

Validitätsprobleme bei der Differenzierung des
agitierten und gehemmt-depressiven Syndroms mit
Hilfe von Selbstbeurteilungsskalen

Wilhelm Kempf, Ferdinand Keller,
Roland Straub

1. Problemstellung und Vorgehensweise

Versuche, die Trennung zwischen dem "agitier-
ten", dem "gehemmten" und eventuell noch ei-
nem "apathischen" depressiven Syndrom auf
Symptomebene vorzunehmen, zeigen, daß es nur
ganz wenige Symptome gibt, die eindeutig ei-
nem der Syndrome zugeordnet werden können.
Bei der Konstruktion der Weissenauer Depres-
sionsskala (WDS; vgl. Straub et al., 1980;
Kempf et al., 1982; Keller et al., 1983)
wurde daher ein methodischer Ansatz gewählt,
der die unmittelbare Symptomebene verläßt,
statt dessen Symptomkombinationen betrach-
tet und der Frage nachgeht, ob sich inner-
halb verschiedener Symptombereiche (z.B.
Schlaf/Tagesrhythmik; Antrieb/Denken) Klas-
sen von Symptomkombinationen (sogenannte
Syndromklassen) identifizieren und den Syn-
dromen zuordnen lassen.

Ausgangspunkt für die Konstruktion der WDS
bildeten die Self-Anxiety-Scale und die
Self-Depression-Scale von Zung (1965, 1971),
beide in einer übersetzten und erweiterten
Fassung von Pichot und Debonis, sowie eine
in Anlehnung an eine frühe Fassung der Ei-

genschaftswörterliste von Janke und Debus (1977) entwickelte Liste körperlicher Symptome zur Erfassung der vegetativen körperlichen Befindlichkeit. Aus diesen Skalen wurden auf Grundlage einer Expertenbefragung jene Items ausgewählt, die den Syndromen "gehemmt" und "agitiert" sicher und weitgehend einheitlich zugeordnet werden konnten.

Nach ersten testtheoretischen Analysen wurden die ausgewählten Items sprachlich überarbeitet und um auf Grundlage einer Lehrbuchanalyse konstruierte zusätzliche Items ergänzt. Die daraus resultierenden insgesamt 50 Items der WDS (vgl. Keller et al., 1983) wurden erneut von 16 im Umgang mit Depressiven erfahrenen Psychiatern und Psychologen den Syndromen agitiert (a) und gehemmt (g) und apathisch (ap) zugeordnet und in 10 Symptombereiche (M = Müdigkeit, T = Trübsal/Wertlosigkeit, D = Antrieb/Denken, H = Herz/Atmung, U = Unruhe, S = Schlaf/Tagesrhythmik, R = Reizbarkeit/Zurückhaltung, A = unbestimmte Angst, Z = Zittern und K = kognitive Items) gruppiert.

Die weitere Vorgehensweise läßt sich in 3 Schritten darstellen:

1. Identifikation von Syndromklassen innerhalb der Symptombereiche mit Hilfe der Latent-Class-Analyse (LCA) nach Lazarsfeld (1950),
2. Untersuchung der internen Validität der identifizierten Syndromklassen mit Hilfe multipler Korrespondenzanalysen (MK)

nach Lebart, Morineau und Tabard (1977)
und

3. Untersuchung der externen Validität der
gewonnenen Ergebnisse.

2. Identifikation von Syndromklassen

Von den insgesamt 10 Symptombereichen wurden 9 einer LCA unterzogen. (Der Symptombereich K blieb zunächst unberücksichtigt, da die darin angesprochenen Symptome nach Experteneinschätzung keinem der Syndrome zugeordnet werden können.) Als Datenmaterial standen 1028 auf der Depressionsstation der PLK Weissenau erhobene Fragebogen zur Verfügung.

Zu brauchbaren Ergebnissen führte die LCA in 6 der 9 Symptombereiche, wobei die Modellanpassung in den Symptombereichen M, T und D sehr gut war. In den Bereichen H, U und S waren die gefundenen Lösungen auf dem 5 %-Niveau knapp signifikant.

In jedem dieser 6 Symptombereiche ergab die LCA jeweils eine symptomreiche und eine symptomarme Klasse sowie eine oder mehrere durch typische Symptomkombinationen charakterisierte Syndromklassen.

Ausgehend von der Experteneinschätzung jener Symptome, die in den Syndromklassen besonders häufig oder besonders selten auftreten, wurden die Syndromklassen D 4 = ziellose Aktivität dem agitierten Syndrom, D 3 = zähes, langsames Denken dem gehemmten Syndrom sowie M 3 = Kraftlosigkeit, Schläffheit, M 4 = Müdigkeit und D 5 = untätiges Grübeln, keine

ziellose Aktivität, dem apathischen Syndrom hypothetisch zugeordnet. Für die übrigen Syndromklassen bot sich eine solche direkte Syndromzuordnung nicht an.

3. Interne Validität

Die Untersuchung der internen Validität der identifizierten Syndromklassen erfolgte auf Grundlage desselben Datenmaterials zum einen unter Berücksichtigung aller 6 Symptombereiche, in denen latente Klassen identifiziert werden konnten (MK 1), und zum anderen unter Beschränkung auf nur jene Symptombereiche, in denen die LCA eine sehr gute Modellanpassung gewährleistete (MK 2). Interpretiert wurden in beiden Analysen nur jene Faktoren, die einen Varianzanteil $> 9\%$ beschreiben.

Übereinstimmend in beiden multiplen Korrespondenzanalysen konnte je ein Faktor (F 1) extrahiert werden, der die symptomreichen von den symptomarmen Syndromklassen trennt, sowie je ein Faktor (F 2), der diese beiden (quantitativen) Extremgruppen von jenen (qualitativen) Klassen unterscheidet, die durch typische Symptomkombinationen charakterisiert sind. Die Übereinstimmung der Faktoren aus den beiden Analysen ist Tabelle 1 zu entnehmen.

MK 2 \ MK 1	- sr/ + sa F 1	- qn/ + ql F 2	- g/ + a&ap F 3	- a/ + ap F 4
	F 1: - sr/+sa	0.94**	-0.02	-0.07*
F 2: - qn/+ql	0.03	0.88**	0.03	0.11**
F 3: - a&ap/ +g&ap	-0.01	0.00	-0.48**	0.46**
.
.
.
F 7: - a /+ap	0.02	0.00	-0.02	0.20**

**) = sehr sign ($\alpha = 1 \%$)

*) = sign ($\alpha = 5 \%$)

Tab. 1 Interkorrelationen der mittels multipler Korrespondenzanalysen extrahierten Faktoren (n = 1028).

Abkürzungen: sr = symptomrein

sa = symptomarm

qn = quantitativ

ql = qualitativ

a = agitiert

g = gehemmt

ap = apathisch

Da sowohl die symptomreichen als auch die symptomarmen Syndromklassen auf keinem der weiteren Faktoren nennenswert laden, lag die Vermutung nahe, daß diese Syndromklassen auch dann keine für die Syndromdiagnose relevante Information liefern, wenn jedes einzelne der darin vertretenen Symptome demselben Syndrom zuzuordnen ist.

Die Interpretation der weiteren in der MK 1 extrahierten Faktoren lautete:

- F 3: gehemmt vs. agitiert und apathisch
(hohe negative Ladungen von D 3, hohe positive Ladungen von M 3 und D 4)

- F 4: agitiert vs. apathisch
(hohe negative Ladung von D 4, hohe positive Ladung von D 5)

Insgesamt deckten die 4 Faktoren jedoch nur 35,9 % der Gesamtvarianz ab, so daß die Vermutung naheliegt, daß die herkömmliche Syndromeinteilung in ein agitiertes, ein gehemmtes und ein apathisches Syndrom - zumindest was das Selbsterleben depressiver Patienten betrifft - nur einen geringen Anteil an der depressiven Symptomatik zu systematisieren vermag. Dieser Eindruck wird durch die Ergebnisse der MK 2 noch verstärkt, in der insgesamt 7 Faktoren einen Varianzanteil von 84,8 % beschrieben. Von diesen erlaubten die Faktoren 3 - 6 eine handlungstheoretische Interpretation, d.h. eine Systematisierung kognitiver und motivationaler Defizite der Handlungsorganisation bei Depressiven, wie sie u.a. bei Beck et al. (1981), Kuhl (1983) und Kastner (1981) beschrieben werden. Lediglich Faktor 3, der zwischen Schwierigkeiten im Bereich der Handlungsorientierung und Schwierigkeiten im Bereich der Handlungsausführung trennt, kann auch im Lichte der herkömmlichen Syndromeinteilung gesehen werden:

- F 3: agitiert & apathisch vs. gehemmt & apathisch
(hohe negative Ladung von D 4 und M 3, hohe positive Ladung von D 3 und M 4)

Diese Interpretation wird auch durch die Korrelationen mit den Faktoren F 3 und F 4 der MK 1 (vgl. Tabelle 1) gestützt.

Ebenfalls im Lichte der herkömmlichen Syndrom-einteilung interpretierbar ist dann erst wieder der 7. Faktor:

- F 7: agitiert vs. apathisch

(hohe negative Ladung von D 4, hohe positive Ladung von M 3 und M 4)

Die Korrelation mit dem gleichlautenden 4. Faktor aus der MK 1 ist allerdings zwar auf 1 %-Niveau signifikant, beträgt jedoch lediglich 0.20 (vgl. Tabelle 1).

4. Externe Validität

Zur weiteren Absicherung der Interpretation der Faktoren wurde schließlich eine externe Validierungsstudie durchgeführt. n = 72 neu aufgenommene Patienten der Depressionsstation des PLK Weissenau wurden vom behandelnden Psychiater auf siebenstufigen Rating-Skalen danach eingeschätzt, wie deutlich die Syndrome agitiert, gehemmt und apathisch bei ihnen ausgeprägt seien.

		Fremdbeurteilung		
		a	g	ap
MK 1	F 1: -sr/+sa	- 0.05	0.23	- 0.11
	F 2: -qn/+ql	0.24	- 0.20	- 0.04
	F 3: -g/+a&ap	0.15	0.20	- 0.32*
	F 4: -a/+ap	0.07	- 0.21	0.02
MK 2	F 1: -sr/+sa	0.06	0.13	- 0.16
	F 2: -qn/+ql	0.23	- 0.24	- 0.14
	F 3: -a&ap/+g&ap	- 0.06	- 0.03	0.22
	.			
	F 7: -a/+ap	0.14	0.06	- 0.25

*)signifikant ($\alpha = 5\%$)

Tab 2 Korrelationen der Faktoren mit den Fremdbeurteilungen (n = 50)

Wie aus Tabelle 2 hervorgeht, sind die Korrelationen der Faktoren mit der psychiatrischen Fremdbeurteilung auf den oben genannten Rating-Skalen durchweg niedrig. Mit einer einzigen Ausnahme sind alle Korrelationen auf dem 5 %-Niveau nicht signifikant. Die Interpretation der Faktoren im Hinblick auf die Syndromeinteilung "agitiert-gehemmt-apathisch" wird durch die Fremdbeurteilung nicht gestützt.

Wie aus Tabelle 3 hervorgeht, sind insbesondere auch jene 5 Syndromklassen, die für die syndrombezogene Interpretation der Faktoren maßgeblich waren, mit der Fremdbeurteilung unkorreliert.

SK	SZ	Fremdbeurteilung		
		a	g	ap
D 4	a	- 0.04	- 0.15	0.02
D 3	g	0.06	- 0.05	- 0.12
M 3	ap	0.17	- 0.05	- 0.24
M 4	ap	0.04	- 0.05	- 0.04
D 5	ap	0.13	- 0.17	- 0.09

Tab 3 Korrelationen der für die syndrombezogene Interpretation der multiplen Korrespondenzanalysen maßgeblichen Syndromklassen mit der Fremdbeurteilung. (n = 63)
 Abkürzungen: SK = Syndromklasse
 SZ = Syndromzuordnung

Auch insgesamt sind die Korrelationen zwischen den identifizierten Syndromklassen und der Fremdbeurteilung mit den in der WDS abgefragten Symptomen selbst, wie aus Tabelle 4 hervorgeht:

Selbstbeurteilung	Fremdbeurteilung		
	a	g	ap
Symptome	8 %	2 %	14 %
Syndromklassen	11 %	4 %	11 %

Tab. 4 Anzahl der signifikanten Korrelationen (in %) der 50 Symptome und der 27 Syndromklassen mit der Fremdbeurteilung ($\alpha = 5\%$; $n = 63$)

Nicht zuletzt ist auch die Experteneinschätzung der Symptome - auf welche die Interpretation der Syndromklassen und Faktoren letztlich zurückgeht - nicht nur mit einiger Unsicherheit behaftet. Selbst jene Symptome, die von den Experten mit guter Übereinstimmung einem der Syndrome zugeordnet wurden, zeigen keineswegs den zu erwartenden Zusammenhang mit der Fremdbeurteilung (vgl. Tabelle 5).

Item Nr.	Syndromzuordnung	Sicherheit der Zuordnung	Fremdbeurteilung		
			a	g	ap
2	a	100 %	0.06	-0.06	0.15
26		100 %	0.21	0.06	-0.04
7		93,75 %	-0.20	-0.13	0.09
17		93,75 %	0.0	0.05	0.15
6		87,5 %	-0.16	-0.02	-0.06
4		75 %	-0.01	-0.20	-0.01
14		75 %	0.19	-0.04	0.19
15		75 %	0.19	-0.19	0.0
24		75 %	0.08	-0.08	-0.06
33		75 %	0.20	-0.02	-0.08

Item Nr.	Syndrom- zuordnung	Sicherheit der Zuord- nung	Fremdbeurteilung		
			a	g	ap
30	g	75 %	-0.12	-0.13	0.22
50	ap	81,25 %	0.11	-0.23	0.17
11		75 %	-0.09	-0.08	0.29

*) signifikant ($\alpha = 5\%$)

Tab. 5 Korrelationen zwischen Symptomen (WDS-Items) und Fremdbeurteilung für jene Items, die von mindestens 75 % der Experten demselben Syndrom und von keinem der Experten einem anderen Syndrom zugeordnet wurden. (n = 63)

Zusammenfassend läßt sich nur konstatieren, daß Selbst- und Fremdbeurteilung praktisch unkorreliert sind, und zwar sowohl auf Item- bzw. Symptomebene als auch auf der Ebene der Syndromklassen und auf Ebene der ihnen zugrundeliegenden Faktoren.

5. Diskussion

Selbst wenn man der Verwendung von psychiatrischen Fremdbeurteilungen als Kriteriumsvariable für die Validierung der WDS eine angemessene Relativierung zuteilwerden läßt, kann nicht übersehen werden: Der Versuch, ausgehend von gängigen psychiatrischen Selbstbeurteilungsskalen zu einer Differenzierung des agitierten, gehemmten und apathischen Syndroms zu gelangen, ist gescheitert. Die Ursache für das Scheitern liegt bereits in der Experteneinschätzung der Symptome. Offensichtlich ist es selbst im Umgang mit Depressiven erfahrenen Psychiatern und Psychologen nicht möglich, sich in die subjektive Erlebnisweise Depressiver hineinzuversetzen.

Als Ausweg aus diesem Dilemma ist auf die unbedingte Notwendigkeit einer genauen phänomenologischen Untersuchung der subjektiven Erlebnisweisen depressiver Menschen hinzuweisen. Tatsächlich beruhen die gebräuchlichen Selbstbeurteilungsskalen der Depression ja nicht auf einer solchen gründlichen und systematischen Kenntnis depressiver Subjektivität, sondern auf der "psychiatrischen" Erfahrung, d.h. auf einer "Sicht von außen", die zudem nicht systematisch gebildet, sondern mehr ad hoc zustande gekommen und zudem stark von theoretischen Vorstellungen und psychiatrischen Schulen gefärbt ist.

Hinzu kommt, daß Fragebogenitems herkömmlicher Art dem Patienten die Zustimmung zu oder die Ablehnung von kontextlosen Statements abverlangen. Eine Beantwortung der Fragen ist dem Patienten folglich nur möglich, indem er sie in seinen eigenen subjektiven Kontext stellt. D.h. dieselbe Frage kann von verschiedenen Patienten je nach ihrer Befindlichkeit sehr unterschiedlich verstanden werden - und ganz anders als vom Konstrukteur des Fragebogens.

Wie Kastner (1983) zu Recht festgestellt hat, wird "sprachliche Kommunikation umso schwieriger (...), je unvergleichbarer die psychischen Zustände der einzelnen Kommunikationsteilnehmer sind". Dies kann u.E. auch durch eine sprachpragmatische Fragebogenvalidierung, wie sie Kastner vorschlägt, nicht behoben werden. Was dadurch erreicht werden kann, ist allenfalls die Absicherung einiger gemeinsamer Sprachgebräuche. Die entscheidende Fehlerquelle bei Fragebögen ist jedoch

weit tiefer verwurzelt: in dem Versuch, sprachliche Kommunikation zu standardisieren und damit des Wesentlichen zu berauben, denn Kommunikation ist ein interaktiver Prozeß.

Gerade weil sprachliche Kommunikation umso schwieriger ist, je unvergleichbarer die psychischen Zustände der einzelnen Kommunikationsteilnehmer sind, kann eine Verständigung über subjektive Befindlichkeiten und Erlebnisweisen depressiver Patienten nur unter optimaler Ausnutzung aller Möglichkeiten, welche die (natürliche) Kommunikation für die Absicherung von Verständnissen bietet, d.h. insbesondere nur in Form einer schrittweisen interaktiven Annäherung erzielt werden.

Zum Zwecke der wissenschaftlichen Systematisierung könnten sich dafür allerdings diagnostische Gesprächsstrategien entwickeln lassen, in denen einerseits Themenbereiche (Symptombereiche) festgelegt werden, die zur Sprache gebracht werden sollen, und in denen andererseits die Form der Gesprächsführung (wie z.B. in der klientenzentrierten Therapie; Rogers, 1972) so gestaltet wird, daß dem Patienten die Chance gegeben wird, Widerstände zu überwinden und dem Gesprächspartner tatsächlich ein Stück weit Einblick in "sein Innerstes" zu gewähren. Ein umfangreiches methodisches Instrumentarium für die Entwicklung solcher Gesprächsstrategien bieten auch die Erkenntnisse des Neurolinguistischen Programmierens (Bandler & Grinder, 1981 a,b), die es ermöglichen, körpersprachliche, sprachstrukturelle und sinnesspezifische Komponenten der Kommunikation bei der Gesprächsführung explizit zu beachten.

In herkömmlichen Fragebogen bleiben nicht nur all diese Möglichkeiten unberücksichtigt. Aus falsch verstandenen Objektivierungs- und Standardisierungsansprüchen heraus wird zudem ein Zerrbild von Kommunikation kreiert, welches das in der Literatur immer wieder konstatierte - und auch in unserer Untersuchung gefundene - Auseinanderklaffen von Selbst- und Fremdbeurteilung geradezu zur Folge haben muß.

LITERATUR

- BANDLER, R.; GRINDER, J.
Metasprache und Psychotherapie. Paderborn:
Jungfermann, 1981 a.
- BANDLER, R.; GRINDER, J.
Neue Wege der Kurzzeit-Therapie. Paderborn:
Jungfermann, 1981 b.
- BECK, A.T.; RUSH, A.Y.; SHAW, B.F.; EMERY, G.
Kognitive Therapie der Depression. München:
Urban & Schwarzenberg, 1981.
- JANKE, W.; DEBUS, G.
Die Eigenschaftswörterliste (EWL-K). Ein Ver-
fahren zur Erfassung der Befindlichkeit.
Göttingen: Hogrefe, 1977.
- KASTNER, M.
Gemeinsamkeiten psychologischer Variablen des
Streß und der Depression. Zeitschrift für
klinische Psychologie, 1981, 10, 177 - 194
- KASTNER, M.
Ein Vorschlag zur Konstruktion von Depressions-
Fragebögen. In: Wolfersdorf, M; Straub, R.;
Hole, G. (Hg.). Der depressive Kranke in der
psychiatrischen Klinik - Theorie und Praxis
der Diagnostik und Therapie. Weinheim: Beltz
(im Druck), 1983
- KELLER, F.
Depressives Syndrom und Handlungsorganisation.
Psychologische Diplomarbeit, Universität Kon-
stanz, 1983.

- KELLER, F.; KEMPF, W.; STRAUB, R.
Zur Differenzierung agitierter und nicht-agitierter depressiver Syndrome anhand von Selbstbeurteilungen. In: Wolfersdorf, M.; Straub, R.; Hole, G. (Hg.). Der depressive Kranke in der psychiatrischen Klinik - Theorie und Praxis der Diagnostik und Therapie. Weinheim: Beltz (im Druck), 1983.
- KEMPF, W.; STRAUB, R.; HELBER, J.; WOLFERSDORF, M.; HOLE, G.
Vorläufige Ergebnisse zur Differenzierung des agitierten und gehemmt-depressiven Syndroms mit Hilfe von Selbstbeurteilungsskalen. Statistische Diskussionsbeiträge Nr. 73, Universität Konstanz, 1982.
- KUHL, J.
Volitional aspects of achievement motivation and learned helplessness: toward a comprehensive theory of action control. Reprint 1/1983. München: MPI für Psychologische Forschung, 1983.
- LAZARSELD, P.F.
Logical and mathematical foundations of latent structure analysis. In: Stouffer, S.A. et al. (Hg.). Studies in social psychology in world war II. Vol. IV, Princeton/N.Y.: Princeton University Press, 1950.
- LEBART, L.; MORINEAU, A.; TABARD, N.
Techniques de la description statistique. Paris: Dunod, 1977.
- ROGERS, C.R.
Die klientenzentrierte Gesprächspsychotherapie. München: Kindler, 1972.
- STRAUB, R.; KEMPF, W.; WOLFERSDORF, M.; HELBER, I.; HOLE, G.
Empirische Untersuchung zur Differenzierung des agitiert- und gehemmt-depressiven Syndroms mit Hilfe von Selbstbeurteilungsskalen. In: Ehlers, W. (Hg.). Diagnostik und Biosignalanalyse bei Kopfschmerz, Depression und Neurose. Stuttgart: Forschungsstelle für Psychotherapie, 1980.
- ZUNG, W.W.K.
A self rating depression scale. Arch. Gen. Psychiat. 1965, 12, 63 - 70
- ZUNG, W.W.K.
A rating instrument for anxiety disorders. Psychosomatics, 1971, 12, 63 - 70