

## **Die diagnostische Reliabilität und Validität des PSE/CATEGO-Systems\***

T. Bronisch, W. Schmid und M. v. Cranach

Max-Planck-Institut für Psychiatrie, Kraepelinstrasse 10, D-8000 München 40, und  
Psychiatrische Klinik der Universität München, D-8000 München, Bundesrepublik Deutschland

### **Diagnostic Reliability and Validity of the PSE/CATEGO-System**

**Summary.** An unselected series of 100 psychiatric inpatient admissions were interviewed at admission by a psychiatrist using the German version of the PSE (Present State Examination), with a second psychiatrist as an observer. The diagnostic agreement between the two project psychiatrists was considerably higher than the agreement between clinicians and CATEGO, a computerised diagnostic system based on PSE data. The disagreements are discussed.

**Key words:** Clinical diagnosis – Computer diagnosis – Diagnostic agreement

**Zusammenfassung.** Mit der deutschen Version des PSE (Present State Examination) wurden 100 unausgewählte, neuaufgenommene psychiatrische Klinikpatienten von einem Psychiater interviewt, mit einem zweiten Psychiater als Beobachter. Die Übereinstimmung der klinischen Diagnosen der beiden Psychiater lag wesentlich höher als die Übereinstimmung zwischen klinischen Diagnosen einerseits und den CATEGO-Diagnosen andererseits (CATEGO ist ein computerisiertes System, das auf der Information von PSE beruht). Die mangelnde Übereinstimmung wird diskutiert.

**Schlüsselwörter:** Klinische Diagnose – Computer-Diagnose – Diagnostische Übereinstimmung

### **Einführung**

Das Problem der geringen diagnostischen Übereinstimmung zwischen klinischen Untersuchern ist in der Psychiatrie seit langem bekannt. Nach den Untersuchungen von Kreitmann (1961) und Beck (1962) sind eine Vielzahl von Studien veröffentlicht worden, die sich mit dem Ausmaß und den Ursachen von Diskrepanzen

---

\* Diese Arbeit erfolgt im Rahmen einer Untersuchung von Schmid W, Bronisch T, Zerssen D v:  
A Comparative Study between PSE/CATEGO and the new Algorithmic Diagnostic System  
DiaSiKa

zwischen psychiatrischen Diagnosen beschäftigt haben. Zubin (1967) und Kendell (1975) haben Literaturübersichten zu diesem Thema veröffentlicht. Mangelnde Interrater-Reliabilität hängt einerseits zusammen mit

Inkongruenzen zwischen verschiedenen diagnostischen Systemen,  
Inkongruenzen innerhalb eines diagnostischen Systems und  
Mangel an eindeutig operationalisierten Zuordnungsregeln,  
andererseits ist sie bedingt durch den Prozeß des Diagnostizierens, u.a. durch Explorationsstil, Symptomwahrnehmung, Symptomgewichtung und ähnliche Variablen, die wiederum von Ausbildung, Ausrichtung und Persönlichkeit des Psychiaters abhängig sind.

Aus diesem Grunde sind mehrfach Versuche unternommen worden, Abschnitte des diagnostischen Prozesses zu standardisieren, um auf diese Weise reliable Diagnosen zu erhalten. Dabei war es Ziel dieser Versuche, „objektive“, d.h. einheitlich zustande gekommene Diagnosen zu erhalten, um für Forschungszwecke homogene und vergleichbare Patientengruppen zu erreichen. Ziel dieser Versuche war es nicht, im klinischen Alltag den Arzt als Diagnostiker durch standardisierte Verfahren zu ersetzen. Spitzer und Endicott beschrieben 1969/1974 ein derartiges Computerprogramm, Diagno (Fassung I, II und III), das — den Entscheidungsprozeß des Klinikers nachahmend — auf der Basis von standardisiert erhobenen Symptomen und anamnestischen Daten Diagnosen stellte, die relativ gut mit klinischen Diagnosen übereinstimmten.

Wing et al. (1974) entwickelten ein Programm, CATEGO, das auf der Grundlage von Symptomen (erhoben mit Hilfe eines standardisierten Interviews, Present State Examination, deutsche Fassung 1978) und Daten zur Krankheitsätiologie (Aetiology Schedule) Diagnosen auf der Grundlage des ICD-Schlüssels stellt. Dieses System hat sich in der Zwischenzeit in zahlreichen Untersuchungen besonders im europäischen Raum etabliert und auch in internationalen Verbundstudien (WHO) Anwendung gefunden (z.B. The International Pilot Study of Schizophrenia 1973).

Die hier vorgelegte Untersuchung hat zwei Ziele:

1. Vergleich der diagnostischen Übereinstimmung zwischen zwei gleich ausgebildeten Psychiatern bei standardisierter Befunderhebung mit Hilfe des PSE und die Übereinstimmung zwischen diesen klinischen Diagnosen und CATEGO-Diagnosen.

2. Überprüfung der Validität der Projekt-Diagnosen und Eindrucksdiagnosen durch Vergleich mit den Abschlußdiagnosen der Krankengeschichten.

## Methodik der Untersuchung

Nach vorangegangenem Training in der Interviewtechnik mit dem Present State Examination (durch v.C.) führten zwei Klinikärzte (B. und S.) 100 PSE-Interviews durch. Es wurden 45 konsekutiv aufgenommene Patienten der Psychiatrischen Abteilung des Klinischen Institutes des Max-Planck-Institutes für Psychiatrie in München sowie 55 Patienten der Aufnahmestationen des Bezirkskrankenhauses Haar untersucht<sup>1</sup>.

1 Wir danken Herrn Direktor Dr. Schulz, den Bereichsdirektoren, Herrn Dr. Achner und Herrn Dr. Öchsner, sowie den Ärzten der Aufnahmestationen für ihre bereitwillige Unterstützung dieser Untersuchung

Hinsichtlich der Einweisungsdiagnosen erfolgte keine Selektion; es wurden lediglich die folgenden im CATEGO-Programm nicht berücksichtigten Diagnosen der International Classification of Diseases, 8. Rev., ausgeschlossen: ICD-Nr. 302 (sexuelle Verhaltensabweichungen), ICD-Nr. 305 (psychosomatische Störungen), ICD-Nr. 306 (besondere Symptome, die nicht anderweitig klassifiziert werden können), ICD-Nr. 307 (vorübergehende kurzfristige psychische Auffälligkeiten, die mit situativen Belastungen im Zusammenhang stehen), ICD-Nr. 308 (Verhaltensstörungen im Kindesalter) sowie die ICD-Nummern 310–315 (Oligophrenien). Die Untersuchung erfolgte in allen Fällen innerhalb der ersten Woche nach der stationären Aufnahme ohne Informationsaustausch mit dem behandelnden Arzt. Das Interview wurde von einem Kliniker geführt, der andere Kliniker konnte zwischendurch und am Ende des Interviews Zusatzfragen stellen; in der Gesprächsführung wechselten sich die beiden Kliniker bei jedem neuen Patienten ab.

Während des Interviews von  $\frac{1}{2}$  h bis zu  $2\frac{1}{2}$  h Dauer füllten die beiden Untersucher unabhängig voneinander den PSE-Fragebogen aus, unmittelbar nach dem Gespräch die Aetiology Schedule; dabei wurde auch die klinische Eindrucksdiagnose festgehalten. Bei etwa 10% der Patienten konnten wegen Mutismus oder nichtkommunikativer Sprache nur Verhalten und Affekt beurteilt werden. PSE und Aetiology Schedule wurden mit Hilfe des CATEGO-Programms ausgewertet<sup>2</sup>. Die Übereinstimmung wurde einerseits durch Kappa ausgedrückt (Cohen 1960), einen Übereinstimmungsindex, der die beobachtete Übereinstimmung um die zu erwartende zufällige Übereinstimmung korrigiert, andererseits in Prozent, bezogen auf die Häufigkeit, mit der die Kategorie mindestens von einem der beiden Diagnostiker gewählt wurde (Schmid et al. 1974). Von Kappa ist dann noch der einseitige Signifikanzwert angegeben.

## Ergebnisse

Zur Berechnung der Übereinstimmung zwischen beiden Klinikern und der Übereinstimmung zwischen klinischer Diagnose und Computer-Diagnose wurden zunächst die einzelnen Diagnosen in Gruppen zusammengefaßt, die in zwei aufeinanderfolgenden Ansätzen immer weiter aufgefächert wurden. Dieses Vorgehen hatte sich schon in früheren Untersuchungen zur Prüfung der diagnostischen Übereinstimmung verschiedener Untersucher untereinander — sowie mit der klinischen Abschlußdiagnose und/oder standardisierten Diagnoseverfahren bewährt (Möller und v. Zerssen 1980; Schmid et al. 1974; Castell et al. 1974). In der ersten Einteilung wird zunächst unterschieden zwischen organischen Syndromen, endogenen Psychosen und Neurosen bzw. Persönlichkeitsstörungen. Endogene und neurotische Depressionen sind dabei ausgeklammert und zu einer vierten Gruppe „Depression“ zusammengefaßt. Dies erscheint sinnvoll, weil bei der nosologischen Differenzierung depressiver Syndrome bekanntlich eine besonders geringe Übereinstimmung zwischen verschiedenen Klinikern besteht.

In der vergleichenden Studie zwischen PSE/CATEGO und dem neuen algorithmischen Diagnose-System DiaSiKa (v. Zerssen in Vorbereitung; Schmid et al. in Vorbereitung) wird in der ersten Einteilung zunächst unterschieden zwischen organisch begründbaren und endogenen Psychosen und Nicht-Psychosen (also Neurosen, Persönlichkeitsstörungen, nichtpsychotischen psychischen Störungen bei körperlichen Erkrankungen). Endogene und neurotische Depressionen sind dabei ebenfalls in einer vierten Gruppe „Depressionen“ zusammengefaßt.

In der zweiten Einteilung werden sowohl bei Schmid et al. als auch bei uns die Alkoholpsychosen von sonstigen organischen Psychosen abgetrennt und die

2 Hierfür möchten wir Herrn Dr. Wegscheider, Psychiatrische Universitätsklinik, und Herrn Dr. Klug, Zentralinstitut für seelische Gesundheit in Mannheim, danken

endogenen Psychosen in Schizophrenien, schizoaffektive Psychosen und Manien unterteilt. Außerdem werden die nichtpsychotischen psychischen Störungen bei körperlichen Erkrankungen, die bei Schmid et al. in der Gruppe der „Nicht-Psychosen“ und bei uns in der Gruppe der organischen Syndrome auftauchten, als eigene Diagnose-Gruppe behandelt. Weiterhin werden paranoide Syndrome, reaktive Psychosen, Neurosen und Persönlichkeitsstörungen gesondert aufgeführt.

Bei unserer Untersuchung an 100 unselektierten Patienten war bei den einzelnen Diagnosen die Fallzahl zu gering, um auf der Ebene vierstelliger ICD-Nummern statistisch relevante Aussagen über die diagnostische Übereinstimmung machen zu können. Bei unserer Einteilung I (Tabelle 1) sind dagegen, wie aus der Tabelle ersichtlich, ausreichend Fälle in jeder diagnostischen Kategorie. Wie man sieht, macht die Gruppe der Schizophrenien und paranoiden Syndrome allein über 50% der Fälle aus.

Die Gesamtübereinstimmung in Einteilung I beträgt 95%;  $\kappa = 0,92$  ( $P < 0,01$ ). Die Kontingenztafel der Einteilung I (Tabelle 1) zeigt, in welche Diagnosegruppen die nicht übereinstimmend diagnostizierten Fälle der beiden Untersucher gelangt sind. Bei den organischen Psychosen ergab sich eine vollständige Übereinstimmung. Ein Patient konnte von beiden Klinikern in keine diagnostische Kategorie eingeordnet werden. In Einteilung II (Tabelle 2) sinkt  $\kappa = 0,88$  ( $P < 0,01$ ) wegen der erheblichen Zunahme der Fehlermöglichkeiten erwartungsgemäß ab. Die prozentuale Übereinstimmung beträgt 91%.

Der nächste Schritt der Untersuchung bezieht sich auf die Übereinstimmung zwischen diesen klinischen Diagnosen und CATEGO-Diagnosen. Zehnmal konnte keine Diagnose gestellt werden. Dadurch sinkt  $\kappa$  erheblich ab (0,56 bzw. 0,60). Vernachlässigt man die zehn nichtdiagnostizierbaren Fälle, so steigt  $\kappa$  wieder auf 0,63 ( $P < 0,01$ ) bzw. 0,70 ( $P < 0,01$ ) an. Die Kontingenztafeln der Einteilung I (Tabellen 3 und 4) zeigen die Übereinstimmung der beiden Kliniker mit den CATEGO-Diagnosen. Auffallend ist dabei, daß von den 10 durch CATEGO nicht diagnostizierbaren Fällen 5 bzw. 7 bei den Klinikern als organische Syndrome, 3 bzw. 2 als endogene Psychosen diagnostiziert wurden. Bei genauerer Untersuchung dieser Fälle zeigte es sich, daß bei diesen Patienten wegen Mutismus oder nichtkommunikativer Sprache lediglich Affekt und Verhalten beurteilt werden konnten. Bei der klinischen Diagnose einer endogenen Psychose geht ein geringer Teil in die CATEGO-Diagnose einer Depression ein. Bei der klinischen Diagnose einer Neurose oder Persönlichkeitsstörung wird von CATEGO häufig die Diagnose einer Depression gestellt. In der Kontingenztafel der Tabelle 5 werden die CATEGO-Diagnosen beider Kliniker in Einteilung I verglichen. Dabei ergibt sich eine hohe Gesamtübereinstimmung von  $\kappa = 0,85$  ( $P < 0,01$ ) oder 90%, die nur wenig unter der Übereinstimmung der beiden Kliniker-Diagnosen ( $\kappa = 0,92$  ( $P < 0,01$ ) oder 95%) liegt (siehe auch Tabelle 1).

Tabelle 6 zeigt die diagnostische Übereinstimmung zwischen Klinikern und CATEGO-Programm in Einteilung II. Auffallend ist dabei, daß CATEGO häufiger Depressionen und Manien, dafür seltener schizophrene Störungen diagnostiziert.  $\kappa$  sinkt erwartungsgemäß auf 0,53 bzw. 0,50 ab. Vernachlässigt man wiederum die zehn nichtdiagnostizierbaren Fälle, so ändert sich an  $\kappa$  nichts (0,52 bzw. 0,50).

Tabelle 1. Kontingenztabelle: Eindrucksdiagnosen beider Kliniker

Einteilung I		Einteilung II					
ICD-8 Kliniker 1		ICD-8 Kliniker 2					
		organ. Syndrome	endog. Psychosen	Depressionen	Neurosen u. Persönlich- keitsstörungen	nichtdiagnosti- zierbare Fälle	Gesamt- zahl
290-294	organ. Syndrome	16 100%	0	0	0	0	16
295/297 298/296.1	endog. Psychosen	0	51 96%	1	1	0	53
296/300.4 (außer 296.1)	Depressionen	0	0	14 78%	0	0	14
300 (außer 300.4) 301/309	Neurosen u. Persön- lichkeitsstörungen	0	0	3	13 76%	0	16
	Nichtdiagnostizierbare Fälle	0	0	0	0	1 100%	1
Gesamtzahl		16	51	18	14	1	
Gesamtübereinstimmung:		95% K = 0,92 (P < 0,01)					

**Tabelle 2.** Diagnostische Übereinstimmung zwischen Klinikern. Es ist jeweils links und rechts angegeben, wie oft jeder die betreffende diagnostische Gruppe wählte; dazwischen in Klammern die Anzahl der Fälle, in denen diese Kategorie übereinstimmend von beiden gewählt wurde. Daneben steht die Übereinstimmung in Prozent, bezogen auf die Häufigkeit, mit der die Kategorie mindestens von einem der beiden Diagnostiker gewählt wurde (s. Schmid et al. 1974). Zusätzlich ist die Übereinstimmung durch  $\kappa$  (Kappa) ausgedrückt (Cohen 1960), einen Übereinstimmungsindex, der die beobachtete Übereinstimmung um die zu erwartende zufällige Übereinstimmung korrigiert

Einteilung II	Kliniker 1	—	Kliniker 2	
Alkoholpsychosen	0	(0)	0	—
Andere organische Psychosen	16	(16)	16	100%
Schizophrenien	45	(43)	44	93%
Schizoaffective Psychosen	1	(1)	1	100%
Manien	2	(2)	3	67%
Depressionen	14	(14)	18	78%
Neurosen	8	(7)	10	64%
Persönlichkeitsstörungen	6	(2)	2	33%
Nichtpsychotische psychische Störungen bei körperlichen Erkrankungen	2	(2)	2	100%
Paranoide Syndrome und reaktive Psychosen	5	(3)	3	60%
Nichtdiagnostizierbare Fälle	1	(1)	1	100%
Gesamt-Übereinstimmung	91%			
	$\kappa = 0,88$ ( $P < 0,01$ )			

Der letzte Teil der Untersuchung betrifft die diagnostische Übereinstimmung der Projektdiagnosen mit den Abschlußdiagnosen der Krankengeschichten. Aus den Krankengeschichten des BKH Haar wurden dabei von einem unabhängigen Psychiater der Psychiatrischen Abteilung des Max-Planck-Institutes für Psychiatrie die Diagnosen nach der ICD, 8. Rev. gestellt<sup>3</sup>. Der Zeitraum zwischen Eindrucks- bzw. CATEGO-Diagnosen einerseits und Abschlußdiagnosen andererseits lag zwischen einigen Wochen und mehreren Monaten und war unabhängig von den Diagnosen. Die Art der Behandlung in diesem Zeitraum richtete sich nach den Diagnose-Gruppen: organische Psychosen wurden — soweit möglich — entsprechend ihrer Ätiologie behandelt, ansonsten symptomatisch mit Psychopharmaka. Die endogenen Psychosen wurden mit Neuroleptika behandelt. Die Gruppe der Depressionen wurde entweder mit Antidepressiva oder psychotherapeutisch behandelt, je nachdem, ob die Depression sich mehr in das Bild einer endogenen oder einer neurotischen Depression einfügen ließ. Patienten mit den Diagnosen „Neurose“ und „Persönlichkeitsstörung“ erhielten eine Psychotherapie.

Tabelle 7a zeigt die Übereinstimmung zwischen Abschlußdiagnosen einerseits und Eindrucksdiagnosen bzw. CATEGO-Diagnosen andererseits. Dabei wurden die zehn nichtdiagnostizierbaren Fälle vernachlässigt. Die Übereinstimmung zwischen Abschlußdiagnosen und Eindrucksdiagnosen ist höher als zwischen Abschlußdiagnosen und CATEGO-Diagnosen.

<sup>3</sup> Wir danken Herrn Dr. Vogel für die Übernahme dieser Aufgabe

Tabelle 3. Kontingenztabel: Eindrucksdiagnose (Kliniker I) – CATEGO-Diagnose

Eindrucksdiagnose	Einteilung I					
	CATEGO-Diagnose					
	organ. Syndrome	endog. Psychosen	Depressionen	Neurosen u. Persönlich- keitsstörungen	nichtdiagnosti- zierbare Fälle	Gesamt- zahl
Organische Syndrome	5 29%	0	0	6	5	16
Endogene Psychosen	0	46 81%	3	1	3	53
Depressionen	0	3	11 48%	0	0	14
Neurosen u. Persönlichkeitsstörungen	1	1	6	7 30%	1	16
Nichtdiagnostizierbare Fälle	0	0	0	0	1 10%	1
Gesamtzahl	6	50	20	14	10	100

ICD-8-Klassifikation siehe Tabelle I

Tabelle 4. Kontingenztafel: Eindrucksdiagnose (Kliniker 2) – CATEGO-Diagnose

Eindrucksdiagnose	Einteilung I					
	CATEGO-Diagnose					
	organ. Syndrome	endog. Psychosen	Depressionen	Neurosen u. Persönlich- keitsstörungen	nichtdiagnosti- zierbare Fälle	Gesamt- zahl
Organische Syndrome	4 24%	0	0	5	7	16
Endogene Psychosen	0	47 82%	2	0	2	51
Depressionen	0	3	15 65%	0	0	18
Neurosen u. Persönlichkeitsstörungen	1	2	3	7 37%	1	14
Nichtdiagnostizierbare Fälle	0	1	0	0	0 0%	1
Gesamtzahl	5	53	20	12	10	100

ICD-8-Klassifikation siehe Tabelle 1



Tabelle 5. Kontingenztafel: CATEGO-Diagnose (Kliniker 1) – CATEGO-Diagnose (Kliniker 2)

CATEGO-Diagnose Kliniker 1	Einteilung 1					
	CATEGO-Diagnose Kliniker 2					
	organ. Syndrome	endog. Psychosen	Depressionen	Neurosen u. Persönlich- keitsstörungen	nichtdiagnosti- zierbare Fälle	Gesamt- zahl
Organische Syndrome	5 83%	0	0	0	1	6
Endogene Psychosen	0	49 98%	0	0	1	50
Depressionen	0	1	18 90%	1	0	20
Neurosen u. Persönlichkeitsstörungen	0	0	2	11 79%	1	14
Nichtdiagnostizierbare Fälle	0	3	0	0	7 70%	10
Gesamtzahl	5	53	20	12	10	100

ICD-Klassifikation siehe Tabelle 1

Tabelle 6. Diagnostische Übereinstimmung zwischen Klinikern und CATEGO-Programm

	Einteilung II					
	Kliniker 1 – CATEGO-Programm			Kliniker 2 – CATEGO-Programm		
Alkoholsychosen	0	(0)	1	0%	0	(0) 1 0%
Andere organische Psychosen	16	(5)	5	45%	16	(4) 4 44%
Schizophrenien	45	(34)	37	76%	44	(31) 35 67%
Schizoaffektive Psychosen	1	(0)	0	0%	1	(0) 0 0%
Manien	2	(2)	8	25%	3	(3) 11 30%
Depressionen	14	(11)	20	48%	18	(15) 20 66%
Neurosen	8	(2)	2	25%	10	(3) 3 30%
Persönlichkeitsstörungen	6	(5)	6	71%	2	(2) 4 50%
Nichtpsychotische psychische Störungen bei körperlichen Erkrankungen	2	(0)	6	0%	2	(0) 5 0%
Paranoide Syndrome und reaktive Psychosen	5	(3)	5	43%	3	(2) 7 25%
Nichtdiagnostizierbare Fälle	1	(1)	10	10%	1	(0) 10 0%
Gesamt-Übereinstimmung	69%				60%	
	$K = 0,53 \ (P < 0,01)$			$K = 0,50 \ (P < 0,01)$		

ICD-Klassifikation siehe Tabelle 1

**Tabelle 7a.** Übereinstimmung zwischen Abschlußdiagnosen einerseits und Eindrucksdiagnosen bzw. CATEGO-Diagnosen andererseits, ausgedrückt in Kappa

	Einteilung I ohne nichtdiagnostizierbare Fälle ( $N=90$ )	
	Abschlußdiagnose (Kliniker 1)	Abschlußdiagnose (Kliniker 2)
Eindrucksdiagnose	0,64 ( $P<0,01$ )	0,63 ( $P<0,01$ )
CATEGO-Diagnose	0,44 ( $P<0,01$ )	0,45 ( $P<0,01$ )

**Tabelle 7b.** Übereinstimmung zwischen CATEGO-Diagnosen einerseits und Eindrucksdiagnosen bzw. Abschlußdiagnosen andererseits, ausgedrückt in Kappa

	Einteilung I ohne nichtdiagnostizierbare Fälle ( $N=90$ )	
	CATEGO-Diagnose (Kliniker 1)	CATEGO-Diagnose (Kliniker 2)
Eindrucksdiagnose	0,63 ( $P<0,01$ )	0,70 ( $P<0,01$ )
Abschlußdiagnose	0,44 ( $P<0,01$ )	0,45 ( $P<0,01$ )

Tabelle 7b zeigt die Übereinstimmung zwischen CATEGO-Diagnosen einerseits und Eindrucks- bzw. Abschlußdiagnosen andererseits: CATEGO- und Eindrucksdiagnosen haben eine höhere Übereinstimmung als CATEGO- und Abschlußdiagnosen.

## Diskussion

Die Ergebnisse zeigen zunächst eine eindrucksvolle Übereinstimmung in den diagnostischen Urteilen („Eindrucksdiagnosen“) der beiden an der Erhebung teilnehmenden Ärzte. Zwei Gründe sind dafür verantwortlich. Als erstes ist die gemeinsame Ausbildung der beiden Ärzte zu nennen. Zugehörigkeit zu verschiedenen Psychiatrie-Schulen ist ein wesentlicher Faktor für die mangelnde Reliabilität psychiatrischer Diagnostik, wie beispielsweise Shepherd et al. (1968) zeigen konnten. Hinzu kommt, daß bei dieser Untersuchung ein zweiter ursächlicher Faktor für die Unzuverlässigkeit des diagnostischen Prozesses in der Psychiatrie ausgeschaltet wurde, nämlich die uneinheitliche Befunderhebung. Wie viele Untersuchungen zeigen (beispielsweise Busch et al. 1980), ist die Exploration und die Übersetzung des Erlebens und Verhaltens des Patienten in Symptome ein Schritt im diagnostischen Prozeß, der zu seiner Unzuverlässigkeit beiträgt. Durch die Verwendung eines standardisierten Interviews ist somit eine weitere Quelle für mangelnde Übereinstimmung weitgehend ausgeschaltet worden. Die Ergebnisse zeigen, daß die systematische Ausschaltung dieser bekannten, den diagnostischen Prozeß beeinflussenden Faktoren (Zugehörigkeit zu verschiedenen psychiatrischen Schulen, unterschiedlicher Explorationsstil, unterschiedliche Symptomdefinitionen) zu einer sehr hohen, praktisch völligen diagnostischen Übereinstimmung

führt. Während die meisten Untersuchungen zur Reliabilität der psychiatrischen Diagnostik eine Beschreibung des Ausmaßes und der Ursachen der mangelnden diagnostischen Übereinstimmung zum Ziel haben, konnte hier gezeigt werden, daß eine gezielte Ausschaltung dieser Faktoren zu einer hohen Reliabilität führt. Wesentlich geringer ist die Übereinstimmung zwischen Eindrucksdiagnosen der Kliniker und den CATEGO-Diagnosen. In einer mit der deutschen Fassung des PSE durchgeführten Untersuchung konnten Wing et al. (1977) zeigen, daß bei einem Vergleich der klinischen Eindrucksdiagnosen und CATEGO-Diagnosen, der *nicht* auf der Ebene der Einzelfälle erfolgte, eine sehr hohe Übereinstimmung besteht. Dazu ist jedoch zu bemerken, daß bei der genannten Untersuchung die klinischen Diagnosen von den Psychiatern gestellt worden waren, die teilweise in England ausgebildet worden waren, das PSE ins Deutsche übertragen und eine genaue Kenntnis der CATEGO-Regeln hatten, mit anderen Worten: sie waren vertraut mit den diagnostischen Grundzügen und „Vorurteilen“ des CATEGO-Systems, wodurch ihre klinischen Diagnosen sicher beeinflußt worden waren. Das war bei der hiesigen Untersuchung nicht der Fall. Unterschiede zwischen den Klinikern und CATEGO sind demnach hier auf unterschiedliche „diagnostische Schulen“ zurückzuführen. Eine nähere Betrachtung der Unterschiede zeigt, daß CATEGO häufiger Depressionen und Manien, dafür seltener schizophrene Störungen diagnostiziert. Damit wird ein bekannter Unterschied zwischen englischer und deutscher Diagnostik sichtbar, der auch anderenorts bereits beschrieben wurde (v. Cranach und Strauss 1977).

Dagegen kann CATEGO offensichtlich Depressionen von nichtdepressiven Neurosen nicht ausreichend trennen. Was die zehn nichtdiagnostizierbaren Fälle betrifft, so handelt es sich um Patienten mit symptomarmen Schizophrenien und organischen Psychosen, die wegen Mutismus oder nichtkommunikativer Sprache nur nach Verhalten und Affekt beurteilt werden konnten. Das PSE-Interview kann offensichtlich diese Symptomatik nicht ausreichend erfassen.

Der dritte Teil der Arbeit beschäftigt sich mit den Übereinstimmungen von Eindrucks-, Projekt- und Abschlußdiagnosen: die Übereinstimmung zwischen CATEGO-Diagnosen und Eindrucksdiagnosen ist höher als zwischen CATEGO-Diagnosen und Abschlußdiagnosen. PSE enthält keine Items zum Verlauf der Erkrankungen. Andererseits gehen in die Abschlußdiagnosen ja auch Fremdinformation, Verlaufsbeobachtung und Ansprechbarkeit auf verschiedene Therapieformen ein. Die Übereinstimmung zwischen Abschlußdiagnosen und Eindrucksdiagnosen ist höher als zwischen Abschlußdiagnosen und CATEGO-Diagnosen: hier spiegeln sich die im letzten Abschnitt erwähnten Schwächen im CATEGO-System wider, die schon in der Übereinstimmung zwischen Eindrucksdiagnosen und CATEGO-Diagnosen zum Ausdruck kamen.

Nimmt man die Abschlußdiagnosen als Kriteriumsdiagnosen, so überrascht es daher nicht, daß die Validität sowohl der Eindrucksdiagnosen als auch der Projektdiagnosen nur einen begrenzten Wert hat.

## Literatur

- Beck AT (1962) Reliability of psychiatric diagnoses: a critique of systematic studies. *Am J Psychiatr* 119:210–216

- Busch H, v. Cranach M, Gulbinat W, Renfordt E, Tegeler J (1980) Reliability of the AMDP-system. *Acta Psychiatr Scand* 62:382-392
- Castell R, Busch H, Hoffmann J, Mittelsten-Scheid H, Waldmann H (1974) Übereinstimmung zwischen konventioneller und Computer-Diagnostik beim ambulanten psychiatrischen Patienten. *Arch Psychiatr Nervenkr* 222:1-11
- Cohen J (1960) A coefficient of agreement of nominal scales. *Educational and Psychological Measurement* 20:37-46
- Cranach M v, Strauß A (1977) Zur Weiterentwicklung des AMP-Systems. *Pharmakopsychiatr* 10:254-261
- Diagnoseschlüssel und Glossar psychiatrischer Krankheiten (1975) In: Degkwitz R, Helmchen H, Kockott G, Mombour W (Hrsg) Deutsche Ausgabe der internationalen Klassifikation der WHO: ICD, 8. Revision und des internationalen Glossars. Springer, Berlin Heidelberg New York
- Kendell RE (1975) The role of diagnosis in psychiatry. Blackwell Scientific Publications, Osney Mead, Oxford. Deutsch: Die Diagnose in der Psychiatrie. Enke, Stuttgart (1978)
- Kreitmann N (1961) The reliability of psychiatric diagnosis. *J Ment Sci* 107:876-886
- PSE Present State Examination (1978) (Wing JK, Cooper E, Sartorius N; deutsche Bearbeitung M. v. Cranach). Beltz Test, Weinheim
- Möller HJ, v Zerssen D (1980) Probleme und Verbesserungsmöglichkeiten der psychiatrischen Diagnostik. In: Biefang S (Hrsg) Evaluationsforschung in der Psychiatrie: Fragestellungen und Methoden. Enke, Stuttgart, S 167-207
- Schmid W, Bronisch T, v Zerssen D (in Vorbereitung) A comparative study between PSE/CATEGO and the new algorithmic diagnostic system DiaSiKa
- Schmid W, Castell R, Mombour W, Mittelsten-Scheid D, v Zerssen D (1974) Die diagnostische Übereinstimmung zwischen Klinikern und DIAL-Programm. *Arch Psychiatr Nervenkr* 218:339-351
- Shepherd M, Brooke EM, Cooper JE, Lin T (1968) An experimental approach to psychiatric diagnosis. *Acta Psychiatr Scand (Suppl)* 201:ad vol 44
- Spitzer RL, Endicott J (1969) Diagno II: Further developments in a computer program for psychiatric diagnosis. *Am J Psychiatr* 125:12-20
- Spitzer RL, Endicott E, Cohen J, Fleiss L (1974) Constraints on the validity of computer diagnosis. *Arch Gen Psychiatr* 31:197-203
- World Health Organization (WHO) (1973) The international pilot study of schizophrenia, IPSS. WHO, Genf
- Wing JK, Coper JE, Sartorius N (1974) Description and classification of psychiatric symptoms. Cambridge University Press, Cambridge
- Wing JK, Nixon J, v Cranach M, Strauß A (1977) Further developments of the "Present State Examination" and CATEGO-system. *Arch Psychiatr Nervenkr* 224:151-160
- Zerssen D v, unter Mitarbeit von Barthelmes H, et al. (in Vorbereitung) Die Diagnostische Sichtlochkartei (DiaSiKa) — ein einfaches Instrument zur Standardisierung der psychiatrischen Diagnostik. Konzeption und Entwicklung
- Zubin J (1967) Classification of the behavior disorders. *Am Rev Psychol* 18:373-406

Eingegangen am 17. September 1981