

## Untersuchungen zur Eignung von Leistungsverfahren für die Messung von Schweregraden endogener Depressionen\*

Adolf Gallwitz und Siegfried Lehrl

Einheit für Medizinische Psychologie und Psychopathometrie (Leiter: Priv.-Doz. Dr. Dr. med. habil. W. Kinzel, Dipl.-Psych.) der Universitäts-Nervenklinik mit Poliklinik Erlangen (Direktor: Prof. Dr. H. H. Wieck)

### Investigation of the Suitability of Efficiency Tests in Evaluating Endogenous Depression

**Summary.** Self-evaluation procedures are usually used to assess the severity of endogenous depression. In this respect, efficiency tests have hitherto been scarcely considered. A literary survey and our own empirical investigations suggest that self-evaluation procedures should mainly be used in mild and moderate endogenous depressions, and that these procedures give false results in severe and very severe depressions. Efficiency procedures seem not to differentiate well between 'normality' and 'mild depression,' but they become more reliable and valid for depressions of increasing severity. Self-evaluation and efficiency procedures should therefore not be substituted for each other; they should complement each other.

**Key words:** Endogenous depression – Efficiency procedures – Self-evaluation procedures – Severity of endogenous depression.

**Zusammenfassung.** Zur Messung der Schweregrade endogener Depressionen werden meistens Selbstbeurteilungsverfahren angewandt. Für diese Zwecke fanden Leistungsverfahren bisher kaum Berücksichtigung. Nach einer Literaturübersicht und eigenen empirischen Untersuchungen drängt sich die Vermutung auf, daß das Hauptanwendungsgebiet von Selbstbeurteilungsverfahren im Abschnitt leichter bis mittelschwerer endogener Depressionen liegen müßte, diese Verfahren jedoch bei schweren und schwersten Depressionen zu falschen Ergebnissen führen. Leistungsverfahren scheinen zwischen psychiatrisch unauffälligen und leicht depressiven Personen nicht so gut zu differenzieren, werden aber bei zunehmendem Schweregrad zuverlässiger und gültiger. Selbstbeurteilungs- und Leistungsverfahren können demnach nicht

\* Mit freundlicher Unterstützung der Firma Dr. R. Pfleger, Bamberg  
Sonderdruckanforderungen an: Dipl.-Psych. S. Lehrl, Universitäts-Nervenklinik Erlangen,  
Schwabachanlage 6, D-8520 Erlangen, Bundesrepublik Deutschland

gegeneinander ausgetauscht, sondern sollten in Ergänzung zueinander angewandt werden.

**Schlüsselwörter:** Endogene Depression – Leistungsverfahren – Schweregrade endogener Depressionen – Selbstbeurteilungsverfahren.

## I. Problemlage

### 1. Gebräuchliche Verfahren zur Depressionsmessung

Die zur Messung psychologischer und psychopathologischer Merkmale verwendeten Testverfahren können nach verschiedenen Klassifikationsgesichtspunkten eingeteilt werden. Eine Einteilung für den psychopathologischen Gebrauch wäre die in *Leistungsverfahren* (z.B. Hawie nach Wechsler, 1956; Benton, 1961), in *Beurteilungsverfahren* (z.B. MMPI nach Hathaway und McKinley, 1963; MPI, MMQ nach Eysenck, 1959 und 1964), in *Gestaltungsverfahren* (z.B. Wartegg-Zeichentest, 1968), in *Wahlverfahren* (z.B. Szondi-Test, 1972; Lüscher-Test, 1971) und in *Entfaltungsverfahren* (z.B. Thematischer Apperzeptionstest nach Murray, 1943).

Für wissenschaftliche Zwecke sind Leistungs- und Beurteilungsverfahren besonders geeignet, da ihr Aufbau und die Vorgänge der Durchführung, Auswertung und Interpretation gut beschreibbar und analysierbar sind. Zur Depressionsforschung werden fast ausschließlich Beurteilungsverfahren herangezogen.

a) *Beurteilungsverfahren*. Neben einigen Fremdbeurteilungsverfahren beherrschen vor allem Selbstbeurteilungsverfahren das Feld der Schweregradmessung endogener Depressionen. Die meisten im Umlauf befindlichen Selbstbeurteilungsverfahren haben leider keine deutschen Normen, wie z.B. der in seiner Art als Verfahren für niedergelassene Ärzte vielleicht vorbildliche „Depressions-Index für die Praxis“ nach Popoff (o.J.) oder das bekannte Depressionsinventar nach Beck (1961). Da für diese Verfahren keine deutschen Normen empirisch erhoben werden, fehlt ihnen ein konstituierendes Merkmal von Tests. Die Erstellung von Normen bedeutet für die Psychopathologie, daß das betreffende Verfahren sowohl an psychiatrisch unauffälligen als auch an depressiven Patienten abgenommen wurde und die Ergebnisse registriert sind.

Legt man den Selbstbeurteilungstests als Kriterium die deutsche Normierung an, schrumpft die Menge der angebotenen Verfahren stark zusammen. Zu den verbleibenden, mehr oder weniger syndromspezifischen deutschen Selbstbeurteilungstests gehören:

- Hamburger Depressions-Skala (v. Kerekjarto und Lienert, 1970)
- Befindlichkeits-Skala (v. Zerssen, 1976a)
- Beschwerden-Liste (v. Zerssen, 1976b)
- Depressivitäts-Skala (v. Zerssen, 1976c)
- Stimmungs-Antriebs-Skala (Supprian, 1976)
- Erlanger Depressions-Skala (Lehrl und Gallwitz, 1977)

Ihren vielleicht wichtigsten Vorzug haben die Selbstbeurteilungsverfahren in der ökonomischen Anwendbarkeit. Gewöhnlich legt man dem Probanden einen

derartigen Testbogen einfach vor und bittet ihn, die am Kopf stehende Anweisung durchzulesen und dann den Bogen auszufüllen. Fremdbeurteilungstests erfordern hingegen vom Untersuchenden mehr Aufwand, da er die Beurteilung anhand einer standardisierten Vorlage selbst vornehmen muß. Darin mag es hauptsächlich begründet sein, daß Fremdbeurteilungsverfahren nicht sehr verbreitet sind.

Im deutschen Sprachraum werden zwar Übersetzungen einiger englischsprachiger Verfahren angewandt, vor allem die Hamilton-Skala (Hamilton, 1960) und die Zung-Skala (Zung, 1972), jedoch sind uns keine syndromspezifischen Fremdeinschätzungsverfahren für Depressionen mit deutschen Normen bekannt.

*b) Leistungsverfahren.* Unseres Wissens wurden außer dem Ansatz von Lutz (1969), Wieck und Lutz (1969) und der Weiterentwicklung dieses Ansatzes durch Straub et al. (1976) und Lehrl et al. (1977) keine Leistungstests für die psychopathometrische Erfassung endogener Depressionen entwickelt.

Der LtD setzt sich aus fünf Subtests zusammen, mit denen Minderungen verschiedener psychischer Funktionen, wie äußere Wahrnehmung, Gedächtnis und Geschwindigkeit psychischer Abläufe gemessen werden. Im ersten Untertest sind abgebildete Gegenstände zu benennen (Zeit) und anschließend zu reproduzieren (Untertest 2). Auch hierbei, wie bei allen folgenden Subtests, wird die Zeit registriert und protokolliert. Subtest 3 besteht darin, daß Klötzchen mit zweistelligen Zahlen möglichst schnell in die richtige Reihenfolge zu bringen sind. Die nächsten beiden Aufgaben entsprechen Subtest 1 und 2 mit veränderten Vorlagen. In Subtest 4 ist aus einer Tafel mit Reihen verschiedener Symbole die Anzahl der Kreise auszuzählen. In Subtest 5 werden zwei Zeilen mit den Buchstaben A und B vorgelegt. Die Anweisung hierfür lautet, statt „A“ immer „B“ und statt „B“ immer „A“ vorzulesen.

Wegen der Zeitbegrenzungen der einzelnen Subtests wird eine Durchführungszeit von acht bis zehn Minuten nicht überschritten. Bei der Zuordnung der Rohpunkte zu Wertpunkten werden das Lebensalter und das prämorbidie Intelligenzniveau berücksichtigt.

## 2. Fragestellung

Nicht nur aufgrund der leichteren Verwissenschaftlichung, sondern auch angeregt durch empirische Befunde (siehe III,1), liegt die Vermutung nahe, daß auch Leistungsverfahren zur Depressionsmessung geeignet sind.

- Zuerst sollte jedoch untersucht werden, ob mit einem solchen Verfahren überhaupt gesunde Personen von endogen depressiven Patienten getrennt werden können.
- Weiter war zu prüfen, ob mit Leistungsverfahren verschiedene Schweregrade unterschieden werden können; dabei werden zwangsläufig die Zusammenhänge von Leistungs- und Beurteilungsverfahren mitbehandelt.
- Schließlich interessierte die kombinierte Trennfähigkeit von Leistungs- und Beurteilungsverfahren.

## III. Methodik

### 1. Allgemeines Vorgehen

Ausgehend von theoretischen Überlegungen und anknüpfend an Publikationen einschlägiger Leistungsuntersuchungen wurden eigene empirische Befunde mit dem Leistungstest für endogene Depression (= LtD) erhoben.

**Tabelle 1.** Beschreibungsparameter der untersuchten Stichprobe

Stichprobe	Anzahl	Alter (Jahre)		Intelligenz (MWT-B IQ)	
		arithmet. Mittelwert	Standard- abweichung	arithmet. Mittelwert	Standard- abweichung
Endogen depressive Patienten					
Motorisch ruhig	65	44,5	9,1	96,1	13,4
Motorisch unruhig	27	46,6	10,6	97,9	12,1
Gesamt	93				
Kontrollpersonen	27	49,8	11,0	96,0	14,2

## 2. Untersuchte Personen

Die wichtigsten Kennwerte zur Beschreibung der untersuchten Stichprobe sind in Tabelle 1 dargestellt.

## 3. Untersuchungsverfahren

Jede Person wurde mit dem Leistungsverfahren LtD (in: Lehrl et al., 1977) und den zwei Selbstbeurteilungsskalen, *Befindlichkeits-Skala* nach v. Zerssen (1976a) und *Erlanger Depressions-Skala* nach Lehrl und Gallwitz (1977), sowie dem *Mehrfachwahl-Wortschatz-Test* (Form B) nach Lehrl (1977) untersucht. Außerdem schätzte der behandelnde Arzt bei 27 depressiven Patienten den Schweregrad der Depression nach sieben Stufen ein.

Die Abnahme der Tests fand in folgender Reihenfolge statt: LtD, MWT-B, EDS, B-S.

Weitere Verfahren wie Tapping, Wörter lesen und Farbwahlverfahren, die nur an Teilstichproben abgenommen wurden, bleiben in dieser Arbeit unberücksichtigt.

## 4. Statistisches Vorgehen

Zur Prüfung der Frage, ob gesunde von endogen depressiven Personen mit dem LtD überhaupt unterschieden werden können, wurde ein möglichst optimaler Trennpunkt (= cutting point) bestimmt. Je nachdem, ob der Meßwert einer Person über oder unter dem Trennpunkt lag, wurde sie der gesunden oder kranken Gruppe zugeordnet. Ziel der Optimierung war es, die Fehlzuordnungen (Gesunde werden laut Testwert als krank eingestuft und umgekehrt) auf ein Minimum zu reduzieren. Da diese Fehlzuordnungen in jeder Stichprobe prozentual reduziert werden sollten, die Stichprobengrößen in unseren Versuchsgruppen jedoch unterschiedlich waren, wurde jede Stichprobe gleich 100% gesetzt. Die Fehlzuordnungen sollten sich auf jeweils beide in Betracht gezogene Stichproben prozentual etwa gleich verteilen. So sollte vermieden werden, daß bei den Kontrollpersonen beispielsweise 0%, bei der Gruppe der Depressiven aber 45% Fehlzuordnungen erfolgten.

Da die verwendeten Selbstbeurteilungsskalen eine Variationsbreite von 0—40 Punkten (EDS) bzw. von 0—56 Punkten (BS) aufweisen, das verwendete

Leistungsverfahren jedoch nur einen Bereich von 0—15 Punkten umfaßt, würde bei der Kombination eines Beurteilungs- mit einem Leistungsverfahren (durch Summierung der erhaltenen Punktwerte zu einem Gesamtpunktwert) das Ergebnis aus dem Beurteilungsverfahren mit größerem Gewicht am Gesamtpunktwert beteiligt sein. Daher hatten wir die Punkte im LtD mit dem Faktor drei in Verbindung mit der EDS und mit dem Faktor vier in Verbindung mit der B-S zum Ausgleich multipliziert. Ein derartiger Gesamtpunktwert ist wichtig für die Überprüfung der Trennfähigkeit zwischen gesunden Personen vs. Patienten der Leistungs- und Selbstbeurteilungsverfahren zusammen. Zur Berechnung des Zusammenhangs der von uns verwendeten Verfahren und des zusätzlichen Arzturteils untereinander berechneten wir Rangkorrelationen mit Tie-Korrektur nach Spearman-Brown (in: Siegel, 1956).

### III. Anderweitige und eigene Ergebnisse

#### *1. Die Trennung von gesunden und depressiven Personen mittels Leistungsverfahren*

a) *Literaturübersicht.* In der Mehrzahl der klinischen Studien zeigten sich Leistungsminderungen während der Phasen zykllothymer Depressionen (Babcock, 1941; Rapaport, 1945; Nelson, 1953; Martin und Rees, 1966; Payne und Hewlett, 1966; Lutz, 1969; Wieck und Lutz, 1969; Polak und Brichcin, 1969; Philippen und Thorwirth, 1971; Blackburn, 1975). Nicht so günstige Ergebnisse erhielten hingegen Shapiro et al. (1958) und Friedman (1964). Der Grund dafür dürfte im nachhinein nur noch schwer zu ermitteln sein.

Die Trennfähigkeit der endogen depressiven Patienten von den unauffälligen Kontrollpersonen ist eine Voraussetzung für die Gültigkeit des verwendeten Verfahrens. Gelingt es nicht, die beiden Stichproben über die Zufallserwartung (50% : 50%) hinaus zu trennen, bringt uns das Verfahren keine neue Information.

b) *Eigene empirische Ergebnisse.* Wie aus den Tabellen 3 und 8 ersichtlich ist, liegt die Trefferrate des Leistungsverfahrens zwischen 70% und 81%. Mit den Beurteilungsverfahren werden Trefferraten von 79% bis 87% erzielt. Diese liegen also etwas höher.

Mit dem LtD kann also eine Trennung zwischen unauffälligen Kontrollpersonen und endogen depressiven Patienten vorgenommen werden, die durchaus

**Tabelle 2.** Rangkorrelationen (Spearman-Brown mit Tie-Korrektur) zwischen LtD und Fremd- und Selbsteinschätzungen der Schwere der zykllothymen Depression.  $n = 27$

	Arzturteil	B-S	EDS
LtD	0,32 *	0,24	0,11
EDS	0,57 *	0,73 *	
B-S	0,64 *		

\* Signifikant von  $r = 0$  verschieden

**Tabelle 3.** Häufigkeitsverteilungen der LtD-Ergebnisse. Außerhalb der Klammern stehen Absolutangaben, in den Klammern Prozentwerte

Wertpunkte im LtD	Kontrollpersonen	Zyklothym depressive Patienten	
		ruhig <sup>a</sup>	unruhig <sup>b</sup>
0	— (—)	— (—)	— (—)
1	2 (7,4)	— (—)	1 (3,7)
2	1 (3,7)	— (—)	2 (7,4)
3	6 (22,2)	5 (7,6)	3 (11,1)
4	3 (11,1)	5 (7,6)	1 (3,7)
B →	9 (33,3)	5 (7,6)	3 (11,1)
A →	5 (18,6)	8 (12,1)	5 (18,6)
6	1 (3,7)	10 (15,1)	3 (11,1)
7	— (—)	13 (19,7)	1 (3,7)
8	— (—)	9 (13,6)	5 (18,5)
9	— (—)	3 (4,5)	1 (3,7)
10	— (—)	5 (7,6)	1 (3,7)
11	— (—)	3 (4,5)	1 (3,7)
12	— (—)	— (—)	— (—)
13	— (—)	— (—)	— (—)
14	— (—)	— (—)	— (—)
15	— (—)	— (—)	— (—)
Summe	27 (100,0)	66 (100,0)	27 (100,0)

<sup>a</sup> Patienten, die nach ihrem Erscheinungsbild vom behandelnden Arzt als motorisch ruhig, „gehemmt“ bezeichnet wurden

<sup>b</sup> Patienten, die nach ihrem Erscheinungsbild vom behandelnden Arzt als motorisch unruhig, „agitiert“ bezeichnet wurden

von ähnlicher praktischer Bedeutung wie die Trennfähigkeit von Selbsteinschätzungs-Skalen sein kann.

## 2. Unterscheidung von Störungsausprägungen

Die Nützlichkeit psychopathologischer Meßverfahren als differentialdiagnostische Hilfe ist sehr umstritten (z.B. Dahl, 1968). Sicherer ist der Einsatz derartiger Verfahren zur Messung von Störungsausprägungen, wie es die Psychopathometrie fordert (Wieck, 1967, 1973).

Daß die Leistungen endogen depressiver Patienten gemindert sind, ist nur eine Voraussetzung für die Störungsgradbestimmung. Aus dem Ausmaß der Minderung soll schließlich auf die Tiefe der endogenen Depression geschlossen werden. Inwieweit die Annahme des parallelen Verlaufes von Leistungsminderung und Depressionstiefe gerechtfertigt ist, muß bei der Entwicklung von Leistungstests geprüft werden.

*a) Literaturübersicht.* In der Literatur sind Untersuchungen zur Gleichartigkeit der Verläufe zwischen Beurteilungs- und Leistungsverfahren kaum bekannt

**Tabelle 4.** Häufigkeitsverteilungen der Ergebnisse in der Befindlichkeits-Skala (B-S). Außerhalb der Klammern stehen Absolutangaben, in den Klammern Prozentwerte

Wertpunkte in der B-S	Kontrollpersonen	Zykllothym depressive Patienten	
		ruhig <sup>a</sup>	unruhig <sup>b</sup>
0—4	6 (22,2)	3 (4,6)	— (—)
5—8	6 (22,2)	4 (6,2)	2 (7,4)
9—12	7 (26,0)	4 (6,2)	1 (3,7)
C → 13—16	4 (14,8)	7 (10,8)	— (—)
17—20	— (—)	3 (4,6)	1 (3,7)
21—24	— (—)	7 (10,8)	5 (18,6)
25—28	2 (7,4)	7 (10,8)	2 (7,4)
29—32	1 (3,7)	9 (13,8)	3 (11,1)
33—36	1 (3,7)	6 (9,2)	4 (14,8)
37—40	— (—)	8 (12,3)	2 (7,4)
41—44	— (—)	5 (7,7)	2 (7,4)
45—49	— (—)	1 (1,5)	3 (11,1)
50—53	— (—)	— (—)	1 (3,7)
54—56	— (—)	1 (1,5)	1 (3,7)
Summe	27 (100,0)	65 (100,0)	27 (100,0)

<sup>a</sup> Patienten, die nach ihrem Erscheinungsbild vom behandelnden Arzt als motorisch ruhig, „gehemmt“ bezeichnet wurden

<sup>b</sup> Patienten, die nach ihrem Erscheinungsbild vom behandelnden Arzt als motorisch unruhig, „agitiert“ bezeichnet wurden

geworden. Ausnahmen bilden die Arbeiten von Lutz (1969) und List (1972), in denen eine Verbesserung der Leistung mit zunehmendem Klinikaufenthalt festgestellt wird.

Man könnte einwenden, daß die erfaßte Leistungsminderung der endogen depressiven Patienten *allein* durch pharmakogene Funktionspsychosen erklärt sind. Zur Überprüfung dieses Einwandes wurden von Straub und Straub (unveröffentlicht) 19 motorisch ruhige bzw. unruhige Patienten unmittelbar nach ihrer Aufnahme in unsere Klinik untersucht. Innerhalb der Klinik hatten sie bis zur psychopathometrischen Untersuchung keine Psychopharmaka erhalten. Der LtD-Median dieser Patienten entsprach mit 6,9 genau dem Median der gesamten endogen depressiven Stichprobe ( $n=93$ ,  $Md=6,9$ ), während die Kontrollpersonen einen Median von 4,6 hatten.

*b) Überlegungen.* Es ist anzunehmen, daß mit zunehmender Depressionstiefe auch die Leistung abnimmt. Sie beträgt beispielsweise für den Extremfall der Katatonie (ausgeprägt schwere Störung) den Wert Null. Bei Leistungsverfahren liegen andere Verhältnisse als bei Selbsteinschätzungs-Skalen vor, denn wie Bech et al. nachweisen, sind die Meßwerte von Beurteilungsverfahren bei schweren Depressionen eher rückläufig. Daher müßten im Abschnitt schwerer Depression Leistungsverfahren den Selbstbeurteilungsverfahren sogar überlegen sein.

**Tabelle 5.** Häufigkeitsverteilungen der Ergebnisse der Erlanger Depressions-Skala (EDS). Außerhalb der Klammern stehen Absolutangaben, in den Klammern Prozentwerte

Wertpunkte EDS	Kontrollpersonen	Zykllothym depressive Patienten	
		ruhig <sup>a</sup>	unruhig <sup>b</sup>
0—1	12 (44,5)	3 (4,6)	2 (7,4)
D→ 2—4	9 (33,3)	4 (6,2)	— (—)
5—7	1 (3,7)	4 (6,2)	1 (3,7)
8—10	2 (7,4)	12 (18,6)	2 (7,4)
11—13	1 (3,7)	8 (12,3)	6 (22,3)
14—16	— (—)	6 (9,2)	5 (18,5)
17—19	2 (7,4)	11 (16,9)	3 (11,1)
20—22	— (—)	6 (9,2)	4 (14,8)
23—25	— (—)	6 (9,2)	2 (7,4)
26—28	— (—)	2 (3,0)	2 (7,4)
29—31	— (—)	— (—)	— (—)
32—34	— (—)	3 (4,6)	— (—)
35—37	— (—)	— (—)	— (—)
38—40	— (—)	— (—)	— (—)
Summe	27 (100,0)	65 (100,0)	27 (100,0)

<sup>a</sup> Patienten, die nach ihrem Erscheinungsbild vom behandelnden Arzt als motorisch ruhig, „gehemmt“ bezeichnet wurden

<sup>b</sup> Patienten, die nach ihrem Erscheinungsbild vom behandelnden Arzt als motorisch unruhig, „agitiert“ bezeichnet wurden

c) *Eigene Befunde.* Um zu untersuchen, ob der LtD auch Schweregrade mißt, benötigen wir ein Kriterium für das Ausmaß der Störung. Selbstbeurteilungstests wie die EDS oder B-S sind dafür aber nur bedingt verwendbar, denn Bech et al. (1975) zeigten, daß Beurteilungsskalen höhere Schweregrade der Depression nicht mehr zu differenzieren vermögen. Beispielsweise ist vielen Psychiatern bekannt, daß sich Patienten mit sehr ausgeprägten Depressionen entleert fühlen, daß ihnen nicht mehr zum Weinen zumute ist und daß sie weder Angst noch Schuld oder Trauer verspüren. D.h. es besteht der Verdacht, daß die Befunde bei höheren Schweregraden wieder abnehmen. Daher sind kaum Rangkorrelationen zwischen Selbstbeurteilungs- und Leistungsverfahren zu erwarten. Aber das Arzturteil, das sich zwar teilweise auf Mitteilungen der Patienten über die Stimmung und das Befinden stützt, müßte dennoch davon teilweise unabhängig sein, da in ihm auch Verhaltensweisen und Ausdruck berücksichtigt werden. Es könnte daher ein besserer Indikator des depressiven Schweregrades sein. In den Korrelationen (Tabelle 2) erhielten wir einen Wert von 0,32 (signifikant von Null verschieden, ( $P < 0,05$ ), bei einseitiger Fragestellung).

Lutz erhielt einen Rangkorrelationskoeffizienten von 0,72 zwischen Arztein-schätzung und Punktewert im Hemmungstest (Lehrl, 1972). Dieser Korrelations-

**Tabelle 6.** Häufigkeitsverteilungen der kombinierten LtD- und B-S-Ergebnisse. Außerhalb der Klammern stehen Absolutangaben, in den Klammern Prozentwerte

Kombination LtD — B-S <sup>a</sup>	Kontrollpersonen	Zykllothym depressive Patienten	
		ruhig <sup>b</sup>	unruhig <sup>c</sup>
0— 9	— (—)	— (—)	— (—)
10— 18	2 (7,4)	2 (3,1)	1 (3,7)
19— 27	12 (44,4)	— (—)	— (—)
E → 28— 36	7 (26,0)	8 (12,3)	2 (7,4)
F → 37— 45	6 (22,2)	9 (13,8)	3 (11,1)
46— 54	— (—)	12 (18,5)	3 (11,1)
55— 63	— (—)	10 (15,4)	9 (33,4)
64— 72	— (—)	14 (21,5)	6 (22,2)
73— 81	— (—)	8 (12,3)	2 (7,4)
82— 90	— (—)	2 (3,1)	1 (3,7)
91— 99	— (—)	— (—)	— (—)
100—109	— (—)	— (—)	— (—)
	— (—)	— (—)	— (—)
Summe	27 (100,0)	65 (100,0)	27 (100,0)

<sup>a</sup> Die kombinierte Wertpunktzahl wurde folgendermaßen gebildet: B-S-Rohpunkte + (4 × LtD-Wertpunkte)

<sup>b</sup> Patienten, die nach ihrem Erscheinungsbild vom behandelnden Arzt als motorisch ruhig, „gehemmt“ bezeichnet wurden

<sup>c</sup> Patienten, die nach ihrem Erscheinungsbild vom behandelnden Arzt als motorisch unruhig, „agitiert“ bezeichnet wurden

koeffizient liegt deutlich höher als der unserer Befunde. Dies ist damit erklärbar, daß unter den Patienten, die von Lutz untersucht wurden, sich mehr schwere Fälle befanden als in unserer Stichprobe (depress. Stupor (3), sehr schwere Depression (8), schwere Depression (13), mittelschwere Depression (7), leichte Depression (2), unauffällig (3), vgl. Lutz (1969)), und die Höhe der Korrelation hängt von der Heterogenität der Stichprobe ab (Kalveram, 1970). Durch die Hinzunahme schwerer und extrem schwerer endogener depressiver Patienten wären unsere Korrelationskoeffizienten vermutlich angestiegen.

### 3. Trennung gesunder und endogen depressiver Personen durch eine Kombination von Beurteilungs- und Leistungsverfahren

Publikationen hierzu sind uns nicht bekannt. Wir können nur eigene Ergebnisse angeben.

Wie aus Tabelle 8 ersichtlich, lassen sich bei einer Kombination von Selbstbeurteilungs- und Leistungsverfahren deutlichere Trennungen als mit einem Verfahrenstyp alleine durchführen.

Die Trefferraten liegen zwischen 86% und 89%.

**Tabelle 7.** Häufigkeitsverteilungen der kombinierten LtD- und EDS-Ergebnisse. Außerhalb der Klammern stehen Absolutangaben, in den Klammern Prozentwerte

Kombination LtD — EDS <sup>a</sup>	Kontrollpersonen	Zykllothym depressive Patienten	
		ruhig <sup>b</sup>	unruhig <sup>c</sup>
0— 4	— (—)	— (—)	— (—)
4— 9	2 (7,4)	— (—)	— (—)
10—14	5 (18,6)	2 (3,1)	1 (3,7)
15—19	12 (44,3)	5 (7,7)	1 (3,7)
G→ 20—24	5 (18,6)	3 (4,6)	3 (11,1)
25—29	3 (11,1)	8 (12,3)	6 (15,4)
30—34	— (—)	8 (12,3)	2 (7,4)
35—39	— (—)	10 (15,4)	8 (29,7)
40—44	— (—)	11 (16,9)	— (—)
45—49	— (—)	10 (15,4)	3 (11,1)
50—54	— (—)	6 (9,2)	3 (11,1)
55—59	— (—)	2 (3,1)	— (—)
60—64	— (—)	— (—)	— (—)
	— (—)	— (—)	— (—)
Summe	27 (100,0)	65 (100,0)	27 (100,0)

<sup>a</sup> Die kombinierte Wertpunktzahl wurde folgendermaßen ermittelt: EDS-Rohpunkte + (3 × LtD-Wertpunkte)

<sup>b</sup> Patienten, die nach ihrem Erscheinungsbild vom behandelnden Arzt als motorisch ruhig, „gehemmt“ bezeichnet wurden

<sup>c</sup> Patienten, die nach ihrem Erscheinungsbild vom behandelnden Arzt als motorisch unruhig, „agitiert“ bezeichnet wurden

**Tabelle 8.** Prozentsatz richtiger Zuordnungen (in Klammern falsche Zuordnungen) bei Bestimmung des optimalen Trennpunktes oder -intervall es. Es wurden jeweils die Kontrollpersonen mit einer Stichprobe der zykllothym depressiven Patienten verglichen

Trennpunkt Verfahren oder -intervall <sup>a</sup>		Kontrollpersonen/ ruhige zykllothym depressive Patienten	Kontrollpersonen/ unruhige zykllothym depressive Patienten	Kontrollpersonen/ alle zykllothym depressive Patienten
LtD	7 <sup>b</sup>	A 80,7 (19,3)	70,3 (29,7)	77,5 (22,5)
	6	B 77,4 (22,6)	70,3 (29,7)	75,2 (24,8)
B-S	17—20	C 78,7 (21,3)	87,0 (13,0)	81,2 (18,8)
EDS	5— 7	D 83,5 (16,5)	86,2 (14,8)	84,0 (16,0)
Kombination LtD — B-S	46—54 <sup>c</sup>	E 85,4 (14,6)	88,9 (11,1)	86,4 (13,6)
	37—45	F 83,2 (16,8)	83,4 (16,6)	81,8 (18,2)
Kombination LtD — EDS	25—29	G 86,7 (13,3)	86,2 (14,8)	86,3 (13,7)

<sup>a</sup> Personen, deren Punktwerte die angegebene Zahl erreichen oder überschreiten, sollen der depressiven Gruppe zugeordnet werden. Die optimalen Trennpunkte oder -intervalle waren für die ruhigen und unruhigen zykllothym depressiven Patienten gleich

<sup>b</sup> Beim Trennpunkt „7“ wird nur eine der Kontrollpersonen fälschlich den Stichproben der zykllothym depressiven Patienten zugeordnet. Dagegen gehen beim Trennpunkt „6“ die Fehlklassifizierungen gleichmäßiger zu Lasten der beteiligten Stichproben

<sup>c</sup> Für das Trennintervall „46—54“ gilt das, was unter „<sup>b</sup>“ ausgeführt wurde

#### IV. Diskussion

Das Leistungsverfahren und die verwendeten Beurteilungsverfahren sind als etwa gleichwertig anzusehen. Dafür sprechen die Ergebnisse zur Validierung beider Verfahrenstypen und die etwa gleichgute Trennfähigkeit zwischen psychiatrisch unauffälligen und endogen depressiven Personen.

In der kombinierten Anwendung trennen beide Verfahrenstypen noch besser als bei einzelner Anwendung. Erstaunlich ist die geringe Korrelation zwischen beiden Verfahrenstypen. So wurden bereits bei Faktorenanalysen von Selbstbeurteilungs- und Leistungsvariablen bei endogen depressiven Patienten weitgehend unabhängige Leistungs- und Selbstbeurteilungsfaktoren gewonnen (Lehrl et al., 1976).

Die nachgewiesene Validität nicht korrelierender Verfahren könnte daran liegen, daß die endogene Depression mehrdimensional ist (wenigstens zwei Dimensionen) und jedes Verfahren eine Dimension mißt. Die klinische Konzeption der endogenen Depression strebt aber eine Eindimensionalität an. So wird von *der* Depression oder von verschiedenen Schweregraden *der* Depression gesprochen. Dies drückt sich auch in Generalfaktormodellen zur Fremdbeurteilung der Schweregrade von Psychosen aus, wie man sie beispielsweise mit der Global Assessment Scale (Endicott et al., 1976) durchführen kann. Aber auch in Selbstbeurteilungsskalen wird oft Eindimensionalität angestrebt, beispielsweise in der Befindlichkeitsskala (v. Zerssen, 1976a) und der Beschwerdeliste (v. Zerssen et al., 1970). Versucht man, sich über die Besonderheiten der Verfahrenstypen klar zu werden, läßt sich das Konzept der Eindimensionalität endogener Psychosen durchaus halten. In Selbstbeurteilungsverfahren werden vorwiegend Angaben über Mißbefinden und Verstimmung gemacht. Es ist aber zu erwarten, daß die Gefühlsbreite mit zunehmender depressiver Störung immer stärker eingeschränkt wird (Wieck, 1974). Analog scheinen von mittleren Schweregraden an Tagesschwankungen zu verschwinden (Waldmann, 1972). Außerdem sind die Patienten mit schweren Depressionen nicht mehr fähig, die Skalen auszufüllen. Daher ist zu erwarten, daß Selbsteinschätzungsskalen von mittelschweren Störungsgraden ab nicht mehr differenzieren (vgl. Befunde von Bech et al.) und daß sie sogar jenseits der mittelschweren Störung rückläufig werden können, weil hier keine Mißgefühle (wie Angst, Trauer) mehr angegeben werden.

Die Leistungsverfahren dürften weniger valide als Beurteilungsverfahren bei der Trennung leichter Depressionen vom Normalzustand sein, da schwache Leistungsminderung oft noch unter den Bereich des Meßfehlers und in die interindividuelle Streuung der Leistungswerte innerhalb des Normalbereiches fallen. Mit zunehmendem Störungsgrad setzen sich die Leistungswerte immer deutlicher vom Normbereich ab. Ab etwa einer mittelschweren Depression ist die Penetranz der Depression so auffällig, daß sie sich vermutlich deutlich in den Ergebnissen widerspiegelt.

Wir nehmen hier an, daß die mangelnde Korrelation von Selbstbeurteilungs- und Leistungsverfahren auf deren Differenzierung unterschiedlicher Schweregrade zurückzuführen ist. Ein Kriterium, das außerhalb der mit diesen Verfahren erhobenen Meßwerten liegt, wäre das Arzturteil. Wir hatten in der vorliegenden

Studie aber nicht genügend schwere und schwerste Depressionen, weshalb die Überprüfung dieser Hypothesen anhand der vorliegenden Daten im nachhinein nicht erfolgversprechend ist. Ihnen sind wir inzwischen andernorts gründlicher nachgegangen (Lehrl et al., 1977; Lehrl et al., 1978). Hätte die Stichprobe mehr derartige Fälle aufgewiesen, so wäre die Trennfähigkeit zwischen psychiatrisch unauffälligen und endogen depressiven Personen vermutlich zugunsten des Leistungsverfahrens angestiegen. Durch diese homogene Auswahl der Stichprobe ist auch zu erklären, weshalb die Korrelationen zwischen Leistungsverfahren und Arzturteil, hier als Kriterium des depressiven Schweregrades verwendet, in der vorliegenden Studie nur bei 0,32, in der Arbeit von Lutz (1969) jedoch bei 0,72 lag. Der Arbeit von Lutz lag eine viel heterogenere Stichprobe zugrunde.

Leistungsverfahren scheinen sich aber noch verbessern zu lassen. Ihre Differenzierungsfähigkeit im Abschnitt leichterer Depressionen müßte zunehmen, wenn monotone Untertests mit aufgenommen werden. Entsprechende Erfahrungen hatten bereits Wieck und Lutz (1969) veröffentlicht.

## Literatur

- Babcock, H.: Time and mind. Cambridge, Mass.: Massachusetts Science and Art Publications 1941
- Bech, P., Gram, L. G., Dein, E., Jacobsen, O., Viteger, J., Bolwing, T. G.: Quantitative rating of depressive states. *Acta Psychiat. Scand.* **51**, 161—170 (1975)
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., Erbaugh, J.: An inventory for measuring depression. *Arch. Gen. Psychiat.* **4**, 561—571 (1961)
- Benton, A. L., Spreen, O.: Der Benton-Test. Bern-Stuttgart: Huber 1961
- Blackburn, I. M.: Mental and psychomotoric speed in depression and mania. *Br. J. Psychiat.* **126**, 329—335 (1975)
- Dahl, G.: Übereinstimmungsvalidität des HAWIE und Entwicklung einer reduzierten Testform. Meisenheim: Hain 1968
- Endicott, J., Spitzer, R. L., Fleiß, J. L., Cohen, J.: The global assessment scale—Procedure for measuring overall severity of psychiatric disturbance. *Arch. Gen. Psychiat.* **33**, 766—771 (1976)
- Eysenck, H. J.: Maudsley Personality Inventory (MPI). Göttingen: Hogrefe 1959
- Eysenck, H. J.: Maudsley-Persönlichkeitsfragebogen (MMQ), 2. Auflage. Göttingen: Hogrefe 1964
- Friedman, A. S.: Minimal effects of severe depression on cognitive functioning. *J. Abnorm. Soc. Psychol.* **69**, 237—243 (1964)
- Hamilton, M.: A rating scale for depression. *J. Neurol. Neurosurg. Psychiat.* **23**, 56—62 (1960)
- Hathaway, S. R., McKinley, J. C., Spreen, O.: MMPI Saarbrücken. Bern: Huber 1963
- Kalveram, K.-T.: Probleme der Selektion in der Faktorenanalyse. *Arch. Psychol.* **122**, 215—222 (1970)
- Kerekjarto, M. v., Lienert, G. A.: Hamburger Depressions-Skala (HDS). *Pharmakopsych.* **3**, 1—21 (1970)
- Lehrl, S.: Ein Hemmungstest zur Messung der zykllothymen Depression. In: *Klinische Psychopathometrie* (Wieck, H. H., ed.). Byk Gulden 1972
- Lehrl, S.: Manual zum MWT-B. Erlangen: Straube 1977
- Lehrl, S., Straub, R., Straub, B.: Ein Vergleich von Beurteilungsskalen und Leistungsverfahren für die Schweregradmessung zykllothymer Depressionen. *Pharmakopsych.* **9**, 247—256 (1976)
- Lehrl, S., Gallwitz, A.: Manual zur Erlanger-Depressions-Skala. Vaterstetten-München: Vless 1977

- Lehrl, S., Gallwitz, A., Straub, R., Straub, B., Kinzel, W.: Manual zum LTD. Vaterstetten-München: Vless 1977
- Lehrl, S., Gallwitz, A., Nusko, G., Burkard, G.: Warum geben Selbstbeurteilungs-Verfahren schwere Depressionen falsch wieder? Therapeut. Umsch. **34**, 687—693 (1977)
- Lehrl, S., Gallwitz, A., Erzigkeit, H., Blaha, L.: Wann liefern Selbstbeurteilungsskalen ein falsches Bild vom Therapieverlauf endogener Depressionen? Therapiewoche **28**, 4150—4159 (1978)
- List, J.: Die Anwendung des Hemmungstestes nach Lutz bei Patienten mit Funktionspsychose. Diss. Erlangen 1972
- Lüscher, M.: Lüscher-Test. Hamburg: Rowolth 1971
- Lutz, H.: Zur Quantifizierung der zykllothymen Hemmung. Diss. Erlangen 1969
- Martin, I., Rees, R.: Reaction times and somatic reactivity in depressed patients. J. Psychosom. Res. **9**, 375—382 (1966)
- Murray, H. A.: Thematic apperception test. Cambridge, Mass.: Harvard University Press 1943
- Nelson, E. H.: An experimental investigation of intellectual speed and power in mental disorders. Unpublished Ph. D. Thesis, University of London Library, 1953
- Payne, R. W., Hewlett, J. H. G.: Thought disorder in psychotic patients. In: Experiments in personality, Volume II (Eysenck, H. J., ed.). London: Routledge and Kegan Paul 1966
- Philippen, N., Thorwirth, V.: Quantitative Untersuchungen über den Einfluß eines Thymoleptikums auf die Antriebshemmung endogener Depressionen. Arch. Psychiat. Nervenkr. **214**, 150—164 (1971)
- Polak, O., Brichcin, M.: Quantitative Analyse der psychomotorischen Hemmung bei endogenen Depressionen. Neuropsychopharmacologie **2**, 251—269 (1969)
- Popoff, L. M.: Depressions-Index für die Praxis. Basel: Ciba Geigy (ohne Jahr)
- Rapaport, D. (Ed.): Diagnostic psychological testing. Chicago: 1945
- Shapiro, M. B., Campell, D., Harris, A., Dewsbury, J. P.: Effects of ECT upon psychomotor speed and the distraction effect in depressed psychiatric patients. J. Ment. Sci. **104**, 681—695 (1958)
- Siegel, S.: Nonparametric statistics for the behavioral sciences. New York: McGraw Hill 1956
- Straub, R., Straub, B., Lehrl, S.: Zyklothyme Depression und Leistung. Psychiat. Neurol. med. Psychol. (Leipzig) **28**, 500—508 (1976)
- Supprian, U.: Die Eppendorfer Stimmungs-Antriebs-Skala (ESTA). Ein Instrument zur Abbildung des Verlaufs manisch-depressiver Psychosen. Pharmakopsychiat. **8**, 8—25 (1975)
- Szondi, L.: Experimentelle Triebdiagnostik (Szondi-Test). 3. Aufl. Bern: Huber 1972
- Waldmann, H.: Die Tagesschwankung in der Depression als rhythmisches Phänomen. Fortschr. Neurol. Psychiat. **40**, 83—104 (1972)
- Wartegg, E.: Wartegg-Zeichen-Test (WZT). 2. Aufl. Göttingen: Hogrefe 1968
- Wechsler, D.: Die Messung der Intelligenz Erwachsener. Bern-Stuttgart: Huber 1956
- Wieck, H. H.: Lehrbuch der Psychiatrie. Stuttgart: Schattauer 1967
- Wieck, H. H.: Angewandte Psychopathometrie. Janssen-Symposion, Bd. 8. Düsseldorf: 1973
- Wieck, H. H.: Neurologie und Psychiatrie in der Praxis. 2. Aufl. Stuttgart-New York: Schattauer 1974
- Wieck, H. H., Lutz, H.: Ein Hemmungstest zur Quantifizierung der zyklothymen Depression. Arzneim.-Forsch. (Drug. Res.) **19**, 886—889 (1969)
- Zerssen, D. v.: Die Beschwerden-Liste als Test. Therapiewoche **21**, 1908—1920 (1971)
- Zerssen, D. v., Koeller, D. M., Rey, E. R.: Die Befindlichkeitsskala (B-S) — ein einfaches Instrument zur Objektivierung von Befindlichkeitsstörungen, insbesondere im Rahmen von Längsschnittuntersuchungen. Arzneim.-Forsch. (Drug. Res.) **7**, 915—918 (1970)
- Zerssen, D. v.: Befindlichkeitsskalen. Weinheim: Beltz 1976a
- Zerssen, D. v.: Beschwerden-Liste. Weinheim: Beltz 1976b
- Zerssen, D. v.: Paranoid-Depressivitäts-Skala. Weinheim: Beltz 1976c
- Zung, W. W. K.: The depression status inventory. An adjunct to the self-rating depression scale. J. Clin. Psychol. **28**, 539—543 (1972)