Arzneimittel-Forsch.* **19**, 399—404 (1969).
- Baumann, U.: Eine Kontrolluntersuchung zur Beschwerdenliste. *Arch. Psychiat. Nervenkr.* **216**, 153—161 (1972).
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., Erbauch, J.: An inventory for measuring depression. *Arch. gen. Psychiat.* **4**, 561—571 (1961).
- Beissert, A.: Trennung klinischer Gruppen im Fragebogenversuch. *Med. Diss.*, Univ. Heidelberg 1967.
- Bleuler, E.: *Lehrbuch der Psychiatrie*. 11. Aufl. von M. Bleuler. Berlin-Heidelberg-New York: Springer 1969.
- Bräutigam, W.: *Reaktionen, Neurosen, Psychopathien*. Stuttgart: Thieme 1968.
- Brengelmann, J. C.: Psychologische Methodik und Psychiatrie. In: H. W. Gruhle, R. Jung, W. Mayer-Gross *u.* M. Müller: *Psychiatrie der Gegenwart*, Bd. I/2: Grundlagen und Methoden der klinischen Psychiatrie, S. 134—175. Berlin-Göttingen-Heidelberg: Springer 1963.
- Bumke, O.: *Lehrbuch der Geisteskrankheiten*, 5. Aufl. München: Bergmann 1942.
- Cohrs, U.: Psychologische Veränderung bei Patienten mit Akne; eine Erkundungsstudie. *Med. Diss.*, Univ. Hamburg 1967.
- Degkwitz, R.: *Leitfaden der Psychopharmakologie*. Stuttgart: Wissenschaftl. Verlagsges. 1967.
- Edwards, A. L.: Social desirability and personality test construction. In: B. M. Bass and I. A. Berg: *Objective approaches to personality assessment*, pp. 100—118. Princeton/N.J.-Toronto-London-New York: Van Nostrand 1959.
- Eysenck, H. J.: Fragebogen als Meßmittel der Persönlichkeit. *Z. exp. angew. Psychol.* **1**, 291—335 (1953).
- Eysenck, H. J.: Das „Maudsley Personality Inventory“ als Bestimmen der neurotischen Tendenz und Extraversion. *Z. exp. angew. Psychol.* **6**, 167—190 (1959).

- Eysenck, H. J.: Funktion und Anwendung der Statistik in der Psychiatrie. In: H. W. Gruhle, R. Jung, W. Mayer-Gross u. M. Müller: Psychiatrie der Gegenwart, Bd. I/2: Grundlagen und Methoden der klinischen Psychiatrie, S. 249—273. Berlin-Göttingen-Heidelberg: Springer 1963.
- Fahrenberg, J.: Ein itemanalysierter Fragebogen funktionell-körperlicher Beschwerden (VELA). *Diagnostica* 11, 141—153 (1965).
- Freedman, A. M., Kaplan, H. J. (eds.): *Comprehensive textbook of psychiatry*. Baltimore: Williams & Wilkins 1967.
- Gilbert, J.: *Clinical psychological tests in psychiatric and medical practice*. Springfield/Ill.: Ch. C. Thomas 1969.
- Haase, H.-J.: Therapie mit Psychopharmaka und anderen psychotropen Medikamenten, 3. Aufl. Stuttgart-New-York: Schattauer 1972.
- Hellpach, W.: Erfahrungen und Grundsätze über Schauverfahren. *Z. Rassenk.* 4, 164—172 (1936).
- Hoeth, F., Gregor, H.: Guter Eindruck und Persönlichkeitsfragebogen. *Psychol. Forsch.* 28, 64—88 (1964).
- Hofstätter, P. R., Wendt, D.: Quantitative Methoden der Psychologie, 3. Aufl. (von: P. R. Hofstätter: *Einführung in die quantitativen Methoden der Psychologie*). München: Barth 1967.
- Hummel, M.: Faktorenanalytische Validitätsprüfung von Fremdanamnesen. *Med. Diss.*, Univ. Heidelberg 1968.
- Jaspers, K.: *Allgemeine Psychopathologie*, 8. Aufl. Berlin-Heidelberg-New York: Springer 1965.
- Jores, A., Kerekjarto, M. v.: *Der Asthmatiker*. Bern: Huber 1967.
- Kalinowsky, L. B., Hippius, H.: *Pharmacological, convulsive and other somatic treatments in psychiatry*. New York-London: Grune & Stratton 1969.
- Kerekjarto, M. v., Lienert, G. A.: Depressionsskalen als Forschungsmittel in der Psychopathologie. *Pharmakopsychiat.* 3, 1—21 (1970).
- Kerekjarto, M. v., Meyer, A.-E., Zerssen, D. v.: Die HHM-Beschwerdenliste bei Patienten einer internistischen Ambulanz. *Z. psychosom. Med. Psychoanal.* 18, 1—16 (1972).
- Kolb, L. C.: *Noyes' modern clinical psychiatry*, 7th ed. Philadelphia-London-Toronto: Saunders 1968.
- Kraepelin, E.: *Psychiatrie*, Bd. IV. 8. Aufl. Leipzig: Barth 1915.
- Lerch, W.: Über die objektive Erfassbarkeit des Krankheitsverlaufes bei Schizophrenen mit Hilfe eines psychiatrischen Fragebogens. *Med. Diss.*, Univ. Heidelberg 1967.
- Lienert, G. A.: *Testaufbau und Testanalyse*, 2. Aufl. Weinheim-Berlin: Beltz 1967.
- Ling, M.: Konstruktion einer Lügenskala. *Psychol. Diplomarbeit*, Univ. Hamburg 1966.
- Lorr, M., Klett, J., McNair, D. M., Lasky, J. J.: Inpatient multidimensional psychiatric scale. Manual. Palo Alto/Calif.: Psychological Press 1963. (Deutsch: Multidimensionale psychiatrische Skala für Klinikpatienten von Maurice Lorr, Douglas, M. McNair, C. James Klett u. Julian J. Lasky, übers. von H. Flegel und Margarete Mäder).
- Meyer, B. C.: Überprüfung eines klinisch-psychiatrischen Persönlichkeitsfragebogens (PNDS, PNDS'). *Med. Diss.*, Univ. Zürich 1973.
- Meyer-Gross, W., Slater, E., Roth, M.: *Clinical Psychiatry*, 3rd ed. London: Baillière, Tindall & Casell 1969.
- Michel, L.: Allgemeine Grundlagen psychometrischer Tests. In: K. Gottschaldt, T. Lersch, F. Sander u. H. Thomae: *Handbuch der Psychologie*, Bd. VI, S. 19—70. Göttingen: Hogrefe 1964.

- Mittenecker, E.: Subjektive Tests zur Messung der Persönlichkeit. In: K. Gottschaldt, T. Lersch, F. Sander u. H. Thomae: Handbuch der Psychologie, Bd. IV, S. 461—487. Göttingen: Hogrefe 1964.
- Mombour, W.: Verfahren zur Standardisierung des psychopathologischen Befundes. *Psychiat. clin. (Basel)* **5**, 73—120, 137—157 (1972).
- Pawlik, K.: Dimensionen des Verhaltens. Bern-Stuttgart: Huber 1968.
- Rautenstrauch, T.: Zur Zeitstabilität und Validität von Fragebogen-Scores für Neurotizismus und Extraversions sowie „Social Desirability“ bei verschiedenen Stichproben. *Med. Diss.*, Univ. Hamburg 1967.
- Schäfer, M.: Faktorenanalytische Überprüfung klinischer Fragebögen. *Med. Diss.*, Univ. Heidelberg 1969.
- Scherf, M.: Psychometrische Untersuchungen bei Herzkranken. *Med. Diss.*, Univ. Hamburg 1968.
- Schneider, C.: Die Psychologie der Schizophrenen und ihre Bedeutung für die Klinik der Schizophrenie. Leipzig: Thieme 1930.
- Schwarz, D., Strian, F.: Psychometrische Untersuchungen zur Befindlichkeit psychiatischer und intern-medizinischer Patienten. *Arch. Psychiat. Nervenkr.* **216**, 70—81 (1972).
- Spreen, O. (Bearb.): MMPI Saarbrücken. Handbuch der deutschen Ausgabe des Minnesota Multiphasic Personality Inventory von S.R. Hathaway and J.C. McKinley. Bern-Stuttgart: Huber 1963.
- Urban, S.: Beitrag zur Entwicklung einer klinischen Testkurzform unter besonderer Berücksichtigung schizophrener Symptomatik. *Med. Diss.*, Univ. Heidelberg 1965.
- Volbehr, U.: Entwicklung einer klinischen Selbstbeurteilungs-Skala zur Erfassung von neurotischer Symptomatik unter Berücksichtigung inhaltlicher Beantwortungsstereotypien. Ein Beitrag zur objektivierenden psychopathologischen Befunderhebung. *Med. Diss.*, Univ. München 1971.
- Wechsler, D.: Die Messung der Intelligenz Erwachsener, 3. Aufl. von C. Bondy. Bern: Huber 1964.
- Weitbrecht, H. J.: Psychiatrie im Grundriß, 2. Aufl. Berlin-Heidelberg-New York: Springer 1968.
- Zerssen, D. v.: Die Beschwerden-Liste als Test. *Therapiewoche* **21**, 1908—1920 (1971a).
- Zerssen, D. v.: Zur standardisierten psychopathologischen Befunderhebung im Rahmen epidemiologischer Untersuchungen. Tagung der Gesellschaft Nord- und Nordwestdeutscher Neurologen und Psychiater, Göttingen 1971; Ref.: *Zbl. ges. Neurol. Psychiat.* **203**, 246 (1971b).
- Zerssen, D. v.: Beschwerden-Skalen bei Depressionen. *Therapiewoche* **23** (im Druck) (1973).
- Zerssen, D. v., unter Mitarbeit von D.-M. Koeller u. E.-R. Rey: Objektivierende Untersuchungen zur prämorbidien Persönlichkeit endogen Depressiver. In: H. Hippius u. H. Selbach: Das depressive Syndrom, S. 183—205. München-Berlin-Wien: Urban & Schwarzenberg 1969.
- Zerssen, D. v., Koeller, D.-M., Rey, E.-R.: Die Befindlichkeits-Skala (B-S) — ein einfaches Instrument zur Objektivierung von Befindlichkeitsstörungen, insbesondere im Rahmen von Längsschnittuntersuchungen. *Arzneimittel-Forsch.* **20**, 915—918 (1970a).
- Zerssen, D. v., Koeller, D.-M., Rey, E.-R.: Die prämorbidie Persönlichkeit von endogen Depressiven. Eine Kreuzvalidierung früherer Untersuchungsergebnisse. *Confin. Psychiat. (Basel)* **18**, 156—179 (1970b).

314 D. v. Zerssen: Selbstbeurteilungs-Skalen zur Abschätzung „subjekt. Befundes“

- Zerssen, D. v., Strian, F., Schwarz, D.: Evaluation of depressive states, especially in longitudinal studies. In: P. Pichot, R. Olivier-Martin: Psychological measurement in pharmacopsychiatry. Basel: Karger 1973 (in press).
- Ziegler, K.: Validierungsuntersuchung mit der HHM-Beschwerdenliste in der Poliklinik. Med. Diss., Univ. Hamburg 1970.
- Zimmermann, J.: Die Messung und Differenzierung von paranoiden und depressiven Tendenzen mit einem item- und faktorenanalysierten Selbstbeurteilungsfragebogen. Med. Diss., Univ. München 1971.
- Zung, W. W. K.: A self-rating depression scale. Arch. gen. Psychiat. 12, 63—70 (1965).

Prof. Dr. med. D. v. Zerssen
Dipl.-Psych.
Max-Planck-Institut für Psychiatrie
D-8000 München 40, Kraepelinstraße 10
Bundesrepublik Deutschland