

Analysen der Antwortprozesse bei der Bearbeitung von Fragebogenitems

Günter Krampen

Exemplarisch für den IPC-Fragebogen zur Erhebung von drei Aspekten generalisierter Kontrollüberzeugungen (Internalität, sozial bedingte Externalität, fatalistische Externalität) wurden die Antwortprozesse von 30 Probanden bei der Bearbeitung der Items mit der Methode des lauten Denkens analysiert. Zusätzlich wurden einfache Selbsteinschätzungen für die mit dem Fragebogen angezielten Konstrukte und freie Item-Skalen-Zuordnungen erhoben, um die „face validity“ zu prüfen. Es zeigt sich, daß bei den Itembeantwortungen generelle und lebensbereichsspezifische Assoziationen in etwa gleich häufig auftreten, daß semantische Probleme bei etwa einem Drittel der Antwortprozesse geäußert werden, daß intra- und interindividuelle Vergleiche nur selten angesprochen werden und daß die angesprochenen Lebensbereiche weitgehend von den Iteminhalten ausgehen. Mit einer (theoretisch stimmigen) Ausnahme konnten keine bedeutsamen Einflüsse dieser subjektiven Iteminterpretationen auf die resultierenden Skalenwerte festgestellt werden. Die Ergebnisse bestätigen die „face validity“ der Skalen und weisen darauf, daß entsprechende Konstrukte auch im Umgangswissen nachweisbar sind. Möglichkeiten der Optimierung von Fragebogen durch Analysen von Antwortprozessen werden angesprochen.

The process of responding to questionnaire items is analysed exemplarily for the German version of the IPC-Scales (measuring internality, powerful others control, chance control) in a sample of 30 adults with the method of loud thinking. To test face validity of the scales, additionally self-ratings of the three aspects of locus of control and free item-scale-classifications were collected after loud thinking. Results show that generalized and scope-specific associations to the items occur in like quantity, that semantical problems are mentioned in a third of the respondings, that intra- and interindividual comparisons occur rarely, and that the contents addressed refer in most cases to the item contents. No significant relations of such subjective iteminterpretations to the scale-scores were detected. Face validity of the three scales is confirmed rather well, which is a hint to the representation of analogous constructs in naive personality theory. Possibilities for optimizing questionnaires, which are based in analyses of the processes of responding to questionnaire items, are discussed.

Ein für eine längere Zeitspanne in der Psychologie vernachlässigtes Problem der Fragebogendiagnostik, das in den 40er Jahren ansatzweise untersucht wurde (vgl. etwa Eisenberg, 1941; Simpson, 1944) und nun langsam wieder in den Blickpunkt des Interesses gelangt, betrifft die Frage, welche kognitiven Prozesse bei einem Individuum ablaufen, das ein Fragebogenitem beantwortet. Diese subjektiven Iteminterpretationen können u.U. mit für geringe Reliabilitätskoeffizienten verantwortlich sein, da sich intraindividuelle Inkonsistenzen in ihnen (sowohl zwischen Items bei einem Erhebungszeitpunkt als auch bei einem Item zu verschiedenen Erhebungszeitpunkten) zunächst auf die Zuverlässigkeit der Messung niederschlagen. Inkonsistente Antwortmuster können etwa dadurch bedingt sein, daß zu einem Item (bezogen auf zwei Erhebungszeitpunkte) oder zu verschiedenen Items eines Fragebogens

(bezogen auf einen oder mehrere Erhebungszeitpunkte) Assoziationen aus unterschiedlichen Lebensbereichen auftreten, die dann zu unterschiedlichen Antworten führen. Durch Formulierungsmängel können solche Inkonsistenzen geradezu herausgefordert werden.

In der traditionellen Persönlichkeitsdiagnostik durch Fragebogenverfahren wird die Interpretation der Iteminhalte durch den Informanden (ganz im behavioristischen Sinne) zumeist nach wie vor als eine Art „black box“ betrachtet. Nach dem Konzept der empirischen, kriterienbezogenen Validität geht es ausschließlich um den Vorhersagewert eines Items für ein empirisches Kriterium im Sinne der konvergenten oder differentiellen Validierung. Der Inhalt einer Itemantwort oder gar die Frage, ob die Antwort auf ein Item phänomenologisch relevant ist (Konstruktvalidität und „face validity“), steht dagegen meist im Hintergrund (vgl. hierzu auch Mittenecker, 1964).

Daß nun Fragebogenverfahren verschiedenen interindividuell variierenden Fehlerquellen unterliegen, ist in der Psychometrie wohl bekannt. Relativ umfangreiche empirische Arbeiten, in denen u. a. auch Gegenmaßnahmen vorgeschlagen werden (etwa der Einsatz von Korrekturskalen), liegen zu Fehlerquellen wie der Zustimmungstendenz, der Tendenz zu extremen Antworten, der Tendenz, in sozial erwünschter oder in unüblicher Weise zu antworten, vor (siehe im Überblick Mittenecker, 1964; Seitz, 1977). Weniger beachtet wurde bisher die Mehrdeutigkeit von Iteminhalten und der Prozeß der inhaltlichen Iteminterpretation durch die Probanden. Die frühen Untersuchungen von Eisenberg (1941) und Simpson (1944) weisen schon darauf, daß erhebliche interindividuelle Differenzen in subjektiven Iteminterpretationen auftreten. In jüngerer Zeit wurden verschiedene Modelle zur Abbildung solcher Antwortprozesse vorgelegt (siehe etwa Cliff, 1977; Fiske, 1978; Nowakowska, 1970; Rogers, 1978), in denen das Antwortverhalten als abhängig von Itemmerkmalen (etwa Itemlänge, Itemmehrdeutigkeit, Itemextremität, Antwortformat) und verschiedenen Kognitionen des Informanden (etwa seine Erfahrungen, die subjektive Repräsentation der Erhebungssituation, die Wertigkeit der Frage etc.) konzipiert wird. Empirische Untersuchungen dazu beziehen sich hauptsächlich auf Analysen der Homogenität in den Formulierungen der Items eines Fragebogens (vgl. etwa Löhr & Angleitner, 1980), der Interdependenzen von Itemmerkmalen (wie Itemlänge, Wortschwierigkeit) und teststatistischen Gütekriterien (vgl. etwa Tränkle, 1982), der Wirkung unterschiedlicher Antwortformate (vgl. etwa Wegener et al., 1982) und der Itembeantwortungszeiten als ein Indikator für die subjektive Itemschwierigkeit (vgl. etwa Rogers, 1973). Seltener sind Studien anzutreffen, in denen subjektive Iteminterpretationen direkt mit den resultierenden Fragebogenwerten in Beziehung gesetzt werden. Effekte von sozialen Vergleichsprozessen und sozialen Orientierungslagen, die durch Iteminhalte (oder Instruktionen) provoziert werden, auf Fragebogenwerte konnten Seitz (1977) und Krampen (1981a) empirisch belegen. Soziale Orientierungslagen können nach diesen Befunden als relevante Determinanten des Antwortprozesses *und* der resultierenden Skalenwerte bezeichnet werden. Soziale Vergleiche, die durch Items herausgefordert werden, sind nun aber nur ein Aspekt subjektiver Iteminterpretationen. Weitere Aspekte betreffen Verständnisprobleme (semantische Probleme), die von den Informanden

assoziierten Lebens- und Handlungsbereiche, intraindividuelle Vergleichsprozesse (also Bezüge zu zeitlich früheren Erfahrungen und Selbstwahrnehmungen oder zu antizipierten Situationen) und den Antworttypus, womit die Frage gemeint ist, ob die Dekodierungsinhalte beim Informanten überhaupt in die Richtung der angezielten Konstrukte gehen. Diese verschiedenen Aspekte des Itembeantwortungsprozesses sollen exemplarisch für den IPC-Fragebogen zur Erhebung von drei Aspekten generalisierter Kontrollüberzeugungen (Krampen, 1981 b) in Verbindung mit der Frage nach seiner „face validity“ untersucht werden. Rein quantitativ orientierte Erhebungsmethoden (wie etwa die Messung von Antwortzeiten) scheiden aufgrund der angestrebten Analyseinhalte aus; es wird daher u. a. die Methode des lauten Denkens eingesetzt (siehe auch Schneider, 1982; Schneider-Düker & Schneider, 1977), die vor allem aus denkpsychologischen Untersuchungen bekannt ist (vgl. im Überblick Dörner, 1976; Huber & Mandl, 1982).

Der IPC-Fragebogen geht auf ein englischsprachiges Forschungsinstrument (Levenson, 1974) zurück, das zunächst mit dem Ziel, interkulturelle Vergleichsstudien zu ermöglichen (siehe Galli, Nigro & Krampen, 1986; Krampen & Wieberg, 1981), ins Deutsche adaptiert (Krampen, 1979), später normiert wurde (Krampen, 1981 b). Mit jeweils acht Items werden die folgenden Aspekte generalisierter Kontrollüberzeugungen erhoben, wobei das klassische eindimensionale Konzept Rotters (1966) verlassen wird: (1.) Internalität (I-Skala), d. h. die subjektiv bei der eigenen Person wahrgenommene Kontrolle über das eigene Leben und über die Ereignisse und Verstärker in der personspezifischen Umwelt; (2.) sozial bedingte Externalität (P-Skala; „powerful others control“), d. h. Externalität, die auf Gefühle der sozialen Abhängigkeit zurückgeht; (3.) fatalistische Externalität (C-Skala; „chance control“), die auf der generalisierten Erwartungshaltung basiert, daß die Welt unstrukturiert und ungeordnet ist, daß das Leben und Ereignisse in ihm weitgehend von Glück, Pech, Zufall und Schicksal abhängen. In Anwendungen der deutschen Version des IPC-Fragebogens (zu Anwendungen der englischen Version und zu interkulturellen Anwendungen siehe Levenson, 1981) konnte bislang seine diskriminative und differentielle Validität recht gut bestätigt werden (siehe etwa Allerdissen et al., 1981; Becker, 1982; Bottenberg & Schade, 1982; Krampen, 1981 b); als problematisch erwies sich jedoch die Stichprobenabhängigkeit der Skalenparameter für die interne Konsistenz und die Testhalbierungsreliabilität (siehe etwa Angleitner et al., 1982; Merz, 1981).

Die Fragestellungen der empirischen Untersuchung können wie folgt zusammengefaßt werden: (1.) In welchem Maße treten Verständnisprobleme, lebensbereichsspezifische Assoziationen sowie inter- und intraindividuelle Vergleichsprozesse bei der Beantwortung der IPC-Skalenitems auf? (2.) Stehen interindividuelle Unterschiede in diesen verschiedenen Aspekten subjektiver Iteminterpretationen mit den IPC-Skalenwerten in Beziehung? (3.) Welche „face validity“ kann den IPC-Skalenitems und den IPC-Skalen zugesprochen werden?

Methode

An den Datenerhebungen¹⁾, die in Einzelsitzungen stattfanden, waren 30 deutsche Erwachsene beteiligt ($M = 28,1$, $SD = 6,7$ Jahre; 15 Frauen; höhere Schulbildung). Die Probanden wurden zunächst gebeten, den IPC-Fragebogen auszufüllen, wobei die Instruktion derart ausgeweitet wurde, daß die Probanden gebeten wurden, alles das, was ihnen beim Ausfüllen des Fragebogens durch den Kopf geht, laut zu sagen. Die Methode des „lauten Denkens“ wurde zuvor erläutert. Die Aussagen wurden auf Tonband mitgeschnitten. Im Anschluß wurden den Probanden die dem IPC-Fragebogen unterliegenden Konstrukte (Internalität, soziale Externalität, fatalistische Externalität) ausführlich beschrieben, und sie wurden um eine Selbsteinschätzung der eigenen Persönlichkeit für die drei Variablen gebeten (6-stufige Likert-Skalen). Es sei angemerkt, daß alle Instruktionen und Erläuterungen während der Datenerhebungen in standardisierter (schriftlich festgelegter) Form gegeben wurden. Danach sollten die Probanden die 24 Items des IPC-Fragebogens den ihnen nun inhaltlich bekannten drei Subskalen (bzw. den Konstrukten) zuordnen (freie Item-Skalen-Zuordnung). Die einzelnen Datenerhebungen dauerten in etwa 45 bis 50 Minuten.

Die Auswertung der Tonbandprotokolle des lautem Denkens erfolgte nach einer Transkription inhaltsanalytisch. Die absoluten Häufigkeiten der folgenden a priori festgelegten Kategorien wurden itemspezifisch ausgezählt:

(1) *Beantwortungsmodus* (generell versus spezifisch): Assoziationen, die sich auf Anwendungen des Iteminhalts in bestimmten Lebens- und Handlungsbereichen beziehen, wurden als spezifischer Antwortmodus, solche, die den Iteminhalt allgemein thematisieren, als genereller Antwortmodus klassifiziert;

(2) *Antworttypus* (I, P, C, IP, IC, PC, IPC): Die geäußerten Denkinhalte wurden danach klassifiziert, welche Konstruktdimensionen assoziativ angesprochen werden, wobei nach „konstrukt reinen“ (I, P, C) und interaktiven Assoziationen (IP, IC, PC, IPC), in denen für ein Item mehrere Konstruktdimensionen angesprochen werden, unterschieden wird. Diese interaktiven Assoziationen gehen in die Richtung der von McKinney (1980) und Hoff (1982) thematisierten Aspekte des „interaktiven Engagement-Stils“ bzw. des „interaktionistisch-flexiblen Kontrollbewußtseins“;

(3) *Semantische Probleme*: Verständnisprobleme der Probanden bei einzelnen Items wurden unter dieser Kategorie ausgezählt. Hierzu zählt u. a. Kritik an Begriffen, an Intensitätsangaben, an Prämissen und an der Allgemeinheit/Spezifität des Iteminhalts;

(4) *Intraindividuelle Vergleiche*: Assoziationen, die zeitliche Vergleiche mit Bezug zum Iteminhalt ansprechen, wurden unter dieser Kategorie ausgezählt;

(5) *Interindividuelle Vergleiche*: Assoziationen, die soziale Vergleiche mit Bezug zum Iteminhalt ansprechen, wurden unter dieser Kategorie ausgezählt;

¹⁾ Für die Hilfe bei der Durchführung der Datenerhebungen und den inhaltsanalytischen Auswertungsarbeiten danke ich Herrn Dipl.-Psych. W.-D. Strinitz herzlich.

(6) *Angesprochene Lebensbereiche*: Diese Kategorien sind eine verfeinerte Analyse der unter (1) „Beantwortungsmodus“ als spezifisch klassifizierten Denkinhalte und beziehen sich inhaltlich auf die bei einzelnen Items angesprochenen Lebensbereiche. Da u. U. bei der Beantwortung eines Items mehrere Lebensbereiche angesprochen werden (was unter der Kategorie „Beantwortungsmodus“ nur mit der Häufigkeit 1 für „spezifisch“ klassifiziert wird), sind hier Mehrfachnennungen möglich.

Alle inhaltsanalytischen Auswertungen wurden unabhängig von zwei Kodierern anhand der Tonbandprotokolle und der Transkriptionen durchgeführt. Die durchschnittliche Übereinstimmung (Auswertungsobjektivität) der Klassifikation für alle Kategorien beträgt $\bar{r}_s = .89$ (Rangkorrelation nach Krueger-Spearman).

Ergebnisse

In Tabelle 1 sind die absoluten Häufigkeiten und die auf die Stichprobengröße bezogenen deskriptiven Kennwerte der relativen Häufigkeiten für die inhaltsanalytischen Kategorien aufgeführt. Beim Antwortmodus zeigt sich nahezu eine Gleichverteilung genereller und bereichsspezifischer Denkinhalte. Im Durchschnitt äußern die Probanden bei 11 der 24 IPC-Skalenitems generelle und bei 13 der IPC-Skalenitems spezifische Inhalte, was auf eine leichte Überlegenheit bereichsspezifischer Assoziationen deutet. Die ebenfalls in Tabelle 1 aufgeführten Befunde zu den inhaltlich angesprochenen Lebens- und Handlungsbereichen weisen darauf, daß vor allem Inhalte aus der Berufstätigkeit und aus zwischenmenschlichen Beziehungen thematisiert werden; erheblich seltener werden Inhalte aus dem Privatleben, dem politisch-öffentlichen Leben, der Freizeit und dem Konsumbereich angesprochen. Nicht tabelliert wurden Inhaltsbereiche, die in weniger als 30 Fällen (bezogen auf ein Minimum von 720 Beobachtungen) auftraten. Sie können als relativ hoch idiosynkratische Denkinhalte bezeichnet werden (etwa Lottospielen, Krankheiten, Wohnungsprobleme). Die im lauten Denken geäußerten Inhalte stehen also in direktem Bezug zu der Variabilität der Iteminhalte des IPC-Fragebogens; lediglich der berufliche Lebensbereich ist leicht überrepräsentiert, was u. U. durch eine Ankerwirkung von Item 1, in dem der Begriff der „Führungsposition“ auftaucht, bedingt sein kann.

Bei einem recht hohen Prozentsatz (30%) der registrierten Inhalte des lauten Denkens werden Verständnisprobleme in irgendeiner Form geäußert. Eine weitere Unterteilung dieser semantischen Probleme in verschiedene Subkategorien erbrachte die Information, daß Schwierigkeiten mit einzelnen Begriffen (z. B. „harte Arbeit“ in Item 21), mit Intensitätsangaben (z. B. „vor allem“ in Item 20) und mit Prämissen (z. B. bei Item 8) dominieren. Interessanter ist die itemspezifische Auszählung der semantischen Probleme, die darauf hinweist, daß die meisten Verständnisprobleme bei den Items mit den Nummern 17, 8 und 16 auftreten. In diesen Items werden Beziehungen zu „wichtigen Leuten“ und die Möglichkeit, selbst Führungspositionen einzunehmen, thematisiert. Hier deuten sich Unterschiede zwischen dem deutschen und dem amerikanischen Kulturkreis an, die bei der deutschsprachigen Bearbeitung der IPC-Skalen wohl noch nicht in ausreichendem Maße beachtet wurden.

Tabelle 1: Statistische Kennwerte der inhaltsanalytischen Kategorien

Kategorie ^a	absolute Häufigkeit ^b f	Werte zu relativer Häufigkeit ^c		
		Median	arithm. Mittel	Streuung
Antwortmodus				
– generell	326	10,8	10,9	0,76
– spezifisch	394	13,2	13,1	0,76
Antworttyp ^d				
– I	256	7,9	8,5	0,53
– P	57	1,6	1,9	0,28
– C	41	1,1	1,4	0,22
– IP	178	5,7	5,9	0,43
– IC	78	2,3	2,6	0,32
– PC	3	0,1	0,1	0,05
– IPC	19	0,3	0,6	0,20
semantische Probleme	216	6,3	7,2	0,62
intraindividuelle Vergleiche	100	2,7	3,3	0,45
interindividuelle Vergleiche	19	0,3	0,6	0,19
assoziierte Lebensbereiche ^e				
– Arbeit	226	7,5	7,5	0,61
– Sozialbez.	118	3,2	3,9	0,48
– Privatleben	47	1,5	1,6	0,24
– Staat/Politik	45	1,0	1,5	0,32
– Hobby	33	0,9	1,1	0,22
– Konsum	30	0,4	1,0	0,25

^a Erläuterung der Kategorien im Text.

^b Summenhäufigkeit jeweils 720 (N = 30 x 24 Items).

^c Bezogen auf N = 30.

^d Nicht klassifizierbare Aussagen: 88.

^e Mehrfachnennungen möglich; Bereiche unter f = 29 sind nicht aufgeführt.

Sowohl intraindividuelle (13,9%) als auch interindividuelle Vergleichsprozesse (2,6%) treten in den im lauten Denken geäußerten Inhalten in geringem Maße auf. An sozialen Normen orientierte Antworten und soziale Orientierungslagen scheinen nach diesen Befunden bei der Beantwortung des IPC-Fragebogens keine, an zeitlichen Vergleichen für die eigene Person orientierte Antworten nur eine geringe Rolle zu spielen. Insgesamt ergibt sich somit als Antwort auf Fragestellung (1), daß sich bei

der Beantwortung des IPC-Fragebogens generelle und lebensbereichsspezifische Assoziationen in etwa die Waage halten, daß semantische Probleme bei etwa einem Drittel der Antwortprozesse geäußert werden, daß inter- und intraindividuelle Vergleiche nur selten angesprochen werden, und daß die angesprochenen Lebensbereiche weitgehend von den Iteminhalten ausgehen.

Zur Prüfung von Fragestellung (2), ob interindividuelle Unterschiede in diesen Aspekten der subjektiven Iteminterpretation mit den IPC-Skalenwerten der Personen in Beziehung stehen, wurde die Gesamtgruppe anhand der Häufigkeit, mit der eine inhaltsanalytische Kategorie von einer Person besetzt wurde, paramedian geteilt. Vorab sei vermerkt, daß die durchschnittlichen IPC-Skalenwerte in der Gesamtgruppe im mittleren Normbereich liegen, daß die Streuungen aber auf eine hinreichende Streubreite der erhobenen Fragebogenwerte deuten. Ohne Hinweise auf Varianzheterogenität in den paramedian geteilten Subgruppen (überprüft mit dem Bartlett-Box-Test), ergaben sich nur für eine der inhaltsanalytischen Kategorien signifikante Mittelwertsunterschiede in einem Skalenwert des IPC-Fragebogens. Personen, die den IPC-Fragebogen eher bereichsspezifisch versus eher generell beantworteten (Beantwortungsmodus; $t(28) \leq 1,29$), die viel versus wenig semantische Kritik üben (semantische Probleme; $t(28) \leq 1,32$) und die häufig versus selten intraindividuelle Vergleiche ziehen ($t(28) \leq 1,21$), unterscheiden sich in den Skalenwerten nicht signifikant. Keine statistisch bedeutsamen Mittelwertsunterschiede konnten auch für die I- und die C-Skala bei Personen festgestellt werden, die selten versus häufig interindividuelle Vergleiche im lauten Denken äußern ($t(28) \leq 1,02$). Lediglich für die P-Skala konnte im Vergleich dieser beiden Gruppen festgestellt werden, daß Probanden, die häufig soziale Vergleiche anstellen ($M_p = 31,5$) erheblich höhere Werte in der sozial bedingten Externalität haben als die, die solche sozialen Vergleiche selten (oder gar nicht) äußern ($M_p = 23,3$; $t(28) = 2,29$, $p < .05$). Dieses Ergebnis kann als ein Validitätsbeleg für die IPC-Skalen betrachtet werden. Ansonsten kann Fragestellung (2) verneint werden. Die im lauten Denken erhobenen Unterschiede in den subjektiven Iteminterpretationen stehen in keinen relevanten Beziehungen zu den im IPC-Fragebogen resultierenden Skalenwerten. Aufgrund dieser Ergebnisse kann davon ausgegangen werden, daß sich die erfaßten Iteminterpretationen nicht verzerrend auf die Messung der Variablen auswirken.

Für die Beantwortung von Fragestellung (3) nach der „face validity“ der IPC-Skalenitems und der IPC-Skalen stehen Informationen aus (a) den Selbsteinschätzungen der Probanden, (b) von im lauten Denken geäußerten Antworttyp und (c) der freien Item-Skalen-Zuordnung zur Verfügung. In Tabelle 2 sind die Korrelationen zwischen den Fragebogenwerten und den Selbsteinschätzungen der Probanden aufgeführt. Die relativ hohen Koeffizientenwerte in der Diagonalen bestätigen die „face validity“ des Erhebungsinstruments. Verstärkt wird dieser Befund noch durch die geringen korrelativen Beziehungen zwischen Selbsteinschätzungen und Fragebogenwerten, die sich auf verschiedene Konstruktdimensionen beziehen. Dies verweist auf die diskriminative Validität der IPC-Skalen und auch der Selbsteinschätzungen der Probanden.

Tabelle 2: Korrelation der IPC-Skalenwerte mit den Selbsteinschätzungen (N = 30)

IPC-Fragebogen	Selbsteinschätzungen		
	I _s	P _s	C _s
I-Skala	.64***	-.34*	-.24
P-Skala	-.15	.68***	.28
C-Skala	-.02	.15	.57***

***p < .001; *p < .05

Auch die in Tabelle 1 aufgeführten Häufigkeiten für die verschiedenen Antworttypen belegen – wenn auch weniger gut – die „face validity“ der IPC-Skalen. Es muß jedoch vermerkt werden, daß neben „konstruktreinen“ auch interaktive Assoziationen, in denen vor allem internale und sozial-externale Inhalte geäußert werden, auftreten. Die relativ geringen Häufigkeiten für P- und C-„reine“ Äußerungen verweisen darauf, daß sozial- und fatalistisch-externale Assoziationen in reiner Form seltener anzutreffen sind, und daß sie vor allem in der Kombination mit internalen (in verneinter Form) auftreten. Auch die hohe Zahl der I-„reinen“ Äußerungen geht u. a. darauf zurück, daß im lauten Denken relativ häufig globale Verneinungen von Internalität im Zusammenhang mit P- oder C-Skalenitems auftraten. Äußerungen, in denen P- und C-Assoziationen gemeinsam auftreten, sind ebenso selten wie solche, in denen alle drei Konstruktaspekte (IPC) thematisiert werden.

Ebenso wie das laute Denken wurde auch die freie Item-Skalen-Zuordnung ohne Kenntnis der tatsächlichen Skalenzugehörigkeit der Items durchgeführt. Für die Items der I-Skalen ergaben sich dabei 7%, für die der P-Skala 25,8% und für die der C-Skala 15,4% falsche Zuordnungen. In den Feinanalysen zeigt sich, daß die relativ hohe Quote der Falschzuordnungen bei den P- und C-Skalenitems vor allem dadurch zustande kommt, daß sie als Negationen der Internalitätsvariablen aufgefaßt werden. Die kreuzweise falsche Zuordnung von P- und C-Skalenitems steht mit der mittleren Interkorrelation dieser beiden Skalen in Einklang (siehe Krampen, 1981 b). Insgesamt zeigt sich aber, daß die Probanden immerhin 84% der Item-Skalen-Zuordnungen ohne Kenntnis der tatsächlichen Zuordnung richtig vornehmen, was ein weiterer Hinweis auf die „face validity“ der IPC-Skalenitems ist.

Diskussion

Exemplarisch für den IPC-Fragebogen zu Kontrollüberzeugungen konnte die relative Häufigkeit subjektiver Iteminterpretationen bei der Beantwortung von Persönlichkeitsfragebogen aufgezeigt werden. Während generelle und bereichsspezifische Assoziationen in einem in etwa ausgewogenen Verhältnis den Antwortprozeß begleiten, wobei die spezifischen in der Mehrzahl der Fälle direkt von den Iteminhalten ausgehen, zeigten sich inter- und intraindividuell orientierte Vergleichsprozesse

nur bei wenigen Probanden und bei wenigen Items. Der Befund zur relativen Seltenheit sozialer Vergleichsprozesse bei der Beantwortung des IPC-Fragebogens steht im Widerspruch zu Ergebnissen von Seitz (1977) und Krampen (1981a), die den Einfluß solcher Prozesse und sozialer Orientierungslagen auf die Beantwortung von Items in Persönlichkeitsfragebogen zu sozialen Einstellungen (Seitz, 1977, S. 81ff.) bzw. zur Leistungsmotivation (Krampen, 1981a) belegen. Die vorliegenden Befunde weisen darauf, daß soziale Vergleiche im Antwortprozeß bei Kontrollüberzeugungen augenscheinlich eine erheblich geringere Bedeutung haben als bei der Erhebung von Variablen aus dem sozial-interaktiven Bereich und aus dem Leistungsbereich. Auch intraindividuelle Vergleichsprozesse, die eigene Persönlichkeitsentwicklungen zum Inhalt haben, konnten nur in geringem Maße in den frei geäußerten Antwortassoziationen festgestellt werden. Häufiger (in etwa einem Drittel der analysierten Antwortprozesse) sind dagegen Verständnisprobleme zu verzeichnen (siehe hierzu auch Löhner & Angleitner, 1980). Die Ergebnisse weisen auf Formulierungsprobleme bei einzelnen Items und auf erhebliche interindividuelle Unterschiede zwischen den Probanden. Dies ist zum einen ein Hinweis auf die Notwendigkeit, bestimmte Itemformulierungen zu verbessern (u. a. durch kulturell adäquatere Itemadaptionen, die bislang wegen der Vergleichbarkeit von Befunden mit denen aus Anwendungen anderssprachiger Versionen der IPC-Skalen vermieden wurden; vgl. etwa Krampen & Wieberg, 1981), zum anderen ein Hinweis auf eventuell den Beantwortungsprozeß beeinflussende Persönlichkeitsmerkmale wie sprachliche Kompetenz, kognitive Komplexität oder auch auf den Einfluß situativer Faktoren auf die emotionale und motivationale Befindlichkeit des Probanden (und seine Bereitschaft zur Kooperation).

Ein zentrales Ergebnis der Untersuchung ist auch der Befund, daß interindividuelle Differenzen in den subjektiven Iteminterpretationen (mit einer Ausnahme, die theoretisch konsistent ist; siehe oben) nicht zu statistisch bedeutsamen Unterschieden in den resultierenden IPC-Skalenwerten führen. Dies spricht für die Robustheit dieses Meßinstruments gegenüber subjektiven, eventuell auch hoch idiosynkratischen Assoziationen bei der Beantwortung. Verzerrungen dieser Art, die sich auf die Fragebogenwerte niederschlagen, können weitgehend ausgeschlossen werden. Dies impliziert jedoch nicht, daß sich die subjektiven Iteminterpretationen, deren Auftreten nachgewiesen wurde, nicht negativ auf die Reliabilitätskennwerte niederschlagen können. Es ist zu vermuten, daß hier eine wesentliche Ursache für die bei Erhebungsinstrumenten zu Kontrollüberzeugungen oftmals erhaltenen relativ geringen Parametern der Testhalbierungsreliabilität und der internen Konsistenz liegt (siehe hierzu im Überblick Krampen, 1982, S. 99ff.).

Die für diese Analysen verwendete Methode des lauten Denkens und der inhaltsanalytischen Auswertung der so produzierten Protokolle hat sich insgesamt gesehen bewährt, was entsprechende methodische Empfehlungen in der Literatur bestätigt (siehe etwa Schneider, 1982). Problematisch bleibt freilich der relativ große Erhebungs- und Auswertungsaufwand, der in der vorliegenden Studie dazu führte, daß man sich mit einer relativ geringen Stichprobengröße (und den damit verbundenen Problemen der Generalisierbarkeit der Befunde) zufrieden geben mußte. Es konnte

aber belegt werden, daß die quantitative Auswertung der qualitativ erhobenen Daten mit dem verwendeten (reduktiven) Kategoriensystem objektiv erfolgen kann.

Die in der Studie parallel überprüfte „face validity“ der IPC-Skalen und ihrer Einzelitems, die eine Beziehung zur phänomenologischen Bedeutung von Antworten auf Items für das angezielte Konstrukt aufweist (siehe Seitz, 1977), führte zu recht positiven Befunden. Einfache Selbsteinschätzungen der Probanden sind relativ hoch und differentiell mit den Skalenwerten aus dem Fragebogen verbunden, die ohne Kenntnis der wahren Skalenzugehörigkeit durchgeführte freie Item-Skalenzuordnung führte in 84% der Beobachtungen zu richtigen Zuordnungen und die im lauten Denken geäußerten Antwortmodi bestätigen schwerpunktmäßig die Konstruktzugehörigkeit der Items. Deutlich wird bei einem Teil dieser Ergebnisse aber auch, daß augenscheinlich das Konstrukt der Internalität bei den Probanden kognitiv stärker verhaftet und ausgeprägt ist als die Konstrukte der sozial und fatalistisch bedingten Externalität. P- und C-Items werden im Antwortprozeß häufiger mit Internalitätsanteilen konfundiert, in der freien Item-Skalenzuordnung werden sie häufiger als negativ gepolte I-Items der I-Skalen zugeordnet. Ohne diese Befunde überbewerten zu wollen, kann gesagt werden, daß solche interaktiven, mehrere Einflußquellen beachtende Kognitionen doch so häufig auftreten, daß die separate Erhebung verschiedener Aspekte generalisierter Kontrollüberzeugungen sinnvoll ist. Genau dies wird mit dem IPC-Fragebogen geleistet, ohne daß dazu die Konzeption eines isolierten „interaktionistisch-dynamischen Kontrollbewußtseins“ sensu Hoff (1982) notwendig erscheint. Dies gilt um so mehr, als die singulären Selbsteinschätzungen der Probanden essentiell mit den entsprechenden Fragebogenskalen korreliert sind und untereinander keine bedeutsamen Beziehungen aufweisen.

Als Fazit kann aus dieser Studie (unter Beachtung der relativ geringen Stichprobengröße) die Folgerung gezogen werden, daß mit dem IPC-Fragebogen ein Erhebungsinstrument vorliegt, das nicht nur über konvergente, diskriminative und differentielle Validität verfügt (siehe Krampen, 1981 b; Levenson, 1981), sondern auch über „face validity“ auf Skalen- und Itemebene. Hinzu tritt der Befund, daß subjektiven Iteminterpretationen für die Generierung der Skalenwerte eine geringe Bedeutung zukommt. Mängel in den Itemformulierungen sind jedoch vorhanden, die auf die Notwendigkeit ihrer Verbesserung (dann unter Verzicht auf die Möglichkeit zu interkulturellen Vergleichsstudien) verweisen.

Literatur

- Allerdissen, R., Florin, I. & Rost, W.: Psychological characteristics of women with bulimia nervosa (bulimarexia). *Behavior Analysis & Modification*, 1981, **4**, 314–317.
- Angleitner, A., Filipp, S.-H. & Braukmann, W.: Testtheoretische Prüfung der Fragebogenverfahren zur Erfassung ausgewählter Personmerkmale. Trier: Universität, Forschungsbericht 20 aus dem Projekt „Entwicklungspsychologie des Erwachsenenalters“, 1982.
- Becker, P.: Interaktions-Angst-Fragebogen. Weinheim: Beltz, 1982.
- Bottenberg, E. H. & Schade, F.-D.: Darstellung alltagsphilosophischer Konzeptionen in einem Bereich selbst- und weltbezogener subjektiver Theorien, Bewertungen. *Psychologie und Praxis*, 1982, **26**, 127–130.

- Cliff, N.: Further study of cognitive processing models for inventory response. *Applied Psychological Measurement*, 1977, **1**, 41–49.
- Dörner, D.: Problemlösen als Informationsverarbeitung. Stuttgart: Kohlhammer, 1976.
- Eisenberg, P.: Individual interpretation of psychoneurotic inventory items. *Journal of General Psychology*, 1941, **25**, 19–40.
- Fiske, D. W.: Strategies for personality research. San Francisco: Jossey-Bass, 1978.
- Galli, I., Nigro, G. & Krampen, G.: Multidimensional locus of control and machiavellianism in Italian and West German students: Similarities and differences. *International Review of Applied Psychology*, 1986, **35**, 453–461.
- Hoff, E.: Kontrollbewußtsein: Grundvorstellungen zur eigenen Person und Umwelt bei jungen Arbeitern. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 1982, **34**, 316–339.
- Huber, G. L. & Mandl, H. (Hrsg.): Verbale Daten. Weinheim: Beltz, 1982.
- Krampen, G.: Differenzierungen des Konstrukts der Kontrollüberzeugungen. *Zeitschrift für experimentelle und angewandte Psychologie*, 1982, **26**, 573–595.
- Krampen, G.: Soziale Orientierungslagen als Prozeßbedingungen bei der Beantwortung von Persönlichkeitsfragebogen-Items. *Diagnostica*, 1981 a, **27**, 127–139.
- Krampen, G.: IPC-Fragebogen zu Kontrollüberzeugungen. Göttingen: Hogrefe, 1981 b.
- Krampen, G.: Differentialpsychologie der Kontrollüberzeugungen. Göttingen: Hogrefe, 1982.
- Krampen, G. & Wieberg, H.-J. W.: Three aspects of locus of control in German, American and Japanese university students. *Journal of Social Psychology*, 1981, **113**, 133–134.
- Levenson, H.: Activism and powerful others. *Journal of Personality Assessment*, 1974, **38**, 377–383.
- Levenson, H.: Differentiating among internality, powerful others, and chance. In H. M. Lefcourt (Hrsg.), *Research within the locus of control construct*. Vol. 1 (S. 15–63). New York: Academic Press, 1981.
- Löhr, F.-J. & Angleitner, A. 1980. Eine Untersuchung zu sprachlichen Formulierungen der Items in deutschen Persönlichkeitsfragebogen. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 1980, **1**, 217–235.
- McKinney, J. P.: Engagement styles (agent versus patient) in childhood and adolescence. *Human Development*, 1980, **23**, 192–209.
- Merz, J.: Testkritische Überprüfung einer deutschen Version des IPC-Fragebogens. *Diagnostica*, 1981, **27**, 253–260.
- Mittenecker, E.: Subjektive Tests zur Messung der Persönlichkeit. In R. Heiss (Hrsg.), *Psychologische Diagnostik* (S. 461–487). (= K. Gottschaldt, P. Lersch, F. Sander & H. Thomae (Hrsg.), *Handbuch der Psychologie*. Band 6.) Göttingen: Hogrefe, 1964.
- Nowakowska, M.: A model of answering to questionnaire items. *Acta Psychologica*, 1970, **34**, 420–439.
- Rogers, T. B.: Towards a definition of the difficulty of personality item. *Psychological Reports*, 1973, **33**, 159–166.
- Rogers, T. B.: Experimental evidence for the similarity of personality and attitude responding. *Acta Psychologica*, 1978, **42**, 21–28.
- Rotter, J. B.: Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 1966, **80** (1, Whole No. 609).
- Schneider, J. F.: Models and empirical approaches to the process of responding to questionnaires. Saarbrücken: Universität des Saarlandes, *Arbeiten der Fachrichtung Psychologie*, 1982, **81**.

- Schneider-Düker, M. & Schneider, J. F.: Untersuchungen zum Beantwortungsprozeß bei psychodiagnostischen Fragebogen. Zeitschrift für experimentelle und angewandte Psychologie, 1977, **24**, 282–302.
- Seitz, W.: Persönlichkeitsbeurteilung durch Fragebogen. Braunschweig: Westermann, 1977.
- Simpson, R. H.: The specific meaning of certain terms indicating different degree of frequency. Quarterly Journal of Speech, 1944, **30**, 328–330.
- Tränkle, U.: Über Zusammenhänge zwischen der sprachlichen Schwierigkeit von Fragebogenitems, teststatistischen Gütekriterien und Antwortungsverhalten. Diagnostica, 1982, **28**, 289–306.
- Wegener, B., Faulbaum, F. & Maag, G.: Die Wirkung von Antwortvorgaben bei Kategoriarskalen. ZUMA-Nachrichten, 1982, **10**, 3–20.

Anschrift des Verfassers:

Priv.-Doz. Dr. Günter Krampen
Universität Trier
FB I – Psychologie
Postfach 38 25
D-5500 Trier