

## ZIELE UND GRENZEN VON PERSÖNLICHKEITSFRAGEBOGEN

Alois Angleitner

Psychologisches Institut der Universität Bonn

In Buros (1965) sind im englischsprachigen Bereich 142 nichtprojektive Charaktertests als im Handel erhältlich zitiert. Ca. 42 Fragebogen geben an, Persönlichkeitsmerkmale zu erfassen. Trotz dieser Prosperität gibt es ein Mißverhältnis zwischen Persönlichkeitstheorie und Empirie, zu dem sich anmerken läßt:

1. Zum einen gibt es Theorien, die überwiegend so wenig explizit formuliert sind, daß das Überführen in konkrete Meßoperationen beinahe schon einer Neukonstruktion dieser jeweiligen Theorien gleichkäme (vgl. u.a. psychoanalytisch orientierte Theorien, Rogers, Maslow etc.). Diese Theorien sind in der Forschung unterrepräsentiert.
2. Auf der anderen Seite gibt es Theorien, für die bislang Tests entwickelt sind, die die dort konzipierten Merkmale erfassen sollen (vgl. Murray: PRF, EPPS, TAT; Kelly: REP-test; faktorenanalytische Theorien: MPI, MMQ, EPI, 16PF, GZTS; Rotter: I-E-Skala, etc.). Diese Theorien sind stark in der empirischen Forschung repräsentiert. Überwiegend handelt es sich bei diesen Tests um Fragebogen. Die ausschließliche Orientierung an verbalem Material wird aber zunehmend als Behinderung empfunden (Fiske 1974). Die Annahme einer möglichen Widerspiegelung von traits in z.B. L-, Qu- und T-Daten enthält aber noch ungeprüfte philosophisch-anthropologische Voraussetzungen.
3. Vor allem ist aber anzumerken, daß viele Persönlichkeitsfragebogen, ohne Bezug auf eine Theorie konstruiert worden sind (MMPI, FPI, PIT). Diese Verfahren scheinen in der empirischen Forschung ebenfalls stark angewandt zu werden.

Die Durchsicht der Untersuchungen verdeutlicht, daß nicht allzusehr darauf geachtet wird, inwieweit die verschiedenen oder keinen Persönlichkeitstheorien verhafteten Fragebogen von ihrer Konzeption aus mit dem Untersuchungsgegenstand in einen vorher postulierten theoretischen Bezug gebracht werden können. Entscheidender für die Auswahl, die oft in der Forschungspraxis getroffen wird, scheinen beispielsweise eher die Länge eines Fragebogens, die Anzahl der gelieferten Merkmale sowie der erste Eindruck bezüglich der Brauchbarkeit der Skalenbenennungen zu sein.

Es sind aber unterschiedliche Ergebnisse zu erwarten, auch wenn die Skalenbenennungen in verschiedenen Fragebogen identisch sind. Dies wird deutlich aus den in Tab. 1, mitgeteilten Korrelationen von 65 männlichen und 59 weiblichen Studenten und Akademikern über Fragebogenskalen verschiedenen Ursprungs.

Tab. 1.: Korrelationen für Männer und Frauen für Skalen in Fragebögen mit gleicher Labelung (Extraversion: FPI11, MPI; Neurotizismus: FPI11, MPI, MMQ; Lügen: MMPI, MMQ; Maskulinität/Femininität: MMPI, FPI12; Depression: MMPI, FPI3; Dominanz: FPI7, 16PF4)

	Extra- version	Neuroti- zismus	Lügen	Maskuli- nität	Depres- sion	Dominanz
Männer	0.65	0.76	0.20	0.29	0.18	0.32
Frauen	0.53	0.74	0.66	0.01	0.50	0.12

Die Neurotizismusskalen ausgenommen zeigen Skalen mit gleicher Labelung unbefriedigende und je nach Geschlechtszugehörigkeit unterschiedliche Übereinstimmungen.

Betrachtet man die Items dieser Skalen, so fällt auf: Ein nicht unerheblicher Anteil der Items ist identisch oder weist beinahe identische Formulierungen auf. Alle berichteten Korrelationen sind deshalb biased, und verweist man auf die inhaltlichen Ähnlichkeiten, so stellen sie allenfalls zum Teil nur mäßige Paralleltestreliabilitätsschätzungen dar.

Für die begrenzte Brauchbarkeit der gegenwärtig benutzten Persönlichkeitsinventare spricht:

1. Uneinigkeit im Gebrauch unterschiedlicher Konstruktionsprinzipien: Die faktorenanalytisch entwickelten Fragebogen überwiegen, aber selten wird über die Überprüfung der Voraussetzungen einer Faktorenanalyse berichtet. Die Festlegung einer bestimmten Anzahl von Faktoren ist bei großen Itemmatrizen relativ willkürlich. Die Überprüfungen der 16PF-Faktoren auf Itemebene erwiesen sich beispielsweise als ziemlich widersprüchlich (Cattell, 1974).
2. Mangelnde Sorgfalt bei der Testkonstruktion: Soweit dieser Schritt aus den Testhandbüchern ersichtlich wird, entsprechen unsere Fragebogen nicht den von der APA (1974) mitgeteilten Standards. So werden keine präzisen Hinweise gegeben, nach welchen Gesichtspunkten die Itemformulierungen vorgenommen bzw. die Auswahl der Items aus anderen Skalen erfolgte und wer die Formulierung vorgenommen hat. Die Samplingprozeduren bezüglich der Personenstichproben, die zur Standardisierung benutzt wurden, werden nur

ungenügend beschrieben. Frauen, ältere Personen, Angehörige niederer sozialer Schichten sind in den Samples unterrepräsentiert.

Die meisten Fragebogen entwickeln Parallelförmigkeiten, deren Äquivalenz nur ungenügend bestätigt ist. Für den 16PF wurden in eigenen Untersuchungen Äquivalenzwerte für die Formen A und B bei mehreren unterschiedlichen Stichproben im Bereich von 0.34 bis 0.44 gefunden. Die in Studien gefundenen niedrigen R-Test-Reliabilitäten werden selten zur Modifikation eines Tests benutzt, sondern eher als "inhaltliche Veränderungen" interpretiert. Dies trifft besonders auf atheoretische im klinisch-diagnostischen Bereich verwendete Fragebogen zu (MMPI, FPI).

3. Die ungelöste Frage der Response-sets: Der meist im späteren Forschungsprozeß ermittelte Varianzbeitrag der Response-sets wird nicht zur Modifikation des Verfahrens benutzt und in Forschung und Praxis weiter nicht beachtet. Ein instruktionsgerechtes, schnelles Beantworten kann auch zu einer Anzahl zufällig gegebener Antworten führen. Es wäre interessant für mittels Zufallsgenerator gegebene Antworten die diesbezüglichen Reliabilitäten zu ermitteln und diese gegen tatsächlich gefundene Reliabilitäten zu prüfen.

4. Trotz Skalenbenennungen: Adjektivattributierungen: Der Diagnostiker ist in der Bemessung der gefundenen Persönlichkeitszüge meist wieder vor die Schwierigkeit gestellt, Stereotypisierungen vorzunehmen. Oft werden wie im 16PF, MMPI, FPI, PIT eine Fülle von Adjektivattributierungen pro Merkmal angeboten und der Diagnostiker steht vor der Wahl, welches Attribut er denn nun dem Probanden zuerkennen will.

Als Folgerungen ergeben sich: Fragebogenkonstruktion hat sorgfältiger und stets im Rückgriff auf eine Theorie zu erfolgen (vgl. Jackson 1970).

#### Literatur

- American Psychological Association (Ed.), Standards for educational and psychological tests. Washington: APA, 1974.
- Buros, O.K. (Ed.) The sixth mental measurement yearbook. Highland Park, New Jersey: Gryphon Press, 1965.
- Cattell, R.B. How good is the modern questionnaire? General principles for evaluation. Journal of personality assessment, 1974, 38, 115-129.
- Fiske, D.W. The limits for the conventional science of personality. Journal of Personality, 1974, 42, 1-11.
- Jackson, D.N. A sequential system for personality scale development. In C.D. Spielberger (Ed.), Current topics in clinical and community psychology. Vol. II New York: Academic Press. 1970. Pp. 62-94.