

Eine Kontrolluntersuchung zur Beschwerdenliste

U. BAUMANN

Psychiatrische Universitätsklinik Zürich, Forschungsdirektion
(Prof. Dr. med. J. Angst)

Eingegangen am 10. Juni 1972

A Controlled Test for the Evaluation of the von Zerssen Symptom List

Summary. Von Zerssen's symptom list (Beschwerdenliste) was analysed with regard to form and content according to the criteria of classical test theory. Data were obtained from a sampling of 427 nineteen year-olds, inducted into the Swiss Army. Factor analysis of the list's 64 items showed 15% of the total variance was attributable to the first factor. The test satisfied the criteria of classical test theory. The content of Zerssen's list more closely covers the FPI scale for psychosomatic complaints (Nervosität) than the FPI scale for neurosis. The results differentiated subjects with or without psychosomatic complaints. The two parallel abbreviated forms of the list proved to be adequate in establishing subjective symptoms.

Key words: Psychological Test — Symptom List — Psychosomatics.

Zusammenfassung. Die Beschwerdenliste von Zerssens (64 Items) wurde formal und inhaltlich nach testmethodischen Gesichtspunkten analysiert. Dazu wurden die Daten einer Stichprobe benutzt, die aus 427 Stellungspflichtigen der Schweizer Armee (Alter: 19 Jahre) besteht. Eine Faktorenanalyse der 64 Items zeigt, daß der erste Faktor 15% der Gesamtvarianz erklärt. Eine Überprüfung der Testform nach den Kriterien der klassischen Testtheorie führt zu befriedigenden Resultaten. Inhaltlich läßt sich die Beschwerdenliste eher mit der FPI-Skala Nervosität und weniger mit der FPI-Skala Neurotizismus in Verbindung bringen. In der Beschwerdenliste unterscheiden sich Probanden mit psychosomatischen Beschwerden und ohne psychosomatische Beschwerden. Die beiden parallelisierten Kurzformen erweisen sich als geeignete Verfahren um den subjektiven Befund zu erfassen.

Schlüsselwörter: Psychologischer Test — Beschwerdenliste — Psychosomatik.

1. Einleitung

Von Zerssen entwickelte anhand von Lehrbüchern und Krankengeschichten einen Persönlichkeitsfragebogen, der 64 Beschwerden¹ zu internmedizinischen Erkrankungen und vegetativen Dystonien umfaßt (von Zerssen, 1971). Der Test wurde besonders im Hinblick auf klinische Fragestellungen entwickelt; nach den Angaben des Testautors ist die Beschwerdenliste bei Patienten leichter anzuwenden als die herkömm-

¹ Item 65 ist nur für Frauen zutreffend und wird im Gesamtscore nicht berücksichtigt.

lichen Persönlichkeitsfragebogen mit Neurotizismusskalen (z. B. MPI, MMQ). Sie ist besonders geeignet für die Erfassung des subjektiven Befundes und der Änderungen im Krankheitsverlauf bei Patienten.

Der Test umfaßt 64 Items, die mit ja/nein beantwortet werden.

Beispiel: Ich leide unter folgenden Beschwerden:

- | | |
|-----------------------|----------|
| 2. Müdigkeit | ja/nein |
| 33. Starkes Schwitzen | ja/nein |
| 54. Schwindelgefühl | ja/nein. |

Daneben liegen 2 parallelisierte Kurzformen (genannt: B-L, B-L') von je 24 Items vor, die mit „ja/nein“ oder den Kategorien „stark-mäßig-kaum-gar nicht“ beantwortet werden.

Nach von Kerekjarto et al. (1972) korreliert die Beschwerdenliste mit der MPI-Neurotizismusskala mit $r = 0,68$ ($N = 98$), mit der MMQ-Neurotizismusskala mit $r = 0,72$ ($N = 126$).

2. Fragestellung

An einer größeren Stichprobe soll die formale Seite der Beschwerdenliste überprüft werden (Trennschärfe, Faktorenstruktur, Konsistenz). In der Publikation zur Beschwerdenliste (von Zerssen, 1971) fehlen exakte Angaben zur Trennschärfe und Faktorenanalyse, so daß mit unserer Arbeit ergänzende Daten geliefert werden sollen. In der Arbeit von Kerekjarto et al. (1972) liegen zwar detaillierte statistische Angaben vor, doch wurde eine modifizierte Form der Beschwerdenliste (67 Items + 1 zusätzliches Item) verwandt.

In einem weiteren Schritt wird abgeklärt, wie die Beschwerdenliste mit anderen Persönlichkeitsdimensionen zusammenhängt. Dazu wird das Freiburger Persönlichkeitsinventar FPI (Fahrenberg, u. Selg 1970) benutzt. Es wurde die ursprüngliche Version (259 Items, Version ALNEV) verwandt; die Auswertung erfolgte aber nach der endgültigen Version (212 Items).

Der FPI ist ein faktorenanalytisch konstruierter, mehrdimensionaler Fragebogen, der 9 Faktoren und 3 übergeordnete Skalen umfaßt.

Die 9 Skalen lauten: Nervosität, Aggressivität, Depressivität, Erregbarkeit, Geselligkeit, Gelassenheit, Dominanzstreben, Gehemmtheit, Offenheit; Extraversion, Emotionale Labilität, Maskulinität.

Die Berechnungen erfolgten im Rechenzentrum der Universität Zürich (IBM 360/50) unter Benützung der Programmbibliothek PSYCHLIB (U. Baumann; DRZ-Programme MIS 1, PAMV, Programm PRINAX von Messner, Hodapp, Freiburg i. Br.; Programm POWER2 von Zimmermann, Freiburg i. Br.).

3. Stichprobe

Im Rahmen einer größeren Untersuchung im Kanton Zürich, die 1969 bei Stellungspflichtigen der Schweizer Armee durchgeführt wurde, fanden u.a. die Beschwerdenliste und das Freiburger Persönlichkeitsinventar FPI Anwendung (Material: Angst²; vgl. Arnold-Krüger, 1971).

Der Stichprobenumfang beträgt $N = 427$, das durchschnittliche Alter 19 Jahre. Zum Teil fehlten die Angaben, so daß in einzelnen Tabellen $N < 427$ ist. In den Tab. 1, 2 und 3 sind Alter, Schulbildung und Wohnort dargestellt. Die Stichprobe ist bezüglich der Variable Alter sehr homogen; die Ergebnisse können aber dennoch Auskunft über die formale und inhaltliche Seite der Beschwerdenliste geben, wobei bei heterogenen Stichproben sehr wahrscheinlich prägnantere Ergebnisse zu erwarten sind.

Tabelle 1. *Alter*

Jahre	<i>f_i</i>	%
22	1	0,2
21	5	1,2
20	20	4,9
19	333	82,0
18	47	11,5
	406	100%

Tabelle 2. *Schulbildung*

Schule	<i>f_i</i>	%
Volksschule	219	51,3
Mittelschule	140	32,8
Höhere Schule	68	15,9
	427	100%

Tabelle 3. *Wohnort*

Wohnort	<i>f_i</i>	%
Großstadt	222	52,4
Kleinstadt	32	7,5
Land	170	40,1
	425	100%

4. Faktorenanalyse

Eine Hauptkomponentenanalyse der 64 Items (Produkt-Moment-Korrelation, Diagonale = 1) ergibt folgende Eigenwerte (Tab. 4).

Von Kerekjarto et al. (1972) führten bei der modifizierten Form der Beschwerdenliste (68 Items) eine Hauptkomponentenanalyse durch; der größte Eigenwert erklärt bei der männlichen Stichprobe 14,8% ($N = 439$), bei der weiblichen Stichprobe 15,5% ($N = 321$) der Gesamtvarianz; bei von Zerssen entfielen 20% der Gesamtvarianz auf den ersten Faktor.

² Herr Prof. Angst stellte das Material für die vorliegende Arbeit zur Verfügung, wofür ihm besonders gedankt sei.

Tabelle 4. Eigenwerte der Hauptkomponentenlösung

Nr.	λ	$\lambda \%$	cum $\lambda \%$
1	9,28	14,50	14,50
2	2,48	3,88	18,38
3	1,99	3,10	21,48
4	1,94	3,04	24,51
5	1,87	2,92	27,43
6	1,74	2,71	30,14
7	1,69	2,64	32,78
8	1,61	2,51	35,29
9	1,56	2,43	37,72
10	1,54	2,41	40,13

Von Kerekjarto et al. (1972) führten bei der Gesamtstichprobe ($N = 760$) eine Faktorenanalyse nach der Centroidmethode durch. Dabei fällt auf, daß die Ergebnisse der Faktorenanalyse nach der Centroidmethode und nach der Hauptkomponentenmethode stark differieren: $\lambda_1 = 41\%$ bei der Centroidmethode, $\lambda_1 = 14,8\%$ resp. $15,5\%$ bei der Hauptkomponentenmethode. Die Ergebnisse sind insofern nicht direkt vergleichbar, als bei der Centroidmethode die gesamte Stichprobe, bei der Hauptkomponentenmethode die männlichen und die weiblichen Stichproben getrennt analysiert worden sind.

Zum Vergleich Centroid-/Hauptachsenmethode meint Ueberla (1968, S. 113), daß die Ergebnisse bei den ersten Faktoren bei geeigneten (!) Daten übereinstimmen. Nach Pawlik (1968, S. 145) erklärt die Centroidmethode meistens weniger Varianz als die Hauptachsenmethode (vgl. auch die ähnlichen Resultate beim Vergleich Centroid-, Hauptachsenmethode; Ueberla, 1968; Pawlik, 1968).

Es ist also zu erwarten, daß sich die beiden Methoden in ihren Ergebnissen nicht sehr stark unterscheiden. Da dennoch Unterschiede bestehen, läßt sich der Schluß ziehen, daß eventuell die Lösungen der Faktorenanalysen von Kerekjarto et al. (1972) stark stichprobenabhängig sind, da einmal die Gesamtstichprobe, das andere Mal zwei Teilstichproben analysiert worden sind. Dagegen sprechen aber die Resultate von Zerssens (1971) (gemischtes Kollektiv von 572 Patienten, erster Eigenwert: 20%) und unsere Ergebnisse (erster Eigenwert: 14,5%); aber auch die Resultate von Kerekjarto et al. selbst, die in den zwei Teilstichproben fast identisch sind, deuten nicht darauf hin, daß die Höhe des ersten Eigenwertes stichprobenabhängig ist. Es ist daher zu vermuten, daß die Lösung der Centroidmethode den Anteil der gemeinsamen Varianz überschätzt.

Auch wenn der erste Faktor nicht 40% der Gesamtvarianz erklärt, kann man dennoch von einem „allgemeinen Faktor“ sprechen, wenn die Mehrzahl der Items auf dem 1. Faktor substantiell laden. Die Bezeichnung „Generalfaktor“ ist weniger zu empfehlen, da damit das eigentliche Generalfaktormodell der Faktorenanalyse angesprochen wird.

Kerekjarto et al. (1972) geben für die 10 ersten Eigenwerte ein kummulatives Prozent von 40,3% resp. 42,2% an. Unsere Ergebnisse stimmen danach sehr gut mit denen von Kerekjarto et al. überein, obwohl es sich bei uns um eine homogene Stichprobe ohne Patienten handelt. Vergleicht man die Ladungen der Items im ersten Faktor bei unserer Untersuchung

mit den Ergebnissen von Kerekjarto et al. (1972), so erhält man eine Produkt-Moment-Korrelation von $r = 0,51$ ($p < 0,01$). Für den Vergleich wurden die 3 zusätzlichen Items der Hamburger Untersuchung gestrichen. Dabei wird vorausgesetzt, daß einzelne Items keinen größeren Einfluß auf die Faktorenlösung haben, eine Annahme, die bei größerer Variblenzahl zu rechtfertigen ist.

In unseren Ergebnissen findet man 40 Items mit einer Ladungshöhe von $0,51 \geq a_{il} \geq 0,35$ (= 63 % der Items), während bei von Kerekjarto et al. 60 % der Items eine Ladung von $a_{il} \geq 0,35$ haben. Der Mittelwert der Itemladungen ist bei Kerekjarto et al. 0,37, bei unserer Stichprobe 0,36 (s. Tab. 5).

Tabelle 5. *Ladungen im ersten Faktor*

a_{i1}	f_i
0,10–0,19	6
0,20–0,29	10
0,30–0,39	20
0,40–0,49	24
0,50–0,59	4
$\bar{a} = 0,36$	

5. Schwierigkeitsindex. Trennschärfenkoeffizient

Die Schwierigkeitsindices sind bedeutend niedriger ausgefallen als bei von Kerekjarto et al. (1972) (s. Tab. 6). Dies ist dadurch erklärbar, daß es sich bei unserem Material um gesunde Probanden handelt, bei von Kerekjarto et al. (1972) aber um Patienten, bei denen die einzelnen Beschwerden häufiger vorkommen als bei Gesunden.

Tabelle 6
Schwierigkeitsindices der Items

p_i	f_i
0,00–0,09	25
0,10–0,19	26
0,20–0,29	8
0,30–0,39	4
0,40–0,49	1
$\bar{p} = 0,14$	

Tabelle 7
Trennschärfenkoeffizienten

r_{it}	f_i
0,10–0,19	9
0,20–0,29	13
0,30–0,39	24
0,40–0,49	18
$\bar{r} = 0,33$	

Die Trennschärfenkoeffizienten (Produkt-Moment-Korrelation; part-whole-korrigiert; Lienert, 1967) sind in Tab. 7 zusammengefaßt aufgeführt. Ein Vergleich mit den Werten von Kerekjarto et al. (1972) zeigt, daß die Koeffizienten relativ stichprobenunabhängig sind (Produkt-

Moment-Korrelation zwischen den Koeffizienten von Kerekjartos und unseren Koeffizienten: $r = 0,60$, $p < 0,01$. 2 der 64 Trennschärfenkoef- fizienten sind nicht signifikant ($p > 0,05$).

Im Vergleich zu den Werten von Kerekjartos et al. (1972) sind in unserer Stichprobe, die keine Patienten umfaßt, die Koeffizienten noch befriedigend (Kerekjarto et al.: $\bar{r} = 0,38$).

6. Reliabilität

Zur Konsistenzschätzung wurde Cronbach α verwandt (Novick, 1967). Die Gesamtskala hat ein $\alpha = 0,90$, ist also sehr homogen. Die Gesamtskala hat eine Halbierungsreliabilität (ABAB) von $r = 0,89$ (aufgewertet nach Spearman-Brown; Lienert, 1967); die Höhe des Koeffizienten stimmt überein mit der verallgemeinerten Halbierungsreliabilität, mit der sog. Konsistenz des Tests.

7. Zusammenhang mit anderen Persönlichkeitsdimensionen

Nach von Zerssen (1971) kann die Beschwerdenliste als eine für klinische Zwecke besonders geeignete Neurotizismusskala bezeichnet werden. Tab. 8 zeigt, daß die Beschwerdenliste aber nur einen Teilaспект des Neurotizismus mißt, nämlich den Aspekt der Nervosität (FPI 1). Andere Skalen wie Depressivität (FPI 3), Erregbarkeit (FPI 4) und Gehemmtheit (FPI 8), die ebenso zum Neurotizismusbereich gehören, korrelieren zwar mit der Beschwerdenliste sehr signifikant, aber numerisch nur gering (Produkt-Moment-Korrelation).

Tabelle 8. Korrelationen Beschwerdenliste/FPI-Skalen, MPI-Skalen

		FPI								
		1	2	3	4	5	6	7	8	9
Beschwerdenliste		0,62*	0,20*	0,38*	0,23*	-0,16*	-0,18*	0,10	0,28*	0,16*
<hr/>										
		FPI			MPI					
		E	N	M	N	E				
Beschwerdenliste		- 0,07	0,39*		- 0,42*	0,44*		- 0,21*		

* $p < 0,01$: $r = 0,13$.

Die Differenz zwischen Korrelation FPI 1/Beschwerdenliste ($r = 0,62$) und FPI N/Beschwerdenliste ($r = 0,39$) ist sehr signifikant ($p < 0,01$, t -Test für abhängige Korrelationen, s. Mittenecker, 1963); die Beschwerdenliste korreliert also höher mit der Skala Nervosität als mit der Skala Neurotizismus.

Die Beschwerdenliste kann von daher bei gesunden Probanden nicht als äquivalent zu Neurotizismuskalen betrachtet werden, sondern nur als Schätzung für den Neurotizismuswert. In klinischem Material ist der Zusammenhang zwischen Beschwerdenliste und Neurotizismuswert nach von Kerekjarto et al. (1972) höher: Beschwerdenliste/Neurotizismus MPI: $r = 0,68$ ($N = 98$), Beschwerdenliste/Neurotizismus MMQ: $r = 0,72$ ($N = 126$). In unserer Stichprobe beträgt dagegen die Korrelation Beschwerdenliste/Neurotizismus MPI $r = 0,44$ (MPI: Eysenck, 1959).

8. Mittelwertvergleiche

Bei der Gesamtstichprobe ($N = 427$) erhält man einen Mittelwert in der Beschwerdenliste von $M = 8,75$ und eine Standardabweichung von $s = 7,83$. Von Zerssen (1971) gibt für 92 gesunde Probanden $M = 10,1$, $s = 7,9$ an. Der Unterschied zwischen den beiden Gruppen ist nicht signifikant (t -Test). Bei der ärztlichen Untersuchung anlässlich der militärischen Aushebung werden die Probanden auch bezüglich psychosomatischer Krankheiten oder Beschwerden (leichteren oder schwereren Grades) exploriert resp. entsprechende ärztliche Zeugnisse miteinbezogen. Unterteilt man die 427 Stellungspflichtigen in Probanden ohne psychosomatische Beschwerden und mit psychosomatischen Beschwerden (Kopfbeschwerden, Herzbeschwerden, Ekzeme, Asthma usw.; Kriterium: ärztliche Untersuchung), so lassen sich die beiden Gruppen in der Beschwerdenliste vergleichen (s. Tab. 9; t -Test).

Tabelle 9. Mittelwerte von Untergruppen

	<i>N</i>	<i>M</i>	<i>s</i>	
Pbn ohne psychosomatische Beschwerden	300	11,55	8,65	
Pbn mit psychosomatischen Beschwerden	127	7,57	7,16	$p < 0,01$

Der sehr signifikante Mittelwertunterschied läßt sich als Hinweis auf die Validität der Beschwerdenliste interpretieren.

9. Kurzform

Von Zerssen (1971) hat aus den 64 Items der Beschwerdenliste zwei parallele Kurzformen von je 24 Items konstruiert (B-L, B-L'). Die beiden Kurzformen korrelieren nach von Zerssen mit $r = 0,87$ ($N = 100$, Auswertung der Gesamtform nach dem Kurzformschlüssel).

An unserem Material wurde ebenso die Gesamtform nach dem Schlüssel der Kurzform ausgewertet. Dies setzt voraus, daß der Kontext der einzelnen Items unwesentlich und auch die Länge der Testform von

geringer Bedeutung ist. Die Ergebnisse sollen nur hinweisenden Charakter haben, da eine genaue Analyse der Kurzform auch eine Verwendung der Kurzform bedingen würde.

Die beiden parallelen Kurzformen korrelieren mit sich selbst und mit der Gesamtform sehr hoch (Tab.10; Produkt-Moment-Korrelation).

Tabelle 10. *Interkorrelationen der Kurzformen, Gesamtform*

	Kurzformen	
	B-L	B-L'
Kurzform, B-L'	0,83	
Gesamtform	0,94	0,94
	<i>r</i> = 0,13	<i>p</i> < 0,01

Tabelle 11. *Mittelwertvergleich bei Kurzformen*

	N	B-L			B-L'		
		M	s	p	M	s	p
Gesamtstichprobe	427	3,47	3,44		3,28	3,39	
Mit psychosomatischen Beschwerden	127	4,65	3,84	<i>p</i> < 0,01	4,52	3,83	<i>p</i> < 0,01
Ohne psychosomatische Beschwerden	300	2,98	3,14		2,76	3,05	

Der Mittelwertvergleich (*t*-Test) zwischen den beiden Gruppen mit/ohne psychosomatische Beschwerden ergab ein sehr signifikantes Resultat wie bei der Gesamtform (s. Tab.11).

10. Diskussion

Die Ergebnisse zeigen, daß die Beschwerdenliste in ihrer Gesamtform, aber auch in den beiden Kurzformen den formalen Kriterien der Testtheorie genügt. Inhaltlich gesehen, erfaßt der Test nur einen Teil der allgemeinen Neurotizismusskalen, nämlich den der Beschwerden und weniger Aspekte der Verstimmtheit, Unsicherheit, Gehemmtheit. Ob der Test bei klinischen Gruppen den gesamten Neurotizismusbereich erfaßt, müßte erst noch geklärt werden.

Die Beschwerdenliste hat gegenüber den bekannten Persönlichkeitsfragebogen (MPI, FPI, MMPI) den Vorteil, daß die Fragen für Patienten leichter zu beantworten sind, da die Angabe von Beschwerden für Patienten in einer Klinik einfacher ist als die Beantwortung von Fragen.

Die beiden Kurzformen, die auch eine abgestufte Beantwortung zulassen, dürften für klinische Untersuchungen von großem Interesse sein, da sie hoch korrelieren und leicht durchzuführen sind.

Eine Überprüfung der Beschwerdenliste mit dem Raschschen Testmodell (Fischer, 1968) war zwar geplant, doch ließ die homogene Stichprobe keine sinnvolle Auswertung zu.

Literatur

- Arnold-Krueger, M. A.: Beziehungen zwischen Rauchen, Unfallquote und Persönlichkeitsaspekten. Phil. Diss., Freiburg i. Br. 1971.
- Eysenck, H. J.: Das „Maudsley Personality Inventory“ MPI. Göttingen: Hogrefe 1959.
- Fahrenberg, J., Selg, H.: Das Freiburger Persönlichkeitsinventar FPI. Göttingen: Hogrefe 1970.
- Fischer, G. H. (Hrsg.): Psychologische Testtheorie. Bern: Huber 1968.
- Kerekjarto, M. v., Meyer, A. E., Zerssen, D.: Die HHM-Beschwerdenliste bei Patienten einer internistischen Ambulanz. *Z. psychosom. Med. Psychoanal.* 18, 1–16 (1972).
- Lienert, G. A.: Testaufbau und Testanalyse, 2. Aufl. Weinheim: Beltz 1967.
- Mittenecker, E.: Planung und statistische Auswertung von Experimenten, 4. Aufl. Wien: Deuticke 1963.
- Novick, M. R.: Coefficient alpha and the reliability of composit measurement. *Psychometrika* 32, 1–13 (1967).
- Pawlak, K.: Dimensionen des Verhaltens. Bern: Huber 1968.
- Ueberla, K.: Faktorenanalyse. Berlin-Heidelberg-New York: Springer 1968.
- Zerssen, D. v.: Die Beschwerdenliste als Test. *Therapiewoche* 21, 1908–1914 (1971).

Dr. U. Baumann, Dipl.-Psych.
Psychiatrische Universitätsklinik
Forschungsdirektion
CH-8029 Zürich, Lenggstraße 31
Schweiz