

## Zur Entscheidungslogik bei Symptom-Beurteilungen

U. Baumann, G. Schneidewind, J. Angst, H. Helmchen und H. Hippius

Psychiatrische Universitätsklinik Zürich, Forschungsdirektion (Prof. Dr. med. J. Angst),  
Psychiatrische Klinik der Freien Universität Berlin (Prof. Dr. med. H. Helmchen)  
Psychiatrische Universitätsklinik München (Prof. Dr. med. H. Hippius)

Eingegangen am 30. Mai 1975

### The Logic of Decision in Symptom Rating

*Summary.* The rating systems used in psychiatry consider sometimes different ways of classifying parameters, such as the valuability and the certainty of decision. Those two parameters have been analysed by means of the AMP-ratings of two samples (Berlin: 1879 patients, Zürich: 786 patients). We have found that the reply "not valuable" appears much less frequently than the reply "questionable valuable". For delusional symptoms and disorders of perception the certainty of decision is lowest. The various analyses carried out lead us to the conclusion that without loosing informative value of any consequence we can renounce any separate documentation on valuability and certainty of decision.

*Key words:* Psychopathology — Rating-Scales — AMP-System — Logic of Decision.

*Zusammenfassung.* Psychiatrische Ratingsysteme berücksichtigen oft verschiedene Einstufungsgesichtspunkte, so auch die Beurteilbarkeit und Entscheidungssicherheit (sicher — fraglich vorhanden). Diese beiden Aspekte werden anhand der AMP-Ratings, die bei zwei Stichproben (Berlin: 1879 Patienten, Zürich: 786 Patienten) erhoben worden sind, analysiert. Es zeigt sich, daß die Kategorie „nicht beurteilbar“ bedeutend seltener vorkommt als die Kategorie „fraglich vorhanden“. Bei den Wahnsymptomen und den Sinnestäuschungen ist die Entscheidungssicherheit am geringsten. Die verschiedenen Analysen führen zum Schluß, daß ohne größeren Informationsverlust auf eine getrennte Dokumentation von Beurteilbarkeit und Entscheidungssicherheit verzichtet werden kann.

*Schlüsselwörter:* Psychopathologie — Rating-Skalen — AMP-System — Entscheidungslogik.

### 1. Einleitung

In der Psychiatrie sind Fremdbeurteilungsverfahren (Ratings) von besonderer Wichtigkeit (Baumann *et al.*, 1975); Ratings, als Methode der systematischen Beobachtung, liegen in verschiedensten Formen vor (Mombour, 1972; Pichot, 1974; Rating allgemein: von Cranach u. Frey, 1969). Während bei einigen Ratingsystemen die Symptome eindimensional eingestuft werden (vgl. z. B. IMPS von Lorr u. Klett, 1967; Hamilton-Scale von Hamilton, 1967; BPRS von Overall u. Groham, 1962), finden wir bei anderen eine komplexere Entscheidungslogik. So werden im AMP-System (Angst *et al.*, 1969; Scharfetter, 1972; Baumann, 1974) nach Bente vier Ebenen der Entscheidungslogik unterschieden:

Beurteilbarkeit (beurteilbar — nicht beurteilbar);

Entscheidungssicherheit über Vorhandensein oder Nichtvorhandensein (sicher — fraglich vorhanden);

Vorhandensein (vorhanden — nicht vorhanden);  
 Quantifizierung (leicht ausgeprägt — mittel ausgeprägt — schwer ausgeprägt).

In der Praxis enthält die Dimension Beurteilbarkeit zwei verschiedene Bedeutungen:

nicht beurteilbar trotz gezielter Untersuchungsbemühungen (gemäß Manual);  
 nicht beurteilbar, weil nicht untersucht.

Die Unterscheidung zwischen Vorhandensein und Quantifizierung ist meßtheoretisch (Sixtl, 1967) als Unterschied zwischen Messen auf Nominal- und Ordinalskalenniveau zu sehen. Da der Begriff Quantifizieren nicht eindeutig ist, wäre es besser von Graduierung zu reden.

Eine ähnliche Einteilung finden wir bei Wing *et al.* (1974, S. 196), die bei der PSE (Present-State-Examination Schedule) folgende vier Einstufungsaspekte aufführen:

- Code 9: nicht beurteilbar (Beurteilbarkeit);
- Code 8: fraglich vorhanden (Entscheidungssicherheit);
- Code 0 (1, 2): Vorhandensein;
- Code 1, 2: Graduierung.

Bei der PSE und beim AMP könnten die beiden Kategorien Beurteilbarkeit und Entscheidungssicherheit unter dem Stichwort *Aussagemöglichkeit* (Aussage über Vorhandensein des Symptoms möglich, respektive unmöglich) zusammengefaßt werden. Während beim AMP-System Intensität und Häufigkeit bei der Graduierung der Symptome implizit vermischt werden, was die Reliabilität des Systems beeinträchtigen dürfte, werden bei Wing *et al.* (1974) diese beiden Gesichtspunkte zwar berücksichtigt, aber danach nicht getrennt codiert.

*Beispiel.* Item 19 (Wing *et al.*, 1974, S. 197): Rate subjectively inefficient thinking

1 = Sympton definitely present during the past month of moderate clinical intensity, or intense less than 50% of the time.

2 = Symptom clinically intense more than 50% of the past month.

Noch weiter gehen z. B. Murphy *et al.* (1974), welche in der Manic-State-Rating-Scale die einzelnen Symptome jeweils getrennt nach Häufigkeit und Intensität einstufen lassen.

Daß ein Entscheidungsbaum — im Gegensatz zur eindimensionalen Einstufung — zu komplexeren Auswertungsstrategien führt, ist klar. Von der Praxis drängt sich daher die Frage auf, ob die Komplexität der Entscheidungslogik zu Recht besteht oder ob eventuell eine Vereinfachung möglich wäre, was für Auswertungsfragen sehr wesentlich ist. Theoretisch ist z. B. die Unterteilung in Beurteilbarkeit und Entscheidungssicherheit sinnvoll, doch ist bisher nicht klar, wie häufig diese beiden Aspekte in der Praxis verwendet werden. Im folgenden soll daher dieser Frage nachgegangen werden; auf das Problem Intensität-Häufigkeit und Graduierung können wir nicht weiter eingehen, da dies den Rahmen sprengen würde (vgl. Cohen *et al.*, 1969; Fahrenberg, 1975).

## 2. Fragestellung

An zwei größeren Stichproben, die eine Kreuzvalidierung der Ergebnisse erlauben (Lienert, 1969), sollen folgende Einzelfragen geklärt werden:

Wie viele Symptome sind bei den einzelnen Patienten nicht beurteilbar oder fraglich vorhanden (Auszählung nach *Patient*)?

Wie ist die Beurteilbarkeit und Entscheidungssicherheit bei den einzelnen Symptomen (Auszählung nach *Symptom*)?

Unterscheiden sich verschiedene Diagnosegruppen bezüglich der Beurteilbarkeit und Entscheidungssicherheit?

Wie sind die Patienten definiert, bei denen die Zahl der Symptome ohne Aussage besonders groß ist?

## 3. Methode

### a) Untersuchungsinstrument

Wie schon in Abschnitt 1 erwähnt, werden im AMP-System, das durch die Arbeitsgemeinschaft für Methodik und Dokumentation in der Psychiatrie (AMP) herausgegeben worden ist, die beiden Aspekte Beurteilbarkeit und Entscheidungssicherheit unterschieden. Es liegt daher nahe, anhand dieses Rating-Verfahrens die entsprechenden Analysen durchzuführen; dies um so mehr, als zwei größere Stichproben bereits vorlagen (vgl. Abschnitt 4).

Das AMP-System dient der Dokumentation des psychiatrischen Befundes und besteht aus vier Markierlesebögen: Allgemeine Anamnese, Krankheitsanamnese, Psychischer Befund (123 Symptome), Somatischer Befund (58 Symptome). Sämtliche 181 Symptome werden nach der in Abschnitt 1 referierten Entscheidungslogik eingestuft.

### b) Datenverarbeitung

Die Datenverarbeitung erfolgte im Universitäts-Rechenzentrum Zürich (IBM 370-155) mit Programmen der Programmbibliothek Psychlib (Baumann) und ad-hoc-Programmen. Zur Aufbereitung der AMP-Bögen wurde das von Scheidegger entwickelte AMP-Wandel-Programm benutzt (Scheidegger, 1973)<sup>1</sup>.

## 4. Stichproben

### a) Stichprobe Berlin

Die Berliner Stichprobe umfaßt 1879 Patienten, die zwischen 1968 und 1972 routinemäßig mit dem AMP-System in der Psychiatrischen Universitätsklinik Berlin dokumentiert worden sind. Sie setzt sich zusammen aus 1067 Männern (57%) und 812 Frauen (43%), die ein durchschnittliches Alter von 37 Jahren haben (Median = 33 Jahre). Die Altersverteilung ist leicht linksschief und geht von 12–99 Jahre. Die diagnostische Zusammensetzung finden wir in Tabelle 1.

Je ein Viertel der Stichprobe besteht aus schizophreneren (ICD 295) und aus neurotischen (ICD 300) Patienten, während Patienten mit affektiven Psychosen (ICD 296) als drittgrößte Gruppe mit 12,6% in Erscheinung treten. Betrachtet man die Untendiagnosen, so bilden ICD 3004 (Depressive Neurose, N = 294), ICD 2953 (paranoide Schizophrenie, N = 191) und ICD 2960 (Involutionsdepression, N = 101) die drei größten Gruppen.

<sup>1</sup> Herrn P. Scheidegger sei für die Mitarbeit bei der Datenaufbereitung besonders gedankt, ebenso Fr. S. Zulauf.

Tabelle 1. Diagnosen (ICD) der Stichproben Berlin und Zürich

Diagnosen (ICD)	Berlin		Zürich	
	Anzahl	%	Anzahl	%
290 Demenzen bei (prä)senilen Hirnkrankheiten	25	1,3	19	2,4
291 Alkoholpsychosen	57	3,0	16	2,0
292 Psychosen bei intrakraniellen Infektionen	8	0,4	6	0,8
293 Psychosen bei anderen organischen Hirnstörungen	43	2,3	65	8,3
294 Psychosen bei anderen körperlichen Krankheiten	46	2,4	3	0,4
295 Schizophrenie	438	23,3	360	45,8
296 Affektive Psychosen	236	12,6	219	27,9
297 Paranoide Syndrome	83	4,4	8	1,0
298, 299 Andere Psychosen	71	3,8	24	3,1
300 Neurosen	470	25,0	44	5,6
301 Persönlichkeitsstörungen	51	2,7	7	0,9
302 Sexuelle Verhaltensabweichungen	6	0,3		
303 Alkoholismus	60	3,2		
304 Medikamentenabhängigkeit	60	3,2	1	0,1
305 Psychosomatische Störungen	6	0,3	1	0,1
306 Besondere Symptome	11	0,6	1	0,1
307 Psychische Auffälligkeiten	15	0,8	1	0,1
308 Verhaltensstörungen im Kindesalter	3	0,2		
309 Andere Störungen	48	2,6		
310—315 Oligophrenie	12	0,7	4	0,5
345 Epilepsie	44	2,3		
950—959 Selbstbeschädigung	12	0,6	3	0,4
< 290, > 316 andere Diagnosen	74	3,9	4	0,5
	1879	99,9	786	100,0

*b) Stichprobe Zürich*

Die Daten stammen aus der Forschungsdirektion der Psychiatrischen Universitätsklinik Zürich. Der größte Teil des Materials wurde zwischen 1967 und 1974 als Ausgangsbefund im Rahmen von Psychopharmakastudien erhoben ( $N = 684$ ). Weitere Unterlagen von 102 Patienten stammen aus einer Dissertation (Britschgi, 1975) zur Psychopathologie von organischen Psychosen; die Daten sind in verschiedenen Kliniken gewonnen worden. Die 102 Patienten sind insofern selegiert, als Britschgi nicht systematisch alle Patienten der entsprechenden Diagnosegruppen untersucht hat, sondern nur Patienten, bei denen ein Interview zumindest teilweise möglich war.

Die insgesamt 786 Patienten sind durchschnittlich 47 Jahre alt (Median = 48); die Spannweite beträgt 11—91 Jahre. Die 349 Männer (44%) und 437 Frauen (56%) verteilen sich vorwiegend auf die Diagnosen Schizophrenie (ICD 295: 46%) und affektive Psychosen (ICD 296: 28%) (vgl. Tabelle 1). Am stärksten vertreten sind die Unterdiagnosen ICD 2953 (paranoide Schizophrenie,  $N = 205$ ), ICD 2962 (monopolare Depression,  $N = 116$ ), ICD 2957 (schizoaffektive Psychosen,  $N = 84$ ) und ICD 2960 (Involutionsdepression,  $N = 66$ ).

*c) Vergleich der beiden Stichproben*

Die beiden Stichproben unterscheiden sich bezüglich Geschlecht, Alter und Diagnose (Chiquaret,  $P < 0,01$ ), was durch die Verschiedenheit der Institutionen (Unterschied zwischen deutschen und schweizerischen Universitätskliniken) und der Dokumentationsziele (Routine, Forschung) bedingt ist.

Die Berliner Stichprobe umfaßt mehr Männer als Frauen, bei der Zürcher Stichprobe ist das Verhältnis umgekehrt. Die Berliner Stichprobe ist durchschnittlich jünger als die Zürcher und diagnostisch breiter gestreut. Für unsere methodische Untersuchung ist die Verschiedenheit der Stichproben aber kein Nachteil, sondern sogar von Vorteil für die Kreuzvalidierung.

### 5. Beurteilbarkeit und Entscheidungssicherheit bei einzelnen Patienten

Im folgenden wird die Zahl der nicht beurteilbaren und die Zahl der fraglich vorhandenen Symptome pro Patient getrennt nach psychischem und somatischem Befund betrachtet. Die meisten Patienten beider Stichproben ( $\geq 91\%$ ) haben alle Symptome des psychischen oder des somatischen Befundes beurteilbar (vgl. Tabelle 2 und 3, Spalte 1 und 4). Daneben gibt es aber in der Berliner Stichprobe einige wenige Patienten mit sehr vielen nicht beurteilbaren Symptomen.

Häufiger findet sich die Kategorie „fraglich vorhanden“, die bei der Mehrzahl der Patienten 0-, 1- oder 2mal vorkommt (vgl. Tabelle 2 und 3, Spalte 2 und 5). Die Häufigkeitsverteilungen sind überall linksschief. Die große Ähnlichkeit der

Tabelle 2. Beurteilbarkeit und Entscheidungssicherheit pro Patient (Stichprobe Berlin)

Anzahl Sym- ptome pro Patient	Psychischer Befund				Somatischer Befund			
	nicht beurteilbar	fraglich vorhanden	nicht beur- teilbar und fraglich vorhanden	nicht beurteilbar	fraglich vorhanden	nicht beur- teilbar und fraglich vorhanden		
	$N_i$	%	$N_i$	%	$N_i$	%	$N_i$	%
0	1721	91,6	762	40,6	721	38,4	1840	97,9
1	44	2,3	440	23,4	423	22,5	4	0,2
2	19	1,0	274	14,6	263	14,0	1	0,1
3	12	0,6	171	9,1	170	9,0	5	0,3
4	13	0,7	76	4,0	81	4,3	3	0,2
5	4	0,2	66	3,5	67	3,6	1	0,1
6	8	0,4	29	1,5	29	1,5	2	0,1
7	5	0,3	19	1,0	23	1,2	2	0,1
8	2	0,1	15	0,8	19	1,0	1	0,1
9	1	0,1	8	0,4	11	0,6		
10–19	13	0,7	19	1,0	30	1,6	7	0,4
20–39	14	0,7			18	1,0	12	0,6
40–59	14	0,7			14	0,7	1	0,1
≥60	9	0,5			10	0,5		
	1879	99,9	1879	99,9	1879	99,9	1879	100,2
	(1)		(2)		(3)		(4)	
	1879	99,9	1879	99,9	1879	100,2	1879	100,0
	(5)		(6)					

**Tabelle 3.** Beurteilbarkeit und Entscheidungssicherheit pro Patient (Stichprobe Zürich)

Anzahl Sym- ptome pro Patient	Psychischer Befund						Somatischer Befund					
	nicht beurteilbar		fraglich vorhanden		nicht beur- teilbar und fraglich vorhanden		nicht beurteilbar		fraglich vorhanden		nicht beur- teilbar und fraglich vorhanden	
	<i>N<sub>i</sub></i>	% <sub>i</sub>	<i>N<sub>i</sub></i>	% <sub>i</sub>	<i>N<sub>i</sub></i>	% <sub>i</sub>	<i>N<sub>i</sub></i>	% <sub>i</sub>	<i>N<sub>i</sub></i>	% <sub>i</sub>	<i>N<sub>i</sub></i>	% <sub>i</sub>
0	732	93,1	506	64,4	484	61,6	761	96,8	723	92,0	701	89,2
1	23	2,9	139	17,7	138	17,6	14	1,8	47	6,0	57	7,3
2	12	1,5	71	9,0	73	9,3	4	0,5	11	1,4	14	1,8
3	7	0,9	33	4,2	37	4,7	3	0,4	4	0,5	9	1,1
4	3	0,4	15	1,9	18	2,3	1	0,1	1	0,1	2	0,3
5	2	0,3	8	1,0	10	1,3						
6	2	0,3	6	0,8	10	1,3	1	0,1			1	0,1
7	1	0,1	2	0,3	6	0,8						
8			1	0,1			1	0,1			1	0,1
9			3	0,4	3	0,4						
10–19	1	0,1	2	0,2	4	0,4						
20–39	3	0,4			2	0,2	1	0,1			1	0,1
≥40					1	0,1						
	786	100,0	786	100,0	786	100,0	786	99,9	786	100,0	786	100,0
	(1)		(2)		(3)		(4)		(5)		(6)	

Verteilungen „fraglich vorhanden“ und „keine Aussage“ (nicht beurteilbar und fraglich vorhanden) deutet darauf hin, daß die Einstufung „nicht beurteilbar“ von geringer Bedeutung ist.

Beide Stichproben zeigen zwar dieselben Trends, doch finden sich bei der Entscheidungssicherheit (fraglich vorhanden) Unterschiede. Während in der Berliner Stichprobe 41% der Patienten alle psychischen Symptome eindeutig diagnostizierbar haben, sind es bei der Zürcher Stichprobe 64%; im somatischen Befund lauten die Zahlen 87% (Berlin) und 92% (Zürich). Beide Differenzen sind auf dem 1%-Niveau signifikant (Chiquadrat-test). Die Zahl der Patienten, die alle Symptome beurteilbar haben, ist in beiden Stichproben — in psychischen und somatischen Befunden — gleich groß (Chiquadrat-test nicht signifikant). Der Unterschied in der Entscheidungssicherheit kann auf die verschiedenartige Stichprobenzusammensetzung, aber auch auf verschiedenen intensiven Rater-training zurückgeführt werden.

#### **6. Beurteilbarkeit und Entscheidungssicherheit beim einzelnen Symptom**

Im Gegensatz zu Abschnitt 5 werden nun die Symptome einzeln analysiert; dabei brauchen wir auf die 58 somatischen Symptome nicht weiter einzugehen, da bei allen Symptomen der Prozentsatz der Patienten mit der Kategorie „nicht beurteilbar“ oder „fraglich vorhanden“ unter 1% liegt. Eine Ausnahme bildet das Symptom „Appetit vermindert“, das in der Berliner Stichprobe bei 1,2% als „fraglich vorhanden“ eingestuft wird!

Beim psychischen Befund sind folgende Symptome am häufigsten nicht beurteilbar:

Berlin: Gedächtnis- (2,7%), Merkfähigkeitsstörungen (2,6%), Konfabulationen (2,1%).

Zürich: Merkfähigkeits- (1,8%), Gedächtnissstörungen (1,4%), Stimmenhören (1,4%), Konzentrationsstörungen (1,3%).

Bedeutend höher ist bei einzelnen Symptomen der Prozentsatz der Kategorie „fraglich vorhanden“ (*unterstrichene Symptome*: nur in einer Stichprobe mit größerem Prozentsatz vertreten):

Berlin: Suicidtendenzen (9,3%), Merkfähigkeits- (6,9%), Gedächtnissstörungen (6,5%), Stimmenhören (6,0%), Konzentrationsstörungen (5,0%). Optische Halluzinationen (3,9%), Auffassungsstörungen (3,8%), Wahnstimmung (3,8%), Körperhalluzinationen/Körperfühlstörungen (3,7%), Beeinträchtigungs-/Verfolgungswahn (3,6%), Wahnwahrnehmung (3,2%), andere akustische Halluzinationen (3,1%).

Zürich: Stimmenhören (8,4%), Optische Halluzinationen (3,4%), Körperhalluzinationen/Körperfühlstörungen (3,4%), Wahnstimmung (2,9%), andere akustische Halluzinationen (2,8%), Beeinträchtigungs-/Verfolgungswahn (2,8%), Wahneinfall/Wahngedanken (2,5%), Wahnwahrnehmung (2,2%), Suicidtendenzen (2,0%).

Acht Symptome werden sowohl in Berlin als auch in Zürich besonders häufig als fraglich vorhanden eingeschätzt: Stimmenhören, andere akustische Halluzinationen, optische Halluzinationen, Körperhalluzinationen/Körperfühlstörungen, Wahnstimmung, Wahnwahrnehmung, Beeinträchtigungs-/Verfolgungswahn; Suicidtendenzen. Es erstaunt nicht, daß vor allem Symptome des Wahnbereiches und der Sinnestäuschungen besonders schwierig diagnostizierbar sind. Daß die Symptome Merkfähigkeits-, Gedächtnis-, Konzentrations- und Auffassungsstörungen bei der Zürcher Stichprobe weniger häufig als beim Berliner Material fraglich sind, dürfte auf die Stichprobenselektion zurückzuführen sein; Patienten mit organischen Psychosen, bei denen diese Merkmale besonders häufig sind, wurden im Rahmen der Zürcher Dissertation so ausgewählt, daß ihre Symptomatik im Rahmen eines Interviews erkennbar war. Unklare Fälle sind dadurch unterrepräsentiert (vgl. auch Abschnitt 4 b).

## 7. Beurteilbarkeit und Entscheidungssicherheit pro Diagnosegruppe

Da verschiedene Diagnosegruppen zu klein waren (vgl. Tabelle 1), konnten nur folgende sieben Gruppen analysiert werden:

ICD 290—294 organische Psychosen;

ICD 295 Schizophrenie;

ICD 296 affektive Psychosen;

ICD 297—299 übrige Psychosen;

ICD 300 Neurosen;

ICD 301—316 Restgruppe psychiatrischer Diagnosen;

andere Diagnosen außerhalb des psychiatrischen Bereiches (< 290, > 316).

In jeder Gruppe wurde pro Patient der durchschnittliche Prozentsatz an nicht beurteilbaren und an fraglich vorhandenen Symptomen berechnet (Tabelle 4).

Bei den meisten Diagnosegruppen ist der psychische Befund weniger klar einzustufen als der somatische. Umgekehrt verhält es sich im Zürcher Material bei der Gruppe ICD 297—299 (übrige Psychosen) und ICD 290—294 (organische Psychosen).

Tabelle 4. Beurteilbarkeit (NB) und Entscheidungssicherheit (FR) pro Patient und Diagnosegruppe (Prozent der Symptome des psychischen, respektive somatischen Befundes mit entsprechender Einstufung)

ICD	B: Berlin Z: Zürich	N	Psychischer Befund			Somatischer Befund		
			NB	FR	NB+FR	NB	FR	NB+FR
290—294	B	179	2,2	1,5	3,6	1,0	0,5	1,5
	Z	109	0,0	0,2	0,2	0,1	0,3	0,4
295	B	438	1,5	1,8	3,3	0,5	0,3	0,8
	Z	360	0,3	1,0	1,3	0,1	0,2	0,3
296	B	236	0,4	1,1	1,5	0,0	0,3	0,3
	Z	219	0,1	0,4	0,5	0,1	0,2	0,3
297—299	B	154	1,5	1,2	2,7	1,5	0,3	1,8
	Z	32	0,6	0,7	1,3	1,3	0,3	1,6
300	B	470	0,2	0,9	1,1	0,6	0,3	0,8
	Z	44	0,0	0,1	0,1	0,0	0,1	0,1
301—316	B	272	0,8	1,0	1,8	0,2	0,5	0,7
	Z	15 <sup>a</sup>	—	—	—	—	—	—
<290, >316	B	130	1,1	1,0	2,1	0,2	0,3	0,5
	Z	7 <sup>a</sup>	—	—	—	—	—	—
	B	1879	1,0	1,2	2,2	0,5	0,3	0,9
	Z	786	0,2	0,6	0,8	0,2	0,2	0,3

<sup>a</sup> N zu klein, Prozente wären sinnlos.

Bei der Sammelgruppe ICD 297—299 (übrige Psychosen) haben wir im psychischen und somatischen Befund in beiden Stichproben relativ viel Symptome ohne Aussage (nicht beurteilbar und fraglich vorhanden). Beim psychischen Befund ist die Symptomatik bei den Diagnosen Schizophrenie (ICD 295) (Berlin, Zürich) und organischen Psychosen (ICD 290—294) (Berlin) relativ schwierig einzustufen, für die letztere Gruppe gilt dies auch für den somatischen Befund. Am klarsten sind die Symptome der Diagnosegruppen Neurosen (ICD 300) und affektive Psychosen (ICD 296) erkennbar. Zur Erklärung der guten Einstufbarkeit bei den organischen Psychosen im Zürcher Material sei auf die Bemerkung im vorherigen Abschnitt verwiesen (Selektion!).

Das Verhältnis der Prozentzahlen „nicht beurteilbar“ zu „fraglich vorhanden“ pro Diagnosegruppe lässt sich nicht interpretieren, da die meisten Differenzen in den beiden Stichproben zu gering und die Ergebnisse zum Teil konträr sind.

#### 8. Patienten mit erhöhter Zahl von Symptomen ohne Aussage

Bei der Auswertung von Fragebogen wird oft eine empirisch festgelegte Schranke für die Zahl tolerierter Missing Data (ausgelassener Fragen) angegeben; wird diese Zahl überschritten, ist der Test auf Skalenebene nicht auswertbar. So werden z. B. bei der routinemäßigen Auswertung des Freiburger Persönlichkeitseinventars FPI (Fahrenberg *et al.*, 1973) maximal 5 Missing Data bei insgesamt 212 Items toleriert. Bei Ratings bietet sich ein ähnliches Vorgehen an, indem die Kategorie „nicht beurteilbar“ und „fraglich vorhanden“ als Missing Data zu betrachten sind, die keine Aussage über das Vorhandensein eines Symptoms zu-

lassen. Ist die Zahl der Symptome ohne Aussage erhöht, wäre eine Auswertung auf Syndromebene nicht mehr möglich. Problematisch wäre dieses Vorgehen, wenn dadurch vermehrt Patienten bestimmter Diagnosegruppen eliminiert würden, was nach Abschnitt 7 zu vermuten ist.

Es wurde daher versuchsweise eine Missing-Data-Schranke eingeführt und geprüft, welche Patienten diese Schranke überschreiten. Würde man beim psychischen Befund 5% aller Symptome ohne Aussage zulassen (= 6 Symptome), so könnte man 125 der 1879 Patienten der Berliner Stichprobe auf Skalenebene nicht auswerten; dies schien uns ein zu hoher Prozentsatz zu sein, weshalb wir diese Schranke mit 15% festsetzen (10% führt fast zum gleichen Resultat wie 15%). 43 Patienten der Berliner Stichprobe überschritten im psychischen ( $\geq 19$  Symptome ohne Aussage) und 20 Patienten im somatischen Befund ( $\geq 10$  Symptome ohne Aussage) die kritische Grenze. Berücksichtigt man, daß 14 Patienten sowohl im psychischen als auch im somatischen Befund zu viele Symptome ohne Aussage hatten, so reduzieren sich die 63 (43 + 20) Fälle auf 49 Patienten (Berliner Material). Diese hatten aus folgendem Grund zu viele Symptome ohne Aussage (n. b.: nicht beurteilbar, f.: fraglich vorhanden).

- 24:  $\geq 19$  n. b. Symptome beim psychischen Befund;
- 14:  $\geq 19$  n. b. Symptome beim psychischen und  $\geq 10$  n. b. Symptome beim somatischen Befund;
- 6:  $\geq 10$  n. b. Symptome beim somatischen Befund;
- 4:  $\geq 19$  n. b. und f. Symptome beim psychischen Befund;
- 1:  $\geq 19$  f. Symptome beim psychischen Befund.

Bei der Zürcher Stichprobe hatten drei Patienten zu viele Symptome ohne Aussage:

- 2:  $\geq 19$  n. b. Symptome beim psychischen Befund;
- 1:  $\geq 19$  n. b. Symptome beim psychischen und  $\geq 10$  n. b. Symptome beim somatischen Befund.

Zwei Punkte sind bemerkenswert: Der kleinere Teil der Patienten würde allein wegen des unklaren somatischen Befundes für die Ausweitung auf Skalenebene ausgeschieden. (Berlin: 6 von 49, Zürich: 0 von 3) und mit 5 Ausnahmen wird die kritische Grenze allein mit einer zu großen Zahl nicht beurteilbarer Symptome überschritten; bei diesen 5 Patienten hat aber nur *einer* zu viele Symptome, die alle fraglich vorhanden waren.

Vergleicht man in der Berliner Stichprobe die 49 Patienten mit den übrigen 1830 bezüglich Alter, Geschlecht und Diagnose (Chiquadratstest), so haben wir beim Alter und Geschlecht keinen Unterschied ( $P > 0,05$ ). Diagnostisch ergibt sich eine Differenz (nur auf dem 5%-Niveau signifikant), indem bei den 49 Patienten die Diagnosen Schizophrenie (ICD 295) und organische Psychosen (ICD 290—294) leicht übervertreten sind.

Ist der kritische Wert nicht erreicht, so schließt das nicht aus, daß die Symptome ohne Aussage vermehrt in *einem* oder zwei Syndrombereichen zu finden sind. Nach unseren bisherigen Erfahrungen in Zürich empfiehlt sich daher ein zweistufiges Vorgehen: Vergleich mit kritischem Wert und Festlegung von kritischen Grenzen bei jeder Syndromskala.

## 9. Diskussion

In unserer Untersuchung haben wir folgende Ergebnisse gefunden:

Die Kategorie „nicht beurteilbar“ wird bedeutend seltener als die Kategorie „fraglich vorhanden“ benutzt. Einige Patienten weisen aber sehr viele nicht beurteilbare Symptome auf.

Die meisten Patienten haben — faßt man beide Kategorien zusammen — im psychischen Befund zwischen 0—2 Symptome ohne Aussage; im somatischen Befund haben nur sehr wenige Patienten ein oder mehr Symptome ohne Aussage.

Bei den beiden Symptomen Gedächtnis- und Merkfähigkeitsstörungen haben relativ viele Patienten (< 3%) das Symptom nicht beurteilbar. Bei den Symptomen aus dem Wahnbereich und der Sinnestäuschungen ist die Entscheidungssicherheit am geringsten (< 9% fraglich vorhanden).

Der psychische Befund ist weniger klar diagnostizierbar als der somatische. Die Diagnosegruppen ICD 297—299 (übrige Psychosen), organische Psychosen (290—294) und Schizophrenie (ICD 295) sind schwieriger einzustufen als die Diagnosegruppen Neurosen (ICD 300) und affektive Psychosen (ICD 296).

Patienten mit stark erhöhter Zahl von Symptomen ohne Aussage haben etwas häufiger die Diagnose Schizophrenie oder organische Psychosen (ICD 290—294) als die übrigen. Der Grund liegt in der *Nicht-Beurteilbarkeit* der Symptome und nicht in der Entscheidungsunsicherheit.

Die Zusammenfassung der Ergebnisse führt zum Schluß, daß die Kategorie Beurteilbarkeit *einzelne Patienten* und die Entscheidungssicherheit *einzelne Symptome* (oder den Rater) charakterisiert. Wir finden kaum Patienten mit sehr vielen fraglich vorhandenen Symptomen, aber einige mit einer großen Zahl von nicht beurteilbaren Symptomen. Abschließend stellt sich die Frage, ob es sinnvoll ist, wegen einer kleinen Gruppe von Patienten zwei formale Merkmale (Beurteilbarkeit und Entscheidungssicherheit) getrennt zu dokumentieren. Eine Zusammenfassung beider Aspekte unter dem Stichwort Aussagemöglichkeit (über Vorhandensein eines Symptoms) hätte verschiedene Vorteile: der Beurteiler müßte nur noch drei Entscheidungsebenen berücksichtigen und nicht mehr vier. Das Dokumentationsblatt würde vereinfacht, indem eine Spalte wegfällt. Gleichzeitig hätte man auch die Realität der statistischen Auswertung berücksichtigt, in der eine getrennte Auswertung beider Aspekte meistens nicht möglich ist (zu kleine Kategorienbesetzung). Wie diese Arbeit zeigt, läßt sich diese Zusammenfassung von Entscheidungssicherheit und Beurteilbarkeit rechtfertigen.

## Literatur

- 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., Poeldinger, W., Schmidlin, P., Schmitt, W., Weis, P.: Das Dokumentationssystem der Arbeitsgemeinschaft für Methodik und Dokumentation in der Psychiatrie (AMP). Arzneimittel-Forsch. (Drug. Res.) **19**, 339—405 (1966)
- Baumann, U.: Diagnostische Differenzierungsfähigkeit von Psychopathologie-Skalen. Arch. Psychiat. Nervenkr. **219**, 89—103 (1974)
- Baumann, U., Rothweiler, R., Scheidegger, P.: Methodische Probleme bei Psychopharmakastudien, unter besonderer Berücksichtigung des AMP-Systems. Arzneimittel-Forsch. (Drug. Res.) **25**, 270—273 (1975)

- Britschgi, A.: Zur Dokumentation der Psychopathologie bei organischen Psychosen. Med. Diss. (in Vorbereitung)
- Cohen, R., Rey, E. R., Sixtl, F.: Die Kombination von „Häufigkeit“ und „Intensität“ im diagnostischen Urteil. *Psychol. Forsch.* **33**, 9–20 (1969)
- v. Cranach, M., Ferenz, H. G.: Systematische Beobachtung. In: C. F. Graumann (Hrsg.), *Sozialpsychologie. Handbuch der Psychologie*. Bd. 7, 1. Halbband. Göttingen: Hogrefe 1969
- Fahrenberg, J.: Die Freiburger Beschwerdeliste FBL. *Z. klin. Psychol.* **4**, 79–100 (1975)
- Fahrenberg, J., Selg, H., Hampel, R.: Das Freiburger Persönlichkeitsinventar FPI, 2. Aufl. Göttingen: Hogrefe 1973
- Hamilton, M.: Development of a rating scale for primary depressive illness. *Brit. J. soc. clin. Psychol.* **6**, 278–296 (1967)
- ICD: Diagnoseschlüssel und Glossar psychiatrischer Krankheiten. Deutsche Übersetzung der internationalen Klassifikation der WHO: ICD, 8. Rev., und des int. Glossars, 2. Aufl. Berlin-Heidelberg-New York: Springer 1971
- Lienert, G. A.: Testaufbau und Testanalyse, 3. Aufl. Weinheim: Beltz 1969
- Lorr, M., Klett, C. J.: Inpatient multidimensional psychiatric scale. Manual. Palo Alto: Consulting Psychologist 1967
- Mombar, W.: Verfahren zur Standardisierung des psychopathologischen Befundes. Teil 1 und 2. *Psychiat. Clin.* **5**, 73–120, 137–157 (1972)
- Murphy, D. L., Beigel, A., Weingartner, H., Bunney, W. E.: The quantitation of manic behavior. In: P. Pichot, S. 203–220. Basel-New York: Karger 1974
- Overall, J. E., Gorham, D. R.: The brief psychiatric rating scale. *Psychol. Rep.* **10**, 799–812 (1962)
- Pichot, P. (Ed.): Psychological measurements in psychopharmacology. Modern problems of pharmacopsychiatry, Vol. 7. Basel-New York: Karger 1974
- Scharfetter, Chr.: Das AMP-System. Manual zur Dokumentation psychiatrischer Befunde, 2. Aufl. Berlin-Heidelberg-New York: Springer 1972
- Scheidegger, P.: AMP-Wandel-Programme. Benutzer-Handbuch. Psychiatr. Universitätsklinik Zürich 1973
- Sixtl, F.: Meßmethoden der Psychologie. Weinheim: Beltz 1967
- Wing, J. K., Cooper, J. E., Sartorius, N.: Measurement and classification of psychiatric symptoms. London: Cambridge University Press 1974

PD Dr. phil. U. Baumann, Dipl.-Psych.  
 Psychiatrische Universitätsklinik  
 Forschungsdirektion  
 Postfach 68  
 CH-8029 Zürich  
 Schweiz