

## EINIGE PROBLEME BEI DER ANWENDUNG DER I-E-SKALA (INTERNE/EXTERNE KONTROLLERWARTUNG)

### 1. Problemstellung

Ziel des vorliegenden Aufsatzes ist es, anhand eines konkreten Beispiels die Probleme aufzuzeigen, die mit der Beurteilung und Anwendung von bereits existierenden Einstellungsskalen in der Feldforschung verbunden sind. Schon eingeführte 'Skalen' oder Fragenbatterien werden in Umfragen häufig verwendet, nicht nur aus Ökonomiegründen, sondern vor allem um Vergleichbarkeit in Bezug auf das intendierte theoretische Konstrukt herzustellen. Bei immer neuen ad-hoc-Konstruktionen ist die Kommensurabilität zumindest ungesichert. Allerdings wirft der Einsatz bereits vorhandener Einstellungsskalen in Befragungen Probleme eigener Art auf. Befragungssituationen erfüllen in der Regel nicht die idealen Bedingungen, die für vollgültige testtheoretische Analysen Voraussetzung sind. Leider ist es oft auch so, daß die 'Skalen' schlecht dokumentiert und entsprechend unterschiedliche Versionen in Umlauf sind, und daß Auswertung und Interpretation von Fall zu Fall variieren.

Diesem Zustand möchte ZUMA mit der Herausgabe seines sog. Skalenhandbuchs (s. ZUMANACHRICHTEN No. 3) in einem ersten Schritt entgegenreten. Das Handbuch wird in Bezug auf eine umfassende Anzahl von Instrumenten zunächst einmal alle Information, die verfügbar ist, dokumentieren. Es soll aber zugleich über die bloße Sammelarbeit hinausgehen, indem für jede aufgenommene 'Skala' individuell die charakteristischen Probleme, die bei der Anwendung in Umfragen auftreten, diskutiert werden. Für eine solche Aufarbeitung kann es kein allgemeines Schema für ein Vorgehen, das auf alle möglichen Instrumente anwendbar wäre, geben. Jeder Fall liegt anders und ist komplex. Um davon einen Eindruck zu vermitteln, werden im folgenden - ohne Anspruch auf Vollständigkeit - wichtige theoretische und pragmatische Probleme beim Umgang mit der Skala zur internen und externen Kontrollerwartung (ROTTER, 1954) dargestellt, und es werden Lösungsvorschläge gemacht. Die I-E-Skala erfreut sich gegenwärtig einer gewissen Beliebtheit. Sie wurde auch in den Nationalen Sozialen Survey (vgl. S.46 in diesem Heft) aufgenommen, so daß der Beitrag ebenfalls die dort gemachten Erfahrungen widerspiegelt.

## 2. Theoretischer Kontext der I-E-Skala

### 2.1 Theoretische Grundlagen

Im Rahmen der lerntheoretisch orientierten Persönlichkeitspsychologie wurde das Konzept der internen/externen Kontrollenerwartung (IEC) eingeführt, um die subjektive Seite in der Beziehung zwischen Bedürfnissen und Verhalten zu repräsentieren (vgl. ROTTER, 1954). Die zugrundeliegende Hypothese ist, daß das zielgerichtete Verhalten von Personen nicht lediglich eine Funktion eines spezifischen Bedürfnisses ist, sondern daß die Erwartung, die Situation im Sinne der eigenen Ziele beeinflussen zu können, als Einflußfaktor ebenfalls eine wichtige Rolle spielt.

Die soziale Lerntheorie von ROTTER postuliert, daß Personen eine generalisierte Kontrollenerwartung bezüglich ihrer Fähigkeit, die Ereignisse in ihrem Leben zu bestimmen, bilden. Die beiden Extreme der internen vs. externen Kontrollenerwartung werden dabei folgendermaßen definiert:

- D1 Interne Kontrollenerwartung liegt vor, wenn Personen erwarten, daß ihr eigenes Verhalten den Verlauf ihres Lebens bestimmt.
- D2 Externe Kontrollenerwartung liegt vor, wenn Personen glauben, durch ihr eigenes Verhalten die Ereignisse in ihrem Leben kaum oder gar nicht beeinflussen zu können.

Dabei sind offenbar eigene 'Fähigkeiten' und 'Anstrengungen' Faktoren, die für die kausale Erklärung von Umweltereignissen eher dann in Frage kommen, wenn eine interne Kontrollenerwartung vorliegt, während 'Zufall', 'Schicksal', 'Glück', 'die sozialen Verhältnisse' usw. Kontrollfaktoren für einen externen Erwartungsstil darstellen (vgl. FISCHER, 1976: 9).

### 2.2 Zusammenhang mit anderen Variablen

Interne/externe Kontrollenerwartung kann in dreierlei Weise mit anderen Variablen verknüpft sein:

- 1. IEC als unabhängige Variable in einem Modell zur Erklärung von Verhalten

2. IEC als abhängige Variable in einem Modell, in dem die Entstehung von Kontrollerwartung erklärt wird
3. IEC als intervenierende Variable in einem komplexen Modell, in dem Kontrollerwartungen als abhängige und als unabhängige Variable auftreten.

Da im Zusammenhang mit der Einstellungsmessung besonders die Beziehung zwischen Einstellung und Verhalten relevant ist, werden wir uns im folgenden exemplarisch nur mit diesem Bereich befassen. Dabei bildet die Erwartung der Konsequenzen des eigenen Verhaltens eine zentrale unabhängige Variable zur Erklärung von Handlungen bzw. Entscheidungen von Personen. Handlungsalternativen können dann im Sinne eines einfachen belief-value-Modells für eine betroffene Person 'relevant' werden (OPP, 1978: 53; KAUFMANN-MALL, 1978: 20f.) :

- D3 Die Relevanz  $R$  einer Handlungsalternative  $j$  ergibt sich aus der Summe der Produkte, die aus der subjektiven Wahrscheinlichkeit (Erwartung  $E$ ) für das Auftreten einzelner Handlungskonsequenzen  $i$  von  $j$  und dem Belohnungswert (Valenz  $V$ ) dieser Handlungskonsequenzen gebildet werden:

$$R_j = \sum_i E_{ij} V_{ij}$$

Die Relevanz einer Handlungsalternative wird damit als eine nicht unabhängige zu messende Variable eingeführt. Auf ihrer Basis kann als mögliche Entscheidungsregel angenommen werden:

- H1 Diejenige Handlungsalternative  $j$ , für die sich die höchste Relevanz errechnet, wird ausgeführt.

Attitudentheoretiker wie z.B. FISHBEIN (1978) oder ROSENBERG (1960) unterscheiden affektive, kognitive und konative Elemente von Attitüden. Dabei wird aber die kausale Struktur der Beziehungen zwischen diesen Elementen meist implizit und unklar gelassen. KAUFMANN-MALL (1978: 110) hat in seiner integrativen kognitiv-hedonistischen Theorie einen Vorschlag zur kausalen Verknüpfung dieser Variablen gemacht. In Schema 1 wird eine Zuordnung seiner Begriffe mit denen der Attitudentheorie vorgenommen:

## Schema 1: Zuordnung der Begriffe der Attitudentheorie und der kognitiv-hedonistischen Theorie

<u>Attitudentheorie</u>	<u>kognitiv-hedonistische Theorie</u>
affektives Element	Valenz der Folgen einer Handlung
kognitives Element	Erwartung der Folgen einer Handlung
konatives Element	Verhalten

---

Unter der kognitiven Komponente verstehen die Attitudentheoretiker Hypothesen (beliefs) über die Realität. Dies ist weitgehend identisch mit dem, was hier unter Erwartungen verstanden wird. Solche subjektiven Hypothesen sind zum Teil Verhaltenshypothesen folgender Art: Verhalten A führt zu Konsequenz X. Dies entspricht dem engeren Erwartungsbegriff. Der weitere Erwartungsbegriff bezieht sich dagegen ganz allgemein auf Hypothesen über Zusammenhänge in der Umwelt, ohne persönlichen Bezug. ROTTER (1954) geht von Erwartungen im engeren Sinne aus, also von Verhaltenshypothesen mit Bezug zum eigenen Verhalten. Die Erwartungen bestehen in subjektiven Wahrscheinlichkeiten dafür, daß eine bestimmte Verhaltensalternative zu einer bestimmten Konsequenz führt. Diese werden wiederum durch die Lerngeschichte einer Person geprägt. ROTTERS Theorie ist also ein Spezialfall der allgemeineren kognitiv-hedonistischen Theorie.

### 2.3 Differenzierungen des theoretischen Konzepts

Allgemeine theoretische Überlegungen sowie empirische Ergebnisse bei der Anwendung der I-E-Skala legen eine Differenzierung des ursprünglichen ROTTERSchen Konzepts nahe. Im folgenden werden stichwortartig wesentliche Aspekte solcher Differenzierungen erläutert. Die Beziehung zwischen der objektiven Situation und den Kontrollerwartungen kann dabei nicht systematisch thematisiert werden.

#### 2.3.1 Art der Lernerfahrung

Erwartungen stellen Kognitionen über die Umwelt dar, die gelernt worden sind. Gelernte Erwartungen beziehen sich dabei nicht nur auf den eigenen Erfahrungsbereich, sondern auch auf räumlich oder sozial entfernte Objekte so-

---

wie auf komplexe und abstrakte Sachverhalte.

Die reichhaltige Literatur über Vorurteile und Ideologien zeigt nun, daß die Übernahme von Kognitionen gerade für über den eigenen Erfahrungsbereich hinausgehende Sachverhalte sozial normiert ist. Als konkretes Beispiel kann hier die Erwartung über die Erfolgchancen von - individueller oder kollektiver - politischer Partizipation gelten, die Bestandteil der politischen Kultur ist und zwischen Gruppen bzw. Nationen stark variiert. Man spricht in diesem Fall von kollektiven Deutungsmustern. Diese können sich sowohl, wie im beschriebenen Beispiel, auf spezifische Lebensbereiche, jedoch auch generell auf die Einschätzung der persönlichen Verantwortlichkeit beziehen, wie z.B. die 'protestantische Ethik', die eine starke interne Kontrollerwartung impliziert (vgl. FISCHER, 1976: 13). Die Verteilungen von IEC können daher von Gruppe zu Gruppe variieren und sich innerhalb einer Gruppe im Zeitverlauf ändern.

### 2.3.2 Soziale Normierung

Ungeachtet der wissenschaftlichen Neutralität des IEC-Konzepts können in Gruppen starke Normen hinsichtlich einer mehr internen oder externen Kontrollerwartung entstehen. So kann unabhängig von der tatsächlich vorhandenen Kontrollerwartung ein starker Konformitätsdruck die Ergebnisse der Messung beeinflussen. ROTTER beschreibt selbst, daß bei Häftlingen und College-Anwärtern die Korrelation zwischen der I-E-Skala und einer Skala, die die Tendenz zu sozial erwünschtem Antwortverhalten mißt (Social Desirability Scale von MARLOWE-CROWNE), höher war als bei anderen Gruppen. Er führt dies darauf zurück, daß hier von den Befragten eine Testsituation wahrgenommen wurde, die von ihnen eine interne Kontrollerwartung forderte (1966: 281).

### 2.3.3 Situationsklasse bzw. Lebensbereich, auf den sich die Kontrollerwartung bezieht

Generalisierte Kontrollerwartungen haben einen spezifischen Geltungsbereich, der anhand einer Ähnlichkeitsdimension (physikalische Ähnlichkeit oder funktionale Äquivalenz; FISCHER, 1976: 7) bestimmt werden kann. Die Erwartung, seinen persönlichen Lebensbereich kontrollieren zu können, kann z.B. durchaus mit einer geringen Kontrollerwartung hinsichtlich der Kontrol-

---

lierbarkeit politischer Ereignisse Hand in Hand gehen.

Es erhebt sich also nicht nur die Frage, ob es so etwas wie eine über alle Situationsklassen generalisierte Kontrollerwartung unabhängig von situationspezifischen überhaupt gibt, sondern auch die, zwischen welchen Situationsklassen eine Person differenziert und wie stark ihre Kontrollerwartungen für diese jeweils variieren. Die Implikationen dieser Frage für die Konstruktion eines Meßinstrumentes liegen auf der Hand: im ersten Fall müßten bereichsunabhängige Fragen gefunden werden, die als Maß für eine Variable 'Allgemeine Kontrollerwartung' zusätzlich zur Messung situationsspezifischer Erwartungen dienen können. Im zweiten Fall müßte dagegen der Score für die Allgemeine Kontrollerwartung als Durchschnittswert aus den bereichsspezifischen Kontrollerwartungen gebildet werden. Er hängt damit stark von der Auswahl der Bereiche ab, zu denen Fragen gestellt werden.

Es liegt daher aus theoretischen Gründen nahe, neben allgemeinen auch situationspezifische Kontrollerwartungen zu erfassen und explizite Hypothesen über deren Wirkungsweise zu formulieren, wie z.B.:

- H2 Der Einfluß bereichsunabhängiger generalisierter Erwartungen ist umso geringer, je mehr bereichsspezifische Erfahrungen eine Person in einem Bereich hat (vgl. dazu FISCHER, 1976: 6f., 54).

#### 2.3.4 Grad der Beeinflußbarkeit der Kontrollfaktoren

In ROTTERs Theorie wird nur die Kontrollebene (intern vs. extern) systematisch variiert, während die Kontrollfaktoren innerhalb der beiden Ebenen als funktional äquivalent betrachtet werden. Die empirischen Ergebnisse legen jedoch die Vermutung nahe, daß die Kontrollfaktoren je nach dem Grad ihrer Beeinflußbarkeit zu unterschiedlichen Verhaltenskonsequenzen führen. Folgende generelle Hypothese läßt sich hier formulieren:

- H3 Geringe wahrgenommene Beeinflußbarkeit von Kontrollfaktoren führt zu Passivität (ROST-SCHAUDE et al., 1979: 16f.).

Bei ROTTER wurden aufgrund seines Experimentaldesigns beide Dimensionen konfundiert, da die Zuordnung zur Gruppe mit interner Kontrollerwartung (rich-

---

tige Lösung möglich) mit dem Grad der Beeinflußbarkeit der Kontrollfaktoren (Anstrengung) korrespondierte.

Die empirischen Ergebnisse von GURIN et al. (1969) zeigten jedoch, daß im Fall der wahrgenommenen Veränderbarkeit der Situation auch bei externer Kontrollorientierung aktives Verhalten auftritt. Als Beispiel führen die Autoren dabei die Mitarbeit schwarzer Jugendlicher in der Civil-Rights-Bewegung an, die in dem Moment auftrat, als persönliche Mißerfolge der Minoritätsangehörigen nicht mehr deren persönlichem Versagen, sondern ungerechten, aber prinzipiell als änderbar wahrgenommenen gesellschaftlichen Bedingungen zugeschrieben wurden.

Das genannte Beispiel betrifft zunächst nur die Kontrollfaktoren der externen Ebene, die Validität von Hypothese 2 für die der internen ist jedoch wahrscheinlich. Die entsprechende Hypothese würde dann lauten, daß Anstrengungen zur Zielerreichung umso eher unternommen werden, als man sich entsprechende Fähigkeiten zuschreibt (vgl. ROST-SCHAUDE et al., 1979: 6)

### 2.3.5 Objekt der Kontrollerwartungen

GURIN et al. stellten bei ihren empirischen Arbeiten fest, daß in der ursprünglichen Konzeptualisierung von ROTTER zwei Ebenen nicht auseinandergelassen wurden, die sie persönliche Kontrolle und Kontrollideologie nannten (1969: 35). Persönliche Kontrolle bezieht sich dabei auf die wahrgenommene Möglichkeit einer Person, ihr eigenes Leben beeinflussen zu können, Kontrollideologie dagegen auf ihre Wahrnehmung, wie diese Möglichkeit durchschnittlich in der Gesellschaft ausgeprägt ist. Neben dem angesprochenen Lebensbereich kann also das Objekt variieren, bezüglich dessen Kontrollerwartungen geäußert werden, das einmal die eigene Person und zum andern die Gesellschaft insgesamt ist. Beide können insoweit unabhängig voneinander variieren, als ein Individuum eine Differenz zwischen der eigenen Situation und der anderer Personen bzw. der durchschnittlichen gesellschaftlichen Situation überhaupt wahrnimmt. Das bedeutet, daß man theoretisch zwischen Erwartungsträger und Objekt der Kontrollerwartungen unterscheiden muß, wobei dieses Objekt auch ein Kollektiv sein kann (vgl. dazu FISCHER, 1976: 59 ff.).

---

## 2.3.6 Theoretische Konsequenzen

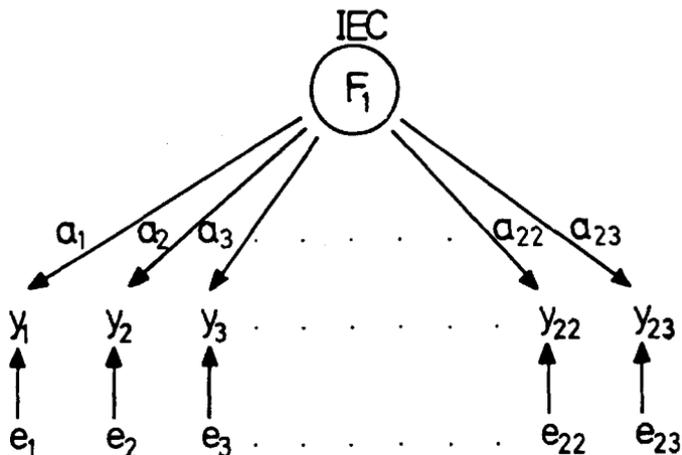
Bei der Formulierung von Hypothesen müssen alle fünf beschriebenen Differenzierungskriterien berücksichtigt werden. Zur Überprüfung dieser Hypothesen in der Feldforschung ist es daher notwendig, daß alle fünf Variablen angemessen repräsentiert sind. Die Art der Lernerfahrung sowie das Ausmaß der sozialen Normierung können jeweils durch geeignete Indikatoren erfaßt werden. Z.B. kann persönliche Vertrautheit mit Situationen, für die Kontrollerwartungen erfragt werden (Kontakt mit Bürokratien, Erfahrung mit politischer Partizipation), als Indikator für Lernerfahrung dienen. Soziale Erwünschtheit wiederum kann über Fragen nach den perzipierten Normen erfaßt werden.

Die drei Variablen Situationsklasse, Beeinflußbarkeit der Kontrollfaktoren und Objekt der Kontrollerwartungen sollten dagegen neben der Variablen Locus of Control (intern/extern) direkt in die Konstruktion einer I-E-Skala eingehen, in der pro Situationsklasse (Lebensbereich) die Kontrollfaktoren für alle Variablenkombinationen systematisch zu berücksichtigen sind (Schema 2).

Schema 2: Kontrollfaktoren

		beeinflußbar		nicht beeinflußbar	
		Selbst	Gesellschaft	Selbst	Gesellschaft
intern	A	eigene Anstrengung	B	C	D
			individuelle Anstrengung		
extern	E	persönlich relevante Machtverhältnisse	F	G	H
			gesellschaftlich relevante Machtverhältnisse		

Abb.1: Meßmodell für die Rotter-Skala



### 3. Methodische Fragen

#### 3.1 Die ursprüngliche Rotter-Skala

Die in Abschnitt 2.1 diskutierten theoretischen Vorstellungen und Annahmen gingen implizit in die Konstruktion der I-E-Skala ein, die ROTTER 1966 veröffentlichte. Diese Skala besteht aus 23 Items und 6 Füllitems. Alle Items sind als forced-choice Items formuliert. Das heißt, die Befragten müssen sich für eine von zwei Antwortvorgaben entscheiden. Der Testrohwert einer Person errechnet sich dann aus der Anzahl der externen Alternativen, die die Person gewählt hat. Ein Item aus der Rotter-Skala lautet z.B.:

Externe Attribuierung: 'I have often found, that what is going to happen will happen.'

Interne Attribuierung: 'Trusting to fate has never turned out as well for me as making a decision to take a definite course

of action.'

Der Skalinhalt der Gesamtskala ist bezüglich a) der Art der angesprochenen Kontrollfaktoren und b) der Art der Situationsklassen heterogen. Dennoch hat ROTTER zunächst postuliert, daß sein Meßinstrument eindimensional sei. Diese Hypothese bzw. das zu Grunde liegende Meßmodell kann anschaulich durch das Pfaddiagramm in Abb. 1 dargestellt werden, wobei  $a_1 \dots a_{23}$  Faktorenladungen,  $F_1$  = latente IEC Variable,  $e_1 \dots e_{23}$  Residuen der Items (Zufallsmeßfehler), und  $y_1 \dots y_{23}$  Items sind.

Das dargestellte Modell entspricht dem einer (konfirmatorischen) Faktorenanalyse und kann bei quantitativer Messung mit LISREL (vgl. JÖRESKOG & SÖR - BOM, 1978) und bei dichotomen Items - wie in der Rotter-Originalskala - mit FADIV (vgl. MUTHEN, 1978) empirisch überprüft werden. Durch die Theorie ist hier zunächst nur die Zahl der Faktoren auf 1 festgelegt, während Annahmen über die Höhe der Faktorenladungen bzw. der Meßfehler fehlen.

### 3.2 Dimensionalität des Konstrukts

Empirische Ergebnisse (vgl. z.B.: GURIN et al., 1978; zusammenfassend CHERLIN & BOURQUE, 1974; ROTTER, 1975; zu deutschen Ergebnissen ROST-SCHAUDE et al., 1978, und SIMON, 1979) führten zu einer Modifikation des ursprünglichen Konzepts der Kontrolliererwartung. Als zentrale Frage erwies sich nämlich die Dimensionalität der Skala und die Stabilität der Faktorenladungen. Folgende Punkte wurden in diesem Zusammenhang diskutiert:

- 1) Kann man Persönliche Kontrolle von Kontrollideologie als eigenen Faktor unterscheiden? (s. Abschnitt 3.2.1)
- 2) Sind Kontrolliererwartungen bereichsspezifisch, z.B. bezogen auf Politik oder Berufserfolg, oder gibt es eine generelle Kontrolliererwartung? (3.2.2)
- 3) Wie verändert sich die Faktorenstruktur durch Hinzunahme anderer Items? (3.3.1)
- 4) Wie wirken sich unterschiedliche Skalen (forced choice vs. Likert) aus? (3.3.3)
- 5) Sind interne und externe Kontrolle verschiedene Faktoren? (3.3.3)

- 6) Wie verändern sich die Zahl der Faktoren und die Faktorenladungen bei variierenden Gruppen bzw. Populationen? (3.4.1)
- 7) Wie wirken sich Response Sets auf die Dimensionalität der Skala und die Faktorenladungen aus? (3.4.2)

### 3.2.1 Persönliche Kontrolle und Kontrollideologie als generelle Kontrollfaktoren

GURIN et al. fanden sowohl in einer Stichprobe farbiger U.S.-Studenten (1969) als auch in einer repräsentativen Stichprobe der U.S.-Bevölkerung (1978) mehr als den einen von Rotter postulierten Faktor. In den Ergebnissen zeigte sich, daß alle Items, in denen das Pronomen ich zur Messung der generalisierten Kontrollerwartung verwendet wurde, auf einem Faktor laden. Ein Beispiel für ein derartiges Item ist:

'Wenn i c h Pläne mache, bin i c h nahezu sicher, daß i c h sie auch durchsetze.'

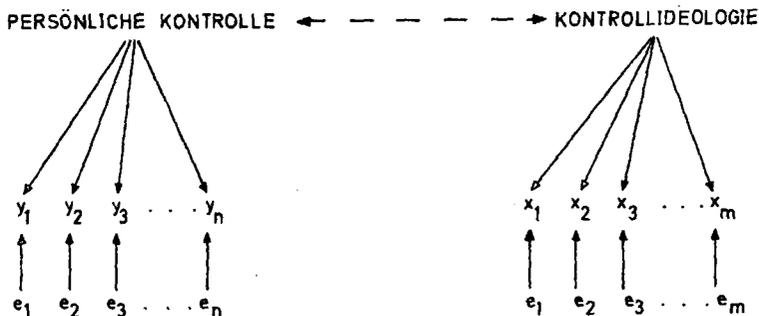
Die latente Dimension, die dieser Art von Items zu Grunde liegt, wird Persönliche Kontrolle genannt. Im Gegensatz dazu luden alle Items mit dem Pronomen man bzw. der Verwendung der dritten Person auf einem weiteren Faktor, der hier Kontrollideologie genannt werden soll. Durch die Verwendung des Pronomen man wird bei den Befragten nicht ihre persönliche Kontrollerwartung, sondern ihre Einschätzung der Kontrollerwartung im gesamten gesellschaftlichen Bereich bzw. für den Durchschnitt von Personen in einer Gesellschaft erfragt. Diese Einschätzung kann natürlich stark von der persönlichen Kontrollerwartung abweichen.

Während nun GURIN et al. (1978) die beiden Teildimensionen Persönliche Kontrolle und Kontrollideologie als Faktoren genereller Kontrolle replizieren konnten, ergab sich in der Untersuchung von CHERLIN & BOURQUE (1974) nur ein Faktor für Generelle Kontrolle. Dieser erklärt sowohl die mit ich als auch mit man formulierten Items.

Eine Illustration der eben gemachten Ausführungen gibt Abb. 2.

---

Abb. 2: Modifiziertes Meßmodell von GURIN et al.



Im Unterschied zu GURIN et al. nehmen wir aber an, daß die beiden latenten Variablen (Faktoren) korreliert sind. Die Verwendung der orthogonalen Faktorenanalyse ohne empirische Prüfung der Orthogonalitätsannahme in den bisherigen Analysen ist nicht überzeugend und dürfte die Ergebnisse der Koeffizientenschätzungen verzerren (vgl. WEEDE & JAGODZINSKI, 1977).

### 3.2.2 Generalisierte Kontrollerwartung oder Kontrollerwartung für spezifizierte Bereiche

In der Analyse von GURIN et al. (1978) ergaben sich zusätzlich zu den zwei generellen Faktoren drei weitere Faktoren für spezielle Bereiche. Einer dieser Faktoren bezieht sich auf die Politische Kontrolle, auf dem vier der fünf Items mit politischem Inhalt hoch luden. Ein Beispiel ist:

'Es ist schwierig für einen zu kontrollieren, was Politiker in ihrem Amt leisten.'

Ein weiterer Faktor bezieht sich auf die Zurechnung der Ursachen von Berufserfolg (success mobility). Zwei Items lassen sich auf Grund ihrer Formulie-

rung und der Ergebnisse der Faktorenanalyse als Indikatoren dieses Konstrukts interpretieren, z.B.:

'Erfolg zu haben ist eine Frage harter Arbeit und hat wenig oder gar nichts mit Glück zu tun.'

Ein dritter Faktor bezieht sich auf Interpersonelle Kontrolle und wird z.B. markiert durch:

'Leute, die es nicht fertigbringen, andere für sich einzunehmen, verstehen einfach nicht, wie man mit anderen umgeht.'

Im Gegensatz zu diesen Ergebnissen fanden CHERLIN & BOURQUE (1974) über den generellen Kontrollfaktor hinaus nur einen weiteren bereichsspezifischen Faktor, und zwar Politische Kontrolle. Wie in Abschnitt 3.2.3 ausgeführt wird, sind aber die Unterschiede zwischen CHERLIN & BOURQUE und GURIN et al. auf alternative Methoden der Faktorenextraktion zurückzuführen.

### 3.2.3 Gründe für die unterschiedlichen Faktorenstrukturen

Betrachtet man die Ergebnisse der verschiedenen Faktorenanalysen, so ist man zunächst versucht, deren Unterschiedlichkeit auf die Verwendung von unterschiedlichen Skalen zurückzuführen. Es zeigt sich jedoch, daß zwar jeweils einige der Original-Items weggelassen wurden, die übrige Skala aber in derselben Form wie bei ROTTER vorgegeben wurde.

Am einfachsten ist noch die Diskrepanz zwischen ROTTERS einfaktorieller und den übrigen mehrdimensionalen Lösungen zu erklären. Da ROTTER nur nach einer Dimension suchte, gab er sich damit zufrieden, daß seine erste Hauptkomponente 53% der Gesamtvarianz erklärte, und hielt die übrigen Faktoren schlicht für irrelevant (1966: 282).

CHERLIN & BOURQUE weisen auf eine zweite mögliche Ursache für die verschiedenen gefundenen Faktorenstrukturen hin, nämlich den unterschiedlichen Fragebogenkontext. Solange jedoch die Items der Rotter-Skala separat analysiert werden, dürften Kontexteffekte kaum als so schwerwiegend anzunehmen

---

sein.

Ein weiteres Argument bezieht sich auf die verschiedenen Populationen, bei denen die Ergebnisse gewonnen wurden. Die Ergebnisse von Gurin et al. dürften am zuverlässigsten sein, da sie an einer repräsentativen Zufallsstichprobe gewonnen wurden (1978: 276). Die Unterschiede zwischen den Populationen andererseits erklären für sich genommen nicht so völlig unterschiedliche Faktorenstrukturen, ohne die Validität des zugrundeliegenden theoretischen Konzepts dadurch in Frage zu stellen.

Neben diesen möglichen inhaltlichen Gründen müssen jedoch auch die Effekte der jeweils verwendeten Methode in Betracht gezogen werden. Betrachtet man die Verfahrensweisen von CHERLIN & BOURQUE einerseits und GURIN et al. andererseits, so springen folgende Punkte ins Auge:

- (1) Da die Rotter-Items dichotom skaliert sind, kann die Verwendung des bei Faktorenanalysen üblichen Produkt-Moment-Korrelationskoeffizienten  $r$  Unterschiede des Fragebogenkontexts und der gewählten Population dramatisieren, da dessen Höhe stark von den Randverteilungen der einzelnen Variablen abhängt.
- (2) Die Kriterien für die Wahl der Faktorenzahl sind letzten Endes willkürlich; CHERLIN & BOURQUE benutzten die theoretische Bedeutsamkeit als Abbruchkriterium bei der Faktorenextraktion (2 Faktoren), GURIN et al. brachen die Extraktion ohne Angaben von Gründen nach dem 5. Faktor ab. Der Anteil der erklärten Varianz - bei GURIN et al. 41% und bei CHERLIN & BOURQUE 29% kann im übrigen als sehr bescheiden angesehen werden. Die Wahl der Faktorenzahl hat jedoch einen nicht zu unterschätzenden Einfluß auf die Faktorenladungen.
- (3) Es wurden unterschiedliche Faktorenextraktionsmethoden verwandt; GURIN et al. arbeiteten mit Kommunalitätsschätzungen in der Hauptdiagonalen, CHERLIN & BOURQUE dagegen mit der unveränderten Korrelationsmatrix (Einsen in der Hauptdiagonalen).
- (4) Generell ist zu den referierten Untersuchungen zu bemerken, daß die Verwendung der orthogonalen Faktorenrotation ohne empirische Prüfung der

Orthogonalitätsannahme nicht überzeugend ist und die Ergebnisse der Koeffizientenschätzungen verzerren dürfte. Vor allem, wenn die Existenz eines Konstrukts "Generalisierte Kontrollerwartung" als ein den bereichsspezifischen Erwartungen zugrundeliegender Faktor zweiter Ordnung angenommen werden soll, liegt die Vermutung korrelierter Faktoren nahe.

Die gefundenen Faktorenstrukturen sind also aufgrund von inhaltlichen und methodischen Unterschieden unvergleichbar, so daß auf der Basis der vorliegenden Informationen keiner von ihnen der Vorzug gegeben werden kann. Der Hauptgrund für die Unterschiedlichkeit der Ergebnisse liegt offensichtlich in der Wahl der Faktorenzahl. GURIN et al. waren nämlich in der Lage, bei einer Reanalyse ihrer Daten eine Faktorenstruktur mit zwei Faktoren entsprechend der von CHERLIN & BOURQUE zu replizieren, wobei sie sogar ein anderes Verfahren der Faktorenextraktion benutzten (Kommunalitätsschätzungen statt Einsen in der Hauptdiagonalen; 1978: 280f.). Das heißt jedoch nicht, daß die übrigen diskutierten möglichen Ursachen keine relevanten Unterschiede hervorzubringen vermögen. Denn auch bei einer ähnlichen Faktorenstruktur können die Faktorenladungen deutlich differieren, was zu unterschiedlichen Factor-Scores und damit zu unterschiedlichen Individual-Scores führt.

### 3.2.4 Bildung von Subskalen

Legen die Ergebnisse von Faktorenanalysen die Mehrdimensionalität einer Skala nahe, so erhebt sich die Frage, ob aus den einzelnen Item-Scores ein Gesamt-Score gebildet werden soll, oder ob man es vorzieht, nur jeweils die Items eines einzelnen Faktors zu einer Subskala zusammenzufassen. Folgende Möglichkeiten sind denkbar:

- (1) Addition der Item-Scores zu einem Gesamt-Score ohne Berücksichtigung der Mehrdimensionalität des Konstrukts. Diese ist jedoch nur empfehlenswert, sofern wie bei ROTTER im Fall einer Hauptkomponentenanalyse der erste Faktor einen großen Teil und zugleich erheblich mehr Varianz erklärt als alle übrigen. Außerdem müssen alle Items auf diesem hinreichend hoch laden. Sind mehrere, deutlich unterscheidbare und varianzbindende Faktoren im Spiel und laden einige der Items eindeutig nur auf einem der Faktoren, so gehen bei einer einfachen Addition die unterschiedlichen Dimensionen je nach der Anzahl der zu ihnen gehörenden

# ZUMA

---

Items mit verschiedenem Gewicht ein, was nicht wünschenswert ist.

- (2) Verzicht auf die Berechnung eines Gesamt-Scores und Aufteilung der Skala in Subskalen. Dies würde inhaltlich bedeuten, daß man die Vorstellung einer einzigen zugrundeliegenden latenten Variablen ganz aufgibt.
- (3) Bildung eines Gesamt-Scores und Berechnung von Scores für die Subskalen. Bei diesem Vorgehen wird trotz Akzeptierung der Mehrdimensionalität der Skala an der Vorstellung eines einheitlichen zugrundeliegenden Konstrukts festgehalten. Methodisch entsprechen diesem Vorgehen zwei unterschiedliche Faktorenstrukturen. Einmal ist eine Struktur denkbar, bei der alle Items auf einem Generalfaktor (erste Hauptkomponente) laden und zusätzlich für jede Itemgruppe noch ein weiterer Faktor (nach Rotation) gefunden wird. Wahrscheinlicher ist jedoch eine schiefwinklige, mehrfaktorielle Struktur, bei der sich der Generalfaktor als Faktor zweiter Ordnung ergibt.

Beide Fälle implizieren, daß der Generalfaktor ein Konstrukt ist, das sich aus einer Reihe von Spezialfaktoren zusammensetzt. In unserem Beispiel würde das bedeuten, daß die bereichsunspezifische Kontrollerwartung nicht unabhängig von den bereichsspezifischen existiert, sondern einen Durchschnittswert aus diesen darstellt. Die Beziehungen zwischen Gesamt-Score und den Scores der einzelnen Subskalen können empirisch untersucht werden. Es ist auch möglich, für spezielle Vorhersagen eine Regressionsgleichung aufzustellen, in der sowohl der Gesamt-Score als auch einer oder mehrere Subskalen-Scores als unabhängige Variablen eingehen.

### 3.3 Effekte bei Abweichungen von der ursprünglichen Skala

#### 3.3.1 Der Einfluß der Item-Auswahl

Es liegt auf der Hand, daß die Auswahl der Items, die in eine Skala eingehen, nicht nur die Faktorenstruktur, sondern auch die resultierenden Skalen-Scores stark beeinflusst. Streng genommen sind daher nur Skalen mit gleichen Items vergleichbar.

Verkürzung einer Skala kommt immer dann in Betracht, wenn der im Fragebogen verfügbare Platz knapp ist, auf das betreffende theoretische Konstrukt aber

---

nicht verzichtet werden soll. Neben der dimensionalen Struktur können hierbei jedoch auch Reliabilität und Validität der Skala beeinträchtigt werden, so daß von diesem Verfahren ohne vorherige gründliche Oberprüfungen dringend abgeraten werden muß. Es ist eigentlich nur in zwei Fällen zulässig; nämlich erstens, wenn es für eine Skala eine speziell entwickelte und getestete Kurzform gibt, oder zweitens, wenn nur die Items einer Subskala verwendet werden.

Die Verlängerung einer Skala ist prinzipiell weniger riskant, da man die zusätzlichen Items bei der Analyse wieder eliminieren kann, sofern sie die Qualitätskriterien der Skala negativ beeinflussen. Beispiel für eine Veränderung der Rotter-Skala durch Hinzufügen von neuen Items ist die erste Untersuchung von GURIN et al. bei schwarzen College-Studenten (1969). Hier wurden zusätzlich zu den 23 Rotter-Items drei weitere Items einer am Survey Research Center (Ann Arbor, Michigan) häufig benutzten Political Efficacy Scale und 13 weitere Items vorgegeben. Letztere sollten die Einstellungen der Studenten zu Wirkungen persönlicher und externer Kräfte bezüglich der Rassen-situation in den USA messen (1969: 34f.). Neben den Dimensionen Persönliche Kontrolle und Kontrollideologie ergaben sich zwei weitere Faktoren 'System Modifiability' und 'Race Ideology', die eindeutig auf die hinzugefügten Items zurückzuführen sind.

### 3.3.2 Formulierungsveränderungen

Um die Anwendbarkeit der Skala bei speziellen Populationen zu erleichtern, werden häufig einzelne oder alle Items einer Skala umformuliert. Durch diese Umformulierung wird versucht, die Items für die betreffende Population verständlich zu machen. Skalen, insbesondere solche, die im Universitätskontext entwickelt und an studentischen Populationen getestet werden, müssen für eine Repräsentativerhebung in der Regel modifiziert werden. Das Problem ist dann, bedeutungsäquivalente Formulierungen zu finden, die die grundlegenden Skaleneigenschaften bewahren. Gleiches gilt natürlich für Übersetzungen in andere Sprachen. Streng genommen müßten die umformulierten bzw. übersetzten Items wie eine neue Skala getestet und anhand derselben Kriterien wie die ursprüngliche Skala validiert werden.

Eine deutsche Übersetzung der Rotter-Skala erbrachte Faktoren, die teils de-

---

nen von CHERLIN & BOURQUE, teils denen von GURIN et al. entsprechen (ROST-SCHAUDE et al., 1978: 328). Der Faktor der politischen Kontrolle, der sowohl bei CHERLIN & BOURQUE als auch bei GURIN et al. zu finden ist, taucht auch hier als eine eigenständige Dimension auf. Er hat sich also bisher als der robusteste erwiesen, da er unabhängig von der gewählten Faktorenzahl und der Sprache offensichtlich weitgehend unverändert wieder auftritt. Ein systematischer Vergleich der deutschen und englischen Version der Rotter-Skala, der notwendig wäre, um methodische Artefakte von substantiellen Abweichungen unterscheiden zu können, steht jedoch noch aus.

### 3.3.3 Veränderungen in den Antwortvorgaben

Die möglichen Effekte einer Veränderung in der Form der Vorgabe der Items werden häufig gar nicht explizit reflektiert, obwohl sie vermutlich ebenfalls sehr bedeutungsvoll sind. Dabei sind Varianten in dreierlei Hinsicht möglich:

- (1) Art der Vorgabe der Items: Forced Choice, d.h. Auswahl zwischen zwei Statements vs. Einzelvorgabe der Statements. In Abschnitt 2.2.1 wurde bereits die Frage aufgeworfen, ob interne bzw. externe Kontrolle evtl. zwei unterschiedliche Konstrukte sind. Dies läßt sich bei der Verwendung des Forced-Choice-Formats natürlich nicht beantworten, da die Versuchspersonen jeweils zwischen einem internen und einem externen Statement wählen müssen. Bei der Einzelvorgabe der Statements läßt sich dagegen anhand der Korrelationen feststellen, ob es sich tatsächlich um ein bipolares Konstrukt handelt oder nicht. Wäre das der Fall, müßten polare ("umgedrehte") Statements stark negativ miteinander korrelieren. Die Fälle, in denen die Items der Rotter-Skala einzeln (Likert-Format) vorgegeben wurden, zeigten, daß IEC offensichtlich kein bipolares Konstrukt ist. GURIN et al. berichten über die Ergebnisse einer Faktorenanalyse von COLLINS, bei der die Statements der beiden Pole überwiegend auf unterschiedlichen Dimensionen laden (1978: 280). Die Ergebnisse von FISCHER (1976: 88f.) und von Simon (1979) bestätigen ebenfalls, daß interne und externe Kontrolle offensichtlich nicht als zwei Pole desselben Konstrukts aufgefaßt werden, sondern eigenständige Dimensionen sind. Dies hat jedoch u.E. keine theoretisch relevanten Ursachen, sondern dürfte auf semantische Probleme (inadäquate Verbalisierung der bei-

den Pole) zurückzuführen sein.

- (2) Art der verwendeten Skala: dichotom (stimmt/stimmt nicht) oder mehrstufig, wobei hier zusätzlich Skalen mit und ohne neutralen Nullpunkt unterschieden werden müssen. Daneben ist auch noch die Möglichkeit einer Magnitude-Skalierung in Betracht zu ziehen (vgl. WEGENER, 1979, 1980), um individuell nicht nur sehr differenziert abzubilden, sondern auch, um ein hohes Skalenniveau zu garantieren.
- (3) Art der zu beurteilenden Eigenschaft der Statements: kognitiv (stimmt/stimmt nicht) oder evaluativ (stimme zu/lehne ab). Es liegt auf der Hand, daß die Dimensionalität durch einen Wechsel der Beurteilungsdimension stark beeinflußt werden kann. Bei der I-E-Skala ist beispielsweise theoretisch eindeutig vorgegeben, daß eine kognitive Dimension erfaßt werden soll, da es um Erwartungen und nicht um Bewertungen geht. In der ursprünglichen Version von RÖTTER kommt dies auch deutlich in den Anweisungen zum Ausdruck: 'Please select the one statement of each pair (and only one) which you more strongly believe to be the case as far as you're concerned' (1966: 295). GURIN et al. sowie CHERLIN & BOURQUE verwendeten im wesentlichen dasselbe Format der Fragestellung.

In beiden deutschen Untersuchungen wurde dagegen die evaluative Dimension angesprochen: Die Versuchspersonen sollten ihre Zustimmung/Ablehnung zu den Items ausdrücken. Dabei waren sich die Autoren offensichtlich der Diskrepanz zwischen theoretischem Konzept und Art der Antwortgabe nicht bewußt. So schreibt FISCHER einerseits, daß die Versuchspersonen je nachdem, ob sie Aussagen für zutreffend halten oder nicht, der Kategorie interne/externe Orientierung zugeordnet wurden, andererseits sollten sie die Items jedoch auch nach dem Grad ihrer Zustimmung beurteilen (1976: 76f.; auch BECK et al., 1979). Effekte dieses Wechsels in den Antwortvorgaben dürften mit Sicherheit nachzuweisen sein, ähnlich wie auch in anderen Bereichen der Psychologie Wahrnehmungs- und Bewertungsstrukturen keinesfalls deckungsgleich sind.

### 3.4. Effekte unterschiedlicher Anwendungsbedingungen

#### 3.4.1 Unterschiedliche Populationen

Sofern das zu messende Konstrukt im Rahmen einer allgemeinen, nicht raumzeitlich beschränkten Theorie verwendet wird, müßte das Meßinstrument auf

---

# ZUMA

---

alle denkbaren Populationen anwendbar sein und dieselben Zusammenhänge mit anderen Variablen aufweisen. Das heißt jedoch nicht, daß alle Populationen dieselbe Verteilung ihrer Skalenwerte auf dem Konstrukt aufweisen müssen. Sie können sich vielmehr deutlich sowohl in ihren Mittelwerten als auch in ihren Varianzen unterscheiden.

Ergeben sich aber Unterschiede in der Faktorenstruktur zwischen verschiedenen Populationen, so deutet das darauf hin, daß diese sich hinsichtlich ihres semantischen Verständnisses der Items, der Beurteilungsaufgabe, der Relevanz des Konstrukts oder ihres Differenzierungsgrades unterscheiden. Im Falle einer vergleichbaren Faktorenstruktur für verschiedene Populationen können die unterschiedlichen Scores dieser Populationen (oder auch ihre Varianzen) miteinander verglichen werden.

Die Evidenz aus den verschiedenen Anwendungen der I-E-Skala zur Klärung dieser Frage ist nicht einheitlich. ROST-SCHAUDE et al. stellten z.B. bei Frauen eine differenziertere dimensionale Struktur fest als bei Männern, während GURIN et al. sowie CHERLIN & BOURQUE dieselbe Faktorenstruktur für verschiedene Populationen fanden, erstere für die drei Gruppen 'weiße Männer', 'weiße Frauen' und 'Schwarze' (1978: 277), letztere für Studenten und einen Bevölkerungsquerschnitt (1974: 572).

### 3.4.2. Response Sets

Die Anwendungssituation oder spezifische Charakteristika von Individuen können durch die Wirksamkeit von Response Sets die Ergebnisse bei der Anwendung eines Meßinstrumentes verzerren. Weiter oben wurde bereits auf die Wirkung von Social Desirability hingewiesen, die RÖTTER selbst bei einigen seiner Untersuchungen feststellte (1966: 281; 1975: 62). Die Wirkung eines zweiten Response Sets, nämlich Acquiescence (Ja-Sage-Tendenz), wird bei der Verwendung des Forced-Choice-Formats durch die Art der Aufgabenstellung ausgeschlossen, ist jedoch bei der Einzelvorgabe von Statements durchaus wahrscheinlich. Es empfiehlt sich, bei der Analyse der Skalenergebnisse solche Response Sets zu berücksichtigen. Im Fall von Social Desirability ist das durch Verwendung der entsprechenden Skala möglich. Acquiescence läßt sich dagegen über den Anteil zustimmend beantworteter Statements kontrollieren. Populationsspezifische Response-Sets können natürlich Auswirkungen auf

---

die Dimensionalität der Skala haben. Hat man eine diesbezügliche Vermutung, so läßt sich dies anhand von subgruppenspezifischen Faktorenanalysen prüfen.

### 3.5 Reliabilität und Validität

Die empirischen Ergebnisse zeigen, daß die Reliabilität der Rotter-Skala relativ hoch ist, sofern von Reliabilität hier als von interner Konsistenz, d.h. der Korrelation der einzelnen Items mit dem Gesamt-Score, die Rede ist. Dies gilt sowohl für die Gesamtskala als auch für ihre Anwendung in verschiedenen Populationen sowie für die verschiedenen Subskalen (GURIN et al., 1978: 279f.; CHERLIN & BOURQUE, 1974: 577f.). Aus den akzeptablen Werten der Reliabilität der Gesamtskala darf aber nicht geschlossen werden, daß nur ein Faktor die Korrelationen erklärt, denn im Fall korrelierter Faktoren können auch Items verschiedener Dimensionen hoch miteinander korrelieren. Man erfährt dadurch also nichts über die dimensionale Struktur der Skala. Auch das Ausmaß der systematischen Meßfehler wird dadurch nicht erfaßt. Wie hoch ist die formale oder Konstruktvalidität der einzelnen Items, gemessen durch ihre Faktorenladungen auf jeweils dem Faktor, den sie messen sollen? Für die Untersuchung von GURIN et al. (1978) zumindest gilt, daß nur eine recht mittelmäßige formale Gültigkeit vorliegt. Nicht nur sind die Faktorladungen der Items auf ihren Zieldimensionen gering, auch die Untergruppen unterscheiden sich in dieser Hinsicht beträchtlich. Außerdem haben alle Items zumindest in einer der drej untersuchten Gruppen auch auf anderen als den Zielfaktoren Ladungen von mindestens .10. Sie messen also mehr als nur die eine intendierte Dimension.

Ein anderer Gültigkeitsaspekt bezieht sich auf die Korrelationen der Gesamtskala und der Subskalen mit externen Variablen, mit denen Zusammenhänge theoretisch postuliert wurden. In Bezug wiederum auf die Studie von GURIN et al. lassen sich in Hinsicht auf diesen Aspekt sog. diskriminanter Validität die folgenden Aussagen machen:

Um die Beziehungen zwischen Kontrollideologie und persönlicher Kontrolle einerseits und politischem Verhalten andererseits zu überprüfen, berechneten GURIN et al. die Korrelationen zwischen diesen beiden Kontrollvariablen sowie konventionellem und unkonventionellem politischem Verhalten in drei

# ZUMA

---

Teilpopulationen, nämlich Schwarzen einerseits, liberalen und konservativen Weißen andererseits.

Dabei ergab sich, daß bei Schwarzen und Liberalen keine Beziehung zwischen dem Ausmaß der persönlichen Kontrollenerwartungen und dem Ausmaß politischer Aktivität bestand. Dagegen fand sich eine Beziehung zwischen externer Kontrollideologie und konventioneller politischer Aktivität bei diesen beiden Subgruppen (.14 und .28). Ähnliche Ergebnisse zeigten sich für nichtkonventionelle politische Aktivitäten (.25 und .26).

Bei der Subgruppe der Konservativen zeigte sich ein ganz anderes Bild. Sowohl Kontrollideologie als auch Persönliche Kontrolle sind in dieser Gruppe signifikant mit konventioneller politischer Aktivität verknüpft (.20 und .27), nicht jedoch mit unkonventioneller. Offensichtlich fassen die Konservativen, im Gegensatz zu Schwarzen und Liberalen, Kontrollideologie und Persönliche Kontrolle nicht als unterschiedliche Konstrukte auf. Bei den Konservativen steigt also mit zunehmender interner Kontrollenerwartung das Ausmaß konventioneller politischer Aktivität. Bei Liberalen und Schwarzen hängt dagegen das Ausmaß politischer Aktivität nur von der Kontrollideologie ab: Mit steigender externer Kontrollideologie steigt bei ihnen das Ausmaß sowohl konventioneller als auch unkonventioneller politischer Aktivität an.

Wir finden also, daß eine Wechselwirkung zwischen der politischen Ausrichtung von Personen und der Kontrollenerwartung hinsichtlich der Voraussage politischen Verhaltens besteht. Zum zweiten wirkt externe Kontrollenerwartung im Sinne subjektiver Machtlosigkeit (Dimension: Persönliche Kontrolle) ganz anders auf politisches Verhalten als Kontrollideologie im Sinne wahrgenommener sozialer Ursachen für Ungleichheit etc. Eine globale Analyse, die nicht mehr nach Dimensionen und Subgruppen differenziert, würde somit zumindest zu stark verzerrten Koeffizienten, wenn nicht zu Scheinwiderlegungen und Scheinbestätigungen von Hypothesen führen.

Der Wert der Reliabilitäts- und Validitätsbestimmungen in Bezug auf Populationen und Subskalen wird generell durch drei Faktoren beeinträchtigt, die bei der Interpretation beachtet werden müssen. Erstens wurde in den genannten Untersuchungen ohne empirische Prüfungen die Annahme orthogonaler Fakto-

---

ren gemacht. MUTHEN, der diese Annahme geprüft hat, erhielt Korrelationen zwischen den Faktoren, die zwischen .20 und .40 lagen (1978). Zum zweiten wurden durchweg standardisierte Reliabilitäts- und Validitätskoeffizienten verwendet. Bei diesen können jedoch unterschiedliche Werte nicht nur durch unterschiedliche Kovarianzen, sondern auch durch unterschiedliche Streuungen in den verschiedenen Stichproben verursacht sein. Daher müßten für den Vergleich unstandardisierte Koeffizienten verwendet werden (vgl. zur entsprechenden Methodik JURESKOG, 1971; SORBOM, 1974). Drittens wurden faktorenanalytische Verfahren angewandt, die Intervallskalenniveau voraussetzen. Es ist nicht zu erwarten, daß eine speziell für dichotome Items geeignete Faktorenanalyse dieselben Ergebnisse erbringen würde.

#### 4. Integratives Modell

Auf der Basis der theoretischen und methodischen Überlegungen, die bis jetzt angestellt wurden, könnte für die empirische Weiterarbeit an der I-E-Skala ein Vorgehen angestrebt werden, bei dem inhaltliche Hypothesen und Gesichtspunkte der Messung in ein gemeinsames Meßmodell eingehen, das sich konfirmatorisch überprüfen läßt. Die Skizze eines solchen möglichen Modells sei abschließend dargestellt.

Wenn man davon ausgeht, daß mindestens drei Indikatoren pro Konstrukt notwendig sind, um alle Arten zufälliger und systematischer Meßfehler erfassen zu können (COSTNER, 1969), dann wären allein 24 Items zur Messung aller in Schema 2 aufgeführten Konstrukte notwendig. Allerdings wären damit erst die generalisierten Kontrollerwartungen als allgemeines Persönlichkeitsmerkmal berücksichtigt. Unerfaßt blieben Kontrollerwartungen in spezifischen Bereichen wie z.B. Bürokratie oder Politik. In dieser Beschränkung ist aber immerhin denkbar, daß sich leicht eine entsprechende Anzahl von Items finden oder formulieren läßt, die die Felder A - H in Schema 2 abdecken. Nachfolgend haben wir für jeden der acht Bereiche ein denkbare Beispiel-Item aufgeführt. Die Items sind z.T. Übersetzungen der ursprünglichen Rotter-Skala und finden sich auch im Nationalen Sozialen Survey (siehe dazu auch S. 46f. dieser ZUMANACHRICHTEN):

A) 'Was in meinem Leben geschieht, bestimme ich selbst.'

B) 'Erfolg hängt von harter Arbeit ab, Glück hat damit wenig zu tun.'

---

# ZUMA

---

- C) 'Meine Mißerfolge sind darauf zurückzuführen, daß mir für bestimmte Dinge einfach die Begabung fehlt.'
- D) 'Wer begabt ist, wird auf die Dauer auch Erfolg haben.'
- E) 'Ob ich meine Ziele erreiche oder nicht, hängt von den politischen Umständen ab.'
- F) 'Was man im Leben bekommt, hängt gar nicht so sehr von den eigenen Anstrengungen ab, sondern von der Wirtschaftslage, der Lage auf dem Arbeitsmarkt, den Tarifabschlüssen und den Sozialleistungen des Staates.'
- G) 'Ich habe oft das Gefühl, nicht beeinflussen zu können, was mit mir geschieht.'
- H) 'Die meisten Menschen sind sich nicht bewußt, wie sehr ihr Leben von Zufällen beeinflußt wird.'

Eine Itemauswahl auf der Grundlage des theoretischen Gefüges von Schema 2 führt mit größerer Wahrscheinlichkeit in der Auswertung zu einer übersichtlichen Faktorenstruktur, die eine konfirmatorische Validierung aussichtsreich macht, als bei einer bunten Mischung heterogener Items. Falls man aus Zeit- oder anderen Gründen gezwungen ist, weniger Items als statistisch wünschenswert zu verwenden, sollte man sich auf wenige wichtige Bereiche beschränken. Umgekehrt ist es aber für spezifische Themenstellungen wünschenswert, neben den generalisierten Erwartungen auch solche für bestimmte einzelne Lebensbereiche zu operationalisieren, z.B. für Bürokratie, Politik, interpersonelle Beziehungen usw.

Neben diesen inhaltlichen Komponenten, die durch die Theorie RÖTTERSs vorgegeben sind, sollte ein zu überprüfendes Meßmodell allerdings auch die methodischen Einflußfaktoren - Response Sets, Nichtentsprechung intern und extern formulierter Items, Kontrollfaktoren für Erfolg vs. Kontrollfaktoren für Mißerfolg etc. - beinhalten. Verglichen mit der ursprünglichen Theorie RÖTTERSs erhalten wir damit eine wesentlich komplexere Theorie und Meßtheorie. Abb. 3 gibt ein Pfaddiagramm als Beispiel für eine komplexe Kausalstruktur für die Itemgruppe A, B, E und F (beeinflussbare Kontrollfaktoren) in Schema 2, einmal für bereichsunspezifische generalisierte Controllerwartungen, zum anderen für Controllerwartungen im Bereich Bürokratie.

Die Zahl der notwendigen Indikatoren ist letzten Endes davon abhängig, wieviele Konstrukte (latente Variablen) im Modell berücksichtigt werden sol-

---

len. Dabei muß man zwischen fünf Arten von Konstrukten unterscheiden:

1. Den Variablen, die mit den Items der Skala gemessen werden sollen; in unserem Fall entsprechen diese den Feldern von Schema 2, wobei zunächst nicht zwischen intern/extern formulierten Items unterschieden wird. In Abb. 3 sind dies die Faktoren erster Ordnung, Persönliche Kontrolle und Kontrollideologie.
2. Den nicht durch eigene Indikatoren gemessenen Faktoren zweiter Ordnung.
3. Den durch Indikatoren gemessenen Methodenfaktoren Soziale Wünschbarkeit und Wahrgenommene Beeinflußbarkeit der Kontrollfaktoren (eigene Anstrengung, gesellschaftliche Machtverhältnisse).
4. Den nicht durch eigene Indikatoren gemessenen Methodenfaktoren. Dabei wird in Abb. 3 nur der Faktor "Interne vs. externe Formulierung des Statements" durch die Indizierung der Indikatoren mit e bzw. i berücksichtigt. Daneben sind möglicherweise jedoch auch die Effekte der Faktoren Acquiescence und Erfolg/Mißerfolg von Bedeutung.
5. Zusätzlichen, im theoretischen Modell enthaltenen Variablen, z.B. Bürokratierfahrung oder zu erklärende Verhaltensvariablen, die als Feld möglicher Vorhersagen auf der Basis bereinigter Kontrollwartungsfaktoren interessant sind. Entsprechende Variablen sind natürlich untersuchungsspezifisch.

Die Zahl der notwendigen Indikatoren ergibt sich aus der Zahl der zu schätzenden statistischen Parameter: Sie muß mindestens so groß sein, daß das mathematische Gleichungssystem lösbar ("identifiziert") ist. Da die genaue Beschreibung des Identifikationsgrades des Modells relativ kompliziert ist, wurde hier der Einfachheit halber davon ausgegangen, daß je vier Indikatoren für die Faktoren erster Ordnung und je drei für die übrigen durch direkte Indikatoren gemessenen Faktoren genügen.

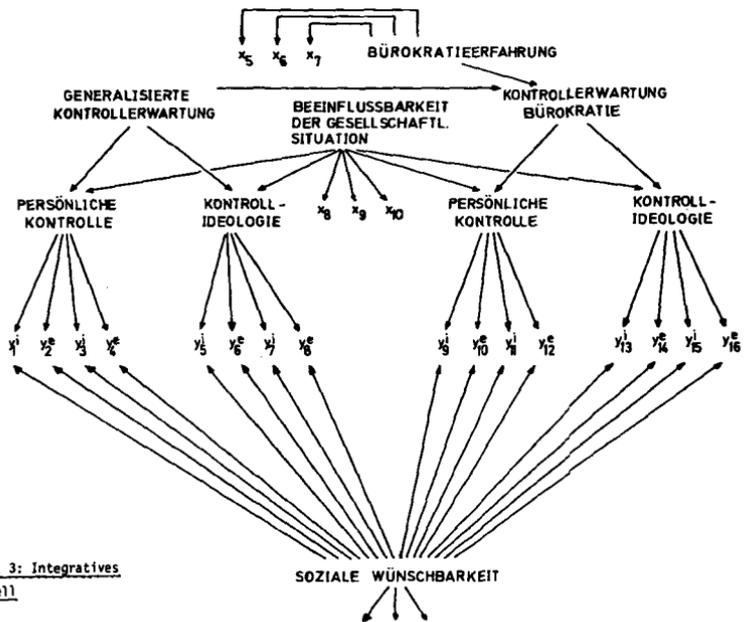


Abb. 3: Integratives  
Modell

Arbeiten in Zusammenhang mit dem Skalenhandbuch und allgemeinen Skalierungsproblemen werden bei ZUMA von Ursula Hoffmann-Lange, Peter Schmidt und Bernd Wegener betreut.

BECK, U., BRATER, M. & WEGENER, B., Berufswahl und Berufszuweisung. Zur sozialen Verwandtschaft von Ausbildungsberufen, Frankfurt: Campus 1979.

CHERLIN, A. & BOURQUE, L.B., Dimensionality and reliability of the Rotter I-E-scale. Sociometry, 37, 1974, 565-582.

COLLINS, B.E., Four components of the Rotter internal-external scale: Belief in a difficult world, a just world, a predictable world and a politically responsive world. Journal of Personality and Social Psychology, 29, 1974, 381-391.

COSTNER, H.L., Theory, deduction and rules of correspondence. American Journal of Sociology, 75, 1969, 245-263.

CROWNE, D.P. & MARLOWE, E.D., New scale of social desirability independent of psychopathology. In: J.B. Rotter, I.E. Chance & E.J. Phares (Hrsg.), Applications of a social learning theory of personality, New York: Holt, Rinehart and Winston, 1972.

FISCHER, J., Generalisierte Kontrollerwartungen. Theoretische Analyse und Entwicklung eines Fragebogens. Bericht aus dem SFB 22, Erlangen-Nürnberg 1976.

FISHBEIN, M. & AJZEN, I., Belief, attitude, intention and behavior, Reading, Mass.: Addison-Wesley, 1975.

GURIN, P., GURIN, G., LAO, R.C. & BEATTICE, M., Internal-external control in the motivational dynamics of Negro youth. Journal of Social Issues, 25, 1969, 29-53.

GURIN, P., GURIN, G., & MORRISON, B. M., Personal and ideological aspects of internal and external control. Social Psychology, 41, 1978, 275-296.

JÖRESKOG, K.G., Statistical analysis of sets of congeneric tests, Psychometrika, 36, 1971, 109-133.

JÖRESKOG, K.G., Simultaneous factor analysis in several populations. Psychometrika, 36, 1971, 409-426.

JÖRESKOG, K.G., Analyzing psychological data by structural analysis of cova-

---

- riance matrices. In: D.H.Krantz, R.C. Atkinson, R.D. Luce & P. Suppes (Hrsg.), Contemporary developments in mathematical psychology, Vol. II, San Francisco: W.H. Freeman and Co., 1974.
- JÖRESKOG, K.G. & SÖRBOM, D., LISREL IV, User's Guide, National Educational Resources, Inc., Chicago, 1978.
- KAUFMANN-MALL, K., Kognitiv-hedonistische Theorie menschlichen Verhaltens, Bern: H. Huber, 1978.
- MUTHEN, B., Contributions to factor analysis of dichotomous variables. Psychometrika, 43, 1978.
- OPP, K.D., Theorie sozialer Krisen, Hamburg: Hoffmann und Campe, 1978.
- PHARES, E.J., locus of control in personality, Morristown N.J.: General Learning Press, 1976.
- ROSENBERG, M.S., An Analysis of affective-cognitive consistency. In: C.I. Hovland & M.J. Rosenberg (Hrsg.), Attitude organization and change, New Haven: Yale University Press, 1960, 15-64.
- ROST-SCHAUDE, E., KUMPF, M. & FREY, D., Untersuchungen zu einer deutschen Fassung der Internal-External Control Skala von Rotter. In: Bericht über den 29. Kongress der Deutschen Gesellschaft für Psychologie, Hogrefe, 1978, 327-329.
- ROST-SCHAUDE, E., KUMPF, M. & FREY, D., Auswirkungen des Verlusts kognizierter Kontrolle bei unterschiedlichen Ursachenattributionen, Bericht aus dem SFB 24, Mannheim, 1979.
- ROTTER, J.B., Social learning and clinical psychology, Englewood Cliffs, N.Y.: Prentice Hall, 1954.
- ROTTER, J.B., Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 80, 1966, (whole No. 609); wieder abgedruckt in: J.B. Rotter, J.E. Chance & E.J. Phares (Hrsg.), Applications of a social learning theory of personality, New York: Holt, Rinehart and Winston, 1972, 260-294.
- ROTTER, J.B., CHANCE, J.E. & PHARES, E.J., Applications of a social learning theory of personality, New York: Holt, Rinehart and Winston, 1972.
- ROTTER, J.B., Some problems and misconceptions related to the construct of internal versus external control of reinforcement. Journal of Consulting and Clinical Psychology, 43, 1975, 56-67.
- SIMON, H.U., Machtlosigkeit von Studenten: Ein Beitrag über sozialpsychologische Machttheorien und ihre empirische Überprüfung. Diss., Mann-
-

heim 1979.

- SORBOM, D., A general method for studying differences in factor means and factor structure between groups. British Journal of Mathematical and Statistical Psychology, 27, 1974, 229-239.
- WEEDE, E. & JAGODZINSKI, W., Einführung in die konfirmatorische Faktorenanalyse. Zeitschrift für Soziologie, 6, 1977, 315-333.
- WEGENER, B., Magnitude-Messung beruflicher Einstellungen. In: U. BECK, M. BRÄTER & B. WEGENER, Berufswahl und Berufszuweisung. Zur sozialen Verwandtschaft von Ausbildungsberufen, Frankfurt: Campus 1979, 125-162.
- WEGENER, B. (Hrsg.), Social attitudes and psychophysical measurement, Hillsdale, N.J.: Erlbaum 1980 (im Druck).